Storage Foundation Cluster File System High Availability 8.0 Configuration and Upgrade Guide - Linux

VERITAS™

Last updated: 2021-12-28

Legal Notice

Copyright © 2021 Veritas Technologies LLC. All rights reserved.

Veritas and the Veritas Logo are trademarks or registered trademarks of Veritas Technologies LLC or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.

This product may contain third-party software for which Veritas is required to provide attribution to the third-party (“Third-Party Programs”). Some of the Third-Party Programs are available under open source or free software licenses. The License Agreement accompanying the Software does not alter any rights or obligations you may have under those open source or free software licenses. Refer to the third-party legal notices document accompanying this Veritas product or available at:
https://www.veritas.com/about/legal/license-agreements

The product described in this document is distributed under licenses restricting its use, copying, distribution, and decompilation/reverse engineering. No part of this document may be reproduced in any form by any means without prior written authorization of Veritas Technologies LLC and its licensors, if any.

THE DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID. VERITAS TECHNOLOGIES LLC SHALL NOT BE LIABLE FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH THE FURNISHING, PERFORMANCE, OR USE OF THIS DOCUMENTATION. THE INFORMATION CONTAINED IN THIS DOCUMENTATION IS SUBJECT TO CHANGE WITHOUT NOTICE.

The Licensed Software and Documentation are deemed to be commercial computer software as defined in FAR 12.212 and subject to restricted rights as defined in FAR Section 52.227-19 "Commercial Computer Software - Restricted Rights" and DFARS 227.7202, et seq. "Commercial Computer Software and Commercial Computer Software Documentation," as applicable, and any successor regulations, whether delivered by Veritas as on premises or hosted services. Any use, modification, reproduction release, performance, display or disclosure of the Licensed Software and Documentation by the U.S. Government shall be solely in accordance with the terms of this Agreement.

Veritas Technologies LLC
2625 Augustine Drive
Santa Clara, CA 95054
http://www.veritas.com

Technical Support

Technical Support maintains support centers globally. All support services will be delivered in accordance with your support agreement and the then-current enterprise technical support policies. For information about our support offerings and how to contact Technical Support, visit our website:
https://www.veritas.com/support

You can manage your Veritas account information at the following URL:
https://my.veritas.com

If you have questions regarding an existing support agreement, please email the support agreement administration team for your region as follows:

Worldwide (except Japan)	CustomerCare@veritas.com
Japan	CustomerCare_Japan@veritas.com

Documentation

Make sure that you have the current version of the documentation. Each document displays the date of the last update on page 2. The latest documentation is available on the Veritas website:
https://sort.veritas.com/documents

Documentation feedback

Your feedback is important to us. Suggest improvements or report errors or omissions to the documentation. Include the document title, document version, chapter title, and section title of the text on which you are reporting. Send feedback to:
infoscaledocs@veritas.com

You can also see documentation information or ask a question on the Veritas community site:
http://www.veritas.com/community/

Veritas Services and Operations Readiness Tools (SORT)

Veritas Services and Operations Readiness Tools (SORT) is a website that provides information and tools to automate and simplify certain time-consuming administrative tasks. Depending on the product, SORT helps you prepare for installations and upgrades, identify risks in your datacenters, and improve operational efficiency. To see what services and tools SORT provides for your product, see the data sheet:
https://sort.veritas.com/data/support/SORT_Data_Sheet.pdf

Contents

Section 1 Introduction to SFCFSHA 15

Chapter 1 Introducing Storage Foundation Cluster File System High Availability 16

About this document 16
About Storage Foundation Cluster File System High Availability 17
About Veritas InfoScale Operations Manager 17
About I/O fencing 18
About Veritas Services and Operations Readiness Tools (SORT) 19
About configuring SFCFSHA clusters for data integrity 20
About I/O fencing for Storage Foundation Cluster File System High Availability in virtual machines that do not support SCSI-3 PR 21
About I/O fencing components 21

Section 2 Configuration of SFCFSHA 24

Chapter 2 Preparing to configure 25

I/O fencing requirements 25
Coordinator disk requirements for I/O fencing 25
CP server requirements 26
Non-SCSI-3 I/O fencing requirements 28

Chapter 3 Preparing to configure SFCFSHA clusters for data integrity 30

About planning to configure I/O fencing 30
Typical SFCFSHA cluster configuration with server-based I/O fencing 34
Recommended CP server configurations 35
Setting up the CP server 38
Planning your CP server setup 38
Installing the CP server using the installer 39
Configuring the CP server cluster in secure mode 40

Setting up shared storage for the CP server database 41
Configuring the CP server using the installer program 41
Configuring the CP server manually 50
Verifying the CP server configuration 55

Chapter 4 Configuring SFCFSHA 57

Overview of tasks to configure SFCFSHA using the product installer 58
Starting the software configuration 58
Specifying systems for configuration 59
Configuring the cluster name 60
Configuring private heartbeat links 60
Configuring the virtual IP of the cluster 67
Configuring SFCFSHA in secure mode 69
Configuring a secure cluster node by node 70
Configuring the first node 70
Configuring the remaining nodes 71
Completing the secure cluster configuration 72
Adding VCS users 74
Configuring SMTP email notification 75
Configuring SNMP trap notification 77
Configuring global clusters 78
Completing the SFCFSHA configuration 79
Verifying the NIC configuration 80
About Veritas License Audit Tool 80
Verifying and updating licenses on the system 81
Checking licensing information on the system 81
Replacing a SFCFSHA keyless license with another keyless license 81
Replacing a SFCFSHA keyless license with a permanent license 82
Configuring SFDB 83

Chapter 5 Configuring SFCFSHA clusters for data integrity 84

Setting up disk-based I/O fencing using installer 84
Configuring disk-based I/O fencing using installer 84
Initializing disks as VxVM disks 87
Checking shared disks for I/O fencing 88
Refreshing keys or registrations on the existing coordination points for disk-based fencing using the installer 92
Setting up server-based I/O fencing using installer 93

Refreshing keys or registrations on the existing coordination points for server-based fencing using the installer 101
Setting the order of existing coordination points for server-based fencing using the installer 103
Setting up non-SCSI-3 I/O fencing in virtual environments using installer 106
Setting up majority-based I/O fencing using installer 108
Enabling or disabling the preferred fencing policy 110

Chapter 6 Performing an automated SFCFSHA configuration using response files 113

Configuring SFCFSHA using response files 113
Response file variables to configure SFCFSHA 114
Sample response file for SFCFSHA configuration 124

Chapter 7 Performing an automated I/O fencing configuration using response files 126

Configuring I/O fencing using response files 126
Response file variables to configure disk-based I/O fencing 127
Sample response file for configuring disk-based I/O fencing 130
Configuring CP server using response files 130
Response file variables to configure CP server 131
Sample response file for configuring the CP server on single node VCS cluster 133
Sample response file for configuring the CP server on SFHA cluster 133
Response file variables to configure server-based I/O fencing 134
Sample response file for configuring server-based I/O fencing 136
Response file variables to configure non-SCSI-3 I/O fencing 137
Sample response file for configuring non-SCSI-3 I/O fencing 138
Response file variables to configure majority-based I/O fencing 139
Sample response file for configuring majority-based I/O fencing 140

Chapter 8 Manually configuring SFCFSHA clusters for data integrity 141

Setting up disk-based I/O fencing manually 141
Identifying disks to use as coordinator disks 142
Setting up coordinator disk groups 142
Creating I/O fencing configuration files 143
Modifying VCS configuration to use I/O fencing 144
Verifying I/O fencing configuration 146

Setting up server-based I/O fencing manually 146
Preparing the CP servers manually for use by the SFCFSHA cluster 147
Generating the client key and certificates manually on the client nodes 149
Configuring server-based fencing on the SFCFSHA cluster manually 151
Configuring CoordPoint agent to monitor coordination points 157
Verifying server-based I/O fencing configuration 159
Setting up non-SCSI-3 fencing in virtual environments manually 160
Sample /etc/vxfenmode file for non-SCSI-3 fencing 162
Setting up majority-based I/O fencing manually 166
Creating I/O fencing configuration files 166
Modifying VCS configuration to use I/O fencing 166
Verifying I/O fencing configuration 168

Section 3 Upgrade of SFCFSHA 170

Chapter 9 Planning to upgrade SFCFSHA 171

About the upgrade 171
Supported upgrade paths 173
Transitioning between the Veritas products 177
Considerations for upgrading SFCFSHA to 8.0 on systems configured with an Oracle resource 177
Preparing to upgrade SFCFSHA 178
Getting ready for the upgrade 178
Creating backups 180
Determining if the root disk is encapsulated 181
Pre-upgrade planning when VVR is configured 181
Preparing to upgrade VVR when VCS agents are configured 184
Upgrading the array support 187
Considerations for upgrading REST server 188
Using Install Bundles to simultaneously install or upgrade full releases (base, maintenance, rolling patch), and individual patches 188

Chapter 10 Performing a full upgrade of SFCFSHA using the installer 191

Performing a full upgrade using the product installer 191
Ensuring the file systems are clean 191
Performing the upgrade 192
Upgrading the operating system 198

Upgrading SFDB 199

Chapter 11 Performing a rolling upgrade of SFCFSHA 200

About rolling upgrade 200
Performing a rolling upgrade of SFCFSHA using the product installer 201
Performing a rolling upgrade of SFCFSHA from 7.4.2 or later to 8.0 205
Performing rolling upgrades in CVR environments 207

Chapter 12 Performing a phased upgrade of SFCFSHA 210

About phased upgrade 210
Prerequisites for a phased upgrade 210
Planning for a phased upgrade 211
Phased upgrade limitations 211
Phased upgrade example 211
Phased upgrade example overview 212
Performing a phased upgrade using the product installer 213
Moving the service groups to the second subcluster 213
Upgrading the operating system on the first subcluster 215
Upgrading the SFCFSHA stack on the first subcluster 215
Preparing the second subcluster 216
Activating the first subcluster 217
Upgrading the operating system on the second subcluster 218
Upgrading the second subcluster 219
Completing the phased upgrade 219

Chapter 13 Performing an automated SFCFSHA upgrade using response files 221

Upgrading SFCFSHA using response files 221
Response file variables to upgrade SFCFSHA 222
Sample response file for full upgrade of SFCFSHA 225
Sample response file for rolling upgrade of SFCFSHA 226
Sample response file for rolling upgrade of SFCFSHA from 7.4.2 or later to 8.0 226

Chapter 14 Upgrading Volume Replicator 227

Upgrading Volume Replicator 227
Upgrading VVR without disrupting replication 227
Upgrading VVR sites for InfoScale 7.3.1 or earlier 228

Upgrading VVR sites with InfoScale 7.4 or later 229
Upgrading VVR sites in VCS control for InfoScale 7.3.1 or earlier 230
Upgrading VVR sites in VCS control for InfoScale 7.4 or later 231
Post-upgrade tasks for VVR sites 233

Chapter 15 Upgrading VirtualStore 234

Supported upgrade paths 234
Upgrading SVS to SFCFSHA 8.0 234

Chapter 16 Performing post-upgrade tasks 236

Resetting DAS disk names to include host name in FSS environments 236
Re-joining the backup boot disk group into the current disk group 237
Reverting to the backup boot disk group after an unsuccessful upgrade 237
CVM master node needs to assume the logowner role for VCS managed VVR resources 238

Section 4 Post-configuration tasks 239

Chapter 17 Performing post configuration tasks 240

Upgrading disk layout versions 240
Switching on Quotas 241
About enabling LDAP authentication for clusters that run in secure mode 242
Enabling LDAP authentication for clusters that run in secure mode 243
About configuring authentication for SFDB tools 247
Configuring vxdbd for SFDB tools authentication 248
Enabling data encryption over wire 248

Section 5 Configuration of disaster recovery environments 251

Chapter 18 Configuring disaster recovery environments 252

Disaster recovery options for SFCFSHA 252
About setting up a campus cluster for disaster recovery 253

About setting up a global cluster environment for SFCFSHA 255
About configuring a parallel global cluster using Volume Replicator (VVR) for replication 256

Section 6 Adding and removing nodes 258

Chapter 19 Adding a node to SFCFSHA clusters 259

About adding a node to a cluster 259
Before adding a node to a cluster 260
Adding a node to a cluster using the Veritas InfoScale installer 263
Adding the node to a cluster manually 266
Starting Veritas Volume Manager (VxVM) on the new node 267
Configuring cluster processes on the new node 268
Setting up the node to run in secure mode 269
Starting fencing on the new node 273
After adding the new node 273
Configuring Cluster Volume Manager (CVM) and Cluster File System (CFS) on the new node 274
Configuring the ClusterService group for the new node 275
Adding a node using response files 276
Response file variables to add a node to a SFCFSHA cluster 277
Sample response file for adding a node to a SFCFSHA cluster 277
Configuring server-based fencing on the new node 278
Adding the new node to the vxfen service group 278
Adding nodes to a cluster that is using authentication for SFDB tools 279
Updating the Storage Foundation for Databases (SFDB) repository after adding a node 280
Sample configuration file for adding a node to the cluster 281

Chapter 20 Removing a node from SFCFSHA clusters 285

About removing a node from a cluster 285
Removing a node from a cluster 286
Modifying the VCS configuration files on existing nodes 287
Modifying the Cluster Volume Manager (CVM) configuration on the existing nodes to remove references to the deleted node 290
Removing the node configuration from the CP server 290
Removing security credentials from the leaving node 291

Updating the Storage Foundation for Databases (SFDB) repository after removing a node 291
Sample configuration file for removing a node from the cluster 292

Section 7 Configuration and Upgrade reference 295

Appendix A Installation scripts 296

Installation script options 296
About using the postcheck option 301

Appendix B Configuration files 304

About the LLT and GAB configuration files 304
About the AMF configuration files 307
About I/O fencing configuration files 308
Sample configuration files for CP server 311
Sample main.cf file for CP server hosted on a single node that runs VCS 312
Sample main.cf file for CP server hosted on a two-node SFHA cluster 314
Sample CP server configuration (/etc/vxcps.conf) file output 317

Appendix C Configuring the secure shell or the remote shell for communications 318

About configuring secure shell or remote shell communication modes before installing products 318
Manually configuring passwordless ssh 319
Setting up ssh and rsh connection using the installer -comsetup command 322
Setting up ssh and rsh connection using the pwdutil.pl utility 324
Restarting the ssh session 327
Enabling rsh for Linux 327

Appendix D High availability agent information 329

About agents 329
VCS agents included within SFCFSHA 330
Enabling and disabling intelligent resource monitoring for agents manually 330
Administering the AMF kernel driver 333

CVMCluster agent 334
Entry points for CVMCluster agent 334
Attribute definition for CVMCluster agent 334
CVMCluster agent type definition 335
CVMCluster agent sample configuration 336
CVMVxconfigd agent 336
Entry points for CVMVxconfigd agent 337
Attribute definition for CVMVxconfigd agent 337
CVMVxconfigd agent type definition 339
CVMVxconfigd agent sample configuration 339
CVMVolDg agent 339
Entry points for CVMVolDg agent 339
Attribute definition for CVMVolDg agent 340
CVMVolDg agent type definition 342
CVMVolDg agent sample configuration 342
CFSMount agent 343
Entry points for CFSMount agent 343
Attribute definition for CFSMount agent 344
CFSMount agent type definition 346
CFSMount agent sample configuration 347
CFSfsckd agent 347
Entry points for CFSfsckd agent 347
Attribute definition for CFSfsckd agent 348
CFSfsckd agent type definition 349
CFSfsckd agent sample configuration 350

Appendix E Sample SFCFSHA cluster setup diagrams for CP server-based I/O fencing 351

Configuration diagrams for setting up server-based I/O fencing 351
Two unique client clusters served by 3 CP servers 351
Client cluster served by highly available CPS and 2 SCSI-3 disks 352
Two node campus cluster served by remote CP server and 2 SCSI-3 disks 354
Multiple client clusters served by highly available CP server and 2 SCSI-3 disks 356

Appendix F Configuring LLT over UDP 358

Using the UDP layer for LLT 358
When to use LLT over UDP 358
Manually configuring LLT over UDP using IPv4 358
Broadcast address in the /etc/lltab file 359

The link command in the /etc/lltab file 360
The set-addr command in the /etc/lltab file 360
Selecting UDP ports 361
Configuring the netmask for LLT 361
Configuring the broadcast address for LLT 362
Sample configuration: direct-attached links 362
Sample configuration: links crossing IP routers 364
Using the UDP layer of IPv6 for LLT 365
When to use LLT over UDP 365
Manually configuring LLT over UDP using IPv6 366
The link command in the /etc/lltab file 366
The set-addr command in the /etc/lltab file 367
Selecting UDP ports 367
Sample configuration: direct-attached links 368
Sample configuration: links crossing IP routers 369
About configuring LLT over UDP multiport 371
Manually configuring LLT over UDP multiport 371
Enabling LLT ports in firewall 373
Disabling the UDP multiport feature 374

Appendix G Using LLT over RDMA 375

Using LLT over RDMA 375
About RDMA over RoCE or InfiniBand networks in a clustering environment 375
How LLT supports RDMA capability for faster interconnects between applications 376
Using LLT over RDMA: supported use cases 377
Configuring LLT over RDMA 377
Choosing supported hardware for LLT over RDMA 378
Installing RDMA, InfiniBand or Ethernet drivers and utilities 379
Configuring RDMA over an Ethernet network 380
Configuring RDMA over an InfiniBand network 382
Tuning system performance 386
Manually configuring LLT over RDMA 388
LLT over RDMA sample /etc/lltab 392
Verifying LLT configuration 392
Troubleshooting LLT over RDMA 393
IP addresses associated to the RDMA NICs do not automatically plumb on node restart 393
Ping test fails for the IP addresses configured over InfiniBand interfaces 394

After a node restart, by default the Mellanox card with Virtual Protocol Interconnect (VPI) gets configured in InfiniBand mode 394
The LLT module fails to start 394

Section 1

Introduction to SFCFSHA

- Chapter 1. Introducing Storage Foundation Cluster File System High Availability

Chapter 1

Introducing Storage Foundation Cluster File System High Availability

This chapter includes the following topics:

- About this document
- About Storage Foundation Cluster File System High Availability
- About Veritas InfoScale Operations Manager
- About I/O fencing
- About Veritas Services and Operations Readiness Tools (SORT)
- About configuring SFCFSHA clusters for data integrity

About this document

This document is applicable to both Storage Foundation Cluster File System High Availability (SFCFSHA) and Storage Foundation Cluster File System (SFCFS).

Note: The commands used for the Red Hat Enterprise Linux (RHEL) operating system in this document also apply to supported RHEL-compatible distributions.

About Storage Foundation Cluster File System High Availability

Storage Foundation Cluster File System High Availability (SFCFSHA) extends Storage Foundation to support shared data in a storage area network (SAN) environment. Using SFCFSHA, multiple servers can concurrently access shared storage and files transparently to applications.

SFCFSHA also provides increased automation and intelligent management of availability and performance.

SFCFSHA includes Cluster Server, which adds high availability functionality to the product.

The Veritas File Replicator feature can also be licensed with this product.

For information on high availability environments, read the Cluster Server documentation.

About Veritas InfoScale Operations Manager

Veritas InfoScale Operations Manager provides a centralized management console for Veritas InfoScale products. You can use Veritas InfoScale Operations Manager to monitor, visualize, and manage storage resources and generate reports.

Veritas recommends using Veritas InfoScale Operations Manager to manage Storage Foundation and Cluster Server environments.

You can download Veritas InfoScale Operations Manager from:

https://www.veritas.com/content/support/en_US/downloads

Refer to the Veritas InfoScale Operations Manager documentation for installation, upgrade, and configuration instructions.

The Veritas Enterprise Administrator (VEA) console is no longer packaged with Veritas InfoScale products. If you want to continue using VEA, a software version is available for download from:

https://www.veritas.com/form/trialware/vcs-utilities

Storage Foundation Management Server is deprecated.

If you want to manage a single cluster using Cluster Manager (Java Console), a version is available for download from:

https://www.veritas.com/form/trialware/vcs-utilities

You cannot manage the new features of this release using the Java Console. Cluster Server Management Console is deprecated.

About I/O fencing

I/O fencing protects the data on shared disks when nodes in a cluster detect a change in the cluster membership that indicates a split-brain condition.

The fencing operation determines the following:

- The nodes that must retain access to the shared storage
- The nodes that must be ejected from the cluster

This decision prevents possible data corruption. The installer installs the I/O fencing driver, part of VRTSvxfen RPM, when you install Veritas InfoScale Enterprise. To protect data on shared disks, you must configure I/O fencing after you install Veritas InfoScale Enterprise and configure Storage Foundation Cluster File System High Availability.

I/O fencing modes - disk-based and server-based I/O fencing - use coordination points for arbitration in the event of a network partition. Whereas, majority-based I/O fencing mode does not use coordination points for arbitration. With majority-based I/O fencing you may experience loss of high availability in some cases. You can configure disk-based, server-based, or majority-based I/O fencing:

Disk-based I/O fencing	I/O fencing that uses coordinator disks is referred to as disk-based I/O fencing. Disk-based I/O fencing ensures data integrity in a single cluster.
Server-based I/O fencing	I/O fencing that uses at least one CP server system is referred to as server-based I/O fencing. Server-based fencing can include only CP servers, or a mix of CP servers and coordinator disks. Server-based I/O fencing ensures data integrity in clusters. In virtualized environments that do not support SCSI-3 PR, Storage Foundation Cluster File System High Availability supports non-SCSI-3 I/O fencing. See "About I/O fencing for Storage Foundation Cluster File System High Availability in virtual machines that do not support SCSI-3 PR" on page 21.

Majority-based I/O fencing	Majority-based I/O fencing mode does not need coordination points to provide protection against data corruption and data consistency in a clustered environment. Use majority-based I/O fencing when there are no additional servers and or shared SCSI-3 disks to be used as coordination points.

See " About planning to configure I/O fencing" on page 30.

Note: Veritas recommends that you use I/O fencing to protect your cluster against split-brain situations.

See the *Storage Foundation Cluster File System High Availability Administrator's Guide*.

About Veritas Services and Operations Readiness Tools (SORT)

Veritas Services and Operations Readiness Tools (SORT) is a Web site that automates and simplifies some of the most time-consuming administrative tasks. SORT helps you manage your datacenter more efficiently and get the most out of your Veritas products.

SORT can help you do the following:

Prepare for your next installation or upgrade	■ List product installation and upgrade requirements, including operating system versions, memory, disk space, and architecture. ■ Analyze systems to determine if they are ready to install or upgrade Veritas products. ■ Download the latest patches, documentation, and high availability agents from a central repository. ■ Access up-to-date compatibility lists for hardware, software, databases, and operating systems.
Manage risks	■ Get automatic email notifications about changes to patches, array-specific modules (ASLs/APMs/DDIs/DDLs), and high availability agents from a central repository. ■ Identify and mitigate system and environmental risks. ■ Display descriptions and solutions for hundreds of Veritas error codes.

Improve efficiency	■ Find and download patches based on product version and platform. ■ List installed Veritas products and license keys. ■ Tune and optimize your environment.

Note: Certain features of SORT are not available for all products. Access to SORT is available at no extra cost.

To access SORT, go to:

https://sort.veritas.com

About configuring SFCFSHA clusters for data integrity

When a node fails, SFCFSHA takes corrective action and configures its components to reflect the altered membership. If an actual node failure did not occur and if the symptoms were identical to those of a failed node, then such corrective action would cause a split-brain situation.

Some example scenarios that can cause such split-brain situations are as follows:

- Broken set of private networks
  If a system in a two-node cluster fails, the system stops sending heartbeats over the private interconnects. The remaining node then takes corrective action. The failure of the private interconnects, instead of the actual nodes, presents identical symptoms and causes each node to determine its peer has departed. This situation typically results in data corruption because both nodes try to take control of data storage in an uncoordinated manner.
- System that appears to have a system-hang
  If a system is so busy that it appears to stop responding, the other nodes could declare it as dead. This declaration may also occur for the nodes that use the hardware that supports a "break" and "resume" function. When a node drops to PROM level with a break and subsequently resumes operations, the other nodes may declare the system dead. They can declare it dead even if the system later returns and begins write operations.

I/O fencing is a feature that prevents data corruption in the event of a communication breakdown in a cluster. SFCFSHA uses I/O fencing to remove the risk that is associated with split-brain. I/O fencing allows write access for members of the active cluster. It blocks access to storage from non-members so that even a node that is alive is unable to cause damage.

After you install Veritas InfoScale Enterprise and configure SFCFSHA, you must configure I/O fencing in SFCFSHA to ensure data integrity.

See " About planning to configure I/O fencing" on page 30.

About I/O fencing for Storage Foundation Cluster File System High Availability in virtual machines that do not support SCSI-3 PR

In a traditional I/O fencing implementation, where the coordination points are coordination point servers (CP servers) or coordinator disks, Clustered Volume Manager (CVM) and Veritas I/O fencing modules provide SCSI-3 persistent reservation (SCSI-3 PR) based protection on the data disks. This SCSI-3 PR protection ensures that the I/O operations from the losing node cannot reach a disk that the surviving sub-cluster has already taken over.

See the *Cluster Server Administrator's Guide* for more information on how I/O fencing works.

In virtualized environments that do not support SCSI-3 PR, Storage Foundation Cluster File System High Availability attempts to provide reasonable safety for the data disks. Storage Foundation Cluster File System High Availability requires you to configure non-SCSI-3 I/O fencing in such environments. Non-SCSI-3 fencing either uses server-based I/O fencing with only CP servers as coordination points or majority-based I/O fencing, which does not use coordination points, along with some additional configuration changes to support such environments.

See "Setting up non-SCSI-3 I/O fencing in virtual environments using installer" on page 106.

See "Setting up non-SCSI-3 fencing in virtual environments manually" on page 160.

About I/O fencing components

The shared storage for SFCFSHA must support SCSI-3 persistent reservations to enable I/O fencing. SFCFSHA involves two types of shared storage:

- Data disks—Store shared data
  See "About data disks" on page 21.
- Coordination points—Act as a global lock during membership changes
  See "About coordination points" on page 22.

About data disks

Data disks are standard disk devices for data storage and are either physical disks or RAID Logical Units (LUNs).

These disks must support SCSI-3 PR and must be part of standard VxVM or CVM disk groups. CVM is responsible for fencing data disks on a disk group basis. Disks that are added to a disk group and new paths that are discovered for a device are automatically fenced.

About coordination points

Coordination points provide a lock mechanism to determine which nodes get to fence off data drives from other nodes. A node must eject a peer from the coordination points before it can fence the peer from the data drives. SFCFSHA prevents split-brain when vxfen races for control of the coordination points and the winner partition fences the ejected nodes from accessing the data disks.

Note: Typically, a fencing configuration for a cluster must have three coordination points. Veritas also supports server-based fencing with a single CP server as its only coordination point with a caveat that this CP server becomes a single point of failure.

The coordination points can either be disks or servers or both.

- Coordinator disks
  Disks that act as coordination points are called coordinator disks. Coordinator disks are three standard disks or LUNs set aside for I/O fencing during cluster reconfiguration. Coordinator disks do not serve any other storage purpose in the SFCFSHA configuration.
  You can configure coordinator disks to use Veritas Volume Manager's Dynamic Multi-pathing (DMP) feature. Dynamic Multi-pathing (DMP) allows coordinator disks to take advantage of the path failover and the dynamic adding and removal capabilities of DMP. So, you can configure I/O fencing to use DMP devices. I/O fencing uses SCSI-3 disk policy that is dmp-based on the disk device that you use.

  Note: The dmp disk policy for I/O fencing supports both single and multiple hardware paths from a node to the coordinator disks. If few coordinator disks have multiple hardware paths and few have a single hardware path, then we support only the dmp disk policy. For new installations, Veritas only supports dmp disk policy for IO fencing even for a single hardware path.

  See the *Storage Foundation Administrator's Guide*.
- Coordination point servers

The coordination point server (CP server) is a software solution which runs on a remote system or cluster. CP server provides arbitration functionality by allowing the SFHA cluster nodes to perform the following tasks:

- Self-register to become a member of an active SFCFSHA cluster (registered with CP server) with access to the data drives
- Check which other nodes are registered as members of this active SFCFSHA cluster
- Self-unregister from this active SFCFSHA cluster
- Forcefully unregister other nodes (preempt) as members of this active SFCFSHA cluster

In short, the CP server functions as another arbitration mechanism that integrates within the existing I/O fencing module.

Note: With the CP server, the fencing arbitration logic still remains on the SFCFSHA cluster.

Multiple SFCFSHA clusters running different operating systems can simultaneously access the CP server. TCP/IP based communication is used between the CP server and the SFCFSHA clusters.

About preferred fencing

The I/O fencing driver uses coordination points to prevent split-brain in a VCS cluster. By default, the fencing driver favors the subcluster with maximum number of nodes during the race for coordination points. With the preferred fencing feature, you can specify how the fencing driver must determine the surviving subcluster.

You can configure the preferred fencing policy using the cluster-level attribute PreferredFencingPolicy for the following:

- Enable system-based preferred fencing policy to give preference to high capacity systems.
- Enable group-based preferred fencing policy to give preference to service groups for high priority applications.
- Enable site-based preferred fencing policy to give preference to sites with higher priority.
- Disable preferred fencing policy to use the default node count-based race policy.

See the *Storage Foundation Cluster File System High Availability Administrator's Guide* for more details.

See "Enabling or disabling the preferred fencing policy" on page 110.

Section 2

Configuration of SFCFSHA

- Chapter 2. Preparing to configure
- Chapter 3. Preparing to configure SFCFSHA clusters for data integrity
- Chapter 4. Configuring SFCFSHA
- Chapter 5. Configuring SFCFSHA clusters for data integrity
- Chapter 6. Performing an automated SFCFSHA configuration using response files
- Chapter 7. Performing an automated I/O fencing configuration using response files
- Chapter 8. Manually configuring SFCFSHA clusters for data integrity

Chapter 2

Preparing to configure

This chapter includes the following topics:

- I/O fencing requirements

I/O fencing requirements

Depending on whether you plan to configure disk-based fencing or server-based fencing, make sure that you meet the requirements for coordination points:

- Coordinator disks
  See "Coordinator disk requirements for I/O fencing" on page 25.
- CP servers
  See "CP server requirements" on page 26.

If you have installed Veritas InfoScale Enterprise in a virtual environment that is not SCSI-3 PR compliant, review the requirements to configure non-SCSI-3 fencing.

See "Non-SCSI-3 I/O fencing requirements" on page 28.

Coordinator disk requirements for I/O fencing

Make sure that the I/O fencing coordinator disks meet the following requirements:

- For disk-based I/O fencing, you must have at least three coordinator disks or there must be odd number of coordinator disks.
- The coordinator disks must be DMP devices.
- Each of the coordinator disks must use a physically separate disk or LUN. Veritas recommends using the smallest possible LUNs for coordinator disks.
- Each of the coordinator disks should exist on a different disk array, if possible.
- The coordinator disks must support SCSI-3 persistent reservations.

- Coordinator devices can be attached over iSCSI protocol but they must be DMP devices and must support SCSI-3 persistent reservations.
- Veritas recommends using hardware-based mirroring for coordinator disks.
- Coordinator disks must not be used to store data or must not be included in disk groups that store user data.
- Coordinator disks cannot be the special devices that array vendors use. For example, you cannot use EMC gatekeeper devices as coordinator disks.
- The coordinator disk size must be at least 128 MB.

CP server requirements

Storage Foundation Cluster File System High Availability 8.0 clusters (application clusters) support coordination point servers (CP servers) that are hosted on the following VCS and SFHA versions:

- VCS 6.2.1 or later single-node cluster
- SFHA 6.2.1 or later cluster

Upgrade considerations for CP servers

- Upgrade VCS or SFHA on CP servers to version 8.0 if the current release version is prior to version 6.2.1.
- You do not need to upgrade CP servers to version 8.0 if the release version is 6.2.1 or later.
- CP servers on version 6.2.1 or later support HTTPS-based communication with application clusters on version 6.2.1 or later.
- CP servers on version 6.2.1 to 7.0 support IPM-based communication with application clusters on versions before 6.2.1.
- You need to configure VIPs for HTTPS-based communication if release version of application clusters is 6.2.1 or later.

Make sure that you meet the basic hardware requirements for the VCS/SFHA cluster to host the CP server.

See the *Veritas InfoScale Installation Guide*.

Note: While Veritas recommends at least three coordination points for fencing, a single CP server as coordination point is a supported server-based fencing configuration. Such single CP server fencing configuration requires that the coordination point be a highly available CP server that is hosted on an SFHA cluster.

Make sure you meet the following additional CP server requirements which are covered in this section before you install and configure CP server:

- Hardware requirements
- Operating system requirements
- Networking requirements (and recommendations)
- Security requirements

Table 2-1 lists additional requirements for hosting the CP server.

Table 2-1 CP server hardware requirements

Hardware required	Description
Disk space	To host the CP server on a VCS cluster or SFHA cluster, each host requires the following file system space: ■ 550 MB in the /opt directory (additionally, the language pack requires another 15 MB) ■ 300 MB in /usr ■ 20 MB in /var ■ 10 MB in /etc (for the CP server database)
Storage	When CP server is hosted on an SFHA cluster, there must be shared storage between the nodes of this SFHA cluster.
RAM	Each CP server requires at least 512 MB.
Network	Network hardware capable of providing TCP/IP connection between CP servers and SFCFSHA clusters (application clusters).

Table 2-2 displays the CP server supported operating systems and versions. An application cluster can use a CP server that runs any of the following supported operating systems.

Table 2-2 CP server supported operating systems and versions

CP server	Operating system and version
CP server hosted on a VCS single-node cluster or on an SFHA cluster	CP server supports any of the following operating systems: ■ Linux: ■ RHEL 7.9, 8.2, 8.4 ■ OL 7.9, 8.2, 8.4 ■ SLES 12, 15 For the list of operating systems that CP server supports, refer to the Software Compatibility List (SCL). Review other details such as supported operating system levels and architecture for the supported operating systems. See the *Veritas InfoScale Release Notes* for that platform.

Following are the CP server networking requirements and recommendations:

- Veritas recommends that network access from the application clusters to the CP servers should be made highly-available and redundant. The network connections require either a secure LAN or VPN.
- The CP server uses the TCP/IP protocol to connect to and communicate with the application clusters by these network paths. The CP server listens for messages from the application clusters using TCP port 443 if the communication happens over the HTTPS protocol. TCP port 443 is the default port that can be changed while you configure the CP server.
  Veritas recommends that you configure multiple network paths to access a CP server. If a network path fails, CP server does not require a restart and continues to listen on all the other available virtual IP addresses.
- When placing the CP servers within a specific network configuration, you must take into consideration the number of hops from the different application cluster nodes to the CP servers. As a best practice, Veritas recommends that the number of hops and network latency from the different application cluster nodes to the CP servers should be equal. This ensures that if an event occurs that results in an I/O fencing scenario, there is no bias in the race due to difference in number of hops or network latency between the CPS and various nodes.

For information about establishing secure communications between the application cluster and CP server, see the *Storage Foundation Cluster File System High Availability Administrator's Guide*.

Non-SCSI-3 I/O fencing requirements

Supported virtual environment for non-SCSI-3 fencing:

- VMware Server ESX 4.0, 5.0, 5.1, and 5.5 on AMD Opteron or Intel Xeon EM64T (x86_64)

  Guest operating system: See the *Veritas InfoScale Release Notes* for the list of supported Linux operating systems.

Make sure that you also meet the following requirements to configure fencing in the virtual environments that do not support SCSI-3 PR:

- Storage Foundation Cluster File System High Availability must be configured with Cluster attribute UseFence set to SCSI3
- For server-based I/O fencing, all coordination points must be CP servers

Chapter 3

Preparing to configure SFCFSHA clusters for data integrity

This chapter includes the following topics:

- About planning to configure I/O fencing
- Setting up the CP server

About planning to configure I/O fencing

After you configure SFCFSHA with the installer, you must configure I/O fencing in the cluster for data integrity. Application clusters on release version 8.0 (HTTPS-based communication) only support CP servers on release version 6.1 and later.

You can configure disk-based I/O fencing, server-based I/O fencing, or majority-based I/O fencing. If your enterprise setup has multiple clusters that use VCS for clustering, Veritas recommends you to configure server-based I/O fencing.

The coordination points in server-based fencing can include only CP servers or a mix of CP servers and coordinator disks.

Veritas also supports server-based fencing with a single coordination point which is a single highly available CP server that is hosted on an SFHA cluster.

Warning: For server-based fencing configurations that use a single coordination point (CP server), the coordination point becomes a single point of failure. In such configurations, the arbitration facility is not available during a failover of the CP server in the SFHA cluster. So, if a network partition occurs on any application cluster during the CP server failover, the application cluster is brought down. Veritas recommends the use of single CP server-based fencing only in test environments.

You use majority fencing mechanism if you do not want to use coordination points to protect your cluster. Veritas recommends that you configure I/O fencing in majority mode if you have a smaller cluster environment and you do not want to invest additional disks or servers for the purposes of configuring fencing.

Note: Majority-based I/O fencing is not as robust as server-based or disk-based I/O fencing in terms of high availability. With majority-based fencing mode, in rare cases, the cluster might become unavailable.

If you have installed Storage Foundation Cluster File System High Availability in a virtual environment that is not SCSI-3 PR compliant, you can configure non-SCSI-3 fencing.

See Figure 3-2 on page 33.

Figure 3-1 illustrates a high-level flowchart to configure I/O fencing for the Storage Foundation Cluster File System High Availability cluster.

Figure 3-1 Workflow to configure I/O fencing

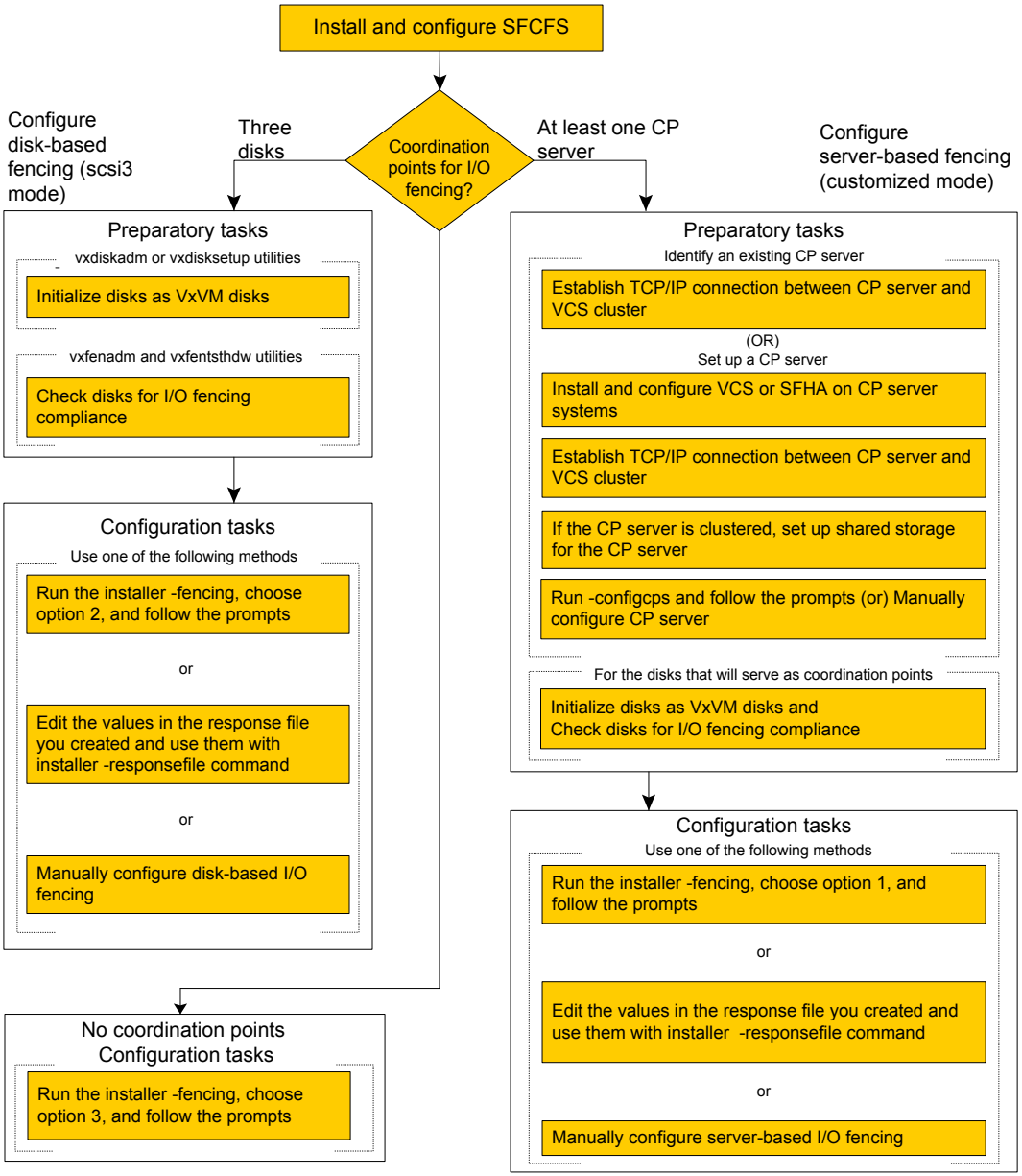
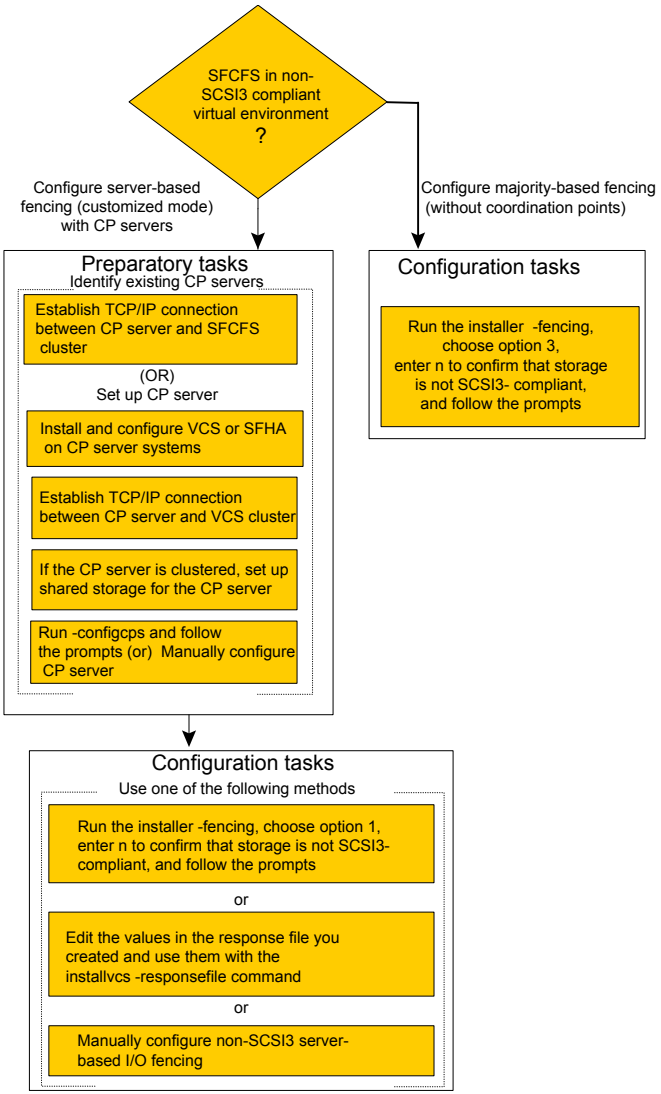


Figure 3-2 illustrates a high-level flowchart to configure non-SCSI-3 I/O fencing for the Storage Foundation Cluster File System High Availability cluster in virtual environments that do not support SCSI-3 PR.

Figure 3-2 Workflow to configure non-SCSI-3 I/O fencing



After you perform the preparatory tasks, you can use any of the following methods to configure I/O fencing:

Using the installer	See "Setting up disk-based I/O fencing using installer" on page 84. See "Setting up server-based I/O fencing using installer" on page 93. See "Setting up non-SCSI-3 I/O fencing in virtual environments using installer" on page 106. See "Setting up majority-based I/O fencing using installer" on page 108.
Using response files	See "Response file variables to configure disk-based I/O fencing" on page 127. See "Response file variables to configure server-based I/O fencing" on page 134. See "Response file variables to configure non-SCSI-3 I/O fencing" on page 137. See "Response file variables to configure majority-based I/O fencing" on page 139. See "Configuring I/O fencing using response files" on page 126.
Manually editing configuration files	See "Setting up disk-based I/O fencing manually" on page 141. See "Setting up server-based I/O fencing manually" on page 146. See "Setting up non-SCSI-3 fencing in virtual environments manually" on page 160. See "Setting up majority-based I/O fencing manually " on page 166.

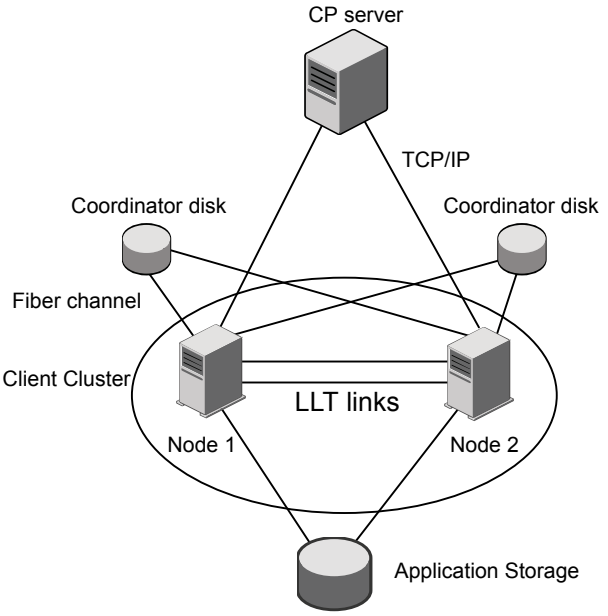
You can also migrate from one I/O fencing configuration to another.

See the *Storage foundation High Availability Administrator's Guide* for more details.

Typical SFCFSHA cluster configuration with server-based I/O fencing

Figure 3-3 displays a configuration using a SFCFSHA cluster (with two nodes), a single CP server, and two coordinator disks. The nodes within the SFCFSHA cluster are connected to and communicate with each other using LLT links.

Figure 3-3 CP server, SFCFSHA cluster, and coordinator disks



Recommended CP server configurations

Following are the recommended CP server configurations:

- Multiple application clusters use three CP servers as their coordination points
  See Figure 3-4 on page 36.
- Multiple application clusters use a single CP server and single or multiple pairs of coordinator disks (two) as their coordination points
  See Figure 3-5 on page 37.
- Multiple application clusters use a single CP server as their coordination point
  This single coordination point fencing configuration must use a highly available CP server that is configured on an SFHA cluster as its coordination point.
  See Figure 3-6 on page 37.

Warning: In a single CP server fencing configuration, arbitration facility is not available during a failover of the CP server in the SFHA cluster. So, if a network partition occurs on any application cluster during the CP server failover, the application cluster is brought down.

Although the recommended CP server configurations use three coordination points, you can use more than three coordination points for I/O fencing. Ensure that the total number of coordination points you use is an odd number. In a configuration where multiple application clusters share a common set of CP server coordination points, the application cluster as well as the CP server use a Universally Unique Identifier (UUID) to uniquely identify an application cluster.

Figure 3-4 displays a configuration using three CP servers that are connected to multiple application clusters.

Figure 3-4 Three CP servers connecting to multiple application clusters

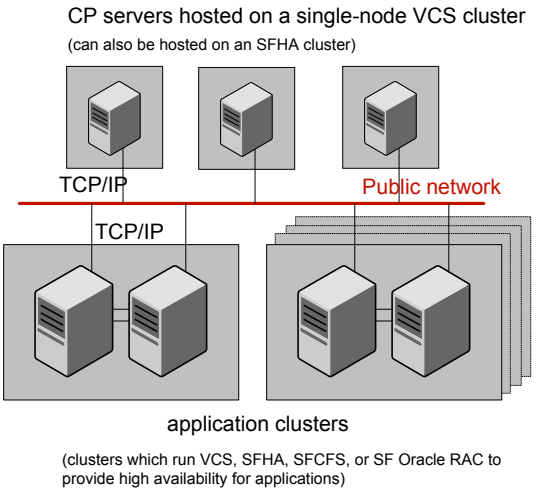


Figure 3-5 displays a configuration using a single CP server that is connected to multiple application clusters with each application cluster also using two coordinator disks.

Figure 3-5 Single CP server with two coordinator disks for each application cluster

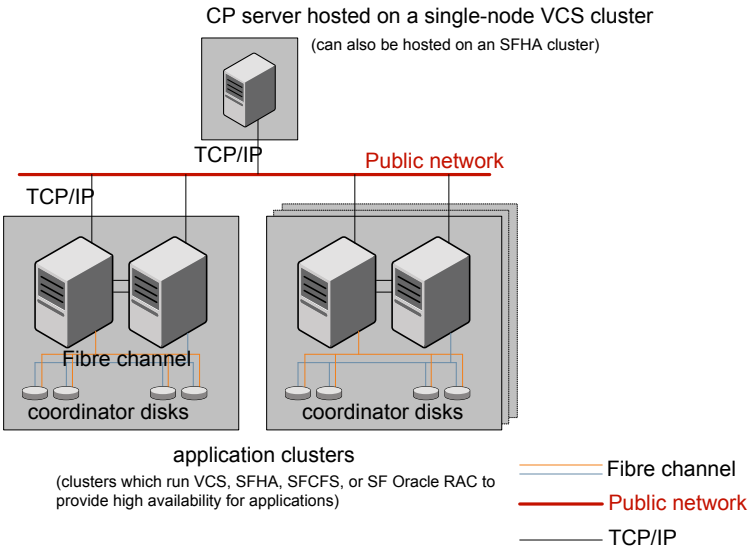
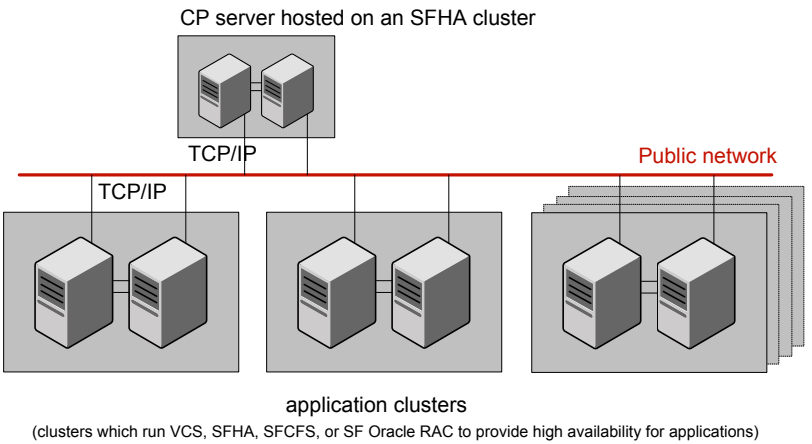


Figure 3-6 displays a configuration using a single CP server that is connected to multiple application clusters.

Figure 3-6 Single CP server connecting to multiple application clusters



See "Configuration diagrams for setting up server-based I/O fencing" on page 351.

Setting up the CP server

Table 3-1 lists the tasks to set up the CP server for server-based I/O fencing.

Table 3-1 Tasks to set up CP server for server-based I/O fencing

Task	Reference
Plan your CP server setup	See "Planning your CP server setup" on page 38.
Install the CP server	See "Installing the CP server using the installer" on page 39.
Set up shared storage for the CP server database	See "Setting up shared storage for the CP server database" on page 41.
Configure the CP server	See " Configuring the CP server using the installer program" on page 41. See "Configuring the CP server manually" on page 50. See "Configuring CP server using response files" on page 130.
Verify the CP server configuration	See "Verifying the CP server configuration" on page 55.

Planning your CP server setup

Follow the planning instructions to set up CP server for server-based I/O fencing.

To plan your CP server setup

1. Decide whether you want to host the CP server on a single-node VCS cluster, or on an SFHA cluster.

   Veritas recommends hosting the CP server on an SFHA cluster to make the CP server highly available.

2. If you host the CP server on an SFHA cluster, review the following information. Make sure you make the decisions and meet these prerequisites when you set up the CP server:

   - You must set up shared storage for the CP server database during your CP server setup.
   - Decide whether you want to configure server-based fencing for the SFCFSHA cluster (application cluster) with a single CP server as coordination point or with at least three coordination points.

Veritas recommends using at least three coordination points.

3 Set up the hardware and network for your CP server.

See "CP server requirements" on page 26.

4 Have the following information handy for CP server configuration:

- Name for the CP server
  The CP server name should not contain any special characters. CP server name can include alphanumeric characters, underscore, and hyphen.
- Port number for the CP server
  Allocate a TCP/IP port for use by the CP server.
  Valid port range is between 49152 and 65535. The default port number for HTTPS-based communication is 443.
- Virtual IP address, network interface, netmask, and networkhosts for the CP server
  You can configure multiple virtual IP addresses for the CP server.

Installing the CP server using the installer

Perform the following procedure to install Veritas InfoScale Enterprise and configure VCS or SFHA on CP server systems.

To install Veritas InfoScale Enterprise and configure VCS or SFHA on the CP server systems

- Depending on whether your CP server uses a single system or multiple systems, perform the following tasks:

CP server setup uses a single system	Install Veritas InfoScale Enterprise or Veritas InfoScale Availability and configure VCS to create a single-node VCS cluster. See the *Veritas InfoScale Installation Guide* for instructions on CP server installation. See the *Cluster Server Configuration and Upgrade Guide* for configuring VCS. Proceed to configure the CP server. See " Configuring the CP server using the installer program" on page 41. See "Configuring the CP server manually" on page 50.
CP server setup uses multiple systems	Install Veritas InfoScale Enterprise and configure SFHA to create an SFHA cluster. This makes the CP server highly available. See the *Veritas InfoScale Installation Guide* for instructions on installing SFHA. See the *Storage Foundation and High Availability Configuration and Upgrade Guide* for configuring SFHA. Proceed to set up shared storage for the CP server database.

Configuring the CP server cluster in secure mode

You must configure security on the CP server only if you want IPM-based (Veritas Product Authentication Service) secure communication between the CP server and the SFHA cluster (CP server clients). However, IPM-based communication enables the CP server to support application clusters till InfoSale release 7.0.

This step secures the HAD communication on the CP server cluster.

Note: If you already configured the CP server cluster in secure mode during the VCS configuration, then skip this section.

To configure the CP server cluster in secure mode

- Run the installer as follows to configure the CP server cluster in secure mode.

```
# /opt/VRTS/install/installer  -security
```

Setting up shared storage for the CP server database

If you configured SFHA on the CP server cluster, perform the following procedure to set up shared storage for the CP server database.

The installer can set up shared storage for the CP server database when you configure CP server for the SFHA cluster.

Veritas recommends that you create a mirrored volume for the CP server database and that you use the VxFS file system type.

To set up shared storage for the CP server database

1. Create a disk group containing the disks. You require two disks to create a mirrored volume.

   For example:

   ```
   # vxdg init cps_dg  disk1 disk2
   ```

2. Create a mirrored volume over the disk group.

   For example:

   ```
   # vxassist -g cps_dg make cps_vol volume_size layout=mirror
   ```

3. Create a file system over the volume.

   The CP server configuration utility only supports vxfs file system type. If you use an alternate file system, then you must configure CP server manually.

   Depending on the operating system that your CP server runs, enter the following command:

   Linux
   ```
   # mkfs -t vxfs /dev/vx/rdsk/cps_dg/cps_volume
   ```

Configuring the CP server using the installer program

Use the configcps option available in the installer program to configure the CP server.

Perform one of the following procedures:

For CP servers on single-node VCS cluster: See "To configure the CP server on a single-node VCS cluster" on page 42.

For CP servers on an SFHA cluster:

See "To configure the CP server on an SFHA cluster" on page 46.

To configure the CP server on a single-node VCS cluster

1. Verify that the `VRTScps` RPM is installed on the node.
2. Run the installer program with the configcps option.

   ```
   # /opt/VRTS/install/installer  -configcps
   ```

3. Installer checks the cluster information and prompts if you want to configure CP Server on the cluster.

   Enter **y** to confirm.

4. Select an option based on how you want to configure Coordination Point server.

   ```
   1) Configure Coordination Point Server on single node VCS system
   2) Configure Coordination Point Server on SFHA cluster
   3) Unconfigure Coordination Point Server
   ```

5. Enter the option: [1-3,q] **1**.

   The installer then runs the following preconfiguration checks:

   - Checks to see if a single-node VCS cluster is running with the supported platform.
     The CP server requires VCS to be installed and configured before its configuration.

   The installer automatically installs a license that is identified as a CP server-specific license. It is installed even if a VCS license exists on the node. CP server-specific key ensures that you do not need to use a VCS license on the single-node. It also ensures that Veritas Operations Manager (VOM) identifies the license on a single-node coordination point server as a CP server-specific license and not as a VCS license.

6. Restart the VCS engine if the single-node only has a CP server-specific license.

   ```
   A single node coordination point server will be configured and
   VCS will be started in one node mode, do you want to
   continue?  [y,n,q]  (y)
   ```

7 Communication between the CP server and application clusters is secured by using the HTTPS protocol from release 6.1.0 onwards.

Enter the name of the CP Server.

```
Enter the name of the CP Server: [b]   cps1
```

8 Enter valid virtual IP addresses for the CP Server with HTTPS-based secure communication. A CP Server can be configured with more than one virtual IP address.

```
Enter Virtual IP(s) for the CP server for HTTPS,
separated by a space:  [b]  10.200.58.231 10.200.58.232
10.200.58.233
```

Note: Ensure that the virtual IP address of the CP server and the IP address of the NIC interface on the CP server belongs to the same subnet of the IP network. This is required for communication to happen between client nodes and CP server.

---

9 Enter the corresponding CP server port number for each virtual IP address or press **Enter** to accept the default value (443).

```
Enter the default port '443' to be used for all the
virtual IP addresses for HTTPS communication or assign the
corresponding port number in the range [49152, 65535] for
each virtual IP address. Ensure that each port number is
separated by a single
space: [b]  (443) 54442 54443 54447
```

10 Enter the absolute path of the CP server database or press **Enter** to accept the default value (/etc/VRTScps/db).

```
Enter absolute path of the database: [b] (/etc/VRTScps/db)
```

11 Verify and confirm the CP server configuration information.

```
CP Server configuration verification:
----------------------------------------------
CP Server Name:  cps1
CP Server Virtual IP(s) for HTTPS: 10.200.58.231, 10.200.58.232,
10.200.58.233
CP Server Port(s) for HTTPS: 54442, 54443, 54447
CP Server Database Dir: /etc/VRTScps/db

----------------------------------------------

Is this information correct? [y,n,q,?]  (y)
```

12 The installer proceeds with the configuration process, and creates a vxcps.conf configuration file.

```
Successfully generated the /etc/vxcps.conf configuration file
Successfully created directory /etc/VRTScps/db on node
```

13 Configure the CP Server Service Group (CPSSG) for this cluster.

```
Enter how many NIC resources you want to configure (1 to 2): 2
```

Answer the following questions for each NIC resource that you want to configure.

14 Enter a valid network interface for the virtual IP address for the CP server process.

```
Enter a valid network interface on sys1 for NIC resource - 1: eth0
Enter a valid network interface on sys1 for NIC resource - 2:  eth1
```

15 Enter the NIC resource you want to associate with the virtual IP addresses.

```
Enter the NIC resource you want to associate with the virtual IP 10.200.58.231 (1 to 2): 1
Enter the NIC resource you want to associate with the virtual IP 10.200.58.232 (1 to 2): 2
```

16 Enter the networkhosts information for each NIC resource.

```
Veritas recommends configuring NetworkHosts attribute to ensure NIC resource
to be always online

Do you want to add NetworkHosts attribute for the NIC device eth0
on system sys1? [y,n,q] y
Enter a valid IP address to configure NetworkHosts for NIC eth0
on system sys1: 10.200.56.22

Do you want to add another Network Host? [y,n,q] n
```

17 Enter the netmask for virtual IP addresses. If you entered an IPv6 address, enter the prefix details at the prompt.

```
Enter the netmask for virtual IP for
HTTPS 192.169.0.220: (255.255.252.0)
```

18 Installer displays the status of the Coordination Point Server configuration. After the configuration process has completed, a success message appears.

```
For example:
Updating main.cf with CPSSG service group.. Done
Successfully added the CPSSG service group to VCS configuration.
Trying to bring CPSSG service group
ONLINE and will wait for upto 120 seconds

The Veritas coordination point server is ONLINE

The Veritas coordination point server has
been configured on your system.
```

19 Run the `hagrp -state` command to ensure that the CPSSG service group has been added.

```
For example:
# hagrp -state CPSSG
#Group Attribute System Value
CPSSG State.... |ONLINE|
```

It also generates the configuration file for CP server (/etc/vxcps.conf). The vxcpserv process and other resources are added to the VCS configuration in the CP server service group (CPSSG).

For information about the CPSSG, refer to the *Cluster Server Administrator's Guide*.

To configure the CP server on an SFHA cluster

1. Verify that the `VRTScps` RPM is installed on each node.
2. Ensure that you have configured passwordless ssh or rsh on the CP server cluster nodes.
3. Run the installer program with the configcps option.

   ```
   # ./installer -configcps
   ```

4. Specify the systems on which you need to configure the CP server.
5. Installer checks the cluster information and prompts if you want to configure CP Server on the cluster.

   Enter **y** to confirm.
6. Select an option based on how you want to configure Coordination Point server.

   ```
   1)  Configure Coordination Point Server on single node VCS system
   2)  Configure Coordination Point Server on SFHA cluster
   3)  Unconfigure Coordination Point Server
   ```

7. Enter **2** at the prompt to configure CP server on an SFHA cluster.

   The installer then runs the following preconfiguration checks:

   - Checks to see if an SFHA cluster is running with the supported platform. The CP server requires SFHA to be installed and configured before its configuration.
8. Communication between the CP server and application clusters is secured by HTTPS from Release 6.1.0 onwards.

   Enter the name of the CP server.

   ```
   Enter the name of the CP Server: [b]  cps1
   ```

9. Enter valid virtual IP addresses for the CP Server. A CP Server can be configured with more than one virtual IP address.

   ```
   Enter Virtual IP(s) for the CP server for HTTPS,
    separated by a space: [b] 10.200.58.231 10.200.58.232 10.200.58.233
   ```

10 Enter the corresponding CP server port number for each virtual IP address or press Enter to accept the default value (443).

```
Enter the default port '443' to be used for all the virtual IP addresses
for HTTPS communication or assign the corresponding port number in the range [49152,
65535] for each virtual IP address. Ensure that each port number is separated by
a single space: [b] (443) 65535 65534 65537
```

11 Enter absolute path of the database.

```
CP Server uses an internal database to store the client information.
As the CP Server is being configured on SFHA cluster, the database should reside
on shared storage with vxfs file system. Please refer to documentation for
information on setting up of shared storage for CP server database.
Enter absolute path of the database: [b] /cpsdb
```

12 Verify and confirm the CP server configuration information.

```
CP Server configuration verification:

CP Server Name: cps1
CP Server Virtual IP(s) for HTTPS: 10.200.58.231, 10.200.58.232,
10.200.58.233
CP Server Port(s) for HTTPS: 65535, 65534, 65537
CP Server Database Dir: /cpsdb

Is this information correct? [y,n,q,?] (y)
```

13 The installer proceeds with the configuration process, and creates a vxcps.conf configuration file.

```
Successfully generated the /etc/vxcps.conf configuration file
Copying configuration file /etc/vxcps.conf to sys0....Done
Creating mount point /cps_mount_data on sys0. ... Done
Copying configuration file /etc/vxcps.conf to sys0. ... Done
Press Enter to continue.
```

14 Configure CP Server Service Group (CPSSG) for this cluster.

```
Enter how many NIC resources you want to configure (1 to 2): 2

Answer the following questions for each NIC resource that you want to configure.
```

15 Enter a valid network interface for the virtual IP address for the CP server process.

```
Enter a valid network interface on sys1 for NIC resource - 1: eth0
Enter a valid network interface on sys1 for NIC resource - 2: eth1
```

16 Enter the NIC resource you want to associate with the virtual IP addresses.

```
Enter the NIC resource you want to associate with the virtual IP 10.200.58.231 (1 to 2): 1
Enter the NIC resource you want to associate with the virtual IP 10.200.58.232 (1 to 2): 2
```

17 Enter the networkhosts information for each NIC resource.

```
Veritas recommends configuring NetworkHosts attribute to ensure NIC resource
to be always online


Do you want to add NetworkHosts attribute for the NIC device eth0
on system sys1? [y,n,q] y
Enter a valid IP address to configure NetworkHosts for NIC eth0
on system sys1: 10.200.56.22


Do you want to add another Network Host? [y,n,q] n
Do you want to apply the same NetworkHosts for all systems? [y,n,q] (y)
```

18 Enter the netmask for virtual IP addresses. If you entered an IPv6 address, enter the prefix details at the prompt.

```
Enter the netmask for virtual IP for
HTTPS 192.168.0.111: (255.255.252.0)
```

19 Configure a disk group for CP server database. You can choose an existing disk group or create a new disk group.

```
Veritas recommends to use the disk group that has at least
two disks on which mirrored volume can be created.
Select one of the options below for CP Server database disk group:

1)  Create a new disk group
2)  Using an existing disk group

Enter the choice for a disk group: [1-2,q]  2
```

20 Select one disk group as the CP Server database disk group.

```
Select one disk group as CP Server database disk group: [1-3,q] 3
1)  mycpsdg
2)  cpsdg1
3)  newcpsdg
```

21 Select the CP Server database volume.

You can choose to use an existing volume or create new volume for CP Server database. If you chose newly created disk group, you can only choose to create new volume for CP Server database.

```
Select one of the options below for CP Server database volume:
 1)  Create a new volume on disk group newcpsdg
 2)  Using an existing volume on disk group newcpsdg
```

22 Enter the choice for a volume: [1-2,q] **2**.

23 Select one volume as CP Server database volume [1-1,q] **1**

```
1) newcpsvol
```

24 After the VCS configuration files are updated, a success message appears.

```
For example:
Updating main.cf with CPSSG service group .... Done
Successfully added the CPSSG service group to VCS configuration.
```

25 If the cluster is secure, installer creates the softlink `/var/VRTSvcs/vcsauth/data/CPSERVER` to `/cpsdb/CPSERVER` and check if credentials are already present at `/cpsdb/CPSERVER`. If not, installer creates credentials in the directory, otherwise, installer asks if you want to reuse exsting credentials.

```
Do you want to reuse these credentials? [y,n,q] (y)
```

26 After the configuration process has completed, a success message appears.

```
For example:
Trying to bring CPSSG service group ONLINE and will wait for upto 120 seconds
The Veritas Coordination Point Server is ONLINE
The Veritas Coordination Point Server has been configured on your system.
```

27 Run the `hagrp -state` command to ensure that the CPSSG service group has been added.

```
For example:
# hagrp -state CPSSG
#Group Attribute System Value
CPSSG State cps1 |ONLINE|
CPSSG State cps2 |OFFLINE|
```

It also generates the configuration file for CP server (`/etc/vxcps.conf`). The vxcpserv process and other resources are added to the VCS configuration in the CP server service group (CPSSG).

For information about the CPSSG, refer to the *Cluster Server Administrator's Guide*.

Configuring the CP server manually

Perform the following steps to manually configure the CP server.

You need to manually generate certificates for the CP server and its client nodes to configure the CP server for HTTPS-based communication.

Table 3-2 Tasks to configure the CP server manually

Task	Reference
Configure CP server manually for HTTPS-communication	See "Configuring the CP server manually for HTTPS-based communication" on page 51. See "Generating the key and certificates manually for the CP server" on page 52. See "Completing the CP server configuration" on page 55.

Note: If a CP server should support pure IPv6 communication, use only IPv6 addresses in the `/etc/vxcps.conf` file. If the CP server should support both IPv6 and IPv4 communications, use both IPv6 and IPv4 addresses in the configuration file.

Configuring the CP server manually for HTTPS-based communication

Perform the following steps to manually configure the CP server in HTTPS-based mode.

To manually configure the CP server

1. Stop VCS on each node in the CP server cluster using the following command:

   ```
   # hastop -local
   ```

2. Edit the `main.cf` file to add the CPSSG service group on any node. Use the CPSSG service group in the sample main.cf as an example:

   See "Sample configuration files for CP server" on page 311.

   Customize the resources under the CPSSG service group as per your configuration.

3. Verify the `main.cf` file using the following command:

   ```
   # hacf -verify /etc/VRTSvcs/conf/config
   ```

   If successfully verified, copy this main.cf to all other cluster nodes.

4. Create the `/etc/vxcps.conf` file using the sample configuration file provided at `/etc/vxcps/vxcps.conf.sample`.

   Veritas recommends enabling security for communication between CP server and the application clusters.

   If you configured the CP server in HTTPS mode, do the following:

- Edit the `/etc/vxcps.conf` file to set `vip_https` with the virtual IP addresses required for HTTPS communication.
- Edit the `/etc/vxcps.conf` file to set `port_https` with the ports used for HTTPS communication.

5 Manually generate keys and certificates for the CP server.

See "Generating the key and certificates manually for the CP server" on page 52.

Generating the key and certificates manually for the CP server

CP server uses the HTTPS protocol to establish secure communication with client nodes. HTTPS is a secure means of communication, which happens over a secure communication channel that is established using the SSL/TLS protocol.

HTTPS uses x509 standard certificates and the constructs from a Public Key Infrastructure (PKI) to establish secure communication between the CP server and client. Similar to a PKI, the CP server, and its clients have their own set of certificates signed by a Certification Authority (CA). The server and its clients trust the certificate.

Every CP server acts as a certification authority for itself and for all its client nodes. The CP server has its own CA key and CA certificate and a server certificate generated, which is generated from a server private key. The server certificate is issued to the Universally Unique Identifier (UUID) of the CP server. All the IP addresses or domain names that the CP server listens on are mentioned in the Subject Alternative Name section of the CP server's server certificate

The OpenSSL library must be installed on the CP server to create the keys or certificates.. If OpenSSL is not installed, then you cannot create keys or certificates. The vxcps.conf file points to the configuration file that determines which keys or certificates are used by the CP server when SSL is initialized. The configuration value is stored in the `ssl_conf_file` and the default value is `/etc/vxcps_ssl.properties`.

To manually generate keys and certificates for the CP server:

1 Create directories for the security files on the CP server.

```
# mkdir -p /var/VRTScps/security/keys /var/VRTScps/security/certs
```

2 Generate an OpenSSL config file, which includes the VIPs.

The CP server listens to requests from client nodes on these VIPs. The server certificate includes VIPs, FQDNs, and host name of the CP server. Clients can reach the CP server by using any of these values. However, Veritas

recommends that client nodes use the IP address to communicate to the CP server.

The sample configuration uses the following values:

- Config file name: *https_ssl_cert.conf*
- VIP: *192.168.1.201*
- FQDN: *cpsone.company.com*
- Host name: *cpsone*

Note the IP address, VIP, and FQDN values used in the [alt_names] section of the configuration file are sample values. Replace the sample values with your configuration values. Do not change the rest of the values in the configuration file.

```
[req]
distinguished_name = req_distinguished_name
req_extensions = v3_req

[req_distinguished_name]
countryName = Country Name (2 letter code)
countryName_default = US
localityName = Locality Name (eg, city)
organizationalUnitName = Organizational Unit Name (eg, section)
commonName = Common Name (eg, YOUR name)
commonName_max = 64
emailAddress = Email Address
emailAddress_max = 40

[v3_req]
keyUsage = keyEncipherment, dataEncipherment
extendedKeyUsage = serverAuth
subjectAltName = @alt_names

[alt_names]
DNS.1 = cpsone.company.com
DNS.2 = cpsone
DNS.3 = 192.168.1.201
```

3 Generate a 4096-bit CA key that is used to create the CA certificate.

The key must be stored at `/var/VRTScps/security/keys/ca.key`. Ensure that only root users can access the CA key, as the key can be misused to create fake certificates and compromise security.

```
# /opt/VRTSperl/non-perl-libs/bin/openssl genrsa -out
/var/VRTScps/security/keys/ca.key 4096
```

4 Generate a self-signed CA certificate.

```
# /opt/VRTSperl/non-perl-libs/bin/openssl req -new -x509 -days
days -sha256 -key /var/VRTScps/security/keys/ca.key -subj \

'/C=countryname/L=localityname/OU=COMPANY/CN=CACERT' -out \

/var/VRTScps/security/certs/ca.crt
```

Where, *days* is the days you want the certificate to remain valid, *countryname* is the name of the country, localityname is the city, *CACERT* is the certificate name.

5 Generate a 2048-bit private key for CP server.

The key must be stored at `/var/VRTScps/security/keys/server_private key`.

```
# /opt/VRTSperl/non-perl-libs/bin/openssl genrsa -out \

/var/VRTScps/security/keys/server_private.key 2048
```

6 Generate a Certificate Signing Request (CSR) for the server certificate.

The Certified Name (CN) in the certificate is the UUID of the CP server.

```
# /opt/VRTSperl/non-perl-libs/bin/openssl req -new -sha256 -key
/var/VRTScps/security/keys/server_private.key \

-config https_ssl_cert.conf -subj \

'/C=CountryName/L=LocalityName/OU=COMPANY/CN=UUID' \

-out /var/VRTScps/security/certs/server.csr
```

Where, *countryname* is the name of the country, *localityname* is the city, *UUID* is the certificate name.

7 Generate the server certificate by using the key certificate of the CA.

```
# /opt/VRTSperl/non-perl-libs/bin/openssl x509 -req -days days
-sha256 -in /var/VRTScps/security/certs/server.csr \
-CA /var/VRTScps/security/certs/ca.crt -CAkey \
/var/VRTScps/security/keys/ca.key \
-set_serial 01 -extensions v3_req -extfile https_ssl_cert.conf \
-out /var/VRTScps/security/certs/server.crt
```

Where, *days* is the days you want the certificate to remain valid, *https_ssl_cert.conf* is the configuration file name.

You successfully created the key and certificate required for the CP server.

8 Ensure that no other user except the root user can read the keys and certificates.

9 Complete the CP server configuration.

See "Completing the CP server configuration" on page 55.

Completing the CP server configuration

To verify the service groups and start VCS perform the following steps:

1 Start VCS on all the cluster nodes.

```
# hastart
```

2 Verify that the CP server service group (CPSSG) is online.

```
# hagrp -state CPSSG
```

Output similar to the following appears:

```
# Group Attribute  System                    Value
  CPSSG State      cps1.example.com  |ONLINE|
```

Verifying the CP server configuration

Perform the following steps to verify the CP server configuration.

To verify the CP server configuration

1 Verify that the following configuration files are updated with the information you provided during the CP server configuration process:

- /etc/vxcps.conf (CP server configuration file)

- /etc/VRTSvcs/conf/config/main.cf (VCS configuration file)
- /etc/VRTScps/db (default location for CP server database for a single-node cluster)
- /cps_db (default location for CP server database for a multi-node cluster)

2 Run the `cpsadm` command to check if the vxcpserv process is listening on the configured Virtual IP.

If the application cluster is configured for HTTPS-based communication, no need to provide the port number assigned for HTTP communication.

```
# cpsadm -s cp_server -a ping_cps
```

where *cp_server* is the virtual IP address or the virtual hostname of the CP server.

Chapter 4

Configuring SFCFSHA

This chapter includes the following topics:

- Overview of tasks to configure SFCFSHA using the product installer
- Starting the software configuration
- Specifying systems for configuration
- Configuring the cluster name
- Configuring private heartbeat links
- Configuring the virtual IP of the cluster
- Configuring SFCFSHA in secure mode
- Configuring a secure cluster node by node
- Adding VCS users
- Configuring SMTP email notification
- Configuring SNMP trap notification
- Configuring global clusters
- Completing the SFCFSHA configuration
- About Veritas License Audit Tool
- Verifying and updating licenses on the system
- Configuring SFDB

Overview of tasks to configure SFCFSHA using the product installer

Table 4-1 lists the tasks that are involved in configuring Storage Foundation Cluster File System High Availability using the script-based installer.

Table 4-1 Tasks to configure SFCFSHA using the script-based installer

Task	Reference
Start the software configuration	See "Starting the software configuration" on page 58.
Specify the systems where you want to configure Storage Foundation Cluster File System High Availability	See "Specifying systems for configuration" on page 59.
Configure the basic cluster	See "Configuring the cluster name" on page 60. See "Configuring private heartbeat links" on page 60.
Configure virtual IP address of the cluster (optional)	See "Configuring the virtual IP of the cluster" on page 67.
Configure the cluster in secure mode (optional)	See "Configuring SFCFSHA in secure mode" on page 69.
Add VCS users (required if you did not configure the cluster in secure mode)	See "Adding VCS users" on page 74.
Configure SMTP email notification (optional)	See "Configuring SMTP email notification" on page 75.
Configure SNMP email notification (optional)	See "Configuring SNMP trap notification" on page 77.
Configure global clusters (optional)	See "Configuring global clusters" on page 78.
Complete the software configuration	See "Completing the SFCFSHA configuration" on page 79.

Starting the software configuration

You can configure Storage Foundation Cluster File System High Availability using the product installer.

Note: If you want to reconfigure Storage Foundation Cluster File System High Availability, before you start the installer you must stop all the resources that are under VCS control using the `hastop` command or the `hagrp -offline` command.

To configure Storage Foundation Cluster File System High Availability using the product installer

1. Confirm that you are logged in as a superuser.
2. Start the configuration using the installer.

   ```
   # /opt/VRTS/install/installer -configure
   ```

   The installer starts the product installation program with a copyright message and specifies the directory where the logs are created.
3. Select the component to configure.
4. Continue with the configuration procedure by responding to the installer questions.

Specifying systems for configuration

The installer prompts for the system names on which you want to configure Storage Foundation Cluster File System High Availability. The installer performs an initial check on the systems that you specify.

To specify system names for configuration

1. Enter the names of the systems where you want to configure Storage Foundation Cluster File System High Availability.

   ```
   Enter the operating_system system names separated
   by spaces:  [q,?] (sys1) sys1 sys2
   ```

2. Review the output as the installer verifies the systems you specify.

   The installer does the following tasks:

   - Checks that the local node running the installer can communicate with remote nodes

     If the installer finds ssh binaries, it confirms that ssh can operate without requests for passwords or passphrases. If ssh binaries cannot communicate with remote nodes, the installer tries rsh binaries. And if both ssh and rsh binaries fail, the installer prompts to help the user to setup ssh or rsh binaries.

- Makes sure that the systems are running with the supported operating system
- Checks whether Veritas InfoScale Enterprise is installed
- Exits if Veritas InfoScale Enterprise8.0 is not installed

3 Review the installer output about the I/O fencing configuration and confirm whether you want to configure fencing in enabled mode.

```
Do you want to configure I/O Fencing in enabled mode? [y,n,q,?] (y)
```

See " About planning to configure I/O fencing" on page 30.

Configuring the cluster name

Enter the cluster information when the installer prompts you.

To configure the cluster

1 Review the configuration instructions that the installer presents.

2 Enter a unique cluster name.

```
Enter the unique cluster name: [q,?] clus1
```

Configuring private heartbeat links

You now configure the private heartbeat links that LLT uses.

VCS provides the option to use LLT over Ethernet or LLT over UDP (User Datagram Protocol) or LLT over RDMA, or LLT over TCP. Veritas recommends that you configure heartbeat links that use LLT over Ethernet or LLT over RDMA for high performance, unless hardware requirements force you to use LLT over UDP. If you want to configure LLT over UDP, make sure you meet the prerequisites.

You must not configure LLT heartbeat using the links that are part of aggregated links. For example, link1, link2 can be aggregated to create an aggregated link, aggr1. You can use aggr1 as a heartbeat link, but you must not use either link1 or link2 as heartbeat links.

See "Using the UDP layer for LLT" on page 358.

See "Using LLT over RDMA: supported use cases " on page 377.

The following procedure helps you configure LLT heartbeat links.

To configure private heartbeat links

1 Choose one of the following options at the installer prompt based on whether you want to configure LLT over Ethernet or LLT over UDP or LLT over TCP or LLT over RDMA.

- Option 1: Configure the heartbeat links using LLT over Ethernet (answer installer questions)
  Enter the heartbeat link details at the installer prompt to configure LLT over Ethernet.
  Skip to step 2.
- Option 2: Configure the heartbeat links using LLT over UDP (answer installer questions)
  Make sure that each NIC you want to use as heartbeat link has an IP address configured. Enter the heartbeat link details at the installer prompt to configure LLT over UDP. If you had not already configured IP addresses to the NICs, the installer provides you an option to detect the IP address for a given NIC.

  Note: Ensure that the interface that is used by the LLT links does not have any other IP in the same subnet were the LLT links are configured. Otherwise, the cluster may behave unpredictably.

  Skip to step 3.
- Option 3: Configure the heartbeat links using LLT over TCP (answer installer questions)
  Make sure that the NIC you want to use as heartbeat link has an IP address configured. Enter the heartbeat link details at the installer prompt to configure LLT over TCP. If you had not already configured IP addresses to the NICs, the installer provides you an option to detect the IP address for a given NIC.
  Skip to step 4.
- Option 4: Configure the heartbeat links using LLT over RDMA (answer installer questions)
  Make sure that each RDMA enabled NIC (RNIC) you want to use as heartbeat link has an IP address configured. Enter the heartbeat link details at the installer prompt to configure LLT over RDMA. If you had not already configured IP addresses to the RNICs, the installer provides you an option to detect the IP address for a given RNIC.
  Skip to step 5.
- Option 5: Automatically detect configuration for LLT over Ethernet

Allow the installer to automatically detect the heartbeat link details to configure LLT over Ethernet. The installer tries to detect all connected links between all systems.

Make sure that you activated the NICs for the installer to be able to detect and automatically configure the heartbeat links.

Skip to step 8.

Note: Option 5 is not available when the configuration is a single node configuration.

---

2 If you chose option 1, enter the network interface card details for the private heartbeat links.

The installer discovers and lists the network interface cards.

You must not enter the network interface card that is used for the public network (typically eth0.)

```
Enter the NIC for the first private heartbeat link on sys1:
[b,q,?] eth1
eth1 has an IP address configured on it. It could be a
public NIC on sys1.
Are you sure you want to use eth1 for the first private
heartbeat link? [y,n,q,b,?] (n) y
Would you like to configure a second private heartbeat link?
[y,n,q,b,?] (y)
Enter the NIC for the second private heartbeat link on sys1:
[b,q,?] eth2
eth2 has an IP address configured on it. It could be a
public NIC on sys1.
Are you sure you want to use eth2 for the second private
heartbeat link? [y,n,q,b,?] (n) y
Would you like to configure a third private heartbeat link?
[y,n,q,b,?](n)

Do you want to configure an additional low priority heartbeat
link? [y,n,q,b,?] (n)
```

3 If you chose option 2, enter the NIC details for the private heartbeat links. This step uses examples such as *private_NIC1* or *private_NIC2* to refer to the available names of the NICs.

```
Enter the NIC for the first private heartbeat link on sys1: [b,q,?]
private_NIC1
Some configured IP addresses have been found on
the NIC private_NIC1 in sys1,
Do you want to choose one for the first private heartbeat link? [y,n,q,?]
Please select one IP address:
     1)  192.168.0.1/24
     2)  192.168.1.233/24
     b)  Back to previous menu

Please select one IP address: [1-2,b,q,?] (1)
Enter the UDP port for the first private heartbeat link on sys1:
[b,q,?] (50000)

Enter the NIC for the second private heartbeat link on sys1: [b,q,?]
private_NIC2
Some configured IP addresses have been found on the
NIC private_NIC2 in sys1,
Do you want to choose one for the second
private heartbeat link? [y,n,q,?] (y)
Please select one IP address:
     1)  192.168.1.1/24
     2)  192.168.2.233/24
     b)  Back to previous menu

Please select one IP address: [1-2,b,q,?] (1) 1
Enter the UDP port for the second private heartbeat link on sys1:
[b,q,?] (50001)

Would you like to configure a third private heartbeat
link? [y,n,q,b,?] (n)

Do you want to configure an additional low-priority heartbeat
link? [y,n,q,b,?] (n) y

Enter the NIC for the low-priority heartbeat link on sys1: [b,q,?]
private_NIC0
Some configured IP addresses have been found on
```

```
the NIC private_NIC0 in sys1,
Do you want to choose one for the low-priority
heartbeat link? [y,n,q,?] (y)
Please select one IP address:
     1)  10.200.59.233/22
     2)  192.168.3.1/22
     b)  Back to previous menu

Please select one IP address: [1-2,b,q,?] (1) 2
Enter the UDP port for the low-priority heartbeat link on sys1:
[b,q,?] (50010)
```

4 If you chose option 3, enter the NIC details for the private heartbeat link.

This step uses an example such as *private_NIC1* to refer to the available name of the NIC.

```
Enter the NIC for the private heartbeat link on sys1: [b,q,?] (eth1)
private_NIC1
Some configured IP addresses have been found on
the NIC private_NIC1 in sys1,
Do you want to choose one for the private
heartbeat link? [y,n,q,?] (y) y
Please select one IP address:
1) 192.168.1.1/24
2) 192.168.2.1/24
b) Back to previous menu

Please select one IP address: [1-2,b,q,?] (1)
Enter the TCP port for the first private heartbeat link on sys1:
[b,q,?] (50000)
```

5 If you chose option 4, choose the interconnect type to configure RDMA.

```
1)  Converged Ethernet (RoCE)
2)  InfiniBand
b)  Back to previous menu

Choose the RDMA interconnect type [1-2,b,q,?] (1) 2
```

The system displays the details such as the required OS files, drivers required for RDMA , and the IP addresses for the NICs.

A sample output of the IP addresses assigned to the RDMA enabled NICs using InfiniBand network. Note that with RoCE, the RDMA NIC values are represented as eth0, eth1, and so on.

```
System        RDMA NIC          IP Address
================================================================
sys1           ib0               192.168.0.1
sys1           ib1               192.168.3.1
sys2           ib0               192.168.0.2
sys2           ib1               192.168.3.2
```

6 If you chose option 4, enter the NIC details for the private heartbeat links. This step uses RDMA over an InfiniBand network. With RoCE as the interconnect type, RDMA NIC is represented as Ethernet (eth).

```
Enter the NIC for the first private heartbeat
link (RDMA) on sys1: [b,q,?] <ib0>

Do you want to use address 192.168.0.1 for the
first private heartbeat link on sys1: [y,n,q,b,?] (y)

Enter the port for the first private heartbeat
link (RDMA) on sys1: [b,q,?] (50000) ?

Would you like to configure a second private
heartbeat link? [y,n,q,b,?] (y)
Enter the NIC for the second private heartbeat link (RDMA) on sys1:
[b,q,?] (ib1)

Do you want to use the address 192.168.3.1 for the second
private heartbeat link on sys1: [y,n,q,b,?] (y)
Enter the port for the second private heartbeat link (RDMA) on sys1:
[b,q,?] (50001)

Do you want to configure an additional low-priority heartbeat link?
[y,n,q,b,?] (n)
```

7 Choose whether to use the same NIC details to configure private heartbeat links on other systems.

```
Are you using the same NICs for private heartbeat links on all
systems? [y,n,q,b,?] (y)
```

If you want to use the NIC details that you entered for sys1, make sure the same NICs are available on each system. Then, enter **y** at the prompt.

If the NIC device names are different on some of the systems, enter **n**. Provide the NIC details for each system as the program prompts.

For LLT over UDP and LLT over RDMA, if you want to use the same NICs on other systems, you must enter unique IP addresses on each NIC for other systems.

8 If you chose option 5, the installer detects NICs on each system and network links, and sets link priority.

If the installer fails to detect heartbeat links or fails to find any high-priority links, then choose option 1 or option 2 to manually configure the heartbeat links.

See step 2 for option 1, or step 3 for option 2, or step 4 for option 3, or step5 for option 4

9 Enter a unique cluster ID:

```
Enter a unique cluster ID number between 0-65535: [b,q,?] (60842)
```

The cluster cannot be configured if the cluster ID 60842 is in use by another cluster. Installer performs a check to determine if the cluster ID is duplicate. The check takes less than a minute to complete.

```
Would you like to check if the cluster ID is in use by another
cluster? [y,n,q] (y)
```

10 Verify and confirm the information that the installer summarizes.

Configuring the virtual IP of the cluster

You can configure the virtual IP of the cluster to use to connect from the Cluster Manager (Java Console), Veritas InfoScale Operations Manager, or to specify in the RemoteGroup resource.

See the *Cluster Server Administrator's Guide* for information on the Cluster Manager.

See the *Cluster Server Bundled Agents Reference Guide* for information on the RemoteGroup agent.

To configure the virtual IP of the cluster

1 Review the required information to configure the virtual IP of the cluster.

2 When the system prompts whether you want to configure the virtual IP, enter `y`.

3 Confirm whether you want to use the discovered public NIC on the first system.

Do one of the following:

- If the discovered NIC is the one to use, press `Enter`.
- If you want to use a different NIC, type the name of a NIC to use and press `Enter`.

```
Active NIC devices discovered on sys1: eth0
Enter the NIC for Virtual IP of the Cluster to use on sys1:
[b,q,?](eth0)
```

4 Confirm whether you want to use the same public NIC on all nodes.

Do one of the following:

- If all nodes use the same public NIC, enter `y`.
- If unique NICs are used, enter `n` and enter a NIC for each node.

```
Is eth0 to be the public NIC used by all systems
[y,n,q,b,?] (y)
```

5 Enter the virtual IP address for the cluster.

You can enter either an IPv4 address or an IPv6 address.

For IPv4:

- Enter the virtual IP address.

  ```
  Enter the Virtual IP address for the Cluster:
  [b,q,?] 192.168.1.16
  ```

- Confirm the default netmask or enter another one:

  ```
  Enter the netmask for IP 192.168.1.16: [b,q,?]
  (255.255.240.0)
  ```

- Verify and confirm the Cluster Virtual IP information.

  ```
  Cluster Virtual IP verification:

        NIC: eth0
        IP: 192.168.1.16
        Netmask: 255.255.240.0

  Is this information correct? [y,n,q] (y)
  ```

For IPv6

- Enter the virtual IP address.

  ```
  Enter the Virtual IP address for the Cluster:
  [b,q,?] 2001:454e:205a:110:203:baff:feee:10
  ```

- Enter the prefix for the virtual IPv6 address you provided. For example:

  ```
  Enter the Prefix for IP
  2001:454e:205a:110:203:baff:feee:10: [b,q,?] 64
  ```

- Verify and confirm the Cluster Virtual IP information.

  ```
  Cluster Virtual IP verification:

        NIC: eth0
        IP: 2001:454e:205a:110:203:baff:feee:10
        Prefix: 64

  Is this information correct? [y,n,q] (y)
  ```

If you want to set up trust relationships for your secure cluster, refer to the following topics:

See "Configuring a secure cluster node by node" on page 70.

Configuring SFCFSHA in secure mode

Configuring SFCFSHA in secure mode ensures that all the communication between the systems is encrypted and users are verified against security credentials. SFCFSHA user names and passwords are not used when a cluster is running in secure mode.

To configure SFCFSHA in secure mode

1. To install and configure SFCFSHA in secure mode, run the command:

   ```
   # ./installer -security
   ```

2. The installer displays the following question before the installer stops the product processes:

   - Do you want to grant read access to everyone? [y,n,q,?]
     - To grant read access to all authenticated users, type **y**.
     - To grant usergroup specific permissions, type **n**.

- Do you want to provide any usergroups that you would like to grant read access?[y,n,q,?]
  - To specify usergroups and grant them read access, type **y**
  - To grant read access only to root users, type **n**. The installer grants read access read access to the root users.
- Enter the usergroup names separated by spaces that you would like to grant read access. If you would like to grant read access to a usergroup on a specific node, enter like 'usrgrp1@node1', and if you would like to grant read access to usergroup on any cluster node, enter like 'usrgrp1'. If some usergroups are not created yet, create the usergroups after configuration if needed. [b]

3 To verify the cluster is in secure mode after configuration, run the command:

```
# haclus -value SecureClus
```

The command returns 1 if cluster is in secure mode, else returns 0.

Configuring a secure cluster node by node

For environments that do not support passwordless ssh or passwordless rsh, you cannot use the `-security` option to enable secure mode for your cluster. Instead, you can use the `-securityonenode` option to configure a secure cluster node by node. Moreover, to enable security in fips mode, use the `-fips` option together with `-securityonenode`.

Table 4-2 lists the tasks that you must perform to configure a secure cluster.

Table 4-2 Configuring a secure cluster node by node

Task	Reference
Configure security on one node	See "Configuring the first node" on page 70.
Configure security on the remaining nodes	See "Configuring the remaining nodes" on page 71.
Complete the manual configuration steps	See "Completing the secure cluster configuration" on page 72.

Configuring the first node

Perform the following steps on one node in your cluster.

To configure security on the first node

1. Ensure that you are logged in as superuser.
2. Enter the following command:

   ```
   # /opt/VRTS/install/installer -securityonenode
   ```

   The installer lists information about the cluster, nodes, and service groups. If VCS is not configured or if VCS is not running on all nodes of the cluster, the installer prompts whether you want to continue configuring security. It then prompts you for the node that you want to configure.

   ```
   VCS is not running on all systems in this cluster. All VCS systems
   must be in RUNNING state. Do you want to continue? [y,n,q] (n) y

   1) Perform security configuration on first node and export
   security configuration files.

   2) Perform security configuration on remaining nodes with
   security configuration files.

   Select the option you would like to perform [1-2,q.?] 1
   ```

   Warning: All VCS configurations about cluster users are deleted when you configure the first node. You can use the `/opt/VRTSvcs/bin/hauser` command to create cluster users manually.

3. The installer completes the secure configuration on the node. It specifies the location of the security configuration files and prompts you to copy these files to the other nodes in the cluster. The installer also specifies the location of log files, summary file, and response file.
4. Copy the security configuration files from the location specified by the installer to temporary directories on the other nodes in the cluster.

Configuring the remaining nodes

On each of the remaining nodes in the cluster, perform the following steps.

To configure security on each remaining node

1. Ensure that you are logged in as superuser.
2. Enter the following command:

```
# /opt/VRTS/install/installer -securityonenode
```

The installer lists information about the cluster, nodes, and service groups. If VCS is not configured or if VCS is not running on all nodes of the cluster, the installer prompts whether you want to continue configuring security. It then prompts you for the node that you want to configure. Enter **2**.

```
VCS is not running on all systems in this cluster. All VCS systems
must be in RUNNING state. Do you want to continue? [y,n,q] (n) y

1) Perform security configuration on first node and export
security configuration files.

2) Perform security configuration on remaining nodes with
security configuration files.

Select the option you would like to perform [1-2,q.?]  2
Enter the security conf file directory: [b]
```

The installer completes the secure configuration on the node. It specifies the location of log files, summary file, and response file.

Completing the secure cluster configuration

Perform the following manual steps to complete the configuration.

To complete the secure cluster configuration

1. On the first node, freeze all service groups except the ClusterService service group.

```
# /opt/VRTSvcs/bin/haconf -makerw

# /opt/VRTSvcs/bin/hagrp -list Frozen=0

# /opt/VRTSvcs/bin/hagrp -freeze groupname -persistent

# /opt/VRTSvcs/bin/haconf -dump -makero
```

2. On the first node, stop the VCS engine.

```
# /opt/VRTSvcs/bin/hastop -all -force
```

3 On all nodes, stop the CmdServer.

```
# systemctl stop CmdServer
```

4 To grant access to all users, add or modify `SecureClus=1` and `DefaultGuestAccess=1` in the cluster definition.

For example:

To grant read access to everyone:

```
Cluster clus1 (
SecureClus=1
DefaultGuestAccess=1
)
```

Or

To grant access to only root:

```
Cluster clus1 (
SecureClus=1
)
```

Or

To grant read access to specific user groups, add or modify SecureClus=1 and GuestGroups={} to the cluster definition.

For example:

```
cluster clus1 (
SecureClus=1
GuestGroups={staff, guest}
```

5 Modify `/etc/VRTSvcs/conf/config/main.cf` file on the first node, and add `-secure` to the WAC application definition if GCO is configured.

For example:

```
Application wac (
                StartProgram = "/opt/VRTSvcs/bin/wacstart -secure"
                StopProgram = "/opt/VRTSvcs/bin/wacstop"
                MonitorProcesses = {"/opt/VRTSvcs/bin/wac -secure"}
                RestartLimit = 3
                )
```

6 On all nodes, create the `/etc/VRTSvcs/conf/config/.secure` file.

```
# touch /etc/VRTSvcs/conf/config/.secure
```

7 On the first node, start VCS. Then start VCS on the remaining nodes.

```
# /opt/VRTSvcs/bin/hastart
```

8 On all nodes, start CmdServer.

```
# systemctl start CmdServer
```

9 On the first node, unfreeze the service groups.

```
# /opt/VRTSvcs/bin/haconf -makerw

# /opt/VRTSvcs/bin/hagrp -list Frozen=1

# /opt/VRTSvcs/bin/hagrp -unfreeze groupname -persistent

# /opt/VRTSvcs/bin/haconf -dump -makero
```

Adding VCS users

If you have enabled a secure VCS cluster, you do not need to add VCS users now. Otherwise, on systems operating under an English locale, you can add VCS users at this time.

To add VCS users

1. Review the required information to add VCS users.
2. Reset the password for the Admin user, if necessary.

```
Do you wish to accept the default cluster credentials of
'admin/password'? [y,n,q] (y) n
Enter the user name: [b,q,?] (admin)
Enter the password:
Enter again:
```

   The password is encrypted using the standard AES-256 algorithm.

3. To add a user, enter **y** at the prompt.

```
Do you want to add another user to the cluster? [y,n,q] (y)
```

4. Enter the user's name, password, and level of privileges.

```
Enter the user name: [b,q,?] smith
Enter New Password:*******

Enter Again:*******
Enter the privilege for user smith (A=Administrator, O=Operator,
G=Guest): [b,q,?] a
```

   The password is encrypted using the standard AES-256 algorithm.

5. Enter **n** at the prompt if you have finished adding users.

```
Would you like to add another user? [y,n,q] (n)
```

6. Review the summary of the newly added users and confirm the information.

Configuring SMTP email notification

You can choose to configure VCS to send event notifications to SMTP email services. You need to provide the SMTP server name and email addresses of people to be notified. Note that you can also configure the notification after installation.

Refer to the *Cluster Server Administrator's Guide* for more information.

To configure SMTP email notification

1. Review the required information to configure the SMTP email notification.
2. Specify whether you want to configure the SMTP notification.

   If you do not want to configure the SMTP notification, you can skip to the next configuration option.

   See "Configuring SNMP trap notification" on page 77.
3. Provide information to configure SMTP notification.

   Provide the following information:

   - Enter the SMTP server's host name.

     ```
     Enter the domain-based hostname of the SMTP server
     (example: smtp.yourcompany.com): [b,q,?] smtp.example.com
     ```

   - Enter the email address of each recipient.

     ```
     Enter the full email address of the SMTP recipient
     (example: user@yourcompany.com): [b,q,?] ozzie@example.com
     ```

   - Enter the minimum security level of messages to be sent to each recipient.

     ```
     Enter the minimum severity of events for which mail should be
     sent to ozzie@example.com  [I=Information, W=Warning,
     E=Error, S=SevereError]: [b,q,?] w
     ```
4. Add more SMTP recipients, if necessary.

   - If you want to add another SMTP recipient, enter `y` and provide the required information at the prompt.

     ```
     Would you like to add another SMTP recipient? [y,n,q,b] (n) y

     Enter the full email address of the SMTP recipient
     (example: user@yourcompany.com): [b,q,?] harriet@example.com

     Enter the minimum severity of events for which mail should be
     sent to harriet@example.com  [I=Information, W=Warning,
     E=Error, S=SevereError]: [b,q,?] E
     ```

   - If you do not want to add, answer **n**.

```
Would you like to add another SMTP recipient? [y,n,q,b] (n)
```

5 Verify and confirm the SMTP notification information.

```
SMTP Address: smtp.example.com
Recipient: ozzie@example.com receives email for Warning or
higher events
Recipient: harriet@example.com receives email for Error or
higher events

Is this information correct? [y,n,q] (y)
```

Configuring SNMP trap notification

You can choose to configure VCS to send event notifications to SNMP management consoles. You need to provide the SNMP management console name to be notified and message severity levels.

Note that you can also configure the notification after installation.

Refer to the *Cluster Server Administrator's Guide* for more information.

To configure the SNMP trap notification

1 Review the required information to configure the SNMP notification feature of VCS.

2 Specify whether you want to configure the SNMP notification.

If you skip this option and if you had installed a valid HA/DR license, the installer presents you with an option to configure this cluster as global cluster. If you did not install an HA/DR license, the installer proceeds to configure SFCFSHA based on the configuration details you provided.

See "Configuring global clusters" on page 78.

3 Provide information to configure SNMP trap notification.

Provide the following information:

- Enter the SNMP trap daemon port.

```
Enter the SNMP trap daemon port: [b,q,?] (162)
```

- Enter the SNMP console system name.

```
Enter the SNMP console system name: [b,q,?] sys5
```

- Enter the minimum security level of messages to be sent to each console.

```
Enter the minimum severity of events for which SNMP traps
should be sent to sys5 [I=Information, W=Warning, E=Error,
S=SevereError]: [b,q,?] E
```

4 Add more SNMP consoles, if necessary.

- If you want to add another SNMP console, enter `y` and provide the required information at the prompt.

```
Would you like to add another SNMP console? [y,n,q,b] (n) y
Enter the SNMP console system name: [b,q,?] sys4
Enter the minimum severity of events for which SNMP traps
should be sent to sys4 [I=Information, W=Warning,
E=Error, S=SevereError]: [b,q,?] S
```

- If you do not want to add, answer `n`.

```
Would you like to add another SNMP console? [y,n,q,b] (n)
```

5 Verify and confirm the SNMP notification information.

```
SNMP Port: 162
Console: sys5 receives SNMP traps for Error or
higher events
Console: sys4 receives SNMP traps for SevereError or
higher events

Is this information correct? [y,n,q] (y)
```

Configuring global clusters

You can configure global clusters to link clusters at separate locations and enable wide-area failover and disaster recovery. The installer adds basic global cluster information to the VCS configuration file. You must perform additional configuration tasks to set up a global cluster.

See the *Cluster Server Administrator's Guide* for instructions to set up Storage Foundation Cluster File System High Availability global clusters.

See the appropriate *Veritas InfoScale* installation guide for instructions to set up Storage Foundation Cluster File System High Availability global clusters.

Note: If you installed a HA/DR license to set up replicated data cluster or campus cluster, skip this installer option.

To configure the global cluster option

1. Review the required information to configure the global cluster option.
2. Specify whether you want to configure the global cluster option.

   If you skip this option, the installer proceeds to configure VCS based on the configuration details you provided.
3. Provide information to configure this cluster as global cluster.

   The installer prompts you for a NIC, a virtual IP address, and value for the netmask.

   You can also enter an IPv6 address as a virtual IP address.

Completing the SFCFSHA configuration

After you enter the SFCFSHA configuration information, the installer prompts to stop the SFCFSHA processes to complete the configuration process. The installer continues to create configuration files and copies them to each system. The installer also configures a cluster UUID value for the cluster at the end of the configuration. After the installer successfully configures SFCFSHA, it restarts SFCFSHA and its related processes.

To complete the SFCFSHA configuration

1. If prompted, press Enter at the following prompt.

   ```
   Do you want to stop InfoScale Enterprise  processes now? [y,n,q,?] (y)
   ```

2. Review the output as the installer stops various processes and performs the configuration. The installer then restarts SFCFSHA and its related processes.

3 Enter y at the prompt to send the installation information to Veritas.

```
Would you like to send the information about this installation
to us  to help improve installation in the future?
[y,n,q,?] (y) y
```

4 After the installer configures Storage Foundation Cluster File System High Availability successfully, note the location of summary, log, and response files that installer creates.

The files provide the useful information that can assist you with the configuration and can also assist future configurations.

summary file	Describes the cluster and its configured resources.
log file	Details the entire configuration.
response file	Contains the configuration information that can be used to perform secure or unattended installations on other systems. See "Configuring SFCFSHA using response files" on page 113.

Verifying the NIC configuration

The installer verifies on all the nodes if all NICs have PERSISTENT_NAME set correctly.

If the persistent interface names are not configured correctly for the network devices, the installer displays the following messages:

```
PERSISTENT_NAME is not set for all the NICs.
You need to set them manually before the next reboot.
```

Set the PERSISTENT_NAME for all the NICs.

Warning: If the installer finds the network interface name to be different from the name in the configuration file, then the installer exits.

About Veritas License Audit Tool

Veritas License Audit Tool by Veritas intelligently scans your organization's network and gives you a comprehensive report of all the Veritas product licenses used at your organization. This robust tool allows your organization to see all the current Veritas products installed your systems. This helps your organization in the following:

- License and maintenance renewal of Veritas products
- Contract renegotiations of Veritas Products
- Re-harvesting and reuse of Veritas Products

Veritas License Audit Tool's robust reporting framework enables you to capture information such as Product name, Product Version, Licensing key, License type, Operating System, Operating System Version and CPU Name.

To download the Veritas License Audit Tool and its Installation and User Guide, click the following link:

https://sort.veritas.com/public/utilities/infoscale/latool/linux/LATool-rhel7.tar

Verifying and updating licenses on the system

After you install Storage Foundation Cluster File System High Availability you can verify and manage the licenses using the vxlicrep program.

See "Checking licensing information on the system" on page 81.

See "Replacing a SFCFSHA keyless license with another keyless license" on page 81.

Checking licensing information on the system

You can use the vxlicrep program to display information about the licenses on a system.

To check licensing information

- Navigate to the `/sbin` folder containing the `vxlicrep` program and enter:

```
# vxlicrep
```

Replacing a SFCFSHA keyless license with another keyless license

You can use the `./installer -license` command or the `vxkeyless` command to replace a SFCFSHA keyless license with another keyless license on each node.

See "Replacing a SFCFSHA keyless license with a permanent license" on page 82.

To update product licenses using the installer command

1. On any one node, enter the following command:

   ```
   # ./installer -license
   ```

2. At the prompt, enter your keyless license text string.

To update product licenses using the vxkeyless command

- On each node, enter the keyless license text string using the command:

  ```
  # vxkeyless set <keyless license text-string>
  ```

  Example:

  ```
  # vxkeyless set ENTERPRISE
  ```

Replacing a SFCFSHA keyless license with a permanent license

Within 60 days of enabling the Storage Foundation Cluster File System High Availability keyless license, you must replace it with a permanent license using the `vxlicinstupgrade` program.

To update product licenses using the vxlicinstupgrade command

1. Make sure you have permissions to log in as root on each of the nodes in the cluster.
2. Enter the permanent license key using the following command on each node:

   ```
   # vxlicinstupgrade -k <key file path>
   ```

   Note: The license key file must not be saved in the root directory (`/`) or the default license directory on the local host (`/etc/vx/licenses/lic`). You can save the license key file inside any other directory on the local host.

3. Make sure keyless licenses are replaced on all cluster nodes before starting Storage Foundation Cluster File System High Availability.

   ```
   # vxlicrep
   ```

To update product licenses using the installer command

1. On any one node, enter the following command:

   ```
   #./installer -license
   ```

2. At the prompt, enter your permanent license key file.

Configuring SFDB

By default, SFDB tools are disabled that is the vxdbd daemon is not configured. You can check whether SFDB tools are enabled or disabled using the`/opt/VRTS/bin/sfae_config status` command.

To enable SFDB tools

1. Log in as root.
2. On SLES 15 systems, install the `insserv-compat` package manually. This package is required to enable the `vxdbd` daemon.
3. Run the following command to configure and start the `vxdbd` daemon. After you perform this step, entries are made in the system startup so that the daemon starts on a system restart.

   `#/opt/VRTS/bin/sfae_config enable`

To disable SFDB tools

1. Log in as root.
2. Run the following command:

   `#/opt/VRTS/bin/sfae_config disable`

For more information, see the *Veritas InfoScale™ Storage and Availability Management for Oracle Databases* guide.

Chapter 5

Configuring SFCFSHA clusters for data integrity

This chapter includes the following topics:

- Setting up disk-based I/O fencing using installer
- Setting up server-based I/O fencing using installer
- Setting up non-SCSI-3 I/O fencing in virtual environments using installer
- Setting up majority-based I/O fencing using installer
- Enabling or disabling the preferred fencing policy

Setting up disk-based I/O fencing using installer

You can configure I/O fencing using the `-fencing` option of the installer.

Configuring disk-based I/O fencing using installer

Note: The installer stops and starts Storage Foundation Cluster File System High Availability to complete I/O fencing configuration. Make sure to unfreeze any frozen VCS service groups in the cluster for the installer to successfully stop Storage Foundation Cluster File System High Availability.

To set up disk-based I/O fencing using the installer

1. Start the installer with `-fencing` option.

   ```
   # /opt/VRTS/install/installer  -fencing
   ```

   The installer starts with a copyright message and verifies the cluster information.

   Note the location of log files which you can access in the event of any problem with the configuration process.

2. Enter the host name of one of the systems in the cluster.

3. Confirm that you want to proceed with the I/O fencing configuration at the prompt.

   The program checks that the local node running the script can communicate with remote nodes and checks whether Storage Foundation Cluster File System High Availability 8.0 is configured properly.

4. Review the I/O fencing configuration options that the program presents. Type **2** to configure disk-based I/O fencing.

   ```
   1. Configure Coordination Point client based fencing
   2. Configure disk based fencing
   3. Configure majority based fencing
   4. Configure fencing in disabled mode
   Select the fencing mechanism to be configured in this
   Application Cluster [1-4,q.?] 2
   ```

5. Review the output as the configuration program checks whether VxVM is already started and is running.

   - If the check fails, configure and enable VxVM before you repeat this procedure.
   - If the check passes, then the program prompts you for the coordinator disk group information.

6. Choose whether to use an existing disk group or create a new disk group to configure as the coordinator disk group.

   The program lists the available disk group names and provides an option to create a new disk group. Perform one of the following:

   - To use an existing disk group, enter the number corresponding to the disk group at the prompt.
     The program verifies whether the disk group you chose has an odd number of disks and that the disk group has a minimum of three disks.

- To create a new disk group, perform the following steps:
  - Enter the number corresponding to the **Create a new disk group** option.
    The program lists the available disks that are in the CDS disk format in the cluster and asks you to choose an odd number of disks with at least three disks to be used as coordinator disks.
    Veritas recommends that you use three disks as coordination points for disk-based I/O fencing.
  - If the available VxVM CDS disks are less than the required, installer asks whether you want to initialize more disks as VxVM disks. Choose the disks you want to initialize as VxVM disks and then use them to create new disk group.
  - Enter the numbers corresponding to the disks that you want to use as coordinator disks.
  - Enter the disk group name.

7 Verify that the coordinator disks you chose meet the I/O fencing requirements.

You must verify that the disks are SCSI-3 PR compatible using the vxfentsthdw utility and then return to this configuration program.

See "Checking shared disks for I/O fencing" on page 88.

8 After you confirm the requirements, the program creates the coordinator disk group with the information you provided.

9 Verify and confirm the I/O fencing configuration information that the installer summarizes.

10 Review the output as the configuration program does the following:

- Stops VCS and I/O fencing on each node.
- Configures disk-based I/O fencing and starts the I/O fencing process.
- Updates the VCS configuration file main.cf if necessary.
- Copies the /etc/vxfenmode file to a date and time suffixed file /etc/vxfenmode-*date-time*. This backup file is useful if any future fencing configuration fails.
- Updates the I/O fencing configuration file /etc/vxfenmode.
- Starts VCS on each node to make sure that the Storage Foundation Cluster File System High Availability is cleanly configured to use the I/O fencing feature.

11 Review the output as the configuration program displays the location of the log files, the summary files, and the response files.

12 Configure the Coordination Point Agent.

```
Do you want to configure Coordination Point Agent on
the client cluster? [y,n,q] (y)
```

13 Enter a name for the service group for the Coordination Point Agent.

```
Enter a non-existing name for the service group for
Coordination Point Agent: [b] (vxfen) vxfen
```

14 Set the level two monitor frequency.

```
Do you want to set LevelTwoMonitorFreq? [y,n,q] (y)
```

15 Decide the value of the level two monitor frequency.

```
Enter the value of the LevelTwoMonitorFreq attribute: [b,q,?] (5)
```

Installer adds Coordination Point Agent and updates the main configuration file.

16 Enable auto refresh of coordination points.

```
Do you want to enable auto refresh of coordination points
if registration keys are missing
on any of them? [y,n,q,b,?]  (n)
```

See "Configuring CoordPoint agent to monitor coordination points" on page 157.

Initializing disks as VxVM disks

Perform the following procedure to initialize disks as VxVM disks.

To initialize disks as VxVM disks

1 List the new external disks or the LUNs as recognized by the operating system. On each node, enter:

```
# fdisk -l
```

2 To initialize the disks as VxVM disks, use one of the following methods:

- Use the interactive vxdiskadm utility to initialize the disks as VxVM disks. For more information, see the *Storage Foundation Administrator's Guide*.
- Use the `vxdisksetup` command to initialize a disk as a VxVM disk.

  ```
  # vxdisksetup -i device_name
  ```

The example specifies the CDS format:

```
# vxdisksetup -i sdr format=cdsdisk
```

Repeat this command for each disk you intend to use as a coordinator disk.

Checking shared disks for I/O fencing

Make sure that the shared storage you set up while preparing to configure SFCFSHA meets the I/O fencing requirements. You can test the shared disks using the vxfentsthdw utility. The two nodes must have `ssh` (default) or `rsh` communication. To confirm whether a disk (or LUN) supports SCSI-3 persistent reservations, two nodes must simultaneously have access to the same disks. Because a shared disk is likely to have a different name on each node, check the serial number to verify the identity of the disk. Use the `vxfenadm` command with the `-i` option. This command option verifies that the same serial number for the LUN is returned on all paths to the LUN.

Make sure to test the disks that serve as coordinator disks.

The vxfentsthdw utility has additional options suitable for testing many disks. Review the options for testing the disk groups (`-g`) and the disks that are listed in a file (`-f`). You can also test disks without destroying data using the `-r` option.

See the *Storage Foundation Cluster File System High Availability Administrator's Guide*.

Checking that disks support SCSI-3 involves the following tasks:

- Verifying the Array Support Library (ASL)
  See "Verifying Array Support Library (ASL)" on page 88.
- Verifying that nodes have access to the same disk
  See "Verifying that the nodes have access to the same disk" on page 89.
- Testing the shared disks for SCSI-3
  See "Testing the disks using vxfentsthdw utility" on page 90.

Verifying Array Support Library (ASL)

Make sure that the Array Support Library (ASL) for the array that you add is installed.

To verify Array Support Library (ASL)

1. If the Array Support Library (ASL) for the array that you add is not installed, obtain and install it on each node before proceeding.

   The ASL for the supported storage device that you add is available from the disk array vendor or Veritas technical support.

2. Verify that the ASL for the disk array is installed on each of the nodes. Run the following command on each node and examine the output to verify the installation of ASL.

   The following output is a sample:

   ```
   # vxddladm listsupport all


   LIBNAME                 VID                 PID
   ============================================================
   libvxhitachi.so         HITACHI         DF350, DF400, DF400F,
                                           DF500, DF500F
   libvxxp1281024.so       HP              All
   libvxxp12k.so           HP              All
   libvxddns2a.so          DDN             S2A 9550, S2A 9900,
                                           S2A 9700
   libvxpurple.so          SUN             T300
   libvxxiotechE5k.so      XIOTECH         ISE1400
   libvxcopan.so           COPANSYS        8814, 8818
   libvxibmds8k.so         IBM             2107
   libvxnvme_nonscsi.so    NVMe            All
   ```

3. Scan all disk drives and their attributes, update the VxVM device list, and reconfigure DMP with the new devices. Type:

   ```
   # vxdisk scandisks
   ```

   See the Veritas Volume Manager documentation for details on how to add and configure disks.

Verifying that the nodes have access to the same disk

Before you test the disks that you plan to use as shared data storage or as coordinator disks using the vxfentsthdw utility, you must verify that the systems see the same disk.

To verify that the nodes have access to the same disk

1. Verify the connection of the shared storage for data to two of the nodes on which you installed Veritas InfoScale Enterprise.
2. Ensure that both nodes are connected to the same disk during the testing. Use the vxfenadm command to verify the disk serial number.

   ```
   # vxfenadm -i diskpath
   ```

   Refer to the `vxfenadm` (1M) manual page.

   For example, an EMC disk is accessible by the /dev/sdx path on node A and the /dev/sdy path on node B.

   From node A, enter:

   ```
   # vxfenadm -i /dev/sdx

   SCSI ID=>Host: 2 Channel: 0 Id: 0 Lun: E

   Vendor id : EMC
   Product id : SYMMETRIX
   Revision : 5567
   Serial Number : 42031000a
   ```

   The same serial number information should appear when you enter the equivalent command on node B using the /dev/sdy path.

   On a disk from another manufacturer, Hitachi Data Systems, the output is different and may resemble:

   ```
   SCSI ID=>Host: 2 Channel: 0 Id: 0 Lun: E

   Vendor id       : HITACHI
   Product id      : OPEN-3
   Revision        : 0117
   Serial Number   : 0401EB6F0002
   ```

Testing the disks using vxfentsthdw utility

This procedure uses the /dev/sdx disk in the steps.

If the utility does not show a message that states a disk is ready, the verification has failed. Failure of verification can be the result of an improperly configured disk array. The failure can also be due to a bad disk.

If the failure is due to a bad disk, remove and replace it. The vxfentsthdw utility indicates a disk can be used for I/O fencing with a message resembling:

```
The disk /dev/sdx is ready to be configured for I/O Fencing on
node sys1
```

For more information on how to replace coordinator disks, refer to the *Storage Foundation Cluster File System High Availability Administrator's Guide*.

To test the disks using vxfentsthdw utility

1. Make sure system-to-system communication functions properly.

   See "About configuring secure shell or remote shell communication modes before installing products" on page 318.

2. From one node, start the utility.

3. The script warns that the tests overwrite data on the disks. After you review the overview and the warning, confirm to continue the process and enter the node names.

   Warning: The tests overwrite and destroy data on the disks unless you use the `-r` option.

   ```
   ******** WARNING!!!!!!!! ********
   THIS UTILITY WILL DESTROY THE DATA ON THE DISK!!

   Do you still want to continue : [y/n] (default: n) y
   Enter the first node of the cluster: sys1
   Enter the second node of the cluster: sys2
   ```

4. Review the output as the utility performs the checks and reports its activities.

5. If a disk is ready for I/O fencing on each node, the utility reports success for each node. For example, the utility displays the following message for the node sys1.

   ```
   The disk is now ready to be configured for I/O Fencing on node
   sys1

   ALL tests on the disk /dev/sdx have PASSED
   The disk is now ready to be configured for I/O fencing on node
   sys1
   ```

6. Run the vxfentsthdw utility for each disk you intend to verify.

Note: Only dmp disk devices can be used as coordinator disks.

Refreshing keys or registrations on the existing coordination points for disk-based fencing using the installer

You must refresh registrations on the coordination points in the following scenarios:

- When the CoordPoint agent notifies VCS about the loss of registration on any of the existing coordination points.
- A planned refresh of registrations on coordination points when the cluster is online without having an application downtime on the cluster.

Registration loss may happen because of an accidental array restart, corruption of keys, or some other reason. If the coordination points lose the registrations of the cluster nodes, the cluster may panic when a network partition occurs.

Warning: Refreshing keys might cause the cluster to panic if a node leaves membership before the coordination points refresh is complete.

To refresh registrations on existing coordination points for disk-based I/O fencing using the installer

1 Start the installer with the `-fencing` option.

```
# /opt/VRTS/install/installer  -fencing
```

The installer starts with a copyright message and verifies the cluster information.

Note down the location of log files that you can access if there is a problem with the configuration process.

2 Confirm that you want to proceed with the I/O fencing configuration at the prompt.

The program checks that the local node running the script can communicate with the remote nodes and checks whether Storage Foundation Cluster File System High Availability 8.0 is configured properly.

3 Review the I/O fencing configuration options that the program presents. Type the number corresponding to refresh registrations or keys on the existing coordination points.

```
Select the fencing mechanism to be configured in this
Application Cluster [1-6,q]
```

4 Ensure that the disk group constitution that is used by the fencing module contains the same disks that are currently used as coordination disks.

5 Verify the coordination points.

```
For example,
Disk Group: fendg
Fencing disk policy: dmp
Fencing disks:
 emc_clariion0_62
 emc_clariion0_65
 emc_clariion0_66
```

Is this information correct? [y,n,q] **(y)**.

```
Successfully completed the vxfenswap operation
```

The keys on the coordination disks are refreshed.

6 Do you want to send the information about this installation to us to help improve installation in the future? [y,n,q,?] **(y)**.

7 Do you want to view the summary file? [y,n,q] **(n)**.

Setting up server-based I/O fencing using installer

You can configure server-based I/O fencing for the Storage Foundation Cluster File System High Availability cluster using the installer.

With server-based fencing, you can have the coordination points in your configuration as follows:

- Combination of CP servers and SCSI-3 compliant coordinator disks
- CP servers only
  Veritas also supports server-based fencing with a single highly available CP server that acts as a single coordination point.

See " About planning to configure I/O fencing" on page 30.

See "Recommended CP server configurations" on page 35.

This section covers the following example procedures:

Mix of CP servers and coordinator disks	See "To configure server-based fencing for the Storage Foundation Cluster File System High Availability cluster (one CP server and two coordinator disks)" on page 94.
Single CP server	See "To configure server-based fencing for the Storage Foundation Cluster File System High Availability cluster" on page 98.

To configure server-based fencing for the Storage Foundation Cluster File System High Availability cluster (one CP server and two coordinator disks)

1. Depending on the server-based configuration model in your setup, make sure of the following:
   - CP servers are configured and are reachable from the Storage Foundation Cluster File System High Availability cluster. The Storage Foundation Cluster File System High Availability cluster is also referred to as the application cluster or the client cluster.
     See "Setting up the CP server" on page 38.
   - The coordination disks are verified for SCSI3-PR compliance.
     See "Checking shared disks for I/O fencing" on page 88.
2. Start the installer with the `-fencing` option.

   ```
   # /opt/VRTS/install/installer  -fencing
   ```

   The installer starts with a copyright message and verifies the cluster information.

   Note the location of log files which you can access in the event of any problem with the configuration process.
3. Confirm that you want to proceed with the I/O fencing configuration at the prompt.

   The program checks that the local node running the script can communicate with remote nodes and checks whether Storage Foundation Cluster File System High Availability 8.0 is configured properly.
4. Review the I/O fencing configuration options that the program presents. Type **1** to configure server-based I/O fencing.

   ```
   Select the fencing mechanism to be configured in this
   Application Cluster [1-3,b,q] 1
   ```
5. Make sure that the storage supports SCSI3-PR, and answer y at the following prompt.

   ```
   Does your storage environment support SCSI3 PR? [y,n,q] (y)
   ```
6. Provide the following details about the coordination points at the installer prompt:
   - Enter the total number of coordination points including both servers and disks. This number should be at least 3.

     ```
     Enter the total number of co-ordination points including both
     Coordination Point servers and disks: [b] (3)
     ```

- Enter the total number of coordinator disks among the coordination points.

```
Enter the total number of disks among these:
[b] (0) 2
```

7 Provide the following CP server details at the installer prompt:

- Enter the total number of virtual IP addresses or the total number of fully qualified host names for each of the CP servers.

```
How many IP addresses would you like to use to communicate
 to Coordination Point Server #1?: [b,q,?] (1) 1
```

- Enter the virtual IP addresses or the fully qualified host name for each of the CP servers. The installer assumes these values to be identical as viewed from all the application cluster nodes.

```
Enter the Virtual IP address or fully qualified host name #1
for the HTTPS Coordination Point Server #1:
[b] 10.209.80.197
```

The installer prompts for this information for the number of virtual IP addresses you want to configure for each CP server.

- Enter the port that the CP server would be listening on.

```
Enter the port that the coordination point server  10.209.80.197
would be listening on or accept the default port
suggested: [b] (443)
```

8 Provide the following coordinator disks-related details at the installer prompt:

- Choose the coordinator disks from the list of available disks that the installer displays. Ensure that the disk you choose is available from all the Storage Foundation Cluster File System High Availability (application cluster) nodes. The number of times that the installer asks you to choose the disks depends on the information that you provided in step 6. For example, if you had chosen to configure two coordinator disks, the installer asks you to choose the first disk and then the second disk:

```
Select disk number 1 for co-ordination point

1) sdx
2) sdy
3) sdz
```

```
Please enter a valid disk which is available from all the
cluster nodes for co-ordination point [1-3,q] 1
```

- If you have not already checked the disks for SCSI-3 PR compliance in step 1, check the disks now.
  The installer displays a message that recommends you to verify the disks in another window and then return to this configuration procedure.
  Press Enter to continue, and confirm your disk selection at the installer prompt.
- Enter a disk group name for the coordinator disks or accept the default.

```
Enter the disk group name for coordinating disk(s):
[b] (vxfencoorddg)
```

9 Verify and confirm the coordination points information for the fencing configuration.

For example:

```
Total number of coordination points being used: 3
Coordination Point Server ([VIP or FQHN]:Port):
    1. 10.209.80.197 ([10.209.80.197]:443)
SCSI-3 disks:
    1. sdx
    2. sdy
Disk Group name for the disks in customized fencing: vxfencoorddg
Disk policy used for customized fencing: dmp
```

The installer initializes the disks and the disk group and deports the disk group on the Storage Foundation Cluster File System High Availability (application cluster) node.

10 Verify and confirm the I/O fencing configuration information.

```
CPS Admin utility location: /opt/VRTScps/bin/cpsadm
Cluster ID: 2122
Cluster Name: clus1
UUID for the above cluster: {ae5e589a-1dd1-11b2-dd44-00144f79240c}
```

11 Review the output as the installer updates the application cluster information on each of the CP servers to ensure connectivity between them. The installer then populates the `/etc/vxfenmode` file with the appropriate details in each of the application cluster nodes.

```
Updating client cluster information on Coordination Point Server 10.209.80.197

Adding the client cluster to the Coordination Point Server 10.209.80.197 .......... Done

Registering client node sys1 with Coordination Point Server 10.209.80.197...... Done
Adding CPClient user for communicating to Coordination Point Server 10.209.80.197 .... Done
Adding cluster clus1 to the CPClient user on Coordination Point Server 10.209.80.197 .. Done

Registering client node sys2 with Coordination Point Server 10.209.80.197 ..... Done
Adding CPClient user for communicating to Coordination Point Server 10.209.80.197 .... Done
Adding cluster clus1 to the CPClient user on Coordination Point Server 10.209.80.197 ..Done

Updating /etc/vxfenmode file on sys1 ................................. Done
Updating /etc/vxfenmode file on sys2 ......... ....................... Done
```

See "About I/O fencing configuration files" on page 308.

12 Review the output as the installer stops and restarts the VCS and the fencing processes on each application cluster node, and completes the I/O fencing configuration.

13 Configure the CP agent on the Storage Foundation Cluster File System High Availability (application cluster). The Coordination Point Agent monitors the registrations on the coordination points.

```
Do you want to configure Coordination Point Agent on
the client cluster? [y,n,q] (y)

Enter a non-existing name for the service group for
Coordination Point Agent: [b] (vxfen)
```

14 Additionally the coordination point agent can also monitor changes to the Coordinator Disk Group constitution such as a disk being accidently deleted from the Coordinator Disk Group. The frequency of this detailed monitoring can be tuned with the LevelTwoMonitorFreq attribute. For example, if you set this attribute to 5, the agent will monitor the Coordinator Disk Group constitution every five monitor cycles.

Note that for the LevelTwoMonitorFreq attribute to be applicable there must be disks as part of the Coordinator Disk Group.

```
Enter the value of the LevelTwoMonitorFreq attribute: (5)
```

15 Enable auto refresh of coordination points.

```
Do you want to enable auto refresh of coordination points
if registration keys are missing
on any of them? [y,n,q,b,?]  (n)
```

16 Note the location of the configuration log files, summary files, and response files that the installer displays for later use.

17 Verify the fencing configuration using:

```
# vxfenadm -d
```

18 Verify the list of coordination points.

```
# vxfenconfig -l
```

To configure server-based fencing for the Storage Foundation Cluster File System High Availability cluster

1 Make sure that the CP server is configured and is reachable from the Storage Foundation Cluster File System High Availability cluster. The Storage Foundation Cluster File System High Availability cluster is also referred to as the application cluster or the client cluster.

2 See "Setting up the CP server" on page 38.

3 Start the installer with `-fencing` option.

```
# /opt/VRTS/install/installer   -fencing
```

The installer starts with a copyright message and verifies the cluster information.

Note the location of log files which you can access in the event of any problem with the configuration process.

4 Confirm that you want to proceed with the I/O fencing configuration at the prompt.

The program checks that the local node running the script can communicate with remote nodes and checks whether Storage Foundation Cluster File System High Availability 8.0 is configured properly.

5 Review the I/O fencing configuration options that the program presents. Type **1** to configure server-based I/O fencing.

```
Select the fencing mechanism to be configured in this
Application Cluster [1-7,q] 1
```

6 Make sure that the storage supports SCSI3-PR, and answer y at the following prompt.

```
Does your storage environment support SCSI3 PR? [y,n,q] (y)
```

7 Enter the total number of coordination points as **1**.

```
Enter the total number of co-ordination points including both
Coordination Point servers and disks: [b] (3) 1
```

Read the installer warning carefully before you proceed with the configuration.

8 Provide the following CP server details at the installer prompt:

- Enter the total number of virtual IP addresses or the total number of fully qualified host names for each of the CP servers.

  ```
  How many IP addresses would you like to use to communicate
  to Coordination Point Server #1? [b,q,?] (1) 1
  ```

- Enter the virtual IP address or the fully qualified host name for the CP server. The installer assumes these values to be identical as viewed from all the application cluster nodes.

  ```
  Enter the Virtual IP address or fully qualified host name
  #1 for the Coordination Point Server #1:
  [b] 10.209.80.197
  ```

  The installer prompts for this information for the number of virtual IP addresses you want to configure for each CP server.

- Enter the port that the CP server would be listening on.

  ```
  Enter the port in the range [49152, 65535] which the
  Coordination Point Server 10.209.80.197
  ```

```
would be listening on or simply accept the default
port suggested: [b] (443)
```

9 Verify and confirm the coordination points information for the fencing configuration.

For example:

```
Total number of coordination points being used: 1
Coordination Point Server ([VIP or FQHN]:Port):
    1. 10.209.80.197 ([10.209.80.197]:443)
```

10 Verify and confirm the I/O fencing configuration information.

```
CPS Admin utility location: /opt/VRTScps/bin/cpsadm
Cluster ID: 2122
Cluster Name: clus1
UUID for the above cluster: {ae5e589a-1dd1-11b2-dd44-00144f79240c}
```

11 Review the output as the installer updates the application cluster information on each of the CP servers to ensure connectivity between them. The installer then populates the `/etc/vxfenmode` file with the appropriate details in each of the application cluster nodes.

The installer also populates the `/etc/vxfenmode` file with the entry `single_cp=1` for such single CP server fencing configuration.

```
Updating client cluster information on Coordination Point Server 10.209.80.197

Adding the client cluster to the Coordination Point Server 10.209.80.197 .......... Done

Registering client node sys1 with Coordination Point Server 10.209.80.197...... Done
Adding CPClient user for communicating to Coordination Point Server 10.209.80.197 .... Done
Adding cluster clus1 to the CPClient user on Coordination Point Server 10.209.80.197 .. Done

Registering client node sys2 with Coordination Point Server 10.209.80.197 ..... Done
Adding CPClient user for communicating to Coordination Point Server 10.209.80.197 .... Done
Adding cluster clus1 to the CPClient user on Coordination Point Server 10.209.80.197 .. Done

Updating /etc/vxfenmode file on sys1 ................................. Done
Updating /etc/vxfenmode file on sys2 ......... ....................... Done
```

See “About I/O fencing configuration files” on page 308.

12 Review the output as the installer stops and restarts the VCS and the fencing processes on each application cluster node, and completes the I/O fencing configuration.

13 Configure the CP agent on the Storage Foundation Cluster File System High Availability (application cluster).

```
Do you want to configure Coordination Point Agent on the
client cluster? [y,n,q] (y)

Enter a non-existing name for the service group for
Coordination Point Agent: [b] (vxfen)
```

14 Enable auto refresh of coordination points.

```
Do you want to enable auto refresh of coordination points
if registration keys are missing
on any of them? [y,n,q,b,?]  (n)
```

15 Note the location of the configuration log files, summary files, and response files that the installer displays for later use.

Refreshing keys or registrations on the existing coordination points for server-based fencing using the installer

You must refresh registrations on the coordination points in the following scenarios:

- When the CoordPoint agent notifies VCS about the loss of registration on any of the existing coordination points.
- A planned refresh of registrations on coordination points when the cluster is online without having an application downtime on the cluster.

Registration loss might occur because of an accidental array restart, corruption of keys, or some other reason. If the coordination points lose registrations of the cluster nodes, the cluster might panic when a network partition occurs.

Warning: Refreshing keys might cause the cluster to panic if a node leaves membership before the coordination points refresh is complete.

To refresh registrations on existing coordination points for server-based I/O fencing using the installer

1. Start the installer with the `-fencing` option.

   ```
   # /opt/VRTS/install/installer  -fencing
   ```

   The installer starts with a copyright message and verifies the cluster information.

   Note the location of log files that you can access if there is a problem with the configuration process.

2. Confirm that you want to proceed with the I/O fencing configuration at the prompt.

   The program checks that the local node running the script can communicate with the remote nodes and checks whether Storage Foundation Cluster File System High Availability 8.0 is configured properly.

3. Review the I/O fencing configuration options that the program presents. Type the number corresponding to the option that suggests to refresh registrations or keys on the existing coordination points.

   ```
   Select the fencing mechanism to be configured in this
   Application Cluster [1-7,q] 6
   ```

4. Ensure that the `/etc/vxfentab` file contains the same coordination point servers that are currently used by the fencing module.

   Also, ensure that the disk group mentioned in the `/etc/vxfendg` file contains the same disks that are currently used by the fencing module as coordination disks.

5. Verify the coordination points.

   ```
   For example,
   Total number of coordination points being used: 3
   Coordination Point Server ([VIP or FQHN]:Port):
         1. 10.198.94.146 ([10.198.94.146]:443)
         2. 10.198.94.144 ([10.198.94.144]:443)
   SCSI-3 disks:
         1. emc_clariion0_61
   Disk Group name for the disks in customized fencing: vxfencoorddg
   Disk policy used for customized fencing: dmp
   ```

6 Is this information correct? [y,n,q] **(y)**

```
Updating client cluster information on Coordination Point Server
 IPaddress

Successfully completed the vxfenswap operation
```

The keys on the coordination disks are refreshed.

7 Do you want to send the information about this installation to us to help improve installation in the future? [y,n,q,?] **(y)**.

8 Do you want to view the summary file? [y,n,q] **(n)**.

Setting the order of existing coordination points for server-based fencing using the installer

This section describes the reasons, benefits, considerations, and the procedure to set the order of the existing coordination points for server-based fencing.

About deciding the order of existing coordination points

You can decide the order in which coordination points can participate in a race during a network partition. In a network partition scenario, I/O fencing attempts to contact coordination points for membership arbitration based on the order that is set in the `vxfentab` file.

When I/O fencing is not able to connect to the first coordination point in the sequence it goes to the second coordination point and so on. To avoid a cluster panic, the surviving subcluster must win majority of the coordination points. So, the order must begin with the coordination point that has the best chance to win the race and must end with the coordination point that has the least chance to win the race.

For fencing configurations that use a mix of coordination point servers and coordination disks, you can specify either coordination point servers before coordination disks or disks before servers.

Note: Disk-based fencing does not support setting the order of existing coordination points.

Considerations to decide the order of coordination points

- Choose the coordination points based on their chances to gain membership on the cluster during the race and hence gain control over a network partition. In effect, you have the ability to save a partition.

- First in the order must be the coordination point that has the best chance to win the race. The next coordination point you list in the order must have relatively lesser chance to win the race. Complete the order such that the last coordination point has the least chance to win the race.

Setting the order of existing coordination points using the installer

To set the order of existing coordination points

1 Start the installer with `-fencing` option.

```
# /opt/VRTS/install/installer  -fencing
```

The installer starts with a copyright message and verifies the cluster information.

Note the location of log files that you can access if there is a problem with the configuration process.

2 Confirm that you want to proceed with the I/O fencing configuration at the prompt.

The program checks that the local node running the script can communicate with remote nodes and checks whether Storage Foundation Cluster File System High Availability 8.0 is configured properly.

3 Review the I/O fencing configuration options that the program presents. Type the number corresponding to the option that suggests to set the order of existing coordination points.

For example:

```
Select the fencing mechanism to be configured in this
Application Cluster [1-7,q] 7


Installer will ask the new order of existing coordination points.
Then it will call vxfenswap utility to commit the
coordination points change.
```

Warning: The cluster might panic if a node leaves membership before the coordination points change is complete.

4 Review the current order of coordination points.

```
Current coordination points order:
(Coordination disks/Coordination Point Server)
Example,
1)  /dev/vx/rdmp/emc_clariion0_65,/dev/vx/rdmp/emc_clariion0_66,
/dev/vx/rdmp/emc_clariion0_62
2)  [10.198.94.144]:443
3)  [10.198.94.146]:443
b)  Back to previous menu
```

5 Enter the new order of the coordination points by the numbers and separate the order by space [1-3,b,q] **3 1 2**.

```
New coordination points order:
(Coordination disks/Coordination Point Server)
Example,
1) [10.198.94.146]:443
2) /dev/vx/rdmp/emc_clariion0_65,/dev/vx/rdmp/emc_clariion0_66,
/dev/vx/rdmp/emc_clariion0_62
3) [10.198.94.144]:443
```

6 Is this information correct? [y,n,q] **(y)**.

```
Preparing vxfenmode.test file on all systems...
Running vxfenswap...
Successfully completed the vxfenswap operation
```

7 Do you want to send the information about this installation to us to help improve installation in the future? [y,n,q,?] **(y)**.

8 Do you want to view the summary file? [y,n,q] **(n)**.

9 Verify that the value of `vxfen_honor_cp_order` specified in the `/etc/vxfenmode` file is set to **1**.

```
For example,
vxfen_mode=customized
vxfen_mechanism=cps
port=443
scsi3_disk_policy=dmp
cps1=[10.198.94.146]
vxfendg=vxfencoorddg
cps2=[10.198.94.144]
vxfen_honor_cp_order=1
```

10 Verify that the coordination point order is updated in the output of the `vxfenconfig -l` command.

```
For example,
I/O Fencing Configuration Information:
======================================

single_cp=0
[10.198.94.146]:443 {e7823b24-1dd1-11b2-8814-2299557f1dc0}
/dev/vx/rdmp/emc_clariion0_65  60060160A38B1600386FD87CA8FDDD11
/dev/vx/rdmp/emc_clariion0_66  60060160A38B1600396FD87CA8FDDD11
/dev/vx/rdmp/emc_clariion0_62  60060160A38B16005AA00372A8FDDD11
[10.198.94.144]:443 {01f18460-1dd2-11b2-b818-659cbc6eb360}
```

Setting up non-SCSI-3 I/O fencing in virtual environments using installer

If you have installed Veritas InfoScale Enterprise in virtual environments that do not support SCSI-3 PR-compliant storage, you can configure non-SCSI-3 fencing.

To configure I/O fencing using the installer in a non-SCSI-3 PR-compliant setup

1. Start the installer with `-fencing` option.

   ```
   # /opt/VRTS/install/installer  -fencing
   ```

   The installer starts with a copyright message and verifies the cluster information.

2. Confirm that you want to proceed with the I/O fencing configuration at the prompt.

   The program checks that the local node running the script can communicate with remote nodes and checks whether Storage Foundation Cluster File System High Availability 8.0 is configured properly.

3. For server-based fencing, review the I/O fencing configuration options that the program presents. Type **1** to configure server-based I/O fencing.

   ```
   Select the fencing mechanism to be configured in this
   Application Cluster
   [1-7,q] 1
   ```

4. Enter **n** to confirm that your storage environment does not support SCSI-3 PR.

   ```
   Does your storage environment support SCSI3 PR?
   [y,n,q] (y) n
   ```

5. Confirm that you want to proceed with the non-SCSI-3 I/O fencing configuration at the prompt.

6. For server-based fencing, enter the number of CP server coordination points you want to use in your setup.

7. For server-based fencing, enter the following details for each CP server:

   - Enter the virtual IP address or the fully qualified host name.
   - Enter the port address on which the CP server listens for connections. The default value is 443. You can enter a different port address. Valid values are between 49152 and 65535.

   The installer assumes that these values are identical from the view of the SFCFSHA cluster nodes that host the applications for high availability.

8. For server-based fencing, verify and confirm the CP server information that you provided.

9. Verify and confirm the SFCFSHA cluster configuration information.

   Review the output as the installer performs the following tasks:

- Updates the CP server configuration files on each CP server with the following details for only server-based fencing, :
  - Registers each node of the SFCFSHA cluster with the CP server.
  - Adds CP server user to the CP server.
  - Adds SFCFSHA cluster to the CP server user.
- Updates the following configuration files on each node of the SFCFSHA cluster
  - `/etc/vxfenmode` file
  - `/etc/vxenviron` file
  - `/etc/sysconfig/vxfen` file
  - `/etc/llttab` file
  - `/etc/vxfentab` (only for server-based fencing)

10 Review the output as the installer stops Storage Foundation Cluster File System High Availability on each node, starts I/O fencing on each node, updates the VCS configuration file main.cf, and restarts Storage Foundation Cluster File System High Availability with non-SCSI-3 fencing.

For server-based fencing, confirm to configure the CP agent on the SFCFSHA cluster.

11 Confirm whether you want to send the installation information to us.

12 After the installer configures I/O fencing successfully, note the location of summary, log, and response files that installer creates.

The files provide useful information which can assist you with the configuration, and can also assist future configurations.

Setting up majority-based I/O fencing using installer

You can configure majority-based fencing for the cluster using the installer .

Perform the following steps to confgure majority-based I/O fencing

1. Start the installer with the -fencing option.

   ```
   # /opt/VRTS/install/installer  -fencing
   ```

   Where *version* is the specific release version. The installer starts with a copyright message and verifies the cluster information.

   Note: Make a note of the log file location which you can access in the event of any issues with the configuration process.

2. Confirm that you want to proceed with the I/O fencing configuration at the prompt. The program checks that the local node running the script can communicate with remote nodes and checks whether SFCFSHA is configured properly.
3. Review the I/O fencing configuration options that the program presents. Type **3** to configure majority-based I/O fencing.

   ```
   Select the fencing mechanism to be configured in this
   Application Cluster [1-7,b,q] 3
   ```

   Note: The installer will ask the following question. Does your storage environment support SCSI3 PR? [y,n,q,?] Input 'y' if your storage environment supports SCSI3 PR. Other alternative will result in installer configuring non-SCSI3 fencing(NSF).

4. The installer then populates the /etc/vxfenmode file with the appropriate details in each of the application cluster nodes.

   ```
   Updating /etc/vxfenmode file on sys1 ..... Done
   Updating /etc/vxfenmode file on sys2 ..... Done
   ```

5. Review the output as the installer stops and restarts the VCS and the fencing processes on each application cluster node, and completes the I/O fencing configuration.
6. Note the location of the configuration log files, summary files, and response files that the installer displays for later use.
7. Verify the fencing configuration.

   ```
   # vxfenadm -d
   ```

Enabling or disabling the preferred fencing policy

You can enable or disable the preferred fencing feature for your I/O fencing configuration.

You can enable preferred fencing to use system-based race policy, group-based race policy, or site-based policy. If you disable preferred fencing, the I/O fencing configuration uses the default count-based race policy.

Preferred fencing is not applicable to majority-based I/O fencing.

See "About preferred fencing" on page 23.

To enable preferred fencing for the I/O fencing configuration

1. Make sure that the cluster is running with I/O fencing set up.

   ```
   # vxfenadm -d
   ```

2. Make sure that the cluster-level attribute UseFence has the value set to SCSI3.

   ```
   # haclus -value UseFence
   ```

3. To enable system-based race policy, perform the following steps:

   - Make the VCS configuration writable.

     ```
     # haconf -makerw
     ```

   - Set the value of the cluster-level attribute PreferredFencingPolicy as System.

     ```
     # haclus -modify PreferredFencingPolicy System
     ```

   - Set the value of the system-level attribute FencingWeight for each node in the cluster.
     For example, in a two-node cluster, where you want to assign sys1 five times more weight compared to sys2, run the following commands:

     ```
     # hasys -modify sys1 FencingWeight 50
     # hasys -modify sys2 FencingWeight 10
     ```

   - Save the VCS configuration.

     ```
     # haconf -dump -makero
     ```

   - Verify fencing node weights using:

     ```
     # vxfenconfig -a
     ```

4 To enable group-based race policy, perform the following steps:

- Make the VCS configuration writable.

  ```
  # haconf -makerw
  ```

- Set the value of the cluster-level attribute PreferredFencingPolicy as Group.

  ```
  # haclus -modify PreferredFencingPolicy Group
  ```

- Set the value of the group-level attribute Priority for each service group. For example, run the following command:

  ```
  # hagrp -modify service_group Priority 1
  ```

  Make sure that you assign a parent service group an equal or lower priority than its child service group. In case the parent and the child service groups are hosted in different subclusters, then the subcluster that hosts the child service group gets higher preference.

- Save the VCS configuration.

  ```
  # haconf -dump -makero
  ```

5 To enable site-based race policy, perform the following steps:

- Make the VCS configuration writable.

  ```
  # haconf -makerw
  ```

- Set the value of the cluster-level attribute PreferredFencingPolicy as Site.

  ```
  # haclus -modify PreferredFencingPolicy Site
  ```

- Set the value of the site-level attribute Preference for each site.

  ```
  For example,
  # hasite -modify Pune Preference 2
  ```

- Save the VCS configuration.

  ```
  # haconf -dump -makero
  ```

6 To view the fencing node weights that are currently set in the fencing driver, run the following command:

```
# vxfenconfig -a
```

To disable preferred fencing for the I/O fencing configuration

1. Make sure that the cluster is running with I/O fencing set up.

```
# vxfenadm -d
```

2. Make sure that the cluster-level attribute UseFence has the value set to SCSI3.

```
# haclus -value UseFence
```

3. To disable preferred fencing and use the default race policy, set the value of the cluster-level attribute PreferredFencingPolicy as Disabled.

```
# haconf -makerw
# haclus -modify PreferredFencingPolicy Disabled
# haconf -dump -makero
```

Chapter 6

Performing an automated SFCFSHA configuration using response files

This chapter includes the following topics:

- Configuring SFCFSHA using response files
- Response file variables to configure SFCFSHA
- Sample response file for SFCFSHA configuration

Configuring SFCFSHA using response files

Typically, you can use the response file that the installer generates after you perform SFCFSHA configuration on one cluster to configure SFCFSHA on other clusters.

To configure SFCFSHA using response files

1. Make sure the Veritas InfoScale Availability or Enterprise RPMs are installed on the systems where you want to configure SFCFSHA.
2. Copy the response file to one of the cluster systems where you want to configure SFCFSHA.

3 Edit the values of the response file variables as necessary.

To configure optional features, you must define appropriate values for all the response file variables that are related to the optional feature.

See "Response file variables to configure SFCFSHA" on page 114.

4 Start the configuration from the system to which you copied the response file. For example:

```
# /opt/VRTS/install/installer -responsefile
/tmp/response_file
```

Where `/tmp/response_file` is the response file's full path name.

Response file variables to configure SFCFSHA

Table 6-1 lists the response file variables that you can define to configure SFCFSHA.

Table 6-1 Response file variables specific to configuring SFCFSHA

Variable	List or Scalar	Description
CFG{opt}{configure}	Scalar	Performs the configuration if the RPMs are already installed. (Required) Set the value to 1 to configure SFCFSHA.
CFG{accepteula}	Scalar	Specifies whether you agree with EULA.pdf on the media. (Required)
CFG{activecomponent}	List	Defines the component to be configured. The value is SFCFSHA80 for SFCFSHA. (Required)

Table 6-1 Response file variables specific to configuring SFCFSHA *(continued)*

Variable	List or Scalar	Description
CFG{keys}{keyless} CFG{keys}{license}	List	`CFG{keys}{keyless}` gives a list of keyless keys to be registered on the system. `CFG{keys}{license}` gives a list of user defined keys to be registered on the system. (Optional)
CFG{systems}	List	List of systems on which the product is to be configured. (Required)
CFG{prod}	Scalar	Defines the product for operations. The value is ENTERPRISE80 for Veritas InfoScale Enterprise. (Required)
CFG{opt}{keyfile}	Scalar	Defines the location of an ssh keyfile that is used to communicate with all remote systems. (Optional)
CFG{opt}{rsh}	Scalar	Defines that *rsh* must be used instead of ssh as the communication method between systems. (Optional)
CFG{opt}{logpath}	Scalar	Mentions the location where the log files are to be copied. The default location is /opt/VRTS/install/logs. **Note:** The installer copies the response files and summary files also to the specified *logpath* location. (Optional)

Table 6-1 Response file variables specific to configuring SFCFSHA *(continued)*

Variable	List or Scalar	Description
CFG{uploadlogs}	Scalar	Defines a Boolean value 0 or 1. The value 1 indicates that the installation logs are uploaded to the Veritas website. The value 0 indicates that the installation logs are not uploaded to the Veritas website. (Optional)

Note that some optional variables make it necessary to define other optional variables. For example, all the variables that are related to the cluster service group (csgnic, csgvip, and csgnetmask) must be defined if any are defined. The same is true for the SMTP notification (smtpserver, smtprecp, and smtprsev), the SNMP trap notification (snmpport, snmpcons, and snmpcsev), and the Global Cluster Option (gconic, gcovip, and gconetmask).

Table 6-2 lists the response file variables that specify the required information to configure a basic Storage Foundation Cluster File System High Availability cluster.

Table 6-2 Response file variables specific to configuring a basic Storage Foundation Cluster File System High Availability cluster

Variable	List or Scalar	Description
CFG{donotreconfigurevcs}	Scalar	Defines if you need to re-configure VCS. (Optional)
CFG{donotreconfigurefencing}	Scalar	Defines if you need to re-configure fencing. (Optional)
CFG{vcs_clusterid}	Scalar	An integer between 0 and 65535 that uniquely identifies the cluster. (Required)
CFG{vcs_clustername}	Scalar	Defines the name of the cluster. (Required)

Table 6-2 Response file variables specific to configuring a basic Storage Foundation Cluster File System High Availability cluster *(continued)*

Variable	List or Scalar	Description
CFG{vcs_allowcomms}	Scalar	Indicates whether or not to start LLT and GAB when you set up a single-node cluster. The value can be 0 (do not start) or 1 (start). (Required)
CFG{fencingenabled}	Scalar	In a Storage Foundation Cluster File System High Availability configuration, defines if fencing is enabled. Valid values are 0 or 1. (Required)

Table 6-3 lists the response file variables that specify the required information to configure LLT over Ethernet.

Table 6-3 Response file variables specific to configuring private LLT over Ethernet

Variable	List or Scalar	Description
CFG{vcs_lltlink#} {"system"}	Scalar	Defines the NIC to be used for a private heartbeat link on each system. At least two LLT links are required per system (lltlink1 and lltlink2). You can configure up to four LLT links. You must enclose the system name within double quotes. (Required)

Table 6-3 Response file variables specific to configuring private LLT over Ethernet *(continued)*

Variable	List or Scalar	Description
CFG{vcs_lltlinklowpri#} {"system"}	Scalar	Defines a low priority heartbeat link. Typically, lltlinklowpri is used on a public network link to provide an additional layer of communication. If you use different media speed for the private NICs, you can configure the NICs with lesser speed as low-priority links to enhance LLT performance. For example, lltlinklowpri1, lltlinklowpri2, and so on. You must enclose the system name within double quotes. (Optional)

Table 6-4 lists the response file variables that specify the required information to configure LLT over UDP.

Table 6-4 Response file variables specific to configuring LLT over UDP

Variable	List or Scalar	Description
CFG{lltoverudp}=1	Scalar	Indicates whether to configure heartbeat link using LLT over UDP. (Required)
CFG{vcs_udplink<n>_address} {<sys1>}	Scalar	Stores the IP address (IPv4 or IPv6) that the heartbeat link uses on node1. You can have four heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective heartbeat links. (Required)

Table 6-4 Response file variables specific to configuring LLT over UDP *(continued)*

Variable	List or Scalar	Description
CFG {vcs_udplinklowpri<n>_address} {<sys1>}	Scalar	Stores the IP address (IPv4 or IPv6) that the low priority heartbeat link uses on node1. You can have four low priority heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective low priority heartbeat links. (Required)
CFG{vcs_udplink<n>_port} {<sys1>}	Scalar	Stores the UDP port (16-bit integer value) that the heartbeat link uses on node1. You can have four heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective heartbeat links. (Required)
CFG{vcs_udplinklowpri<n>_port} {<sys1>}	Scalar	Stores the UDP port (16-bit integer value) that the low priority heartbeat link uses on node1. You can have four low priority heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective low priority heartbeat links. (Required)
CFG{vcs_udplink<n>_netmask} {<sys1>}	Scalar	Stores the netmask (prefix for IPv6) that the heartbeat link uses on node1. You can have four heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective heartbeat links. (Required)

Table 6-4 Response file variables specific to configuring LLT over UDP *(continued)*

Variable	List or Scalar	Description
CFG {vcs_udplinklowpri<n>_netmask} {<sys1>}	Scalar	Stores the netmask (prefix for IPv6) that the low priority heartbeat link uses on node1. You can have four low priority heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective low priority heartbeat links. (Required)

Table 6-5 lists the response file variables that specify the required information to configure virtual IP for Storage Foundation Cluster File System High Availability cluster.

Table 6-5 Response file variables specific to configuring virtual IP for Storage Foundation Cluster File System High Availability cluster

Variable	List or Scalar	Description
CFG{vcs_csgnic} {system}	Scalar	Defines the NIC device to use on a system. You can enter 'all' as a system value if the same NIC is used on all systems. (Optional)
CFG{vcs_csgvip}	Scalar	Defines the virtual IP address for the cluster. (Optional)
CFG{vcs_csgnetmask}	Scalar	Defines the Netmask of the virtual IP address for the cluster. (Optional)

Table 6-6 lists the response file variables that specify the required information to configure the Storage Foundation Cluster File System High Availability cluster in secure mode.

Table 6-6 Response file variables specific to configuring Storage Foundation Cluster File System High Availability cluster in secure mode

Variable	List or Scalar	Description
CFG{vcs_eat_security}	Scalar	Specifies if the cluster is in secure enabled mode or not.
CFG{opt}{securityonenode}	Scalar	Specifies that the securityonenode option is being used.
CFG{securityonenode_menu}	Scalar	Specifies the menu option to choose to configure the secure cluster one at a time. ■ 1—Configure the first node ■ 2—Configure the other node
CFG{secusrgrps}	List	Defines the user groups which get read access to the cluster. List or scalar: list Optional or required: optional
CFG{rootsecusrgrps}	Scalar	Defines the read access to the cluster only for root and other users or user groups which are granted explicit privileges in VCS objects. (Optional)
CFG{security_conf_dir}	Scalar	Specifies the directory where the configuration files are placed.
CFG{opt}{security}	Scalar	Specifies that the security option is being used.
CFG{defaultaccess}	Scalar	Defines if the user chooses to grant read access to everyone. Optional or required: optional
CFG{vcs_eat_security_fips}	Scalar	Specifies that the enabled security is FIPS compliant.

Table 6-7 lists the response file variables that specify the required information to configure VCS users.

Table 6-7 Response file variables specific to configuring VCS users

Variable	List or Scalar	Description
CFG{vcs_userenpw}	List	List of encoded passwords for VCS users The value in the list can be "Administrators Operators Guests" **Note:** The order of the values for the vcs_userenpw list must match the order of the values in the vcs_username list. (Optional)
CFG{vcs_username}	List	List of names of VCS users (Optional)
CFG{vcs_userpriv}	List	List of privileges for VCS users **Note:** The order of the values for the vcs_userpriv list must match the order of the values in the vcs_username list. (Optional)

Table 6-8 lists the response file variables that specify the required information to configure VCS notifications using SMTP.

Table 6-8 Response file variables specific to configuring VCS notifications using SMTP

Variable	List or Scalar	Description
CFG{vcs_smtpserver}	Scalar	Defines the domain-based hostname (example: smtp.example.com) of the SMTP server to be used for web notification. (Optional)
CFG{vcs_smtprecp}	List	List of full email addresses (example: user@example.com) of SMTP recipients. (Optional)

Table 6-8 Response file variables specific to configuring VCS notifications using SMTP *(continued)*

Variable	List or Scalar	Description
CFG{vcs_smtprsev}	List	Defines the minimum severity level of messages (Information, Warning, Error, SevereError) that listed SMTP recipients are to receive. Note that the ordering of severity levels must match that of the addresses of SMTP recipients. (Optional)

Table 6-9 lists the response file variables that specify the required information to configure VCS notifications using SNMP.

Table 6-9 Response file variables specific to configuring VCS notifications using SNMP

Variable	List or Scalar	Description
CFG{vcs_snmpport}	Scalar	Defines the SNMP trap daemon port (default=162). (Optional)
CFG{vcs_snmpcons}	List	List of SNMP console system names (Optional)
CFG{vcs_snmpcsev}	List	Defines the minimum severity level of messages (Information, Warning, Error, SevereError) that listed SNMP consoles are to receive. Note that the ordering of severity levels must match that of the SNMP console system names. (Optional)

Table 6-10 lists the response file variables that specify the required information to configure Storage Foundation Cluster File System High Availability global clusters.

Table 6-10 Response file variables specific to configuring Storage Foundation Cluster File System High Availability global clusters

Variable	List or Scalar	Description
CFG{vcs_gconic} {system}	Scalar	Defines the NIC for the Virtual IP that the Global Cluster Option uses. You can enter 'all' as a system value if the same NIC is used on all systems. (Optional)
CFG{vcs_gcovip}	Scalar	Defines the virtual IP address to that the Global Cluster Option uses. (Optional)
CFG{vcs_gconetmask}	Scalar	Defines the Netmask of the virtual IP address that the Global Cluster Option uses. (Optional)

Sample response file for SFCFSHA configuration

The following example shows a response file for configuring SFCFSHA.

```
##############################################
#Auto generated sfcfsha responsefile #
##############################################


our %CFG;
$CFG{accepteula}=1;
$CFG{opt}{rsh}=1;
$CFG{opt}{trace}=0;
$CFG{vcs_allowcomms}=1;
$CFG{opt}{gco}=1;
$CFG{opt}{vr}=1;
$CFG{opt}{configure}=1;
$CFG{prod}="ENTERPRISE80";
$CFG{systems}=[ qw( sys1 sys2 ) ];
$CFG{activecomponent}=[ qw(SFCFSHA80) ];
$CFG{fencingenabled}=0;
$CFG{vm_newnames_file}{sys1}=0;
$CFG{vm_restore_cfg}{sys1}=0;
```

```
$CFG{vm_newnames_file}{sys2}=0;
$CFG{vm_restore_cfg}{sys2}=0;
$CFG{vcs_clusterid}=127;
$CFG{vcs_clustername}="uxrt6_lin";
$CFG{vcs_username}=[ qw(admin operator) ];
$CFG{vcs_userenpw}=[ qw(JlmElgLimHmmKumGlj
bQOsOUnVQoOUnTQsOSnUQuOUnPQtOS) ];
$CFG{vcs_userpriv}=[ qw(Administrators Operators) ];
$CFG{vcs_lltlink1}{sys1}="eth1";
$CFG{vcs_lltlink2}{sys1}="eth2";
$CFG{vcs_lltlink1}{sys2}="eth1";
$CFG{vcs_lltlink2}{sys2}="eth2";
$CFG{vcs_enabled}=1;
$CFG{opt}{logpath}="/opt/VRTS/install/logs/installer-xxxxxx/";

1;
```

Chapter 7

Performing an automated I/O fencing configuration using response files

This chapter includes the following topics:

- Configuring I/O fencing using response files
- Response file variables to configure disk-based I/O fencing
- Sample response file for configuring disk-based I/O fencing
- Configuring CP server using response files
- Response file variables to configure server-based I/O fencing
- Sample response file for configuring server-based I/O fencing
- Response file variables to configure non-SCSI-3 I/O fencing
- Sample response file for configuring non-SCSI-3 I/O fencing
- Response file variables to configure majority-based I/O fencing
- Sample response file for configuring majority-based I/O fencing

Configuring I/O fencing using response files

Typically, you can use the response file that the installer generates after you perform I/O fencing configuration to configure I/O fencing for Storage Foundation Cluster File System High Availability.

To configure I/O fencing using response files

1. Make sure that Storage Foundation Cluster File System High Availability is configured.
2. Based on whether you want to configure disk-based or server-based I/O fencing, make sure you have completed the preparatory tasks.

   See " About planning to configure I/O fencing" on page 30.
3. Copy the response file to one of the cluster systems where you want to configure I/O fencing.

   See "Sample response file for configuring disk-based I/O fencing" on page 130.

   See "Sample response file for configuring server-based I/O fencing" on page 136.

   See "Sample response file for configuring non-SCSI-3 I/O fencing" on page 138.

   See "Sample response file for configuring majority-based I/O fencing" on page 140.
4. Edit the values of the response file variables as necessary.

   See "Response file variables to configure disk-based I/O fencing" on page 127.

   See "Response file variables to configure server-based I/O fencing" on page 134.

   See "Response file variables to configure non-SCSI-3 I/O fencing" on page 137.

   See "Response file variables to configure majority-based I/O fencing" on page 139.
5. Start the configuration from the system to which you copied the response file. For example:

   ```
   # /opt/VRTS/install/installer
   -responsefile /tmp/response_file
   ```

   Where `/tmp/response_file` is the response file's full path name.

Response file variables to configure disk-based I/O fencing

Table 7-1 lists the response file variables that specify the required information to configure disk-based I/O fencing for SFCFSHA.

Table 7-1 Response file variables specific to configuring disk-based I/O fencing

Variable	List or Scalar	Description
CFG{opt}{fencing}	Scalar	Performs the I/O fencing configuration. (Required)
CFG{fencing_option}	Scalar	Specifies the I/O fencing configuration mode. ■ 1—Coordination Point Server-based I/O fencing ■ 2—Coordinator disk-based I/O fencing ■ 3—Disabled-based I/O fencing ■ 4—Online fencing migration ■ 5—Refresh keys/registrations on the existing coordination points ■ 6—Change the order of existing coordination points ■ 7—Majority-based fencing (Required)
CFG{fencing_dgname}	Scalar	Specifies the disk group for I/O fencing. (Optional) **Note:** You must define the fencing_dgname variable to use an existing disk group. If you want to create a new disk group, you must use both the fencing_dgname variable and the fencing_newdg_disks variable.
CFG{fencing_newdg_disks}	List	Specifies the disks to use to create a new disk group for I/O fencing. (Optional) **Note:** You must define the fencing_dgname variable to use an existing disk group. If you want to create a new disk group, you must use both the fencing_dgname variable and the fencing_newdg_disks variable.

Table 7-1 Response file variables specific to configuring disk-based I/O fencing *(continued)*

Variable	List or Scalar	Description
CFG{fencing_cpagent_monitor_freq}	Scalar	Specifies the frequency at which the Coordination Point Agent monitors for any changes to the Coordinator Disk Group constitution. **Note:** Coordination Point Agent can also monitor changes to the Coordinator Disk Group constitution such as a disk being accidently deleted from the Coordinator Disk Group. The frequency of this detailed monitoring can be tuned with the LevelTwoMonitorFreq attribute. For example, if you set this attribute to 5, the agent will monitor the Coordinator Disk Group constitution every five monitor cycles. If LevelTwoMonitorFreq attribute is not set, the agent will not monitor any changes to the Coordinator Disk Group. 0 means not to monitor the Coordinator Disk Group constitution.
CFG {fencing_config_cpagent}	Scalar	Enter '1' or '0' depending upon whether you want to configure the Coordination Point agent using the installer or not. Enter "0" if you do not want to configure the Coordination Point agent using the installer. Enter "1" if you want to use the installer to configure the Coordination Point agent.
CFG {fencing_cpagentgrp}	Scalar	Name of the service group which will have the Coordination Point agent resource as part of it. **Note:** This field is obsolete if the **fencing_config_cpagent** field is given a value of '0'.

Table 7-1 Response file variables specific to configuring disk-based I/O fencing *(continued)*

Variable	List or Scalar	Description
CFG{fencing_auto_refresh_reg}	Scalar	Enable the auto refresh of coordination points variable in case registration keys are missing on any of CP servers.

Sample response file for configuring disk-based I/O fencing

Review the disk-based I/O fencing response file variables and their definitions.

See "Response file variables to configure disk-based I/O fencing" on page 127.

```
# Configuration Values:
#
our %CFG;
$CFG{fencing_config_cpagent}=1;
$CFG{fencing_auto_refresh_reg}=1;
$CFG{fencing_cpagent_monitor_freq}=5;
$CFG{fencing_cpagentgrp}="vxfen";
$CFG{fencing_dgname}="fencingdg1";
$CFG{fencing_newdg_disks}=[ qw(emc_clariion0_155
 emc_clariion0_162 emc_clariion0_163) ];
$CFG{fencing_option}=2;
$CFG{opt}{configure}=1;
$CFG{opt}{fencing}=1;

$CFG{prod}="ENTERPRISE80";

$CFG{activecomponent}="SFRAC80";
$CFG{systems}=[ qw(sys1sys2)];
$CFG{vcs_clusterid}=32283;
$CFG{vcs_clustername}="clus1";
1;
```

Configuring CP server using response files

You can configure a CP server using a generated responsefile.

On a single node VCS cluster:

- Run the `installer` command with the responsefile option to configure the CP server on a single node VCS cluster.

```
# /opt/VRTS/install/installer -responsefile '/tmp/sample1.res'
```

On a SFHA cluster:

- Run the `installer` command with the responsefile option to configure the CP server on a SFHA cluster.

```
# /opt/VRTS/install/installer -responsefile '/tmp/sample1.res'
```

Response file variables to configure CP server

Table 7-2 describes the response file variables to configure CP server.

Table 7-2 describes response file variables to configure CP server

Variable	List or Scalar	Description
CFG{opt}{configcps}	Scalar	This variable performs CP server configuration task
CFG{cps_singlenode_config}	Scalar	This variable describes if the CP server will be configured on a singlenode VCS cluster
CFG{cps_sfha_config}	Scalar	This variable describes if the CP server will be configured on a SFHA cluster
CFG{cps_unconfig}	Scalar	This variable describes if the CP server will be unconfigured
CFG{cpsname}	Scalar	This variable describes the name of the CP server
CFG{cps_db_dir}	Scalar	This variable describes the absolute path of CP server database
CFG{cps_reuse_cred}	Scalar	This variable describes if reusing the existing credentials for the CP server
CFG{cps_https_vips}	List	This variable describes the virtual IP addresses for the CP server configured for HTTPS-based communication

Table 7-2 describes response file variables to configure CP server *(continued)*

Variable	List or Scalar	Description
CFG{cps_https_ports}	List	This variable describes the port number for the virtual IP addresses for the CP server configured for HTTPS-based communication
CFG{cps_nic_list}{cpsvip<n>}	List	This variable describes the NICs of the systems for the virtual IP address
CFG{cps_netmasks}	List	This variable describes the netmasks for the virtual IP addresses
CFG{cps_prefix_length}	List	This variable describes the prefix length for the virtual IP addresses
CFG{cps_network_hosts}{cpsnic<n>}	List	This variable describes the network hosts for the NIC resource
CFG{cps_vip2nicres_map}{<vip>}	Scalar	This variable describes the NIC resource to associate with the virtual IP address
CFG{cps_diskgroup}	Scalar	This variable describes the disk group for the CP server database
CFG{cps_volume}	Scalar	This variable describes the volume for the CP server database
CFG{cps_newdg_disks}	List	This variable describes the disks to be used to create a new disk group for the CP server database
CFG{cps_newvol_volsize}	Scalar	This variable describes the volume size to create a new volume for the CP server database
CFG{cps_delete_database}	Scalar	This variable describes if deleting the database of the CP server during the unconfiguration
CFG{cps_delete_config_log}	Scalar	This variable describes if deleting the config files and log files of the CP server during the unconfiguration
CFG{cps_reconfig}	Scalar	This variable defines if the CP server will be reconfigured

Sample response file for configuring the CP server on single node VCS cluster

Review the response file variables and their definitions.

See Table 7-2 on page 131.

```
# Configuration Values:
#
our %CFG;

$CFG{cps_db_dir}="/etc/VRTScps/db";
$CFG{cps_https_ports}=[ 443 ];
$CFG{cps_https_vips}=[ "192.168.59.77" ];
$CFG{cps_netmasks}=[ "255.255.248.0" ];
$CFG{cps_network_hosts}{cpsnic1}=
[ "10.200.117.70" ];
$CFG{cps_nic_list}{cpsvip1}=[ "en0" ];
$CFG{cps_singlenode_config}=1;
$CFG{cps_vip2nicres_map}{"192.168.59.77"}=1;
$CFG{cpsname}="cps1";
$CFG{opt}{configcps}=1;
$CFG{opt}{configure}=1;
$CFG{opt}{noipc}=1;
$CFG{opt}{redirect}=1;
$CFG{prod}="AVAILABILITY80";
$CFG{systems}=[ "linux1" ];
$CFG{vcs_clusterid}=23172;
$CFG{vcs_clustername}="clus72";

1;
```

Sample response file for configuring the CP server on SFHA cluster

Review the response file variables and their definitions.

See Table 7-2 on page 131.

```
#
# Configuration Values:
#
our %CFG;
```

```
$CFG{cps_db_dir}="/cpsdb";
$CFG{cps_diskgroup}="cps_dg1";
$CFG{cps_https_ports}=[ qw(50006 50007) ];
$CFG{cps_https_vips}=[ qw(10.198.90.6 10.198.90.7) ];
$CFG{cps_netmasks}=[ qw(255.255.248.0 255.255.248.0 255.255.248.0) ];
$CFG{cps_network_hosts}{cpsnic1}=[ qw(10.198.88.18) ];
$CFG{cps_network_hosts}{cpsnic2}=[ qw(10.198.88.18) ];
$CFG{cps_newdg_disks}=[ qw(emc_clariion0_249) ];
$CFG{cps_newvol_volsize}=10;
$CFG{cps_nic_list}{cpsvip1}=[ qw(eth0 eth0) ];
$CFG{cps_sfha_config}=1;
$CFG{cps_vip2nicres_map}{"10.198.90.6"}=1;
$CFG{cps_volume}="volcps";
$CFG{cpsname}="cps1";
$CFG{opt}{configcps}=1;
$CFG{opt}{configure}=1;
$CFG{opt}{noipc}=1;

$CFG{prod}="ENTERPRISE80";

$CFG{systems}=[ qw(sys1 sys2) ];
$CFG{vcs_clusterid}=49604;
$CFG{vcs_clustername}="sfha2233";

1;
```

Response file variables to configure server-based I/O fencing

You can use a coordination point server-based fencing response file to configure server-based customized I/O fencing.

Table 7-3 lists the fields in the response file that are relevant for server-based customized I/O fencing.

Table 7-3 Coordination point server (CP server) based fencing response file definitions

Response file field	Definition
CFG {fencing_config_cpagent}	Enter '1' or '0' depending upon whether you want to configure the Coordination Point agent using the installer or not. Enter "0" if you do not want to configure the Coordination Point agent using the installer. Enter "1" if you want to use the installer to configure the Coordination Point agent.
CFG {fencing_cpagentgrp}	Name of the service group which will have the Coordination Point agent resource as part of it. **Note:** This field is obsolete if the `fencing_config_cpagent` field is given a value of '0'.
CFG {fencing_cps}	Virtual IP address or Virtual hostname of the CP servers.
CFG {fencing_reusedg}	This response file field indicates whether to reuse an existing DG name for the fencing configuration in customized fencing (CP server and coordinator disks). Enter either a "1" or "0". Entering a "1" indicates reuse, and entering a "0" indicates do not reuse. When reusing an existing DG name for the mixed mode fencing configuration. you need to manually add a line of text , such as "$CFG{fencing_reusedg}=0" or "$CFG{fencing_reusedg}=1" before proceeding with a silent installation.
CFG {fencing_dgname}	The name of the disk group to be used in the customized fencing, where at least one disk is being used.

Table 7-3 Coordination point server (CP server) based fencing response file definitions *(continued)*

Response file field	Definition
CFG {fencing_disks}	The disks being used as coordination points if any.
CFG {fencing_ncp}	Total number of coordination points being used, including both CP servers and disks.
CFG {fencing_ndisks}	The number of disks being used.
CFG {fencing_cps_vips}	The virtual IP addresses or the fully qualified host names of the CP server.
CFG {fencing_cps_ports}	The port that the virtual IP address or the fully qualified host name of the CP server listens on.
CFG{fencing_option}	Specifies the I/O fencing configuration mode. ■ 1—Coordination Point Server-based I/O fencing ■ 2—Coordinator disk-based I/O fencing ■ 3—Disabled-based I/O fencing ■ 4—Online fencing migration ■ 5—Refresh keys/registrations on the existing coordination points ■ 6—Change the order of existing coordination points ■ 7—Majority-based fencing (Required)
CFG{fencing_auto_refresh_reg}	Enable this variable if registration keys are missing on any of the CP servers.

Sample response file for configuring server-based I/O fencing

The following is a sample response file used for server-based I/O fencing:

```
$CFG{fencing_config_cpagent}=0;
$CFG{fencing_cps}=[ qw(10.200.117.145) ];
```

```
$CFG{fencing_cps_vips}{"10.200.117.145"}=[ qw(10.200.117.145) ];
$CFG{fencing_dgname}="vxfencoorddg";
$CFG{fencing_disks}=[ qw(emc_clariion0_37 emc_clariion0_13) ];
$CFG{fencing_scsi3_disk_policy}="dmp";
$CFG{fencing_ncp}=3;
$CFG{fencing_ndisks}=2;
$CFG{fencing_cps_ports}{"10.200.117.145"}=443;
$CFG{fencing_reusedg}=1;
$CFG{opt}{configure}=1;
$CFG{opt}{fencing}=1;
$CFG{prod}="ENTERPRISE80";
$CFG{systems}=[ qw(sys1 sys2) ];
$CFG{vcs_clusterid}=1256;
$CFG{vcs_clustername}="clus1";
$CFG{fencing_option}=1;
```

Response file variables to configure non-SCSI-3 I/O fencing

Table 7-4 lists the fields in the response file that are relevant for non-SCSI-3 I/O fencing.

See "About I/O fencing for Storage Foundation Cluster File System High Availability in virtual machines that do not support SCSI-3 PR" on page 21.

Table 7-4 Non-SCSI-3 I/O fencing response file definitions

Response file field	Definition
CFG{non_scsi3_fencing}	Defines whether to configure non-SCSI-3 I/O fencing. Valid values are 1 or 0. Enter 1 to configure non-SCSI-3 I/O fencing.
CFG {fencing_config_cpagent}	Enter '1' or '0' depending upon whether you want to configure the Coordination Point agent using the installer or not. Enter "0" if you do not want to configure the Coordination Point agent using the installer. Enter "1" if you want to use the installer to configure the Coordination Point agent. **Note:** This variable does not apply to majority-based fencing.

Table 7-4 Non-SCSI-3 I/O fencing response file definitions *(continued)*

Response file field	Definition
CFG {fencing_cpagentgrp}	Name of the service group which will have the Coordination Point agent resource as part of it. **Note:** This field is obsolete if the `fencing_config_cpagent` field is given a value of '0'. This variable does not apply to majority-based fencing.
CFG {fencing_cps}	Virtual IP address or Virtual hostname of the CP servers. **Note:** This variable does not apply to majority-based fencing.
CFG {fencing_cps_vips}	The virtual IP addresses or the fully qualified host names of the CP server. **Note:** This variable does not apply to majority-based fencing.
CFG {fencing_ncp}	Total number of coordination points (CP servers only) being used. **Note:** This variable does not apply to majority-based fencing.
CFG {fencing_cps_ports}	The port of the CP server that is denoted by *cps* . **Note:** This variable does not apply to majority-based fencing.
CFG{fencing_auto_refresh_reg}	Enable this variable if registration keys are missing on any of the CP servers.

Sample response file for configuring non-SCSI-3 I/O fencing

The following is a sample response file used for non-SCSI-3 I/O fencing :

```
# Configuration Values:
# our %CFG;
$CFG{fencing_config_cpagent}=0;
$CFG{fencing_cps}=[ qw(10.198.89.251 10.198.89.252 10.198.89.253) ];
```

```
$CFG{fencing_cps_vips}{"10.198.89.251"}=[ qw(10.198.89.251) ];
$CFG{fencing_cps_vips}{"10.198.89.252"}=[ qw(10.198.89.252) ];
$CFG{fencing_cps_vips}{"10.198.89.253"}=[ qw(10.198.89.253) ];
$CFG{fencing_ncp}=3;
$CFG{fencing_ndisks}=0;
$CFG{fencing_cps_ports}{"10.198.89.251"}=443;
$CFG{fencing_cps_ports}{"10.198.89.252"}=443;
$CFG{fencing_cps_ports}{"10.198.89.253"}=443;
$CFG{non_scsi3_fencing}=1;
$CFG{opt}{configure}=1;
$CFG{opt}{fencing}=1;
$CFG{prod}="ENTERPRISE80";
$CFG{systems}=[ qw(sys1 sys2) ];
$CFG{vcs_clusterid}=1256;
$CFG{vcs_clustername}="clus1";
$CFG{fencing_option}=1;
```

Response file variables to configure majority-based I/O fencing

Table 7-5 lists the response file variables that specify the required information to configure disk-based I/O fencing for SFCFSHA.

Table 7-5 Response file variables specific to configuring majority-based I/O fencing

Variable	List or Scalar	Description
CFG{opt}{fencing}	Scalar	Performs the I/O fencing configuration. (Required)

Table 7-5 Response file variables specific to configuring majority-based I/O fencing *(continued)*

Variable	List or Scalar	Description
CFG{fencing_option}	Scalar	Specifies the I/O fencing configuration mode. ■ 1—Coordination Point Server-based I/O fencing ■ 2—Coordinator disk-based I/O fencing ■ 3—Disabled-based fencing ■ 4—Online fencing migration ■ 5—Refresh keys/registrations on the existing coordination points ■ 6—Change the order of existing coordination points ■ 7—Majority-based fencing (Required)

Sample response file for configuring majority-based I/O fencing

```
# Configuration Values:
# our %CFG;
$CFG{fencing_option}=7;
$CFG{config_majority_based_fencing}=1;
$CFG{opt}{configure}=1;
$CFG{opt}{fencing}=1;
$CFG{prod}="ENTERPRISE80";
$CFG{systems}=[ qw(sys1  sys2) ];
$CFG{vcs_clusterid}=59082;
$CFG{vcs_clustername}="clus1";
```

Chapter 8

Manually configuring SFCFSHA clusters for data integrity

This chapter includes the following topics:

- Setting up disk-based I/O fencing manually
- Setting up server-based I/O fencing manually
- Setting up non-SCSI-3 fencing in virtual environments manually
- Setting up majority-based I/O fencing manually

Setting up disk-based I/O fencing manually

Table 8-1 lists the tasks that are involved in setting up I/O fencing.

Table 8-1

Task	Reference
Initializing disks as VxVM disks	See "Initializing disks as VxVM disks" on page 87.
Identifying disks to use as coordinator disks	See "Identifying disks to use as coordinator disks" on page 142.
Checking shared disks for I/O fencing	See "Checking shared disks for I/O fencing" on page 88.
Setting up coordinator disk groups	See "Setting up coordinator disk groups" on page 142.

Table 8-1 *(continued)*

Task	Reference
Creating I/O fencing configuration files	See "Creating I/O fencing configuration files" on page 143.
Modifying Storage Foundation Cluster File System High Availability configuration to use I/O fencing	See "Modifying VCS configuration to use I/O fencing" on page 144.
Configuring CoordPoint agent to monitor coordination points	See "Configuring CoordPoint agent to monitor coordination points" on page 157.
Verifying I/O fencing configuration	See "Verifying I/O fencing configuration" on page 146.

Identifying disks to use as coordinator disks

Make sure you initialized disks as VxVM disks.

See "Initializing disks as VxVM disks" on page 87.

Review the following procedure to identify disks to use as coordinator disks.

To identify the coordinator disks

1 List the disks on each node.

For example, execute the following commands to list the disks:

```
# vxdisk -o alldgs list
```

2 Pick three SCSI-3 PR compliant shared disks as coordinator disks.

See "Checking shared disks for I/O fencing" on page 88.

Setting up coordinator disk groups

From one node, create a disk group named vxfencoorddg. This group must contain three disks or LUNs. You must also set the coordinator attribute for the coordinator disk group. VxVM uses this attribute to prevent the reassignment of coordinator disks to other disk groups.

Note that if you create a coordinator disk group as a regular disk group, you can turn on the coordinator attribute in Volume Manager.

Refer to the *Storage Foundation Administrator's Guide* for details on how to create disk groups.

The following example procedure assumes that the disks have the device names sdx, sdy, and sdz.

To create the vxfencoorddg disk group

1. On any node, create the disk group by specifying the device names:

```
# vxdg init vxfencoorddg sdx sdy sdz
```

2. Set the coordinator attribute value as "on" for the coordinator disk group.

```
# vxdg -g vxfencoorddg set coordinator=on
```

3. Deport the coordinator disk group:

```
# vxdg deport vxfencoorddg
```

4. Import the disk group with the `-t` option to avoid automatically importing it when the nodes restart:

```
# vxdg -t import vxfencoorddg
```

5. Deport the disk group. Deporting the disk group prevents the coordinator disks from serving other purposes:

```
# vxdg deport vxfencoorddg
```

Creating I/O fencing configuration files

After you set up the coordinator disk group, you must do the following to configure I/O fencing:

- Create the I/O fencing configuration file /etc/vxfendg
- Update the I/O fencing configuration file /etc/vxfenmode

To update the I/O fencing files and start I/O fencing

1. On each nodes, type:

```
# echo "vxfencoorddg" > /etc/vxfendg
```

   Do not use spaces between the quotes in the "vxfencoorddg" text.

   This command creates the /etc/vxfendg file, which includes the name of the coordinator disk group.

2. On all cluster nodes specify the use of DMP disk policy in the `/etc/vxfenmode` file.

- # **cp /etc/vxfen.d/vxfenmode_scsi3_dmp /etc/vxfenmode**

3 To check the updated /etc/vxfenmode configuration, enter the following command on one of the nodes. For example:

```
# more /etc/vxfenmode
```

4 Ensure that you edit the following file on each node in the cluster to change the values of the VXFEN_START and the VXFEN_STOP environment variables to 1:

/etc/sysconfig/vxfen

Modifying VCS configuration to use I/O fencing

After you add coordination points and configure I/O fencing, add the UseFence = SCSI3 cluster attribute to the VCS configuration file /etc/VRTSvcs/conf/config/main.cf.

If you reset this attribute to UseFence = None, VCS does not make use of I/O fencing abilities while failing over service groups. However, I/O fencing needs to be disabled separately.

To modify VCS configuration to enable I/O fencing

1 Save the existing configuration:

```
# haconf -dump -makero
```

2 Stop VCS on all nodes:

```
# hastop -all
```

3 To ensure High Availability has stopped cleanly, run:

```
gabconfig -a
```

In the output of the commands, check that Port h is not present.

4 If the I/O fencing driver vxfen is already running, stop the I/O fencing driver.

For systemd environments with supported Linux distributions:

```
# systemctl stop vxfen
```

For other supported Linux distributions:

```
# /etc/init.d/vxfen stop
```

5 Make a backup of the main.cf file on all the nodes:

```
# cd /etc/VRTSvcs/conf/config
# cp main.cf main.orig
```

6 On one node, use vi or another text editor to edit the main.cf file. To modify the list of cluster attributes, add the UseFence attribute and assign its value as SCSI3.

```
cluster clus1(
UserNames = { admin = "cDRpdxPmHpzS." }
Administrators = { admin }
HacliUserLevel = COMMANDROOT
CounterInterval = 5
UseFence = SCSI3
)
```

Regardless of whether the fencing configuration is disk-based or server-based, the value of the cluster-level attribute UseFence is set to SCSI3.

7 Save and close the file.

8 Verify the syntax of the file /etc/VRTSvcs/conf/config/main.cf:

```
# hacf -verify /etc/VRTSvcs/conf/config
```

9 Start the I/O fencing driver and VCS. Perform the following steps on each node:

- Start the I/O fencing driver.
  The vxfen startup script also invokes the `vxfenconfig` command, which configures the vxfen driver to start and use the coordination points that are listed in /etc/vxfentab.
  For systemd environments with supported Linux distributions:
  ```
  # systemctl start vxfen
  ```
  For other supported Linux distributions:
  ```
  # /etc/init.d/vxfen start
  ```
- Start VCS on the node where main.cf is modified.
  ```
  # /opt/VRTS/bin/hastart
  ```
- Start VCS on all other nodes once VCS on first node reaches RUNNING state.
  ```
  # /opt/VRTS/bin/hastart
  ```

Verifying I/O fencing configuration

Verify from the vxfenadm output that the SCSI-3 disk policy reflects the configuration in the /etc/vxfenmode file.

To verify I/O fencing configuration

1 On one of the nodes, type:

```
# vxfenadm -d
```

Output similar to the following appears if the fencing mode is SCSI3 and the SCSI3 disk policy is dmp:

```
I/O Fencing Cluster Information:
================================

Fencing Protocol Version: 201
Fencing Mode: SCSI3
Fencing SCSI3 Disk Policy: dmp
Cluster Members:

   * 0 (sys1)
   1 (sys2)

RFSM State Information:
   node 0 in state 8 (running)
   node 1 in state 8 (running)
```

2 Verify that the disk-based I/O fencing is using the specified disks.

```
# vxfenconfig -l
```

Setting up server-based I/O fencing manually

Tasks that are involved in setting up server-based I/O fencing manually include:

Table 8-2 Tasks to set up server-based I/O fencing manually

Task	Reference
Preparing the CP servers for use by the Storage Foundation Cluster File System High Availability cluster	See "Preparing the CP servers manually for use by the SFCFSHA cluster" on page 147.

Table 8-2 Tasks to set up server-based I/O fencing manually *(continued)*

Task	Reference
Generating the client key and certificates on the client nodes manually	See "Generating the client key and certificates manually on the client nodes " on page 149.
Modifying I/O fencing configuration files to configure server-based I/O fencing	See "Configuring server-based fencing on the SFCFSHA cluster manually" on page 151.
Modifying Storage Foundation Cluster File System High Availability configuration to use I/O fencing	See "Modifying VCS configuration to use I/O fencing" on page 144.
Configuring Coordination Point agent to monitor coordination points	See "Configuring CoordPoint agent to monitor coordination points" on page 157.
Verifying the server-based I/O fencing configuration	See "Verifying server-based I/O fencing configuration" on page 159.

Preparing the CP servers manually for use by the SFCFSHA cluster

Use this procedure to manually prepare the CP server for use by the SFCFSHA cluster or clusters.

Table 8-3 displays the sample values used in this procedure.

Table 8-3 Sample values in procedure

CP server configuration component	Sample name
CP server	cps1
Node #1 - SFCFSHA cluster	sys1
Node #2 - SFCFSHA cluster	sys2
Cluster name	clus1
Cluster UUID	{f0735332-1dd1-11b2}

To manually configure CP servers for use by the SFCFSHA cluster

1 Determine the cluster name and uuid on the SFCFSHA cluster.

For example, issue the following commands on one of the SFCFSHA cluster nodes (sys1):

```
# grep cluster /etc/VRTSvcs/conf/config/main.cf

cluster clus1

# cat /etc/vx/.uuids/clusuuid

{f0735332-1dd1-11b2-bb31-00306eea460a}
```

2 Use the `cpsadm` command to check whether the SFCFSHA cluster and nodes are present in the CP server.

For example:

```
# cpsadm -s cps1.example.com -a list_nodes

ClusName  UUID                                    Hostname(Node ID) Registered
clus1  {f0735332-1dd1-11b2-bb31-00306eea460a} sys1(0)         0
clus1  {f0735332-1dd1-11b2-bb31-00306eea460a} sys2(1)         0
```

If the output does not show the cluster and nodes, then add them as described in the next step.

For detailed information about the `cpsadm` command, see the *Storage Foundation Cluster File System High Availability Administrator's Guide.*

3 Add the SFCFSHA cluster and nodes to each CP server.

For example, issue the following command on the CP server (cps1.example.com) to add the cluster:

```
# cpsadm -s cps1.example.com -a add_clus\
 -c clus1  -u {f0735332-1dd1-11b2}

Cluster clus1 added successfully
```

Issue the following command on the CP server (cps1.example.com) to add the first node:

```
# cpsadm -s cps1.example.com -a add_node\
 -c clus1 -u {f0735332-1dd1-11b2} -h sys1 -n0

Node 0 (sys1) successfully added
```

Issue the following command on the CP server (cps1.example.com) to add the second node:

```
# cpsadm -s cps1.example.com -a add_node\
 -c clus1 -u {f0735332-1dd1-11b2} -h sys2 -n1

Node 1 (sys2) successfully added
```

See "Generating the client key and certificates manually on the client nodes " on page 149.

Generating the client key and certificates manually on the client nodes

The client node that wants to connect to a CP server using HTTPS must have a private key and certificates signed by the Certificate Authority (CA) on the CP server

The client uses its private key and certificates to establish connection with the CP server. The key and the certificate must be present on the node at a predefined location. Each client has one client certificate and one CA certificate for every CP server, so, the certificate files must follow a specific naming convention. Distinct certificate names help the `cpsadm` command to identify which certificates have to be used when a client node connects to a specific CP server.

The certificate names must be as follows: ca_*cps-vip*.crt and client _*cps-vip*.crt

Where, *cps-vip* is the VIP or FQHN of the CP server listed in the `/etc/vxfenmode` file. For example, for a sample VIP, *192.168.1.201*, the corresponding certificate name is ca_*192.168.1.201*.

To manually set up certificates on the client node

1 Create the directory to store certificates.

```
# mkdir -p /var/VRTSvxfen/security/keys
/var/VRTSvxfen/security/certs
```

Note: Since the openssl utility might not be available on client nodes, Veritas recommends that you access the CP server using SSH to generate the client keys or certificates on the CP server and copy the certificates to each of the nodes.

---

2 Generate the private key for the client node.

```
# /opt/VRTSperl/non-perl-libs/bin/openssl genrsa -out
client_private.key 2048
```

3 Generate the client CSR for the cluster. CN is the UUID of the client's cluster.

```
# /opt/VRTSperl/non-perl-libs/bin/openssl req -new -key -sha256
client_private.key\

-subj '/C=countryname/L=localityname/OU=COMPANY/CN=CLUS_UUID'\

-out client_192.168.1.201.csr
```

Where, *countryname* is the country code, *localityname* is the city, *COMPANY* is the name of the company, and *CLUS_UUID* is the certificate name.

4 Generate the client certificate by using the CA key and the CA certificate. Run this command from the CP server.

```
# /opt/VRTSperl/non-perl-libs/bin/openssl x509 -req -days days
-sha256 -in client_192.168.1.201.csr\

-CA /var/VRTScps/security/certs/ca.crt -CAkey\

/var/VRTScps/security/keys/ca.key -set_serial 01 -out
client_192.168.10.1.crt
```

Where, *days* is the days you want the certificate to remain valid, *192.168.1.201* is the VIP or FQHN of the CP server.

5 Copy the client key, client certificate, and CA certificate to each of the client nodes at the following location.

Copy the client key at `/var/VRTSvxfen/security/keys/client_private.key`. The client is common for all the client nodes and hence you need to generate it only once.

Copy the client certificate at `/var/VRTSvxfen/security/certs/client_`*`192.168.1.201`*`.crt`.

Copy the CA certificate at `/var/VRTSvxfen/security/certs/ca_`*`192.168.1.201`*`.crt`

Note: Copy the certificates and the key to all the nodes at the locations that are listed in this step.

6 If the client nodes need to access the CP server using the FQHN and or the host name, make a copy of the certificates you generated and replace the VIP with the FQHN or host name. Make sure that you copy these certificates to all the nodes.

7 Repeat the procedure for every CP server.

8 After you copy the key and certificates to each client node, delete the client keys and client certificates on the CP server.

Configuring server-based fencing on the SFCFSHA cluster manually

The configuration process for the client or SFCFSHA cluster to use CP server as a coordination point requires editing the `/etc/vxfenmode` file.

You need to edit this file to specify the following information for your configuration:

- Fencing mode
- Fencing mechanism
- Fencing disk policy (if applicable to your I/O fencing configuration)
- CP server or CP servers
- Coordinator disk group (if applicable to your I/O fencing configuration)
- Set the order of coordination points

Note: Whenever coordinator disks are used as coordination points in your I/O fencing configuration, you must create a disk group (vxfencoorddg). You must specify this disk group in the `/etc/vxfenmode` file.

See "Setting up coordinator disk groups" on page 142.

The customized fencing framework also generates the `/etc/vxfentab` file which has coordination points (all the CP servers and disks from disk group specified in `/etc/vxfenmode` file).

To configure server-based fencing on the SFCFSHA cluster manually

1. Use a text editor to edit the following file on each node in the cluster:

   `/etc/sysconfig/vxfen`

   You must change the values of the `VXFEN_START` and the `VXFEN_STOP` environment variables to 1.

2. Use a text editor to edit the `/etc/vxfenmode` file values to meet your configuration specifications.

   - If your server-based fencing configuration uses a single highly available CP server as its only coordination point, make sure to add the `single_cp=1` entry in the `/etc/vxfenmode` file.
   - If you want the `vxfen` module to use a specific order of coordination points during a network partition scenario, set the `vxfen_honor_cp_order` value to be **1**. By default, the parameter is disabled.

   The following sample file output displays what the `/etc/vxfenmode` file contains:

   See "Sample vxfenmode file output for server-based fencing" on page 152.

3. After editing the `/etc/vxfenmode` file, run the vxfen init script to start fencing.

   For example:

   For systemd environments with supported Linux distributions:

   ```
   # systemctl start vxfen
   ```

   For other supported Linux distributions:

   ```
   # /etc/init.d/vxfen start
   ```

Sample vxfenmode file output for server-based fencing

The following is a sample vxfenmode file for server-based fencing:

```
#
```

```
# vxfen_mode determines in what mode VCS I/O Fencing should work.
#
# available options:
# scsi3      - use scsi3 persistent reservation disks
# customized - use script based customized fencing
# disabled   - run the driver but don't do any actual fencing
#
vxfen_mode=customized

# vxfen_mechanism determines the mechanism for customized I/O
# fencing that should be used.
#
# available options:
# cps      - use a coordination point server with optional script
#            controlled scsi3 disks
#
vxfen_mechanism=cps

#
# scsi3_disk_policy determines the way in which I/O fencing
# communicates with the coordination disks. This field is
# required only if customized coordinator disks are being used.
#
# available options:
# dmp - use dynamic multipathing
#
scsi3_disk_policy=dmp

#
# vxfen_honor_cp_order determines the order in which vxfen
# should use the coordination points specified in this file.
#
# available options:
# 0 - vxfen uses a sorted list of coordination points specified
# in this file,
# the order in which coordination points are specified does not matter.
#   (default)
# 1 - vxfen uses the coordination points in the same order they are
#    specified in this file

# Specify 3 or more odd number of coordination points in this file,
# each one in its own line. They can be all-CP servers,
# all-SCSI-3 compliant coordinator disks, or a combination of
```

```
# CP servers and SCSI-3 compliant coordinator disks.
# Please ensure that the CP server coordination points
# are numbered sequentially and in the same order
# on all the cluster nodes.
#
# Coordination Point Server(CPS) is specified as follows:
#
#  cps<number>=[<vip/vhn>]:<port>
#
# If a CPS supports multiple virtual IPs or virtual hostnames
# over different subnets, all of the IPs/names can be specified
# in a comma separated list as follows:
#
# cps<number>=[<vip_1/vhn_1>]:<port_1>,[<vip_2/vhn_2>]:<port_2>,
...,[<vip_n/vhn_n>]:<port_n>
#
# Where,
# <number>
#  is the serial number of the CPS as a coordination point; must
#  start with 1.
# <vip>
#  is the virtual IP address of the CPS, must be specified in
#  square brackets ("[]").
# <vhn>
#  is the virtual hostname of the CPS, must be specified in square
#  brackets ("[]").
# <port>
#  is the port number bound to a particular <vip/vhn> of the CPS.
#  It is optional to specify a <port>. However, if specified, it
#  must follow a colon (":") after <vip/vhn>. If not specified, the
#  colon (":") must not exist after <vip/vhn>.
#
# For all the <vip/vhn>s which do not have a specified <port>,
# a default port can be specified as follows:
#
# port=<default_port>
#
#  Where <default_port> is applicable to all the <vip/vhn>s for
#  which a <port> is not specified. In other words, specifying
#  <port> with a <vip/vhn> overrides the <default_port> for that
#  <vip/vhn>. If the <default_port> is not specified, and there
#  are <vip/vhn>s for which <port> is not specified, then port
#  number 14250 will be used for such <vip/vhn>s.
```

```
#
# Example of specifying CP Servers to be used as coordination points:
# port=57777
# cps1=[192.168.0.23],[192.168.0.24]:58888,[cps1.company.com]
#  cps2=[192.168.0.25]
# cps3=[cps2.company.com]:59999
#
# In the above example,
# - port 58888 will be used for vip [192.168.0.24]
# - port 59999 will be used for vhn [cps2.company.com], and
# - default port 57777 will be used for all remaining <vip/vhn>s:
#    [192.168.0.23]
#    [cps1.company.com]
#    [192.168.0.25]
# - if default port 57777 were not specified, port 14250
# would be used for all remaining <vip/vhn>s:
#    [192.168.0.23]
#    [cps1.company.com]
#    [192.168.0.25]
#
# SCSI-3 compliant coordinator disks are specified as:
#
#  vxfendg=<coordinator disk group name>
# Example:
#  vxfendg=vxfencoorddg
#
# Examples of different configurations:
#  1. All CP server coordination points
# cps1=
# cps2=
# cps3=
#
# 2. A combination of CP server and a disk group having two SCSI-3
# coordinator disks
# cps1=
# vxfendg=
# Note: The disk group specified in this case should have two disks
#
# 3. All SCSI-3 coordinator disks
# vxfendg=
# Note: The disk group specified in case should have three disks
# cps1=[cps1.company.com]
# cps2=[cps2.company.com]
```

```
# cps3=[cps3.company.com]
# port=443
```

Table 8-4 defines the vxfenmode parameters that must be edited.

Table 8-4 vxfenmode file parameters

vxfenmode File Parameter	Description
vxfen_mode	Fencing mode of operation. This parameter must be set to “customized”.
vxfen_mechanism	Fencing mechanism. This parameter defines the mechanism that is used for fencing. If one of the three coordination points is a CP server, then this parameter must be set to “cps”.
scsi3_disk_policy	Configure the vxfen module to use DMP devices, "dmp". **Note:** The configured disk policy is applied on all the nodes.
cps1, cps2, or vxfendg	Coordination point parameters. Enter either the virtual IP address or the FQHN (whichever is accessible) of the CP server. *cps<number>=[virtual_ip_address/virtual_host_name]:port* Where *port* is optional. The default port value is 443. If you have configured multiple virtual IP addresses or host names over different subnets, you can specify these as comma-separated values. For example: `cps1=[192.168.0.23],[192.168.0.24]:58888, [cps1.company.com]` **Note:** Whenever coordinator disks are used in an I/O fencing configuration, a disk group has to be created (vxfencoorddg) and specified in the /etc/vxfenmode file. Additionally, the customized fencing framework also generates the /etc/vxfentab file which specifies the security setting and the coordination points (all the CP servers and the disks from disk group specified in /etc/vxfenmode file).

Table 8-4 vxfenmode file parameters *(continued)*

vxfenmode File Parameter	Description
port	Default port for the CP server to listen on. If you have not specified port numbers for individual virtual IP addresses or host names, the default port number value that the CP server uses for those individual virtual IP addresses or host names is 443. You can change this default port value using the port parameter.
single_cp	Value 1 for single_cp parameter indicates that the server-based fencing uses a single highly available CP server as its only coordination point. Value 0 for single_cp parameter indicates that the server-based fencing uses at least three coordination points.
vxfen_honor_cp_order	Set the value to 1 for vxfen module to use a specific order of coordination points during a network partition scenario. By default the parameter is disabled. The default value is 0.

Configuring CoordPoint agent to monitor coordination points

The following procedure describes how to manually configure the CoordPoint agent to monitor coordination points.

The CoordPoint agent can monitor CP servers and SCSI-3 disks.

See the *Storage Foundation Cluster File System High Availability Bundled Agents Reference Guide* for more information on the agent.

To configure CoordPoint agent to monitor coordination points

1. Ensure that your SFCFSHA cluster has been properly installed and configured with fencing enabled.
2. Create a parallel service group vxfen and add a coordpoint resource to the vxfen service group using the following commands:

```
# haconf -makerw
# hagrp -add vxfen
# hagrp -modify vxfen SystemList sys1 0 sys2 1
# hagrp -modify vxfen AutoFailOver 0
# hagrp -modify vxfen Parallel 1
# hagrp -modify vxfen SourceFile "./main.cf"
# hares -add coordpoint CoordPoint vxfen
# hares -modify coordpoint FaultTolerance 0
# hares -override coordpoint LevelTwoMonitorFreq
# hares -modify coordpoint LevelTwoMonitorFreq 5
# hares -modify coordpoint Enabled 1
# haconf -dump -makero
```

3. Configure the Phantom resource for the vxfen disk group.

```
# haconf -makerw
# hares -add RES_phantom_vxfen Phantom vxfen
# hares -modify RES_phantom_vxfen Enabled 1
# haconf -dump -makero
```

4 Verify the status of the agent on the SFCFSHA cluster using the `hares` commands. For example:

```
# hares -state coordpoint
```

The following is an example of the command and output::

```
# hares -state coordpoint

# Resource    Attribute   System   Value
coordpoint   State        sys1   ONLINE
coordpoint   State        sys2   ONLINE
```

5 Access the engine log to view the agent log. The agent log is written to the engine log.

The agent log contains detailed CoordPoint agent monitoring information; including information about whether the CoordPoint agent is able to access all the coordination points, information to check on which coordination points the CoordPoint agent is reporting missing keys, etc.

To view the debug logs in the engine log, change the dbg level for that node using the following commands:

```
# haconf -makerw

# hatype -modify Coordpoint LogDbg 10

# haconf -dump -makero
```

The agent log can now be viewed at the following location:

/var/VRTSvcs/log/engine_A.log

Note: The Coordpoint agent is always in the online state when the I/O fencing is configured in the majority or the disabled mode. For both these modes the I/O fencing does not have any coordination points to monitor. Thereby, the Coordpoint agent is always in the online state.

Verifying server-based I/O fencing configuration

Follow the procedure described below to verify your server-based I/O fencing configuration.

To verify the server-based I/O fencing configuration

1 Verify that the I/O fencing configuration was successful by running the `vxfenadm` command. For example, run the following command:

```
# vxfenadm -d
```

Note: For troubleshooting any server-based I/O fencing configuration issues, refer to the *Storage Foundation Cluster File System High Availability Administrator's Guide*.

2 Verify that I/O fencing is using the specified coordination points by running the `vxfenconfig` command. For example, run the following command:

```
# vxfenconfig -l
```

If the output displays `single_cp=1`, it indicates that the application cluster uses a CP server as the single coordination point for server-based fencing.

Setting up non-SCSI-3 fencing in virtual environments manually

To manually set up I/O fencing in a non-SCSI-3 PR compliant setup

1 Configure I/O fencing either in majority-based fencing mode with no coordination points or in server-based fencing mode only with CP servers as coordination points.

See "Setting up server-based I/O fencing manually" on page 146.

See "Setting up majority-based I/O fencing manually " on page 166.

2 Make sure that the Storage Foundation Cluster File System High Availability cluster is online and check that the fencing mode is customized mode or majority mode.

```
# vxfenadm -d
```

3 Make sure that the cluster attribute UseFence is set to SCSI-3.

```
# haclus -value UseFence
```

4 On each node, edit the /etc/vxenviron file as follows:

```
data_disk_fencing=off
```

5 On each node, edit the /etc/sysconfig/vxfen file as follows:

```
vxfen_vxfnd_tmt=25
```

6 On each node, edit the /etc/vxfenmode file as follows:

```
loser_exit_delay=55
vxfen_script_timeout=25
```

Refer to the sample /etc/vxfenmode file.

7 On each node, set the value of the LLT sendhbcap timer parameter value as follows:

- Run the following command:

```
lltconfig -T sendhbcap:3000
```

- Add the following line to the /etc/llttab file so that the changes remain persistent after any reboot:

```
set-timer senhbcap:3000
```

8 On any one node, edit the VCS configuration file as follows:

- Make the VCS configuration file writable:

```
# haconf -makerw
```

- For each resource of the type DiskGroup, set the value of the MonitorReservation attribute to 0 and the value of the Reservation attribute to NONE.

```
# hares -modify <dg_resource> MonitorReservation 0

# hares -modify <dg_resource> Reservation "NONE"
```

- Run the following command to verify the value:

```
# hares -list Type=DiskGroup MonitorReservation!=0

# hares -list Type=DiskGroup Reservation!="NONE"
```

The command should not list any resources.

- Modify the default value of the Reservation attribute at type-level.

```
# haattr -default DiskGroup Reservation "NONE"
```

- Make the VCS configuration file read-only

```
# haconf -dump -makero
```

9 Make sure that the UseFence attribute in the VCS configuration file main.cf is set to SCSI-3.

10 To make these VxFEN changes take effect, stop and restart VxFEN and the dependent modules

- On each node, run the following command to stop VCS:
  For systemd environments with supported Linux distributions:
  `# systemctl stop vcs`
  For other supported Linux distributions:
  `# /etc/init.d/vcs stop`
- After VCS takes all services offline, run the following command to stop VxFEN:
  For systemd environments with supported Linux distributions:
  `# systemctl stop vxfen`
  For other supported Linux distributions:
  `# /etc/init.d/vxfen stop`
- On each node, run the following commands to restart VxFEN and VCS:
  For systemd environments with supported Linux distributions:
  `# systemctl start vxfen`
  `# systemctl start vcs`
  For other supported Linux distributions:
  `# /etc/init.d/vxfen start`
  `# /etc/init.d/vcs start`

Sample /etc/vxfenmode file for non-SCSI-3 fencing

```
#
# vxfen_mode determines in what mode VCS I/O Fencing should work.
#
# available options:
# scsi3      - use scsi3 persistent reservation disks
# customized - use script based customized fencing
# disabled   - run the driver but don't do any actual fencing
#
vxfen_mode=customized
```

```
# vxfen_mechanism determines the mechanism for customized I/O
# fencing that should be used.
#
# available options:
# cps      - use a coordination point server with optional script
#            controlled scsi3 disks
#
vxfen_mechanism=cps

#
# scsi3_disk_policy determines the way in which I/O fencing
# communicates with the coordination disks. This field is
# required only if customized coordinator disks are being used.
#
# available options:
# dmp - use dynamic multipathing
#
scsi3_disk_policy=dmp

#
# Seconds for which the winning sub cluster waits to allow for the
# losing subcluster to panic & drain I/Os. Useful in the absence of
# SCSI3 based data disk fencing loser_exit_delay=55
#
# Seconds for which vxfend process wait for a customized fencing
# script to complete. Only used with vxfen_mode=customized
# vxfen_script_timeout=25

#
# vxfen_honor_cp_order determines the order in which vxfen
# should use the coordination points specified in this file.
#
# available options:
# 0 - vxfen uses a sorted list of coordination points specified
# in this file, the order in which coordination points are specified
# does not matter.
#   (default)
# 1 - vxfen uses the coordination points in the same order they are
#    specified in this file

# Specify 3 or more odd number of coordination points in this file,
# each one in its own line. They can be all-CP servers, all-SCSI-3
```

```
# compliant coordinator disks, or a combination of CP servers and
# SCSI-3 compliant coordinator disks.
# Please ensure that the CP server coordination points are
# numbered sequentially and in the same order on all the cluster
# nodes.
#
# Coordination Point Server(CPS) is specified as follows:
#
#  cps<number>=[<vip/vhn>]:<port>
#
# If a CPS supports multiple virtual IPs or virtual hostnames
# over different subnets, all of the IPs/names can be specified
# in a comma separated list as follows:
#
# cps<number>=[<vip_1/vhn_1>]:<port_1>,[<vip_2/vhn_2>]:<port_2>,
# ...,[<vip_n/vhn_n>]:<port_n>
#
# Where,
# <number>
#  is the serial number of the CPS as a coordination point; must
#  start with 1.
# <vip>
#  is the virtual IP address of the CPS, must be specified in
#  square brackets ("[]").
# <vhn>
#  is the virtual hostname of the CPS, must be specified in square
#  brackets ("[]").
# <port>
#  is the port number bound to a particular <vip/vhn> of the CPS.
#  It is optional to specify a <port>. However, if specified, it
#  must follow a colon (":") after <vip/vhn>. If not specified, the
#  colon (":") must not exist after <vip/vhn>.
#
# For all the <vip/vhn>s which do not have a specified <port>,
# a default port can be specified as follows:
#
# port=<default_port>
#
#  Where <default_port> is applicable to all the <vip/vhn>s for which a
#  <port> is not specified. In other words, specifying <port> with a
#  <vip/vhn> overrides the <default_port> for that <vip/vhn>.
# If the <default_port> is not specified, and there are <vip/vhn>s for
# which <port> is not specified, then port number 14250 will be used
```

```
# for such <vip/vhn>s.
#
# Example of specifying CP Servers to be used as coordination points:
# port=57777
# cps1=[192.168.0.23],[192.168.0.24]:58888,[cps1.company.com]
#  cps2=[192.168.0.25]
# cps3=[cps2.company.com]:59999
#
# In the above example,
# - port 58888 will be used for vip [192.168.0.24]
# - port 59999 will be used for vhn [cps2.company.com], and
# - default port 57777 will be used for all remaining <vip/vhn>s:
#    [192.168.0.23]
#    [cps1.company.com]
#    [192.168.0.25]
# - if default port 57777 were not specified, port 14250 would be
#  used for all remaining <vip/vhn>s:
#    [192.168.0.23]
#    [cps1.company.com]
#    [192.168.0.25]
#
# SCSI-3 compliant coordinator disks are specified as:
#
#  vxfendg=<coordinator disk group name>
# Example:
#  vxfendg=vxfencoorddg
#
# Examples of different configurations:
#  1. All CP server coordination points
# cps1=
# cps2=
# cps3=
#
# 2. A combination of CP server and a disk group having two SCSI-3
# coordinator disks
# cps1=
# vxfendg=
# Note: The disk group specified in this case should have two disks
#
# 3. All SCSI-3 coordinator disks
# vxfendg=
# Note: The disk group specified in case should have three disks
# cps1=[cps1.company.com]
```

```
# cps2=[cps2.company.com]
# cps3=[cps3.company.com]
# port=443
```

Setting up majority-based I/O fencing manually

Table 8-5 lists the tasks that are involved in setting up I/O fencing.

Task	Reference
Creating I/O fencing configuration files	Creating I/O fencing configuration files
Modifying VCS configuration to use I/O fencing	Modifying VCS configuration to use I/O fencing
Verifying I/O fencing configuration	Verifying I/O fencing configuration

Creating I/O fencing configuration files

To update the I/O fencing files and start I/O fencing

1. On all cluster nodes, run the following command

   ```
   # cp /etc/vxfen.d/vxfenmode_majority /etc/vxfenmode
   ```

2. To check the updated /etc/vxfenmode configuration, enter the following command on one of the nodes.

   ```
   # cat /etc/vxfenmode
   ```

3. Ensure that you edit the following file on each node in the cluster to change the values of the `VXFEN_START` and the `VXFEN_STOP` environment variables to 1.

   ```
   /etc/sysconfig/vxfen
   ```

Modifying VCS configuration to use I/O fencing

After you configure I/O fencing, add the UseFence = SCSI3 cluster attribute to the VCS configuration file /etc/VRTSvcs/conf/config/main.cf.

If you reset this attribute to UseFence = None, VCS does not make use of I/O fencing abilities while failing over service groups. However, I/O fencing needs to be disabled separately.

To modify VCS configuration to enable I/O fencing

1. Save the existing configuration:

```
# haconf -dump -makero
```

2. Stop VCS on all nodes:

```
# hastop -all
```

3. To ensure High Availability has stopped cleanly, run `gabconfig -a`.

   In the output of the commans, check that Port h is not present.

4. If the I/O fencing driver vxfen is already running, stop the I/O fencing driver.

   For systemd environments with supported Linux distributions:

```
# systemctl stop vxfen
```

   For other supported Linux distributions:

```
# /etc/init.d/vxfen stop
```

5. Make a backup of the main.cf file on all the nodes:

```
# cd /etc/VRTSvcs/conf/config
# cp main.cf main.orig
```

6. On one node, use vi or another text editor to edit the main.cf file. To modify the list of cluster attributes, add the UseFence attribute and assign its value as SCSI3.

```
cluster clus1(
UserNames = { admin = "cDRpdxPmHpzS." }
Administrators = { admin }
HacliUserLevel = COMMANDROOT
CounterInterval = 5
UseFence = SCSI3
)
```

   For fencing configuration in any mode except the disabled mode, the value of the cluster-level attribute UseFence is set to SCSI3.

7. Save and close the file.

8. Verify the syntax of the file /etc/VRTSvcs/conf/config/main.cf:

```
# hacf -verify /etc/VRTSvcs/conf/config
```

9 Using rcp or another utility, copy the VCS configuration file from a node (for example, sys1) to the remaining cluster nodes.

For example, on each remaining node, enter:

```
# rcp sys1:/etc/VRTSvcs/conf/config/main.cf \
/etc/VRTSvcs/conf/config
```

10 Start the I/O fencing driver and VCS. Perform the following steps on each node:

- Start the I/O fencing driver.
  The vxfen startup script also invokes the `vxfenconfig` command, which configures the vxfen driver.
  For systemd environments with supported Linux distributions:
  ```
  # systemctl start vxfen
  ```
  For other supported Linux distributions:
  ```
  # /etc/init.d/vxfen start
  ```
- Start VCS on the node where main.cf is modified.
  ```
  # /opt/VRTS/bin/hastart
  ```
- Start VCS on all other nodes once VCS on first node reaches RUNNING state.
  ```
  # /opt/VRTS/bin/hastart
  ```

Verifying I/O fencing configuration

Verify from the vxfenadm output that the fencing mode reflects the configuration in the /etc/vxfenmode file.

To verify I/O fencing configuration

- On one of the nodes, type:

```
# vxfenadm -d
```

Output similar to the following appears if the fencing mode is majority:

```
I/O Fencing Cluster Information:
================================

 Fencing Protocol Version: 201
 Fencing Mode: MAJORITY
 Cluster Members:

        * 0 (sys1)
          1 (sys2)

 RFSM State Information:
        node   0 in state  8 (running)
        node   1 in state  8 (running)
```

Section 3

Upgrade of SFCFSHA

- Chapter 9. Planning to upgrade SFCFSHA
- Chapter 10. Performing a full upgrade of SFCFSHA using the installer
- Chapter 11. Performing a rolling upgrade of SFCFSHA
- Chapter 12. Performing a phased upgrade of SFCFSHA
- Chapter 13. Performing an automated SFCFSHA upgrade using response files
- Chapter 14. Upgrading Volume Replicator
- Chapter 15. Upgrading VirtualStore
- Chapter 16. Performing post-upgrade tasks

Chapter 9

Planning to upgrade SFCFSHA

This chapter includes the following topics:

- About the upgrade
- Supported upgrade paths
- Transitioning between the Veritas products
- Considerations for upgrading SFCFSHA to 8.0 on systems configured with an Oracle resource
- Preparing to upgrade SFCFSHA
- Considerations for upgrading REST server
- Using Install Bundles to simultaneously install or upgrade full releases (base, maintenance, rolling patch), and individual patches

About the upgrade

This release supports upgrades from 6.2.1 and later versions. If your existing installation is from a pre-6.2.1 version, you must first upgrade to version 6.2.1, then follow the procedures mentioned in this document to upgrade the product.

The installer supports the following types of upgrade:

- Full upgrade
- Automated upgrade using response files

Table 9-1 describes the product mapping after an upgrade.

Table 9-1 Veritas InfoScale product mapping after upgrade

Product (6.2.x and earlier)	Product (7.0 and later)	Component (7.0 and later)
SFCFSHA (with High Availability)	Veritas InfoScale Enterprise	SFCFSHA
SVS	Veritas InfoScale Enterprise	SFCFSHA

Note: From 7.0 onwards, the existing Veritas InfoScale product upgrades to the higher version of the same product. For example, Veritas InfoScale Enterprise 7.4.1 gets upgraded to Veritas InfoScale Enterprise 7.4.2.

During the upgrade, the installation program performs the following tasks:

1. Stops the product before starting the upgrade
2. Upgrades the installed packages and installs additional packages

   Slf license key files are required while upgrading to version 7.4 and later. The text-based license keys that are used in previous product versions are not supported when upgrading to version 7.4 and later. If you plan to upgrade any of the InfoScale products from a version earlier than 7.4, first contact Customer Care for your region to procure an applicable slf license key file. Refer to the following link for contact information of the Customer Care center for your region: https://www.veritas.com/content/support/en_US/contact-us.html.

   If your current installation uses a permanent license key, you will be prompted to update the license to 8.0. Ensure that the license key file is downloaded on the local host, where you want to upgrade the product. The license key file must not be saved in the root directory (`/`) or the default license directory on the local host (`/etc/vx/licenses/lic`). You can save the license key file inside any other directory on the local host.

   If you choose not to update your license, you will be registered with a keyless license. Within 60 days of choosing this option, you must install a valid license key file corresponding to the entitled license level.
3. You must configure the Veritas Telemetry Collector while upgrading, if you have do not already have it configured. For more information, refer to the *About telemetry data collection in InfoScale* section in the *Veritas Installation guide*.
4. Restores the existing configuration.

   For example, if your setup contains an SFCFSHA installation, the installer upgrades and restores the configuration to SFCFSHA. If your setup included

multiple components, the installer upgrades and restores the configuration of the components.

5. Starts the configured components.

Supported upgrade paths

If you are on an unsupported operating system version, ensure that you first upgrade to a supported verison of the operating system. Also, upgrades between major operating system versions are not supported, for example, from RHEL 7 to RHEL 8. If you plan to move between major operating system versions, you need to reinstall the product. For supported operating system versions, see the *Veritas InfoScale Release Notes*.

Table 9-2 lists the supported upgrade paths for upgrades on RHEL and Oracle Linux.

Table 9-2 Supported upgrade paths on RHEL and Oracle Linux

From product version	From OS version	To OS version	To product version	To Component
6.2.1	RHEL 6 Update 4, 5, 6, 7, 8, 9, 10 RHEL 7 GA, Update 1, 2, 3, 4, 5, 6, 7, 8, 9 Oracle Linux 6 Update 4, 5, 6, 7 Oracle Linux 7 GA, Update 2, 3	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 9 Oracle Linux 8 Update 2, 4	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.1	RHEL 6 update 4, 5, 6, 7, 8, 9 RHEL 7 Update 1, 2, 3, 4 Oracle Linux 6 Update 6, 7, 8 Oracle Linux 7 Update 1, 2	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 9 Oracle Linux 8 Update 2, 4	Veritas InfoScale Enterprise 8.0	SFCFSHA

Table 9-2 Supported upgrade paths on RHEL and Oracle Linux *(continued)*

From product version	From OS version	To OS version	To product version	To Component
7.2	RHEL 6 update 6, 7, 8, 9 RHEL 7 Update 1, 2, 3, 4 Oracle Linux 6 Update 4, 5, 6, 7, 8, 9 Oracle Linux 7 Update 1, 2, 3, 4	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 9 Oracle Linux 8 Update 2, 4	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.3	RHEL 6 update 7, 8, 9 RHEL 7 Update 1,2, 3, 4 Oracle Linux 6 Update 7, 8, 9 Oracle Linux 7 Update 1, 2, 3	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 9 Oracle Linux 8 Update 2, 4	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.3.1	RHEL 6 update 7, 8, 9 RHEL 7 Update 3, 4, 5 Oracle Linux 6 Update 7, 8, 9 Oracle Linux 7 Update 3, 4	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 9 Oracle Linux 8 Update 2, 4	Veritas InfoScale Enterprise 8.0	SFCFSHA

Table 9-2 Supported upgrade paths on RHEL and Oracle Linux *(continued)*

From product version	From OS version	To OS version	To product version	To Component
7.4	RHEL 6 Update 7, 8, 9, 10 RHEL 7 Update 3, 4,5 Oracle Linux 6 Update 7, 8, 9 Oracle Linux 7 Update 3, 4	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 9 Oracle Linux 8 Update 2, 4	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.4.1	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 4, 5, 6, 7, 8, 9 Oracle Linux 8 Update 1, 2	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 9 Oracle Linux 8 Update 2, 4	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.4.2	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 7 Oracle Linux 8 Update 1	RHEL 7 Update 9 RHEL 8 Update 2, 4 Oracle Linux 7 Update 9 Oracle Linux 8 Update 2, 4	Veritas InfoScale Enterprise 8.0	SFCFSHA

Table 9-3 lists the supported upgrade paths for upgrades on SLES.

Table 9-3 Supported upgrade paths on SLES

From product version	From OS version	To OS version	To product version	To component
6.2.1	SLES11 SP2, SP3, SP4 SLES 12 GA, SP1, SP2, SP3, SP4	SLES 12 SP5 SLES 15 SP2	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.1, 7.2	SLES 11 SP3, SP4 SLES 12 GA, SP1, SP2	SLES 12 SP5 SLES 15 SP2	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.3	SLES 11 SP3, SP4 SLES 12 SP1, SP2	SLES 12 SP5 SLES 15 SP2	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.3.1	SLES 11 SP3, SP4 SLES 12 SP2, SP4, SP5	SLES 12 SP5 SLES 15 SP2	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.4	SLES 11 SP3, SP4 SLES 12 SP2, SP3	SLES 12 SP5 SLES 15 SP2	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.4.1	SLES 11 SP3, SP4 SLES 12 SP2, SP3, SP4, SP5 SLES 15 GA, SP1, SP2	SLES 12 SP5 SLES 15 SP2	Veritas InfoScale Enterprise 8.0	SFCFSHA
7.4.2	SLES 12 SP 4, SP5 SLES 15 SP1, SP2	SLES 12 SP5 SLES 15 SP2	Veritas InfoScale Enterprise 8.0	SFCFSHA

Transitioning between the Veritas products

You can transition between the products of the Veritas family.

To transition to a product on a system where you already have an InfoScale product installed, run the product installer for the InfoScale product you want to transition to.

Note: The base product version and the version of the product to transition to must be same.

Supported paths:

The following table lists the supported paths to transition from one Veritas product to another Veritas product.

Table 9-4 Supported transition paths

Base product	Transition to			
	InfoScale Foundation	InfoScale Availability	InfoScale Storage	InfoScale Enterprise
InfoScale Foundation	X	X	✓	✓
InfoScale Availability	X	X	X	✓
InfoScale Storage	✓	X	X	✓
InfoScale Enterprise	X	X	X	X

Considerations for upgrading SFCFSHA to 8.0 on systems configured with an Oracle resource

If you plan to upgrade SFCFSHA running on systems configured with an Oracle resource, set the `MonitorOption` attribute to 0 (zero) before you start the upgrade.

For more information on enabling the Oracle health check, see the *Cluster Server Agent for Oracle Installation and Configuration Guide*.

Preparing to upgrade SFCFSHA

Before you upgrade, you need to prepare the systems and storage. Review the following procedures and perform the appropriate tasks.

Getting ready for the upgrade

Complete the following tasks before you perform the upgrade:

- Review the *Veritas InfoScale 8.0 Release Notes* for any late-breaking information on upgrading your system.
- Review the Veritas Technical Support website for additional information: https://www.veritas.com/support/en_US.html
- You can configure the Veritas Telemetry Collector while upgrading, if you have do not already have it configured. For more information, refer to the *About telemetry data collection in InfoScale* section in the *Veritas Installation guide*.
- Make sure that the administrator who performs the upgrade has root access and a good knowledge of the operating system's administration.
- Make sure that all users are logged off and that all major user applications are properly shut down.
- Make sure that you have created a valid backup.
  See "Creating backups" on page 180.
- Ensure that you have enough file system space to upgrade. Identify where you want to copy the RPMs, for example `/packages/Veritas` when the root file system has enough space or `/var/tmp/packages` if the `/var` file system has enough space.
  Do not put the files under `/tmp`, which is erased during a system restart.
  Do not put the files on a file system that is inaccessible before running the upgrade script.
  You can use a Veritas-supplied disc for the upgrade as long as modifications to the upgrade script are not required.
  If `/usr/local` was originally created as a slice, modifications are required.
- Comment out any application commands or processes that are known to hang if their file systems are not present in the startup scripts.
  In case of RHEL 7 and SLES 12 systems, some startup scripts are located at `/etc/vx/`, and the startup scripts of the following services are located at:

Service name	Startup script location and file name
amf.service	`/opt/VRTSamf/bin/amf`

Service name	Startup script location and file name
gab.service	`/opt/VRTSgab/gab`
llt.service	`/opt/VRTSllt/llt`
vcs.service	`/opt/VRTSvcs/bin/vcs`
vcsmm.service	`/opt/VRTSvcs/rac/bin/vcsmm`
vxfen.service	`/opt/VRTSvcs/vxfen/bin/vxfen`

The remaining startup scripts for RHEL 7 and SLES 12 are located at `/etc/init.d/`, like all the other startup scripts for the other supported RHEL distributions.

- Make sure that the current operating system supports version 8.0 of the product. If the operating system does not support it, plan for a staged upgrade.

Note: Before you upgrade RHEL 7.7 OS on a virtual machine, you need to first upgrade Veritas InfoScale 8.0. Later upgrade RHEL 7.7 OS, else the virtual machine may go in an unstable state.

Use `-ignorechecks` CPI option on RHEL 7.0 to RHEL 7.6 version to successfully upgrade Veritas InfoScale product.

- Schedule sufficient outage time and downtime for the upgrade and any applications that use the Veritas InfoScale products. Depending on the configuration, the outage can take several hours.
- Any swap partitions not in `rootdg` must be commented out of `/etc/fstab`. If possible, swap partitions other than those on the root disk should be commented out of `/etc/fstab` and not mounted during the upgrade. The active swap partitions that are not in `rootdg` cause `upgrade_start` to fail.
- Make sure that the file systems are clean before upgrading.
- Upgrade arrays (if required).
See "Upgrading the array support" on page 187.
- To reliably save information on a mirrored disk, shut down the system and physically remove the mirrored disk. Removing the disk in this manner offers a failback point.
- Make sure that DMP support for native stack is disabled (`dmp_native_support=off`). If DMP support for native stack is enabled

(`dmp_native_support=on`), the installer may detect it and ask you to restart the system.

- If you want to upgrade the application clusters that use CP server based fencing to version 6.1 and later, make sure that you first upgrade VCS or SFHA on the CP server systems to version 6.1 and later. And then, from 7.0.1 onwards, CP server supports only HTTPS based communication with its clients and IPM-based communication is no longer supported. CP server needs to be reconfigured if you upgrade the CP server with IPM-based CP server configured.
  For instructions to upgrade VCS or SFHA on the CP server systems, refer to the relevant Configuration and Upgrade Guides.

Creating backups

Save relevant system information before the upgrade.

To create backups

1. Log in as superuser.
2. Before the upgrade, ensure that you have made backups of all data that you want to preserve.
3. Back up information in files such as `/boot/grub/menu.lst`, `/etc/grub.conf` or `/etc/lilo.conf`, and `/etc/fstab`.
4. Installer verifies that recent backups of configuration files in VxVM private region have been saved in `/etc/vx/cbr/bk`.

   If not, a warning message is displayed.

   Warning: Backup `/etc/vx/cbr/bk` directory.

   ---

5. Copy the `fstab` file to `fstab.orig`:

   ```
   # cp /etc/fstab /etc/fstab.orig
   ```

6. Run the `vxlicrep`, `vxdisk list`, and `vxprint -ht` commands and record the output. Use this information to reconfigure your system after the upgrade.
7. If you install Veritas InfoScale Enterprise 8.0 software, follow the guidelines that are given in the *Cluster Server Configuration and Upgrade Guide* for information on preserving your VCS configuration across the installation procedure.

8 Back up the external `quotas` and `quotas.grp` files.

If you are upgrading from 6.2.1, you must also back up the `quotas.grp.64` and `quotas.64` files.

9 Verify that quotas are turned off on all the mounted file systems.

Determining if the root disk is encapsulated

Note: Root Disk Encapsulation (RDE) on Linux Distribution is not supported from 7.3.1 release onwards.

Check if the system's root disk is under VxVM control by running this command:

```
# df -v /
```

The root disk is under VxVM control if `/dev/vx/dsk/rootdg/rootvol` is listed as being mounted as the root (`/`) file system.

If the root disk is encapsulated, follow the appropriate upgrade procedures.

Pre-upgrade planning when VVR is configured

Before installing or upgrading Volume Replicator (VVR):

- Confirm that your system has enough free disk space to install VVR.
- Make sure you have root permissions. You must have root permissions to perform the install and upgrade procedures.
- If replication using VVR is configured, Veritas recommends that the disk group version is at least 110 prior to upgrading.
  You can check the Disk Group version using the following command:

  ```
  # vxdg list diskgroup
  ```

- If replication using VVR is configured, make sure the size of the SRL volume is greater than 110 MB.
  Refer to the *Veritas InfoScale™ Replication Administrator's Guide*.
- If replication using VVR is configured, verify that all the Primary RLINKs are up-to-date on all the hosts.

  ```
  # /usr/sbin/vxrlink -g diskgroup status rlink_name
  ```

  Note: Do not continue until the primary RLINKs are up-to-date.

- If VCS is used to manage VVR replication, follow the preparation steps to upgrade VVR and VCS agents.

See the *Veritas InfoScale™ Replication Administrator's Guide* for more information.

See the *Getting Started Guide* for more information on the documentation.

Planning an upgrade from the previous VVR version

If you plan to upgrade VVR from the previous VVR version, you can upgrade VVR with reduced application downtime by upgrading the hosts at separate times. While the Primary is being upgraded, the application can be migrated to the Secondary, thus reducing downtime. The replication between the (upgraded) Primary and the Secondary, which have different versions of VVR, will still continue. This feature facilitates high availability even when the VVR upgrade is not complete on both the sites. Veritas recommends that the Secondary hosts be upgraded before the Primary host in the Replicated Data Set (RDS).

For information regarding VVR support for replicating across Storage Foundation versions, refer to the *Veritas InfoScale Release Notes*.

Replicating between versions is intended to remove the restriction of upgrading the Primary and Secondary at the same time. VVR can continue to replicate an existing RDS with Replicated Volume Groups (RVGs) on the systems that you want to upgrade. When the Primary and Secondary are at different versions, VVR does not support changing the configuration with the `vradmin` command or creating a new RDS.

Also, if you specify TCP as the network protocol, the VVR versions on the Primary and Secondary determine whether the checksum is calculated. As shown in Table 9-5, if either the Primary or Secondary are running a version of VVR prior to 8.0, and you use the TCP protocol, VVR calculates the checksum for every data packet it replicates. If the Primary and Secondary are at VVR 8.0, VVR does not calculate the checksum. Instead, it relies on the TCP checksum mechanism.

Table 9-5 VVR versions and checksum calculations

VVR prior to 8.0 (DG version <= 140)	VVR 8.0 (DG version >= 310)	VVR calculates checksum TCP connections?
Primary	Secondary	Yes
Secondary	Primary	Yes
Primary and Secondary		Yes
	Primary and Secondary	No

Note: When replicating between versions of VVR, avoid using commands associated with new features. The earlier version may not support new features and problems could occur.

If you do not need to upgrade all the hosts in the RDS simultaneously, you can use replication between versions after you upgrade one host. You can then upgrade the other hosts in the RDS later at your convenience.

Note: If you have a cluster setup, you must upgrade all the nodes in the cluster at the same time.

Planning and upgrading VVR to use IPv6 as connection protocol

SFCFSHA supports using IPv6 as the connection protocol.

This release supports the following configurations for VVR:

- VVR continues to support replication between IPv4-only nodes with IPv4 as the internet protocol
- VVR supports replication between IPv4-only nodes and IPv4/IPv6 dual-stack nodes with IPv4 as the internet protocol
- VVR supports replication between IPv6-only nodes and IPv4/IPv6 dual-stack nodes with IPv6 as the internet protocol
- VVR supports replication between IPv6 only nodes
- VVR supports replication to one or more IPv6 only nodes and one or more IPv4 only nodes from a IPv4/IPv6 dual-stack node
- VVR supports replication of a shared disk group only when all the nodes in the cluster that share the disk group are at IPv4 or IPv6

Considerations for upgrading SFCFSHA to 7.4 or later on systems with an ongoing or a paused replication

Typically, you can upgrade SFCFSHAin a setup where VVR is configured. However, InfoScale does not support upgrade from version 7.3.1 or earlier to version 7.4 or later with an ongoing or a paused replication. To upgrade InfoScale from these earlier versions to 7.4 or later, perform the following steps:

1. Stop replication to the Secondary using the `vradmin stoprep` command for all RVGs.
2. Upgrade InfoScale to version 7.4 or later at the primary and the secondary sites.

3. Upgrade the disk group version at the primary and the secondary sites.
4. Start replication using the `vradmin -a startrep` command.

Preparing to upgrade VVR when VCS agents are configured

To prepare to upgrade VVR when VCS agents for VVR are configured, perform the following tasks sequentially:

- See "Freezing the service groups and stopping all the applications" on page 184.
- See "Preparing for the upgrade when VCS agents are configured" on page 186.

Freezing the service groups and stopping all the applications

This section describes how to freeze the service groups and stop all applications.

To freeze the service groups and stop applications

Perform the following steps for the Primary and Secondary clusters:

1 Log in as the superuser.

2 Make sure that `/opt/VRTS/bin` is in your PATH so that you can execute all the product commands.

3 Before the upgrade, cleanly shut down all applications.

In a shared disk group environment:

- OFFLINE all application service groups that do not contain RVGShared resources. Do not OFFLINE the ClusterService, cvm and RVGLogowner groups.
- If the application resources are part of the same service group as an RVGShared resource, then OFFLINE only the application resources.

In a private disk group environment:

- OFFLINE all application service groups that do not contain RVG resources. Do not OFFLINE the service groups containing RVG resources.
- If the application resources are part of the same service group as an RVG resource, then OFFLINE only the application resources. In other words, ensure that the RVG resource remains ONLINE so that the private disk groups containing these RVG objects do not get deported.

Note: You must also stop any remaining applications not managed by VCS.

---

4 On any node in the cluster, make the VCS configuration writable:

```
# haconf -makerw
```

5 On any node in the cluster, list the groups in your configuration:

```
# hagrp -list
```

6 On any node in the cluster, freeze all service groups except the ClusterService group by typing the following command for each group name displayed in the output from step 5.

```
# hagrp -freeze group_name -persistent
```

Note: Make a note of the list of frozen service groups for future use.

7 On any node in the cluster, save the configuration file (`main.cf`) with the groups frozen:

```
# haconf -dump -makero
```

Note: Continue only after you have performed steps 3 to step 7 for each node of the cluster.

8 Display the list of service groups that have RVG resources and the nodes on which each service group is online by typing the following command on any node in the cluster:

```
# hares -display -type RVG -attribute State
Resource        Attribute        System     Value
VVRGrp           State           sys2    ONLINE
ORAGrp           State           sys2    ONLINE
```

Note: For the resources that are ONLINE, write down the nodes displayed in the System column of the output.

9 Repeat step 8 for each node of the cluster.

10 For private disk groups, determine and note down the hosts on which the disk groups are imported.

See "Determining the nodes on which disk groups are online" on page 186.

11 For shared disk groups, run the following command on any node in the CVM cluster:

```
# vxdctl -c mode
```

Note the master and record it for future use.

Determining the nodes on which disk groups are online

For private disk groups, determine and note down the hosts on which the disk groups containing RVG resources are imported. This information is required for restoring the configuration after the upgrade.

To determine the online disk groups

1 On any node in the cluster, list the disk groups in your configuration, and note down the disk group names listed in the output for future use:

```
# hares -display -type RVG -attribute DiskGroup
```

Note: Write down the list of the disk groups that are under VCS control.

2 For each disk group listed in the output in step 1, list its corresponding disk group resource name:

```
# hares -list DiskGroup=diskgroup Type=DiskGroup
```

3 For each disk group resource name listed in the output in step 2, get and note down the node on which the disk group is imported by typing the following command:

```
# hares -display dg_resname -attribute State
```

The output displays the disk groups that are under VCS control and nodes on which the disk groups are imported.

Preparing for the upgrade when VCS agents are configured

If you have configured the VCS agents, it is recommended that you take backups of the configuration files, such as `main.cf` and `types.cf`, which are present in the `/etc/VRTSvcs/conf/config` directory.

To prepare a configuration with VCS agents for an upgrade

1. List the disk groups on each of the nodes by typing the following command on each node:

   ```
   # vxdisk -o alldgs list
   ```

   The output displays a list of the disk groups that are under VCS control and the disk groups that are not under VCS control.

   Note: The disk groups that are not locally imported are displayed in parentheses.

2. If any of the disk groups have not been imported on any node, import them. For disk groups in your VCS configuration, you can import them on any node. For disk groups that are not under VCS control, choose an appropriate node on which to import the disk group. Enter the following command on the appropriate node:

   ```
   # vxdg -t import diskgroup
   ```

3. If a disk group is already imported, then recover the disk group by typing the following command on the node on which it is imported:

   ```
   # vxrecover -bs
   ```

4. Verify that all the Primary RLINKs are up to date.

   ```
   # vxrlink -g diskgroup status rlink_name
   ```

   Note: Do not continue until the Primary RLINKs are up-to-date.

Upgrading the array support

The Veritas InfoScale 8.0 release includes all array support in a single RPM, `VRTSaslapm`. The array support RPM includes the array support previously included in the `VRTSvxvm` RPM. The array support RPM also includes support previously packaged as external Array Support Libraries (ASLs) and array policy modules (APMs).

See the 8.0 Hardware Compatibility List for information about supported arrays.

When you upgrade Storage Foundation products with the product installer, the installer automatically upgrades the array support. If you upgrade Storage Foundation products with manual steps, you should remove any external ASLs or APMs that were installed previously on your system. Installing the `VRTSvxvm` RPM exits with an error if external ASLs or APMs are detected.

After you have installed Veritas InfoScale 8.0, Veritas provides support for new disk arrays through updates to the `VRTSaslapm` RPM.

For more information about array support, see the *Storage Foundation Cluster File System High Availability Administrator's Guide*.

Considerations for upgrading REST server

If you upgrade from an earlier InfoScale release, perform the following tasks sequentially:

- Upgrade the InfoScale cluster to version 8.0.
- Run the `# /opt/VRTS/install/installer -rest_server` command and follow the prompts to initiate the REST server configuration.

Using Install Bundles to simultaneously install or upgrade full releases (base, maintenance, rolling patch), and individual patches

Beginning with version 6.2.1, you can easily install or upgrade your systems directly to a base, maintenance, patch level or a combination of multiple patches and packages together in one step using Install Bundles. With Install Bundles, the installer has the ability to merge so that customers can install or upgrade directly to maintenance or patch levels in one execution. The various scripts, RPMs, and patch components are merged, and multiple releases are installed together as if they are one combined release. You do not have to perform two or more install actions to install or upgrade systems to maintenance levels or patch levels.

Releases are divided into the following categories:

Table 9-6 Release Levels

Level	Content	Form factor	Applies to	Release types	Download location
Base	Features	RPMs	All products	Major, minor, Service Pack (SP), Platform Release (PR)	FileConnect
Maintenance	Fixes, new features	RPMs	All products	Maintenance Release (MR), Rolling Patch (RP)	Veritas Services and Operations Readiness Tools (SORT)
Patch	Fixes	RPMs	Single product	P-Patch, Private Patch, Public patch	SORT, Support site

When you install or upgrade using Install Bundles:

- InfoScale products are discovered and assigned as a single version to the maintenance level. Each system can also have one or more patches applied.
- Base releases are accessible from FileConnect that requires customer serial numbers. Maintenance and patch releases can be automatically downloaded from SORT.
- Patches can be installed using automated installers.
- Patches can now be detected to prevent upgrade conflict. Patch releases are not offered as a combined release. They are only available from Veritas Technical Support on a need basis.

You can use the `-base_path` and `-patch_path` options to import installation code from multiple releases. You can find RPMs and patches from different media paths, and merge RPM and patch definitions for multiple releases. You can use these options to use new task and phase functionality to correctly perform required operations for each release component. You can install the RPMs and patches in defined phases using these options, which helps you when you want to perform a single start or stop process and perform pre and post operations for all level in a single operation.

Four possible methods of integration exist. All commands must be executed from the highest base or maintenance level install script.

In the example below:

- 8.0 is the base version

- 8.0.1 is the maintenance version
- 8.0.1.1000 is the patch version for 8.0.1
- 8.0.0.1000 is the patch version for 8.0

1. Base + maintenance:

   This integration method can be used when you install or upgrade from a lower version to 8.0.1.

   Enter the following command:

   ```
   # installmr -base_path <path_to_base>
   ```

2. Base + patch:

   This integration method can be used when you install or upgrade from a lower version to 8.0.0.100.

   Enter the following command:

   ```
   # installer -patch_path <path_to_patch>
   ```

3. Maintenance + patch:

   This integration method can be used when you upgrade from version 8.0 to 8.0.1.100.

   Enter the following command:

   ```
   # installmr -patch_path <path_to_patch>
   ```

4. Base + maintenance + patch:

   This integration method can be used when you install or upgrade from a lower version to 8.0.1.100.

   Enter the following command:

   ```
   # installmr -base_path <path_to_base>
   -patch_path <path_to_patch>
   ```

   Note: You can add a maximum of five patches using *-patch_path <path_to_patch> -patch2_path <path_to_patch> ... -patch5_path <path_to_patch>*

Chapter 10

Performing a full upgrade of SFCFSHA using the installer

This chapter includes the following topics:

- Performing a full upgrade using the product installer
- Upgrading SFDB

Performing a full upgrade using the product installer

Performing a full upgrade involves the following tasks:

- Ensuring that the file systems are clean
- Performing the upgrade
- Updating the configuration and confirming startup
- Updating the operating system, if required

For details on installing and upgrading Veritas InfoScale using the installer with the -yum option, refer to the *Veritas InfoScale Installation guide*.

Ensuring the file systems are clean

Before upgrading to SFCFSHA 8.0, ensure that the file systems are clean. To ensure that the logs have been replayed and the file systems are marked clean:

To ensure the file systems are clean

1. Log in as superuser onto any node in the cluster.
2. Take the service group offline on each node of the cluster, which contains VxFS and CFS resources:

```
# hagrp -offline group -sys sys1
# hagrp -offline group -sys sys2
# hagrp -offline group -sys sys3
# hagrp -offline group -sys sys4
```

   where *group* is the VCS service group that has the CVMVolDg and CFSMount resource.

   Repeat this step for each SFCFSHA service group.

   Note: This unmounts the CFS file systems.

3. Unmount all VxFS file systems not under VCS control:

```
# umount mount_point
```

4. Check and repair each VxFS file system:

```
# fsck -t vxfs /dev/vx/dsk/diskgroup/volume
```

   The `fsck` command in `/opt/VRTS/bin` accepts either the block or character device (`/dev/vx/dsk/dg/vol`) or (`/dev/vx/rdsk/dg/vol`). The operating system version of `fsck` may limit the device types it accepts.

   For more information, see the `fsck` and `fsck_vxfs` man pages.

   Repeat this step for each file system.

Performing the upgrade

If you plan to upgrade the operating system, perform the following steps:

To upgrade the operating system

1. Rename the `/etc/llttab` file to prevent LLT from starting automatically when the node starts:

   ```
   # mv /etc/llttab /etc/llttab.save
   ```

2. Create `install-db` file to prevent VxVM daemons or processes from starting automatically when the node starts.

   ```
   # touch /etc/vx/reconfig.d/state.d/install-db
   ```

3. Upgrade the operating system on all nodes in the cluster.

   For instructions, see the operating system documentation.

4. If you upgraded the operating system, restart the nodes if required:

   ```
   # shutdown -r now
   ```

5. After the system restarts, rename the `/etc/llttab` file to its original name:

   ```
   # mv /etc/llttab.save /etc/llttab
   ```

6. Enable VxVM to start after system restarts.

   ```
   # rm /etc/vx/reconfig.d/state.d/install-db
   ```

To perform the upgrade

Note: Root Disk Encapsulation (RDE) is not supported on Linux from 7.3.1 onwards.

---

1. Log in as superuser.
2. Insert the appropriate media disc per your distribution and architecture into your system's DVD-ROM drive.
3. If volume management software is running on your system, the software disc automatically mounts as `/mnt/cdrom`.

   If volume management software is not available to mount the disc, you must mount it manually, enter:

   ```
   # mount -o ro /dev/cdrom /mnt/cdrom
   ```

4 Change to the top-level directory on the disc:

```
# cd /mnt/cdrom
```

5 Verify there are no VxFS file systems mounted on the nodes being upgraded:

```
# mount -t vxfs
```

If any VxFS file systems are mounted, offline the group on each node of the cluster:

```
# hagrp -offline group -sys sys1
# hagrp -offline group -sys sys2
# hagrp -offline group -sys sys3
# hagrp -offline group -sys sys4
```

where *group* is the VCS service group that has the CVMVolDg and CFSMount resource.

If VxFS are not managed by VCS then unmount them manually:

```
# umount mount_point
```

Repeat this step for each SFCFSHA service group.

6 If a cache area is online, you must take the cache area offline before you upgrade the VxVM RPM. Use the following command to take the cache area offline:

```
# sfcache offline cachename
```

7 Start the upgrade from any node in the cluster. Enter the following command, and then press **y** to upgrade the cluster configuration.

```
# ./installer -upgrade
```

8 You are prompted to enter the system names (in the following example, "sys1" and "sys2") on which the software is to be upgraded. Enter the system name or names and then press Return.

```
Enter the system names separated by spaces:
[q, ?] sys1 sys2
```

9 During the initial system check, the installer verifies that communication between systems has been set up.

If the installer hangs or asks for a login password, you have the option to let the installer configure SSH or RSH communications between the systems. If you choose to allow this configuration, select the communication type and provide the root passwords for each system.

10 At the prompt, specify whether you accept the terms of the End User License Agreement (EULA).

```
Do you agree with the terms of the End User License Agreement
as specified in the EULA/en/EULA_ENTERPRISE_Ux_8.0.pdf file
present on media? [y,n,q,?] y
```

11 The installer discovers if any of the systems that you are upgrading have mirrored and encapsulated boot disks. For each system that has a mirrored boot disk, you have the option to create a backup of the system's boot disk group before the upgrade proceeds. If you want to split the boot disk group to create a backup, answer **y**.

12 The installer then prompts you to name the backup boot disk group. Enter the name for it or press **Enter** to accept the default.

13 If you are prompted to start the split operation. Press **y** to continue.

Note: The split operation can take some time to complete.

14 Output shows information that SFCFSHA must be stopped on a running system. Enter **y** to continue.

15 The installer displays the following question before the installer stops the product processes if the current cluster is secured and version is prior to 6.2.:

- Do you want to grant read access to everyone? [y,n,q,?]
  - To grant read access to all authenticated users, type **y**.
  - To grant usergroup specific permissions, type **n**.
- Do you want to provide any usergroups that you would like to grant read access?[y,n,q,?]
  - To specify usergroups and grant them read access, type **y**
  - To grant read access only to root users, type **n**. The installer grants read access read access to the root users.

Note: Separate the usergroup names with spaces. To grant read access to a usergroup on a specific node, specify usergroup as <usergroup>@<node_name>. You can also specify usergroups here and create them later.

16 Enter **y** for summary information and reboots if the boot disk is encapsulated before the upgrade.

Do not remove the log files until the Veritas InfoScale product is working properly on your system. Technical Support will need these log files for debugging purposes.

17 Only perform this step if you have split the mirrored root disk to back it up. After a successful reboot, verify the upgrade and re-join the backup disk group. If the upgrade fails, revert to the backup disk group.

See "Re-joining the backup boot disk group into the current disk group" on page 237.

See "Reverting to the backup boot disk group after an unsuccessful upgrade" on page 237.

Updating the configuration and confirming startup

Perform the following steps on each upgraded node.

To update the configuration and confirm startup

1 Remove the `/etc/VRTSvcs/conf/config/.stale` file, if it exists.

```
# rm -f /etc/VRTSvcs/conf/config/.stale
```

2 Verify that LLT is running:

```
# lltconfig
  LLT is running
```

3 Verify GAB is configured:

```
# gabconfig -l | grep 'Driver.state' | \
  grep Configured
  Driver state : Configured
```

4 Verify VxVM daemon is started and enabled:

```
# /opt/VRTS/bin/vxdctl mode
  mode: enabled
```

5 Confirm all upgraded nodes are in a running state.

```
# gabconfig -a
```

6 After the configuration is complete, the CVM and SFCFSHA groups may come up frozen. To find out the frozen CVM and SFCFSHA groups, enter the following command:

```
# /opt/VRTS/bin/hastatus -sum
```

If the groups are frozen, unfreeze CVM and SFCFSHA groups using the following commands for each group:

- Make the configuration read/write.

  ```
  # /opt/VRTS/bin/haconf -makerw
  ```

- Unfreeze the group.

  ```
  # /opt/VRTS/bin/hagrp -unfreeze group_name -persistent
  ```

- Save the configuration.

  ```
  # /opt/VRTS/bin/haconf -dump -makero
  ```

7 If VVR is configured, and the CVM and SFCFSHA groups are offline, bring the groups online in the following order:

Bring online the CVM groups on all systems.

```
# /opt/VRTS/bin/hagrp -online group_name -sys sys1
# /opt/VRTS/bin/hagrp -online group_name -sys sys2
```

where *group_name* is the VCS service group that has the CVMVolDg resource.

Bring online the RVGShared groups and the virtual IP on the master node using the following commands:

```
# hagrp -online RVGShared -sys masterhost
# hares -online ip_name -sys masterhost
```

Bring online the SFCFSHA groups on all systems.

```
# /opt/VRTS/bin/hagrp -online group_name -sys sys1
# /opt/VRTS/bin/hagrp -online group_name -sys sys2
```

where *group_name* is the VCS service group that has the CFSMount resource.

If the SFCFSHA service groups do not come online then your file system could be dirty.

Note: If you upgrade to Veritas InfoScale Enterprise 8.0 and the file systems are dirty, you have to deport the shared disk group and import it as non-shared. After the import, run `fsck`. `fsck` should succeed. Then deport the disk group and import it back as shared.

---

8 Find out which node is the CVM master. Enter the following:

```
# vxdctl -c mode
```

9 On the CVM master node, upgrade the CVM protocol. Enter the following:

```
# vxdctl upgrade
```

Upgrading the operating system

You can upgrade the operating system, if required.

- Rename the `/etc/lltab` file to prevent LLT from starting automatically when the node starts:

```
# mv /etc/llttab /etc/llttab.save
```

- Upgrade the operating system.
  Refer to the operating system's documentation for more information.
- After you have upgraded the operating system, restart the nodes:

  ```
  # shutdown -r now
  ```

- Rename the `/etc/llttab` file to its original name:

  ```
  # mv /etc/llttab.save /etc/llttab
  ```

Upgrading SFDB

While upgrading to 8.0, the SFDB tools are enabled by default, which implies that the vxdbd daemon is configured. You can enable the SFDB tools, if they are disabled.

To enable SFDB tools

1. Log in as root.
2. Run the following command to configure and start the vxdbd daemon.

```
# /opt/VRTS/bin/sfae_config enable
```

Note: If any SFDB installation with authentication setup is upgraded to 8.0, the commands fail with an error. To resolve the issue, setup the SFDB authentication again. For more information, see the *Veritas InfoScale™ Storage and Availability Management for Oracle Databases* or *Veritas InfoScale™ Storage and Availability Management for DB2 Databases*.

Chapter 11

Performing a rolling upgrade of SFCFSHA

This chapter includes the following topics:

- About rolling upgrade
- Performing a rolling upgrade of SFCFSHA using the product installer
- Performing a rolling upgrade of SFCFSHA from 7.4.2 or later to 8.0
- Performing rolling upgrades in CVR environments

About rolling upgrade

The rolling upgrade process minimizes the downtime of a cluster during an upgrade to the amount of time that it takes to fail over a service group. InfoScale 7.4.2 or later versions let you configure clusters with nodes that run different versions of the VCS engine. To support such configurations, the installer provides a rolling upgrade option for InfoScale components that upgrades the kernel RPMs and the VCS agent-related RPMs in the same process. The rolling upgrade may involve some application downtime while a node is upgraded, but there is zero cluster downtime. The rolling upgrade has two main phases where the installer upgrades kernel RPMs in phase 1 and VCS agent related RPMs in phase 2.

Note: You need to perform a rolling upgrade on a completely configured cluster.

The following is an overview of the flow for a rolling upgrade:

1. The installer performs prechecks on the cluster.

2. The installer moves service groups to free nodes for the first phase of the upgrade as is needed.

   Application downtime occurs during the first phase as the installer moves service groups to free nodes for the upgrade. The only downtime that is incurred is the normal time required for the service group to failover. The downtime is limited to the applications that are failed over and not the entire cluster.

3. The installer performs the second phase of the upgrade on all of the nodes in the cluster. The second phase of the upgrade includes downtime of the Cluster Server (VCS) engine HAD, but does not include application downtime.

[Unresolved xref] illustrates an example of the installer performing a rolling upgrade for three service groups on a two node cluster.

The following limitations apply to rolling upgrades:

- Rolling upgrades are not compatible with phased upgrades. Do not mix rolling upgrades and phased upgrades.
- You can perform a rolling upgrade from 6.2.1 or later versions.
- The rolling upgrade procedures support only minor operating system upgrades.
- The rolling upgrade procedure requires the product to be started before and after upgrade. If the current release does not support your current operating system version and the installed old release version does not support the operating system version that the current release supports, then rolling upgrade is not supported.

Performing a rolling upgrade of SFCFSHA using the product installer

Note: Root Disk Encapsulation (RDE) is not supported on Linux from 7.3.1 onwards.

Before you start the rolling upgrade, make sure that Cluster Server (VCS) is running on all the nodes of the cluster.

Stop all activity for all the VxVM volumes that are not under VCS control. For example, stop any applications such as databases that access the volumes, and unmount any file systems that have been created on the volumes. Then stop all the volumes.

Unmount all VxFS file systems that are not under VCS control.

To perform a rolling upgrade

1. Phase 1 of rolling upgrade begins on the first subcluster. Complete the preparatory steps on the first subcluster.

   Unmount all VxFS file systems not under VCS control:

   ```
   # umount mount_point
   ```

2. Complete updates to the operating system, if required. For instructions, see the operating system documentation.

   Make sure that the existing version of SFCFSHA supports the operating system update you apply. If the existing version of SFCFSHA does not support the operating system update, first upgrade SFCFSHA to a version that supports the operating system update.

   Switch applications to remaining subcluster and upgrade the operating system of the fist subcluster.

   The nodes are restarted after the operating system update.

3. If a cache area is online, you must take the cache area offline before you upgrade the VxVM RPM. Use the following command to take the cache area offline:

   ```
   # sfcache offline cachename
   ```

4. Log in as superuser and mount the SFCFSHA 8.0 installation media.

5. From root, start the installer.

   ```
   # ./installer
   ```

6. From the menu, select `Upgrade a Product` and from the sub menu, select `Rolling Upgrade`.

7. The installer suggests system names for the upgrade. Press **Enter** to upgrade the suggested systems, or enter the name of any one system in the cluster on which you want to perform a rolling upgrade and then press **Enter**.

8. The installer checks system communications, release compatibility, version information, and lists the cluster name, ID, and cluster nodes. Type **y** to continue.

9. The installer inventories the running service groups and determines the node or nodes to upgrade in phase 1 of the rolling upgrade. Type **y** to continue. If you choose to specify the nodes, type **n** and enter the names of the nodes.

10 The installer performs further prechecks on the nodes in the cluster and may present warnings. You can type **y** to continue or quit the installer and address the precheck's warnings.

11 Review the end-user license agreement, and type **y** if you agree to its terms.

12 If the boot disk is encapsulated and mirrored, you can create a backup boot disk.

If you choose to create a backup boot disk, type **y**. Provide a backup name for the boot disk group or accept the default name. The installer then creates a backup copy of the boot disk group.

13 After the installer detects the online service groups, the installer prompts the user to do one of the following:

- Manually switch service groups
- Use the CPI to automatically switch service groups

The downtime is the time that it normally takes for the service group's failover.

Note: It is recommended that you manually switch the service groups. Automatic switching of service groups does not resolve dependency issues if any dependent resource is not under VCS control.

---

14 The installer prompts you to stop the applicable processes. Type **y** to continue.

The installer evacuates all service groups to the node or nodes that are not upgraded at this time. The installer stops parallel service groups on the nodes that are to be upgraded.

15 The installer stops relevant processes, uninstalls old kernel RPMs, and installs the new RPMs. The installer asks if you want to update your licenses to the current version. Select **Yes** or **No**. Veritas recommends that you update your licenses to fully use the new features in the current release.

16 If the cluster has configured Coordination Point Server based fencing, then during upgrade, installer may ask the user to provide the new HTTPS Coordination Point Server.

The installer performs the upgrade configuration and starts the processes. If the boot disk is encapsulated before the upgrade, installer prompts the user to reboot the node after performing the upgrade configuration.

17 Complete the preparatory steps on the nodes that you have not yet upgraded.

Unmount all VxFS file systems not under VCS control on all the nodes.

```
# umount mount_point
```

18 If operating system updates are not required, skip this step.

Go to step 19.

Else, complete updates to the operating system on the nodes that you have not yet upgraded. For instructions, see the operating system documentation.

Repeat steps 3 to 16 for each node.

Phase 1 of rolling upgrade is complete on the first subcluster. Phase 1 of rolling upgrade begins on the second subcluster.

19 Offline all cache areas on the remaining node or nodes:

```
# sfcache offline cachename
```

20 The installer begins phase 1 of the upgrade on the remaining node or nodes. Type **y** to continue the rolling upgrade. If the installer was invoked on the upgraded (rebooted) nodes, you must invoke the installer again.

Note: In case of an FSS environment, phase 1 of the rolling upgrade is performed on one node at a time.

The installer repeats step 9 through step 16.

For clusters with larger number of nodes, this process may repeat several times. Service groups come down and are brought up to accommodate the upgrade.

21 When Phase 1 of the rolling upgrade completes, mount all the VxFS file systems that are not under VCS control manually. Begin Phase 2 of the upgrade. Phase 2 of the upgrade includes downtime for the VCS engine (HAD), which does not include application downtime. Type **y** to continue. Phase 2 of the rolling upgrade begins here.

22 The installer determines the remaining RPMs to upgrade. Press **Enter** to continue.

23 The installer displays the following question before the installer stops the product processes. If the cluster was configured in secure mode and version is prior to 6.2 before the upgrade, these questions are displayed.

- Do you want to grant read access to everyone? [y,n,q,?]
  - To grant read access to all authenticated users, type **y**.
  - To grant usergroup specific permissions, type **n**.
- Do you want to provide any usergroups that you would like to grant read access?[y,n,q,?]

- To specify usergroups and grant them read access, type **y**
- To grant read access only to root users, type **n**. The installer grants read access read access to the root users.

- Enter the usergroup names separated by spaces that you would like to grant read access. If you would like to grant read access to a usergroup on a specific node, enter like 'usrgrp1@node1', and if you would like to grant read access to usergroup on any cluster node, enter like 'usrgrp1'. If some usergroups are not created yet, create the usergroups after configuration if needed. [b]

24 The installer stops Cluster Server (VCS) processes but the applications continue to run. Type **y** to continue.

The installer performs prestop, uninstalls old RPMs, and installs the new RPMs. It performs post-installation tasks, and the configuration for the upgrade.

25 If you have network connection to the Internet, the installer checks for updates.

If updates are discovered, you can apply them now.

26 A prompt message appears to ask if the user wants to read the summary file. You can choose **y** if you want to read the install summary file.

Performing a rolling upgrade of SFCFSHA from 7.4.2 or later to 8.0

To perform a rolling upgrade

1 Log in as superuser and mount the product installation media.

2 Start the installer from the root folder.

```
# ./installer
```

3 From the menu, select **Upgrade a Product** and from the submenu, select **Rolling Upgrade**.

4 The installer suggests system names for the upgrade. Press **Enter** to upgrade the suggested systems, or enter the name of any one system in the cluster on which you want to perform a rolling upgrade and then press **Enter**.

The installer checks system communications, release compatibility, version information, and lists the cluster name, ID, and cluster nodes. Press **y** to continue.

5 The installer lists the running service groups and determines the nodes to upgrade during the rolling upgrade. Press **y** to continue. If you choose to specify the nodes, press **n** and enter the names of the nodes.

6 The installer performs further prechecks on the nodes in the cluster and may present warnings. You can press **y** to continue or quit the installer and address the warnings from the precheck.

7 Review the end-user license agreement, and press **y** if you agree to its terms.

8 If the boot disk is encapsulated and mirrored, you can create a backup boot disk.

If you choose to create a backup boot disk, press **y**. Provide a backup name for the boot disk group or accept the default name. The installer then creates a backup copy of the boot disk group.

9 After the installer detects the online service groups, it prompts you to choose whether to:

- Manually switch service groups
- Let the product installer automatically switch service groups

The downtime is equivalent to the time that it typically takes to fail over a service group.

Note: Veritas recommends that you manually switch the service groups. Automatic switching of service groups does not resolve dependency issues if any dependent resource is not under VCS control.

---

10 The installer prompts you to stop the applicable processes. Press **y** to continue.

The installer evacuates all the service groups to the node or the nodes that are not upgraded at this time.

It stops the parallel service groups on the nodes that are yet to be upgraded.

11 The installer stops the relevant processes, uninstalls the old kernel RPMs, and installs the new RPMs. It asks if you want to update your licenses to the current version; select **Yes** or **No**. Veritas recommends that you update your licenses to fully use the new features in the current release.

12 If the cluster has configured coordination point server-based fencing, the installer may ask you to provide the new HTTPS coordination point server. Provide the value if your are prompted for this input.

13 The installer asks if you want to perform the rolling upgrade. Select **Yes** to continue.

The installer performs the upgrade configuration and starts the processes. If the boot disk is encapsulated before the upgrade, the installer prompts you to reboot the node after performing the upgrade configuration.

14 Complete the preparatory steps on the nodes that you have not yet upgraded.

15 On all the nodes, unmount all VxFS file systems that are not under VCS control.

```
# umount mount_point
```

16 If operating system updates are not required, skip this step. Proceed to step 17.

Otherwise, complete the updates to the operating system on the nodes that you have not yet upgraded. For instructions, see the operating system documentation.

Repeat steps 1 to 12 for each node.

When the rolling upgrade is complete on the first subcluster, the rolling upgrade begins on the second subcluster.

17 Offline all cache areas on the remaining node or nodes.

```
# sfcache offline cachename
```

18 The installer begins the rolling upgrade on the remaining nodes. Press **y** to continue the rolling upgrade. If the installer was invoked on the upgraded (rebooted) nodes, you must invoke the installer again.

Note: In case of an FSS environment, rolling upgrade is performed on one node at a time.

The installer repeats step 5 through step 15.

For clusters with a large number of nodes, this process may repeat several times. Service groups are taken down and brought up to accommodate the upgrade.

19 When the rolling upgrade completes, manually mount all the VxFS file systems that are not under VCS control.

20 If you have network connection to the Internet, the installer checks for updates.

If updates are discovered, you can apply them now.

21 The installer asks you whether you want to read the installation summary file. Press **y** if you want to read the file.

Performing rolling upgrades in CVR environments

In a CVR environment, you perform the rolling upgrade on the secondary site first and then on the primary site.

To perform a rolling upgrade in a CVR environment

1 Perform a rolling upgrade on the secondary sites first, without stopping the replication.

Check whether the cluster nodes have joined the cluster back after the upgrade.

Continue to monitor the replication status.

```
# vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

2 Perform a rolling upgrade for all the nodes on the primary site by following these steps:

At the primary site, the VxFS file systems are always mounted. Therefore, if you directly begin the upgrade, the installer returns the following error:

```
CPI ERROR V-9-40-1480 Some VxFS file systems are mounted
on mount points /<volume_mount_point> on node <node_name>
and need to be unmounted before upgrade.
```

In this event, take the applications offline and unmount the file system manually on the indicated nodes, before you upgrade them.

Additionally, you may have to take the following actions on these nodes:

- If any parallel service groups exist in the hierarchy, take them offline.

  ```
  # /opt/VRTS/bin/hagrp -offline <global_application_name> -sys
  <node_name>
  # /opt/VRTS/bin/cfsumount /<volume_mount_point> <node_name>
  ```

  Make sure that CVM is the final child in the hierarchy, that it is left online, and rest of the hierarchy is taken offline on that node.

- If any parent service groups exist in the hierarchy, take them offline.

  ```
  # /opt/VRTS/bin/cfsumount /<volume_mount_point> <node_name>
  # /opt/VRTS/bin/hagrp -dep cvm

  Parent Child Relationship
  <global_application_name> cvm online local firm
  ```

  <global_application_name> is the parent service group of the CVM; take it offline.

  ```
  # /opt/VRTS/bin/hagrp -offline <global_application_name> -sys
  <node_name>
  ```

- If any local service groups exist in the hierarchy, switch them to another cluster node that is not being upgraded at the same time.
- After you upgrade these nodes, mount the file system again and bring the VCS application online on the targeted node.

```
# /opt/VRTS/bin/cfsmount /<volume_mount_point> <node_name>
# /opt/VRTS/bin/hagrp -online <global_application_name> -sys
<node_name>
```

- Consequently, you may have to take the following actions on these upgraded nodes:
  - Bring the parallel service groups online.
  - Switch the local VCS service groups back.
  - Make sure to bring all the service groups, which were taken offline before the upgrade, online again.
  - Check the replication status and the system or the service group state, and only then proceed to perform the upgrade on the next node.

3 Check the replication status on the primary site.

Post-upgrade tasks for VVR sites

To upgrade disk group and disk layout versions on replication hosts

1 Upgrade the disk group version on all the Secondaries for all the disk groups.

```
# /usr/sbin/vxdg upgrade <disk_group_name>
```

2 Upgrade the disk group version on the Primary for all the disk groups.

```
# /usr/sbin/vxdg upgrade <disk_group_name>
```

3 Upgrade the disk layout version (DLV) on the Primary for all the VxFS file systems.

```
# /opt/VRTS/bin/vxupgrade -n 17 <vxfs_mount_point_name>
# /opt/VRTS/bin/fstyp -v <disk_path_for_mount_point_volume>
```

The DLV upgrade is automatically replicated to the Secondaries.

Chapter 12

Performing a phased upgrade of SFCFSHA

This chapter includes the following topics:

- About phased upgrade
- Performing a phased upgrade using the product installer

About phased upgrade

Perform a phased upgrade to minimize the downtime for the cluster.

Depending on the situation, you can calculate the approximate downtime as follows:

Table 12-1

Fail over condition	Downtime
You can fail over all your service groups to the nodes that are up.	Downtime equals the time that is taken to offline and online the service groups.
You have a service group that you cannot fail over to a node that runs during upgrade.	Downtime for that service group equals the time that is taken to perform an upgrade and restart the node.

Prerequisites for a phased upgrade

Before you start the upgrade, confirm that you have licenses for all the nodes that you plan to upgrade.

Planning for a phased upgrade

Plan out the movement of the service groups from node-to-node to minimize the downtime for any particular service group.

Some rough guidelines follow:

- Split the cluster into two subclusters of equal or near equal size.
- Split the cluster so that your high priority service groups remain online during the upgrade of the first subcluster.
- Before you start the upgrade, back up the VCS configuration files `main.cf` and `types.cf` which are in the `/etc/VRTSvcs/conf/config/` directory.
- Before you start the upgrade make sure that all the disk groups have the latest backup of configuration files in the `/etc/vx/cbr/bk` directory. If not, then run the following command to take the latest backup.

  ```
  # /etc/vx/bin/vxconfigbackup -| [dir] [dgname|dgid]
  ```

Phased upgrade limitations

The following limitations primarily describe not to tamper with configurations or service groups during the phased upgrade:

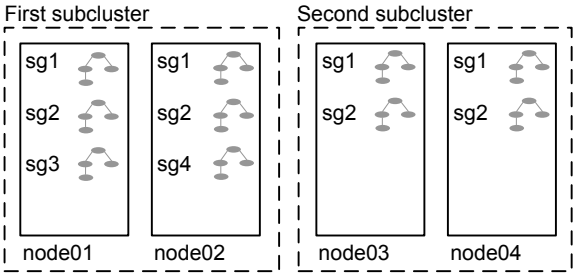
- While you perform the upgrades, do not start any modules.
- When you start the installer, only select SFCFSHA.
- While you perform the upgrades, do not add or remove service groups to any of the nodes.
- After you upgrade the first half of your cluster (the first subcluster), you need to set up password-less ssh or rsh. Create the connection between an upgraded node in the first subcluster and a node from the other subcluster. The node from the other subcluster is where you plan to run the installer and also plan to upgrade.
- Depending on your configuration, you may find that you cannot upgrade multiple nodes at the same time. You may only be able to upgrade one node at a time.
- For very large clusters, you might have to repeat these steps multiple times to upgrade your cluster.

Phased upgrade example

In this example, you have a secure cluster that you have configured to run on four nodes: node01, node02, node03, and node04. You also have four service groups:

sg1, sg2, sg3, and sg4. For the purposes of this example, the cluster is split into two subclusters. The nodes node01 and node02 are in the first subcluster, which you first upgrade. The nodes node03 and node04 are in the second subcluster, which you upgrade last.

Figure 12-1 Example of phased upgrade set up



Each service group is running on the nodes as follows:

- sg1 and sg2 are parallel service groups and run on all the nodes.
- sg3 and sg4 are failover service groups. sg3 runs on node01 and sg4 runs on node02.

In your system list, you have each service group that fails over to other nodes as follows:

- sg1 and sg2 are running on all the nodes.
- sg3 and sg4 can fail over to any of the nodes in the cluster.

Phased upgrade example overview

This example's upgrade path follows:

- Move all the failover service groups from the first subcluster to the second subcluster.
- Take all the parallel service groups offline on the first subcluster.
- Upgrade the operating system on the first subcluster's nodes, if required.
- On the first subcluster, start the upgrade using the installation program.
- Get the second subcluster ready.
- Activate the first subcluster. After activating the first cluster, switch the service groups online on the second subcluster to the first subcluster.
- Upgrade the operating system on the second subcluster's nodes, if required.
- On the second subcluster, start the upgrade using the installation program.

- Activate the second subcluster.

Performing a phased upgrade using the product installer

Performing a phased upgrade involves the following tasks:

- Moving the service groups to the second subcluster
- Upgrading the operating system on the first subcluster, if required
- Upgrading the SFCFSHA stack on the first subcluster
- Preparing the second subcluster
- Activating the first subcluster
- Upgrading the operating system on the second subcluster, if required
- Upgrading the second subcluster
- Finishing the phased upgrade

Before you start the upgrade on the first half of the cluster, back up the VCS configuration files `main.cf` and `types.cf` which are in the directory `/etc/VRTSvcs/conf/config/`.

Moving the service groups to the second subcluster

To move the service groups to the second subcluster

1. Switch failover groups from the first half of the cluster to one of the nodes in the second half of the cluster. In this procedure, `sys1` is a node in the first half of the cluster and `sys4` is a node in the second half of the cluster. Enter the following:

   ```
   # hagrp -switch failover_group -to sys4
   ```

2. On the first half of the cluster, stop all applications that are not configured under VCS. Use native application commands to stop the applications.

3 On the first half of the cluster, unmount the VxFS or CFS file systems that are not managed by VCS.

```
# mount | grep vxfs
```

Verify that no processes use the VxFS or CFS mount point. Enter the following:

```
# fuser -c mount_point
```

Stop any processes using a VxFS or CFS mount point with the mechanism provided by the application.

Unmount the VxFS or CFS file system. Enter the following:

```
# umount /mount_point
```

4 On the nodes in the first subcluster, use the following command to take all the cache area offline:

```
# sfcache offline cachename
```

5 On the first half of the cluster, bring all the VCS service groups offline including CVM group. Enter the following:

```
# hagrp -offline group_name -sys sys1
```

When the CVM group becomes OFFLINE, all the parallel service groups such as the CFS file system will also become OFFLINE on the first half of the cluster nodes.

6 Verify that the VCS service groups are offline on all the nodes in first half of the cluster. Enter the following:

```
# hagrp -state group_name
```

7 Freeze the nodes in the first half of the cluster. Enter the following:

```
# haconf -makerw
# hasys -freeze -persistent sys1
# haconf -dump -makero
```

8 Verify that no volumes remain open. Enter the following:

```
# vxprint -Aht -e v_open
```

Upgrading the operating system on the first subcluster

You can perform the operating system upgrade on the first subcluster, if required.

Before performing operating system upgrade, it is better to prevent LLT from starting automatically when the node starts. For example, you can do the following:

```
# mv /etc/lltab /etc/lltab.save
```

or you can change the `/etc/default/llt` file by setting LLT_START = **0**.

After you finish upgrading the OS, remember to change the LLT configuration to its original configuration.

Refer to the operating system's documentation for more information.

Upgrading the SFCFSHA stack on the first subcluster

To upgrade the SFCFSHA stack on the first subcluster

- On the first half of the cluster, upgrade SFCFSHA by using the installer script. For example use the installer script as shown below:

  ```
  # ./installer -upgrade sys1
  ```

  where *<sys1>* is the node on the first subcluster.

  After the upgrade for first half of the cluster is complete, no GAB ports will be shown in `gabconfig -a` output.

  Note: Do not reboot the nodes in the first subcluster until you complete preparing the second subcluster. See "Preparing the second subcluster" on page 216.

Preparing the second subcluster

To prepare the second subcluster

1. On the second half of the cluster, stop all applications that are not configured under VCS. Use native application commands to stop the application. [Downtime starts now.]
2. On the second half of the cluster, unmount the VxFS and CFS file systems that are not managed by VCS. Enter the following:

   ```
   # mount | grep vxfs
   ```

   Verify that no processes use the VxFS and CFS mount point. Enter the following:

   ```
   # fuser -c mount_point
   ```

   Stop any processes using a VxFS and CFS mount point with the mechanism provided by the application.

   Unmount the VxFS and CFS file system. Enter the following:

   ```
   # umount /mount_point
   ```

3. On the nodes in the second subcluster, use the following command to take all the cache area offline:

   ```
   # sfcache offline cachename
   ```

4. On the second half of the cluster, unfreeze all the VCS service groups on all the nodes using the following commands:

   ```
   # haconf -makerw
   # hagrp -unfreeze group_name -persistent
   # haconf -dump -makero
   ```

5. On the second half of the cluster, bring all the VCS service groups offline, including CVM group. Enter the following:

   ```
   # hagrp -offline group_name -sys sys4
   ```

6. On the second half of the cluster, verify that the VCS service groups are offline. Enter the following:

   ```
   # hagrp -state group_name
   ```

7 Stop VCS on the second half of the cluster. Enter the following:

```
# hastop -local
```

8 On the second half of the cluster, stop the following SFCFSHA modules: GLM, ODM, GMS, VxFEN, GAB, and LLT. Enter the following:

For systemd environments with supported Linux distributions:

```
# /etc/init.d/vxglm stop
# /etc/init.d/vxodm stop
# /etc/init.d/vxgms stop
# systemctl stop vxfen
# systemctl stop gab
# systemctl stop llt
```

For other supported Linux distributions:

```
# /etc/init.d/vxglm stop
# /etc/init.d/vxodm stop
# /etc/init.d/vxgms stop
# /etc/init.d/vxfen stop
# /etc/init.d/gab stop
# /etc/init.d/llt stop
```

Activating the first subcluster

To activate the first subcluster

1 On any node in the subcluster, edit the `/etc/gabtab` file to remove "-n *numericValue*" and add "-x" instead.

Note: This steps needs to be reversed after the upgrade process is successfully completed.

---

2 Restart the node.

3 On the first half of the cluster, start SFCFSHA:

```
# cd /opt/VRTS/install

# ./installer -start sys1 sys2
```

4 Unfreeze the nodes in the first half of the cluster. Enter the following:

```
# haconf -makerw
# hasys -unfreeze -persistent node_name
# haconf -dump -makero
```

5 On the first half of the cluster, bring the VCS service groups online. Enter the following:

```
# hagrp -online group_name -sys node_name
```

After you bring the CVM service group ONLINE, all the GAB ports u, v, w and f come ONLINE and all the CFS mounts service groups also come ONLINE automatically. Only failover service groups need to be brought ONLINE manually.

6 Manually mount the VxFS and CFS file systems that are not managed by VCS. [Downtime ends now.]

Upgrading the operating system on the second subcluster

To upgrade the operating system on the second subcluster

- You can perform the operating system upgrade on the second subcluster, if required.

  Refer to the *Operating system's documentation* for more information.

  Before performing OS upgrade, ensure that LLT does not start automatically when the node starts.

  You can do the following to prevent LLT from starting automatically:

  ```
  # mv /etc/lltab /etc/lltab.save
  ```

  Or you can change the `/etc/sysconfig/llt` file by setting LLT_START = 0.

  After you finish upgrading the OS, remember to change the LLT configuration to its original configuration.

Upgrading the second subcluster

To upgrade the second subcluster

1 Enter the following:

```
# ./installer -upgrade node_name
```

2 On the second half of the cluster, start SFCFSHA (if reboot is not required):

```
# cd /opt/VRTS/install

# ./installer -start sys3 sys4
```

Completing the phased upgrade

To complete the phased upgrade

1 Upgrade the cluster protocol version by performing the following tasks sequentially:

- Identify the current cluster protocol version.

  ```
  haclus -version -info
  ```

- Check whether the current version is compatible with the newer cluster version and whether it can be upgraded successfully.

  ```
  haclus -version -verify <newer-cluster-version>
  ```

  For example:

  ```
  # /opt/VRTSvcs/bin/haclus -version -verify 8.0.0.0000
  ```

- Upgrade the cluster to the newer protocol version.

  ```
  haclus -version -update <newer-cluster-version>
  ```

  For example:

  ```
  # /opt/VRTSvcs/bin/haclus -version -update 8.0.0.0000
  ```

2 Verify that the cluster UUID on the nodes in the second subcluster is the same as the cluster UUID on the nodes in the first subcluster. Run the following command to display the cluster UUID:

```
# /opt/VRTSvcs/bin/uuidconfig.pl [-rsh] -clus -display nodename
```

If the cluster UUID differs, manually copy the cluster UUID from a node in the first subcluster to the nodes in the second subcluster. For example:

```
# /opt/VRTSvcs/bin/uuidconfig.pl [-rsh] -clus -copy -from_sys \
node01 -to_sys node03 node04
```

3. On the nodes in the second subcluster, all the GAB ports a, b, d, h, m, u, v, w and f are ONLINE. Also all the CFS mounts service groups come online automatically.
4. Manually mount the VxFS and CFS file systems that are not managed by VCS in the second half of the cluster.
5. Find out which node is the CVM master. Enter the following:

   ```
   # vxdctl -c mode
   ```

6. On the CVM master node, upgrade the CVM protocol. Enter the following:

   ```
   # vxdctl upgrade
   ```

7. On any node in the first subcluster, edit the `/etc/gabtab` file to remove "-x" and add the original value "-n *numericValue*" instead.

   Note: This step reverses the first step that you perform to activate the first subcluster.

Chapter 13

Performing an automated SFCFSHA upgrade using response files

This chapter includes the following topics:

- Upgrading SFCFSHA using response files
- Response file variables to upgrade SFCFSHA
- Sample response file for full upgrade of SFCFSHA
- Sample response file for rolling upgrade of SFCFSHA
- Sample response file for rolling upgrade of SFCFSHA from 7.4.2 or later to 8.0

Upgrading SFCFSHA using response files

Typically, you can use the response file that the installer generates after you perform SFCFSHA upgrade on one system to upgrade SFCFSHA on other systems.

To perform automated SFCFSHA upgrade

1. Make sure the systems where you want to upgrade SFCFSHA meet the upgrade requirements.
2. Make sure the pre-upgrade tasks are completed.
3. Copy the response file to the system where you want to upgrade SFCFSHA.
4. Edit the values of the response file variables as necessary.

5 Mount the product disc and navigate to the folder that contains the installation program.

6 Start the upgrade from the system to which you copied the response file. For example:

```
# ./installer -responsefile /tmp/response_file
```

Where /tmp/*response_file* is the response file's full path name.

Response file variables to upgrade SFCFSHA

Table 13-1 lists the response file variables that you can define to configure SFCFSHA.

Table 13-1 Response file variables for upgrading SFCFSHA

Variable	Description
CFG{accepteula}	Specifies whether you agree with the EULA.pdf file on the media. List or scalar: scalar Optional or required: required
CFG{systems}	List of systems on which the product is to be installed or uninstalled. List or scalar: list Optional or required: required
CFG{upgrade}	Upgrades all RPMs installed. List or scalar: list Optional or required: required
CFG{keys}{keyless} CFG{keys}{licensefile}	`CFG{keys}{keyless}` gives a list of keyless keys to be registered on the system. `CFG{keys}{licensefile}` gives the absolute file path to the permanent license key to be registered on the system. List or scalar: list Optional or required: required

Table 13-1 Response file variables for upgrading SFCFSHA *(continued)*

Variable	Description
CFG{opt}{keyfile}	Defines the location of an ssh keyfile that is used to communicate with all remote systems. List or scalar: scalar Optional or required: optional
CFG{opt}{tmppath}	Defines the location where a working directory is created to store temporary files and the RPMs that are needed during the install. The default location is /opt/VRTStmp. List or scalar: scalar Optional or required: optional
CFG{opt}{logpath}	Mentions the location where the log files are to be copied. The default location is /opt/VRTS/install/logs. List or scalar: scalar Optional or required: optional
CFG{mirrordgname}{system}	If the root dg is encapsulated and you select split mirror is selected: Splits the target disk group name for a system. List or scalar: scalar Optional or required: optional
CFG{splitmirror}{system}	If the root dg is encapsulated and you select split mirror is selected: Indicates the system where you want a split mirror backup disk group created. List or scalar: scalar Optional or required: optional

Table 13-1 Response file variables for upgrading SFCFSHA *(continued)*

Variable	Description
CFG{opt}{disable_dmp_native_support}	If it is set to 1, Dynamic Multi-pathing support for the native LVM volume groups and ZFS pools is disabled after upgrade. Retaining Dynamic Multi-pathing support for the native LVM volume groups and ZFS pools during upgrade increases RPM upgrade time depending on the number of LUNs and native LVM volume groups and ZFS pools configured on the system. List or scalar: scalar Optional or required: optional
CFG{opt}{patch_path}	Defines the path of a patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed . List or scalar: scalar Optional or required: optional
CFG{opt}{patch2_path}	Defines the path of a second patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed. List or scalar: scalar Optional or required: optional
CFG{opt}{patch3_path}	Defines the path of a third patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed. List or scalar: scalar Optional or required: optional
CFG{opt}{patch4_path}	Defines the path of a fourth patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed. List or scalar: scalar Optional or required: optional

Table 13-1 Response file variables for upgrading SFCFSHA *(continued)*

Variable	Description
CFG{opt}{patch5_path}	Defines the path of a fifth patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed. List or scalar: scalar Optional or required: optional
CFG{rootsecusrgrps}	Defines if the user chooses to grant read access to the cluster only for root and other users/usergroups which are granted explicit privileges on VCS objects. List or scalar: scalar Optional or required: optional
CFG{secusrgrps}	Defines the usergroup names that are granted read access to the cluster. List or scalar: scalar Optional or required: optional

Sample response file for full upgrade of SFCFSHA

The following example shows a response file for upgrading SFCFSHA with keyless key.

```
our %CFG;

$CFG{accepteula}=1;
$CFG{keys}{keyless}=[ "ENTERPRISE" ];
$CFG{opt}{gco}=1;
$CFG{opt}{redirect}=1;
$CFG{opt}{upgrade}=1;
$CFG{opt}{vr}=1;
$CFG{prod}="ENTERPRISE80";
$CFG{systems}=[ "sys01","sys02" ];
$CFG{vcs_allowcomms}=1;

1;
```

Sample response file for rolling upgrade of SFCFSHA

```
our %CFG;
$CFG{accepteula}=1;
$CFG{opt}{gco}=1;
$CFG{opt}{redirect}=1;
$CFG{opt}{rolling_upgrade}=1;
$CFG{opt}{rollingupgrade_phase1}=1;
$CFG{phase1}{"0"}=[ qw( node1 ) ];
## change to the systems of the first sub-cluster
$CFG{phase1}{"1"}=[ qw( node2 ) ];
## change to the systems of the second sub-cluster
$CFG{opt}{rollingupgrade_phase2}=1;
$CFG{reuse_config}=1;
$CFG{systems}=[ qw( node1 node2 ) ];
## change to all the systems of the whole cluster
$CFG{vcs_allowcomms}=1;
1;
```

Sample response file for rolling upgrade of SFCFSHA from 7.4.2 or later to 8.0

```
our %CFG;
$CFG{accepteula}=1;
$CFG{opt}{gco}=1;
$CFG{opt}{redirect}=1;
$CFG{opt}{rolling_upgrade}=1;
$CFG{subcluster}{"0"}=[ qw( node1 ) ];
## change to the systems of the first sub-cluster
$CFG{subcluster}{"1"}=[ qw( node2 ) ];
## change to the systems of the second sub-cluster
$CFG{reuse_config}=1;
$CFG{systems}=[ qw( node1 node2 ) ];
## change to all the systems of the whole cluster
$CFG{vcs_allowcomms}=1;
1;
```

Chapter 14

Upgrading Volume Replicator

This chapter includes the following topics:

- Upgrading Volume Replicator
- Upgrading VVR without disrupting replication

Upgrading Volume Replicator

If a previous version of Volume Replicator (VVR) is configured, the product installer upgrades VVR automatically when you upgrade the Storage Foundation products.

You have the option to upgrade without disrupting replication.

Upgrading VVR without disrupting replication

This section describes the upgrade procedure from an earlier version of VVR to the current version of VVR when replication is in progress, assuming that you do not need to upgrade all the hosts in the RDS simultaneously.

You may also need to set up replication between versions.

See "Planning an upgrade from the previous VVR version" on page 182.

When both the Primary and the Secondary have the previous version of VVR installed, the upgrade should be performed on the secondary site first and then on the primary site.

Note: If you have a cluster setup, you must upgrade all the nodes in the cluster at the same time.

Upgrading VVR sites for InfoScale 7.3.1 or earlier

Use the product installer to first upgrade VVR on the Secondaries and then on the Primary.

To upgrade a Secondary

1. Stop the replication to a Secondary by initiating `stoprep` on the Primary.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> stoprep <RVG_name>
   <secondary_hostname>
   ```

2. Verify that the replication has stopped.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
   ```

3. Upgrade VVR from any version between 6.2.1 and 7.3.1 to VVR 8.0 on the Secondary.

4. Start the replication to the Secondary host by initiating `startrep` on the Primary.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> startrep <RVG_name>
   <secondary_hostname>
   ```

5. Verify that the replication has started.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
   ```

To upgrade the Primary

1. Verify that the replication status is consistent and up-to-date.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
   ```

2. Take the applications and the mount points down.

3. Stop the replication to a Secondary by initiating `stoprep` on the Primary.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> stoprep <RVG_name>
   <secondary_hostname>
   ```

4. Verify that the replication has stopped.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
   ```

5. Upgrade VVR from any version between 6.2.1 and 7.3.1 to VVR 8.0 on the Primary.

6. Start the replication to the Secondary host by initiating `startrep` on the Primary.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> startrep <RVG_name>
   <secondary_hostname>
   ```

7 Verify that the replication has started.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

8 Mount all the file systems and start all the applications on the Primary.

Upgrading VVR sites with InfoScale 7.4 or later

Use the product installer to first upgrade VVR on the Secondaries and then on the Primary.

To upgrade a Secondary

1 Pause the replication to a Secondary by initiating `pauserep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> pauserep <RVG_name>
<secondary_hostname>
```

2 Verify that the replication has paused.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

3 Upgrade from VVR 7.4 or later to VVR 8.0 on the Secondary.

4 Resume the replication to the Secondary by initiating `resumerep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> resumerep <RVG_name>
<secondary_hostname>
```

5 Verify that the replication has resumed.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

To upgrade the Primary

1 Verify that the replication status is consistent and up-to-date.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

2 Take the applications and the mount points down.

3 Pause the replication to the Secondary by initiating `pauserep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> pauserep <RVG_name>
<secondary_hostname>
```

4 Verify that the replication has paused.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

5 Upgrade from VVR 7.4 or later to VVR 8.0 on the Primary.

6 Resume the replication to the Secondary by initiating `resumerep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> resumerep <RVG_name>
<secondary_hostname>
```

7 Verify that the replication has resumed.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

8 Mount all the file systems and start all the applications on the Primary.

Upgrading VVR sites in VCS control for InfoScale 7.3.1 or earlier

Use the product installer to first upgrade VVR on the Secondaries and then on the Primary.

To upgrade a Secondary

1 Stop the replication to a Secondary by initiating `stoprep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> stoprep <RVG_name>
<secondary_hostname>
```

2 Verify that the replication has stopped.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

3 Stop VCS on the Secondary.

```
# /opt/VRTS/bin/hastop -all
```

4 Upgrade VVR from any version between 6.2.1 and 7.3.1 to VVR 8.0 on the Secondary.

VCS starts automatically after the upgrade.

5 Start the replication to the Secondary host by initiating `startrep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> startrep <RVG_name>
<secondary_hostname>
```

6 Verify that the replication has started.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

To upgrade the Primary

1 Verify that the replication status is consistent and up-to-date.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

2 Take the applications and the mount points down by using the VCS Application or the VCS Mount service groups.

3 Stop the replication to a Secondary by initiating `stoprep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> stoprep <RVG_name>
<secondary_hostname>
```

4 Verify that the replication has stopped.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

5 Upgrade VVR from any version between 6.2.1 and 7.3.1 to VVR 8.0 on the Primary.

VCS starts automatically after the upgrade.

6 Start the replication to the Secondary host by initiating `startrep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> startrep <RVG_name>
<secondary_hostname>
```

7 Verify that the replication has started.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

8 Mount all the file systems and start all the applications on the Primary.

Upgrading VVR sites in VCS control for InfoScale 7.4 or later

Use the product installer to first upgrade VVR on the Secondaries and then on the Primary.

To upgrade a Secondary

1 Pause the replication to a Secondary by initiating `pauserep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> pauserep <RVG_name>
<secondary_hostname>
```

2 Verify that the replication has paused.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

3 Stop VCS on the Secondary.

```
# /opt/VRTS/bin/hastop -all
```

4 Upgrade from VVR 7.4 or later to VVR 8.0 on the Secondary.

VCS starts automatically after the upgrade.

5 Resume the replication to the Secondary by initiating `resumerep` on the Primary.

```
# /usr/sbin/vradmin -g <disk_group_name> resumerep <RVG_name>
<secondary_hostname>
```

6 Verify that the replication has resumed.

```
# /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
```

To upgrade the Primary

1. Verify that the replication status is consistent and up-to-date.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
   ```

2. Take the applications and the mount points down by using the VCS Application or the VCS Mount service groups.

3. Pause the replication to the Secondary by initiating `pauserep` on the Primary.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> pauserep <RVG_name>
   <secondary_hostname>
   ```

4. Verify that the replication has paused.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
   ```

5. Stop VCS on the Secondary.

   ```
   # /opt/VRTS/bin/hastop -all
   ```

6. Upgrade from VVR 7.4 or later to VVR 8.0 on the Primary.

   VCS starts automatically after the upgrade.

7. Resume the replication to the Secondary by initiating `resumerep` on the Primary.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> resumerep <RVG_name>
   <secondary_hostname>
   ```

8. Verify that the replication has resumed.

   ```
   # /usr/sbin/vradmin -g <disk_group_name> -l repstatus <RVG_name>
   ```

9. Mount all the file systems and start all the applications on the Primary.

Post-upgrade tasks for VVR sites

To upgrade disk group and disk layout versions on replication hosts

1. Upgrade the disk group version on all the Secondaries for all the disk groups.

   ```
   # /usr/sbin/vxdg upgrade <disk_group_name>
   ```

2. Upgrade the disk group version on the Primary for all the disk groups.

   ```
   # /usr/sbin/vxdg upgrade <disk_group_name>
   ```

3. Upgrade the disk layout version (DLV) on the Primary for all the VxFS file systems.

   ```
   # /opt/VRTS/bin/vxupgrade -n 17 <vxfs_mount_point_name>
   ```

   ```
   # /opt/VRTS/bin/fstyp -v <disk_path_for_mount_point_volume>
   ```

   The DLV upgrade changes are automatically replicated to the secondaries.

Chapter 15

Upgrading VirtualStore

This chapter includes the following topics:

- Supported upgrade paths
- Upgrading SVS to SFCFSHA 8.0

Supported upgrade paths

The following tables describe upgrading VirtualStore (SVS) to Storage Foundation Cluster File System High Availability (SFCFSHA) 8.0.

- 6.0
- 6.0 RP1
- 6.0.1

Upgrading SVS to SFCFSHA 8.0

This section describes how to upgrade from VirtualStore (SVS) to Storage Foundation Cluster File System High Availability (SFCFSHA) 8.0.

To upgrade SVS to SFCFSHA 8.0

1 Unregisters SVS plug-in at VMware vCenter Server:

```
# svsvmwadm -a unregister -v vcip -u user -p pass
```

For example:

```
# svsvmwadm -a unregister -v 10.143.007.132 -u admin -p xxxxxx
```

2 Stop the SVS server:

```
# /opt/VRTS/bin/svsweb stop
```

3 Choose your method of upgrade and then upgrade to Enterprise SFCFSHA

Chapter 16

Performing post-upgrade tasks

This chapter includes the following topics:

- Resetting DAS disk names to include host name in FSS environments
- Re-joining the backup boot disk group into the current disk group
- Reverting to the backup boot disk group after an unsuccessful upgrade
- CVM master node needs to assume the logowner role for VCS managed VVR resources

Resetting DAS disk names to include host name in FSS environments

If you are on a version earlier than 7.1, the VxVM disk names in the case of DAS disks in FSS environments, must be regenerated to use the host name as a prefix. The host prefix helps to uniquely identify the origin of the disk. For example, the device name for the disk *disk1* on the host *sys1* is now displayed as *sys1_disk1*.

To regenerate the disk names, run the following command:

```
# vxddladm -c assign names
```

The command must be run on each node in the cluster.

Re-joining the backup boot disk group into the current disk group

Note: Root Disk Encapsulation (RDE) is not supported on Linux from 7.3.1 onwards.

Perform this procedure to rejoin the backup boot disk if you split the mirrored boot disk during upgrade. After a successful upgrade and reboot, you no longer need to keep the boot disk group backup.

To re-join the backup boot disk group

- Re-join the *backup_bootdg* disk group to the boot disk group.

  ```
  # /etc/vx/bin/vxrootadm -Y join backup_bootdg
  ```

  where the `-Y` option indicates a silent operation, and *backup_bootdg* is the name of the backup boot disk group that you created during the upgrade.

Reverting to the backup boot disk group after an unsuccessful upgrade

Note: Root Disk Encapsulation (RDE) is not supported on Linux from 7.3.1 onwards.

Perform this procedure if your upgrade was unsuccessful and you split the mirrored boot disk to back it up during upgrade. You can revert to the backup that you created when you upgraded.

To revert the backup boot disk group after an unsuccessful upgrade

1. To determine the boot disk groups, look for the *rootvol* volume in the output of the `vxprint` command.

   ```
   # vxprint
   ```

2. Use the `vxdg` command to find the boot disk group where you are currently booted.

   ```
   # vxdg bootdg
   ```

3 Boot the operating system from the backup boot disk group.

4 Join the original boot disk group to the backup disk group.

```
# /etc/vx/bin/vxrootadm -Y join original_bootdg
```

where the `-Y` option indicates a silent operation, and *original_bootdg* is the boot disk group that you no longer need.

CVM master node needs to assume the logowner role for VCS managed VVR resources

If you use VCS to manage RVGLogowner resources in an SFCFSHA environment or an SF Oracle RAC environment, Veritas recommends that you perform the following procedures. These procedures ensure that the CVM master node always assumes the logowner role. Not performing these procedures can result in unexpected issues that are due to a CVM slave node that assumes the logowner role.

For a service group that contains an RVGLogowner resource, change the value of its `TriggersEnabled` attribute to PREONLINE to enable it.

To enable the TriggersEnabled attribute from the command line on a service group that has an RVGLogowner resource

- On any node in the cluster, perform the following command:

```
# hagrp -modify RVGLogowner_resource_sg TriggersEnabled PREONLINE
```

Where *RVGLogowner_resource_sg* is the service group that contains the RVGLogowner resource.

To enable the preonline_vvr trigger, do one of the following:

- If preonline trigger script is not already present, copy the preonline trigger script from the sample triggers directory into the triggers directory:

```
# cp /opt/VRTSvcs/bin/sample_triggers/VRTSvcs/preonline_vvr
/opt/VRTSvcs/bin/triggers/preonline
```

Change the file permissions to make it executable.

- If preonline trigger script is already present, create a directory such as `/preonline` and move the existing preonline trigger as T0preonline to that directory. Copy the preonline_vvr trigger as T1preonline to the same directory.
- If you already use multiple triggers, copy the preonline_vvr trigger as TNpreonline, where TN is the next higher TNumber.

Section 4

Post-configuration tasks

- Chapter 17. Performing post configuration tasks

Chapter 17

Performing post configuration tasks

This chapter includes the following topics:

- Upgrading disk layout versions
- Switching on Quotas
- About enabling LDAP authentication for clusters that run in secure mode
- About configuring authentication for SFDB tools
- Enabling data encryption over wire

Upgrading disk layout versions

In this release, you can create and mount only file systems with disk layout version 13, 14, 15, 16, and 17. You can local mount disk layout version 7, 8, 9, 10, 11, and 12 to upgrade to a later disk layout version.

Note: If you plan to use 64-bit quotas, you must upgrade to the disk layout version 10 or later.

Disk layout version 7, 8, 9, 10, 11, and 12 are deprecated and you cannot cluster mount an existing file system that has any of these versions. To upgrade a cluster file system from any of these deprecated versions, you must local mount the file system and then upgrade it using the `vxupgrade` utility or the `vxfsconvert` utility.

The `vxupgrade` utility enables you to upgrade the disk layout while the file system is online. However, the `vxfsconvert` utility enables you to upgrade the disk layout while the file system is offline.

If you use the `vxupgrade` utility, you must incrementally upgrade the disk layout versions. However, you can directly upgrade to a desired version, using the `vxfsconvert` utility.

For example, to upgrade from disk layout version 7 to a disk layout version 17, using the `vxupgrade` utility:

```
# vxupgrade -n 8 /mnt
# vxupgrade -n 9 /mnt
# vxupgrade -n 10 /mnt
# vxupgrade -n 11 /mnt
# vxupgrade -n 12 /mnt
# vxupgrade -n 13 /mnt
# vxupgrade -n 14 /mnt
# vxupgrade -n 15 /mnt
# vxupgrade -n 16 /mnt
# vxupgrade -n 17 /mnt
```

See the `vxupgrade`(1M) manual page.

See the `vxfsconvert`(1M) manual page.

Note: Veritas recommends that before you begin to upgrade the product version, you must upgrade the existing file system to the highest supported disk layout version. Once a disk layout version has been upgraded, it is not possible to downgrade to the previous version.

Use the following command to check your disk layout version:

```
# fstyp -v /dev/vx/dsk/dg1/vol1 | grep -i version
```

For more information about disk layout versions, see the *Cluster File System High Availability Administrator's Guide*.

Switching on Quotas

This turns on the group and user quotas once all the nodes are upgraded to 8.0, if it was turned off earlier.

To turn on the group and user quotas

- Switch on quotas:

  ```
  # vxquotaon -av
  ```

About enabling LDAP authentication for clusters that run in secure mode

Veritas Product Authentication Service (AT) supports LDAP (Lightweight Directory Access Protocol) user authentication through a plug-in for the authentication broker. AT supports all common LDAP distributions such as OpenLDAP and Windows Active Directory.

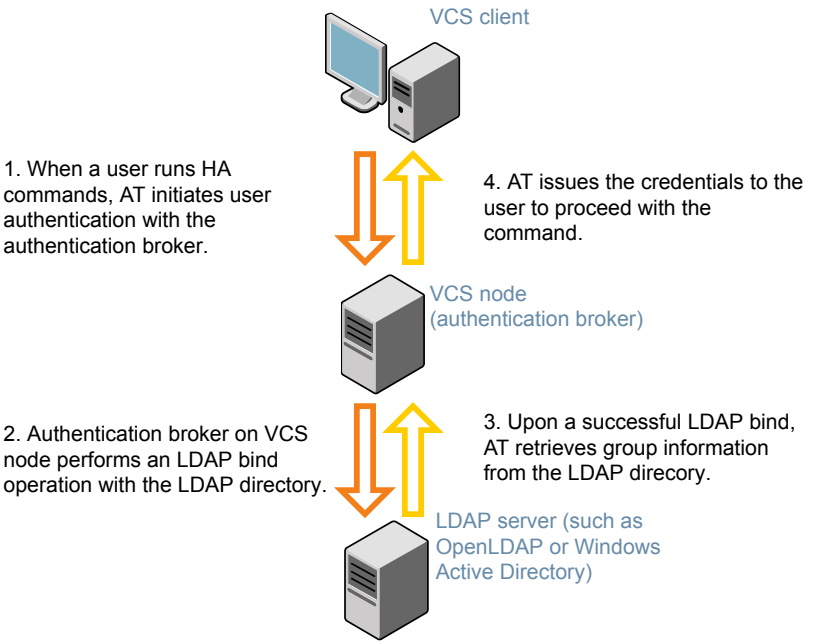
For a cluster that runs in secure mode, you must enable the LDAP authentication plug-in if the VCS users belong to an LDAP domain.

If you have not already added VCS users during installation, you can add the users later.

See the *Cluster Server Administrator's Guide* for instructions to add VCS users.

Figure 17-1 depicts the SFCFSHA cluster communication with the LDAP servers when clusters run in secure mode.

Figure 17-1 Client communication with LDAP servers



The LDAP schema and syntax for LDAP commands (such as, ldapadd, ldapmodify, and ldapsearch) vary based on your LDAP implementation.

Before adding the LDAP domain in Veritas Product Authentication Service, note the following information about your LDAP environment:

- The type of LDAP schema used (the default is RFC 2307)
  - UserObjectClass (the default is posixAccount)
  - UserObject Attribute (the default is uid)
  - User Group Attribute (the default is gidNumber)
  - Group Object Class (the default is posixGroup)
  - GroupObject Attribute (the default is cn)
  - Group GID Attribute (the default is gidNumber)
  - Group Membership Attribute (the default is memberUid)
- URL to the LDAP Directory
- Distinguished name for the user container (for example, UserBaseDN=ou=people,dc=comp,dc=com)
- Distinguished name for the group container (for example, GroupBaseDN=ou=group,dc=comp,dc=com)

Enabling LDAP authentication for clusters that run in secure mode

The following procedure shows how to enable the plug-in module for LDAP authentication. This section provides examples for OpenLDAP and Windows Active Directory LDAP distributions.

Before you enable the LDAP authentication, complete the following steps:

- Make sure that the cluster runs in secure mode.

  ```
  # haclus -value SecureClus
  ```

  The output must return the value as 1.

- Make sure that the AT version is 6.1.6.0 or later.

  ```
  # /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/vssat showversion
  vssat version: 6.1.14.18
  ```

To enable OpenLDAP authentication for clusters that run in secure mode

1 Run the LDAP configuration tool `atldapconf` using the `-d` option. The `-d` option discovers and retrieves an LDAP properties file which is a prioritized attribute list.

```
# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/atldapconf \
-d -s domain_controller_name_or_ipaddress -u domain_user

Attribute list file name not provided, using AttributeList.txt

Attribute file created.
```

You can use the `catatldapconf` command to view the entries in the attributes file.

2 Run the LDAP configuration tool using the `-c` option. The `-c` option creates a CLI file to add the LDAP domain.

```
# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/atldapconf \
-c -d LDAP_domain_name

Attribute list file not provided, using default AttributeList.txt

CLI file name not provided, using default CLI.txt

CLI for addldapdomain generated.
```

3 Run the LDAP configuration tool `atldapconf` using the `-x` option. The `-x` option reads the CLI file and executes the commands to add a domain to the AT.

```
# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/atldapconf -x

Using default broker port 14149

CLI file not provided, using default CLI.txt

Looking for AT installation...

AT found installed at ./vssat

Successfully added LDAP domain.
```

4 Check the AT version and list the LDAP domains to verify that the Windows Active Directory server integration is complete.

```
# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/vssat showversion

vssat version: 6.1.12.8

# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/vssat listldapdomains

Domain Name : mydomain.com

Server URL : ldap://192.168.20.32:389

SSL Enabled : No

User Base DN : CN=people,DC=mydomain,DC=com

User Object Class : account

User Attribute : cn

User GID Attribute : gidNumber

Group Base DN : CN=group,DC=domain,DC=com

Group Object Class : group

Group Attribute : cn

Group GID Attribute : cn

Auth Type : FLAT

Admin User :

Admin User Password :

Search Scope : SUB
```

5 Check the other domains in the cluster.

```
# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/vssat showdomains -p vx
```

The command output lists the number of domains that are found, with the domain names and domain types.

6 Generate credentials for the user.

```
# unset EAT_LOG

# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/vssat authenticate \
-d ldap:LDAP_domain_name -p user_name -s user_password -b \
localhost:14149
```

7 Add non-root users as applicable.

```
# useradd user1

# passwd pw1

Changing password for "user1"

user1's New password:

Re-enter user1's new password:

# su user1

# bash

# id

uid=204(user1) gid=1(staff)

# pwd

# mkdir /home/user1

# chown user1 /home/ user1
```

8 Add the non-root user to the VCS configuration.

```
# haconf -makerw
# hauser -add user1
# haconf -dump -makero
```

9 Log in as non-root user and run VCS commands as LDAP user.

```
# cd /home/user1

# ls

# cat .vcspwd

101 localhost mpise LDAP_SERVER ldap

# unset VCS_DOMAINTYPE

# unset VCS_DOMAIN

# /opt/VRTSvcs/bin/hasys -state

   #System        Attribute          Value

cluster1:sysA     SysState           FAULTED

cluster1:sysB     SysState           FAULTED

cluster2:sysC     SysState           RUNNING

cluster2:sysD     SysState           RUNNING
```

About configuring authentication for SFDB tools

To configure authentication for Storage Foundation for Databases (SFDB) tools, perform the following tasks:

Configure the `vxdbd` daemon to require authentication	See "Configuring vxdbd for SFDB tools authentication" on page 248.
Add a node to a cluster that is using authentication for SFDB tools	See "Adding nodes to a cluster that is using authentication for SFDB tools" on page 279.

Configuring vxdbd for SFDB tools authentication

To configure vxdbd, perform the following steps as the root user

1. Run the `sfae_auth_op` command to set up the authentication services.

```
# /opt/VRTS/bin/sfae_auth_op -o setup
Setting up AT
Starting SFAE AT broker
Creating SFAE private domain
Backing up AT configuration
Creating principal for vxdbd
```

2. Stop the `vxdbd` daemon.

```
# /opt/VRTS/bin/sfae_config disable
vxdbd has been disabled and the daemon has been stopped.
```

3. Enable authentication by setting the `AUTHENTICATION` key to `yes` in the `/etc/vx/vxdbed/admin.properties` configuration file.

   If `/etc/vx/vxdbed/admin.properties` does not exist, then use`cp /opt/VRTSdbed/bin/admin.properties.example /etc/vx/vxdbed/admin.properties`.

4. Start the `vxdbd` daemon.

```
# /opt/VRTS/bin/sfae_config enable
vxdbd has been enabled and the daemon has been started.
It will start automatically on reboot.
```

The `vxdbd` daemon is now configured to require authentication.

Enabling data encryption over wire

Post VVR upgrade, if you want to set up replication with encrypted data volumes, ensure that you meet the following prerequisites:

- Install the user certificates on all nodes, in the `/etc/vx/vvr/` directory, to establish SSL socket communication between user space utilities (`vxrsyncd` and `vradmind`). Secure communication between user space utilities is created using certificate-based SSL socket.

Certificate location	Description
/etc/vx/vvr/key.pem	Specifies the location of the private key. The user space utilities uses the private key in a Privacy Enhanced Mail (PEM) format.
/etc/vx/vvr/cert.pem	Specifies the location of the certificates. The user space utilities uses the certificates in PEM format.
/etc/vx/vvr/cacert.pem	Specifies the location of the root certificate. The user space utilities uses the root certificate for mutual authentication in Privacy Enhanced Mail (PEM) format.

- Restart the VVR utilities using the following commands for the certificate to take effect:
  - `/usr/sbin/vxstart_vvr stop`
  - `/usr/sbin/vxstart_vvr start`

After you install the user certificates and restart the VVR utilities, perform the following steps:

Note: Ensure that you have upgraded VVR at both the sites. Failing this, the `vradmin startrep` and the `addsec` commands may fail with **-encrypted** option.

---

1. Stop replication to the Secondary using the following command:

   `vradmin -g` *`diskgroup`* `-f stoprep` *`rvgname`*

2. Dissociate the Primary and Secondary Rlinks using the following command:

   `vradmin -g` *`diskgroup`* `delsec` *`local_rvgname sec_hostname/ip`*

   The `vradmin delsec` command performs the following by default:

   - Dissociates the data volumes and SRL from the Secondary RVG.
   - Removes the Secondary RVG from its RDS, deletes the Secondary RVG, and deletes the associated Primary and Secondary RLINKs.

3. Add the Secondary to the RDS using the following command on the Primary:

   `vradmin -g` *`diskgroup`* `-encrypted addsec` *`local_rvgname pri_hostname \ sec_hostname`*

   It also creates an encrypted Rlink between primary and secondary host.

4. Synchronize the Secondary and start replication using automatic synchronization, using the following command:

```
vradmin -g diskgroup -encrypted -a startrep local_rvgname
sec_hostname
```

For more information about enabling data encryption over the wire, see, *Veritas InfoScale Replication Administrator's Guide*

Section 5

Configuration of disaster recovery environments

- Chapter 18. Configuring disaster recovery environments

Chapter 18

Configuring disaster recovery environments

This chapter includes the following topics:

- Disaster recovery options for SFCFSHA
- About setting up a campus cluster for disaster recovery
- About setting up a global cluster environment for SFCFSHA
- About configuring a parallel global cluster using Volume Replicator (VVR) for replication

Disaster recovery options for SFCFSHA

SFCFSHA supports configuring a disaster recovery environment using:

- Campus cluster
- Global clustering option (GCO) with replication
- Global clustering using Volume Replicator (VVR) for replication

For more about planning for disaster recovery environments:

You can install and configure clusters for your disaster recovery environment as you would for any cluster using the procedures in this installation guide.

For a high level description of the tasks for implementing disaster recovery environments:

See “About setting up a campus cluster for disaster recovery” on page 253.

See “About setting up a global cluster environment for SFCFSHA” on page 255.

See "About configuring a parallel global cluster using Volume Replicator (VVR) for replication" on page 256.

For complete details for configuring your disaster recovery environment once clusters are installed and configured:

See the *Veritas InfoScale™ Disaster Recovery Implementation Guide*.

About setting up a campus cluster for disaster recovery

Campus clusters:

- Are connected using a high speed cable that guarantees network access between the nodes
- Provide local high availability and disaster recovery functionality in a single cluster
- Employ shared disk groups mirrored across sites with Veritas Volume Manager (VxVM)
- Are supported by Storage Foundation Cluster File System High Availability (SFCFSHA)

The following high-level tasks illustrate the setup steps for a campus cluster in a parallel cluster database environment. The example values are given for SF for Oracle RAC and should be adapted for an SFCFSHA cluster using another database application.

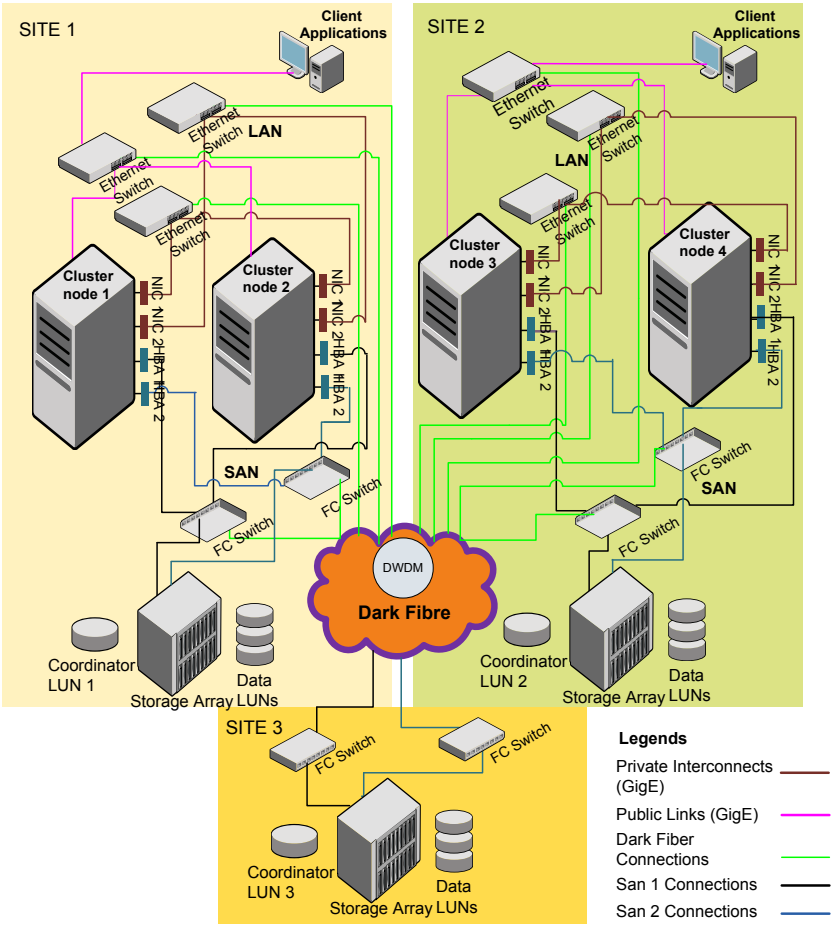
Table 18-1 Tasks for setting up a parallel campus cluster for disaster recovery

Task	Description
Prepare to set up campus cluster configuration	See the *Veritas InfoScale™ Disaster Recovery Implementation Guide*.
Configure I/O fencing to prevent data corruption	See the *Veritas InfoScale™ Disaster Recovery Implementation Guide*.
Prepare to install Oracle RAC Clusterware and database binaries	See the *Storage Foundation for Oracle RAC Configuration and Upgrade Guide*.
Prepare to install your database software.	See your database documentation.
Configure VxVM disk groups for campus cluster	See the *Veritas InfoScale™ Disaster Recovery Implementation Guide*.

Table 18-1 Tasks for setting up a parallel campus cluster for disaster recovery *(continued)*

Task	Description
Install Oracle RAC Clusterware and database binaries	For Oracle RAC, see the *Storage Foundation for Oracle RAC Configuration and Upgrade Guide*. For SFCFSHA, see your database documentation.
Install your database software.	See your database documentation.
Configure VCS service groups	See the *Veritas InfoScale™ Disaster Recovery Implementation Guide*.

Figure 18-1 Sample SF Oracle RAC configuration



Although a Coordination Point (CP) server is not used in the current example, it can also be used instead of a third site for a coordinator disk.

About setting up a global cluster environment for SFCFSHA

Configuring a global cluster for environment with parallel clusters requires the coordination of many component setup tasks. The procedures provided here are

guidelines. You will need the *Veritas InfoScale Disaster Recovery Implementation Guide* to install and configure SFCFSHA on each cluster.

- Configure a SFCFSHA cluster at the primary site
- Configure an SFCFSHA cluster at the secondary site
- Configure a global cluster environment
- Test the HA/DR configuration

Upon successful testing, you can bring the environment into production

About configuring a parallel global cluster using Volume Replicator (VVR) for replication

Configuring a global cluster for environment with SFCFSHA and Volume Replicator requires the coordination of many component setup tasks. The tasks listed below are guidelines.

Before configuring two clusters for global clustering, you must verify that:

- You have the correct installation options enabled for SFCFSHA, whether you are using keyless licensing or installing keys manually. You must have the GCO option for a global cluster and VVR enabled.
  Review SFCFSHA requirements and licensing information.
- Both clusters have SFCFSHA software installed and configured.

Note: You can install and configure both clusters at the same time, or you can configure the second cluster at a later time than the first.

You can use this guide to install and configure SFCFSHA on each cluster. For details for configuring a global cluster environment and replication between the clusters using VVR:

See the *Veritas InfoScale Disaster Recovery Implementation Guide*.

With two clusters installed and configured , you are ready to configure a global cluster environment using VVR. You must perform the following tasks to modify both cluster configurations to support replication in the global cluster environment.

Once the global clusters and replication with VVR are configured, the following replication use cases are supported for it:

- Migration of the role of the primary site to the remote site
- Takeover of the primary site role by the secondary site

- Migrate the role of primary site to the secondary site
- Migrate the role of new primary site back to the original primary site
- Take over after an outage
- Resynchronize after an outage
- Update the rlink to reflect changes

For details on the replication use cases:

See the *Veritas InfoScale Disaster Recovery Implementation Guide*.

Section 6

Adding and removing nodes

- Chapter 19. Adding a node to SFCFSHA clusters
- Chapter 20. Removing a node from SFCFSHA clusters

Chapter 19

Adding a node to SFCFSHA clusters

This chapter includes the following topics:

- About adding a node to a cluster
- Before adding a node to a cluster
- Adding a node to a cluster using the Veritas InfoScale installer
- Adding the node to a cluster manually
- Adding a node using response files
- Configuring server-based fencing on the new node
- Adding nodes to a cluster that is using authentication for SFDB tools
- Updating the Storage Foundation for Databases (SFDB) repository after adding a node
- Sample configuration file for adding a node to the cluster

About adding a node to a cluster

After you install Veritas InfoScale and create a cluster, you can add and remove nodes from the cluster.You can create clusters of up to 64 nodes.

You can add a node:

- Using the product installer
- Manually

The following table provides a summary of the tasks required to add a node to an existing SFCFSHA cluster.

Table 19-1 Tasks for adding a node to a cluster

Step	Description
Complete the prerequisites and preparatory tasks before adding a node to the cluster.	See "Before adding a node to a cluster" on page 260.
Add a new node to the cluster.	See "Adding a node to a cluster using the Veritas InfoScale installer" on page 263. See "Adding the node to a cluster manually" on page 266.
Complete the configuration of the new node after adding it to the cluster.	See "Configuring Cluster Volume Manager (CVM) and Cluster File System (CFS) on the new node" on page 274.
If you are using the Storage Foundation for Databases (SFDB) tools, you must update the repository database.	See "Adding nodes to a cluster that is using authentication for SFDB tools" on page 279. See "Updating the Storage Foundation for Databases (SFDB) repository after adding a node" on page 280.

The example procedures describe how to add a node to an existing cluster with two nodes.

Before adding a node to a cluster

Before preparing to add the node to an existing SFCFSHA cluster, perform the required preparations.

- Verify hardware and software requirements are met.
- Set up the hardware.
- Prepare the new node.

To verify hardware and software requirements are met

1. Review hardware and software requirements for SFCFSHA.
2. Verify the new system has the same identical operating system versions and patch levels as that of the existing cluster
3. Verify the existing cluster is installed with Enterprise and that SFCFSHA is running on the cluster.

4 If the cluster is upgraded from the previous version, you must check the cluster protocol version to make sure it has the same version as the node to be added. If there is a protocol mismatch, the node is unable to join the existing cluster.

Check the cluster protocal version using:

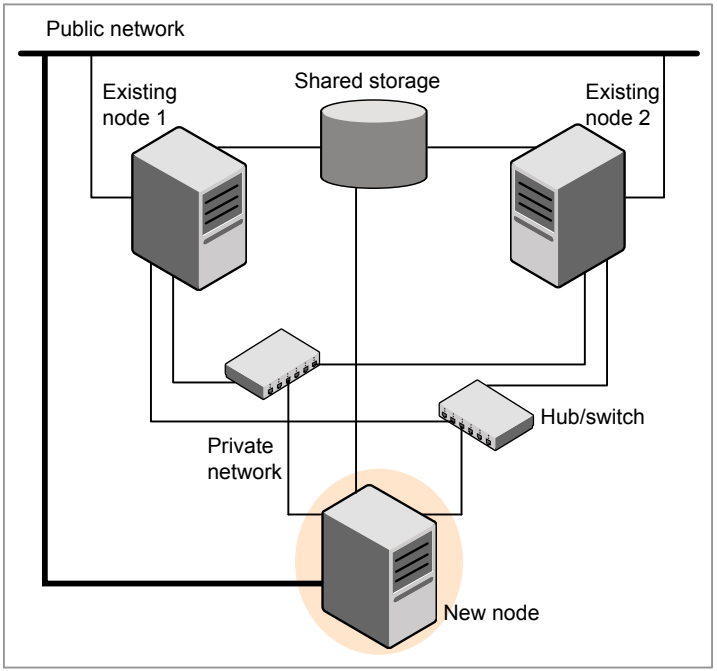
```
# vxdctl protocolversion
Cluster running at protocol 260
```

5 If the cluster protocol on the master node is below 260, upgrade it using:

```
# vxdctl upgrade newProtocolVersion
```

Before you configure a new system on an existing cluster, you must physically add the system to the cluster as illustrated in Figure 19-1.

Figure 19-1 Adding a node to a two-node cluster using two switches



To set up the hardware

1 Connect the SFCFSHA private Ethernet controllers.

Perform the following tasks as necessary:

- When you add nodes to a cluster, use independent switches or hubs for the private network connections. You can only use crossover cables for a two-node cluster, so you might have to swap out the cable for a switch or hub.
- If you already use independent hubs, connect the two Ethernet controllers on the new node to the independent hubs.

Figure 19-1 illustrates a new node being added to an existing two-node cluster using two independent hubs.

2 Make sure that you meet the following requirements:

- The node must be connected to the same shared storage devices as the existing nodes.
- The node must have private network connections to two independent switches for the cluster.
  For more information, see the *Cluster Server Configuration and Upgrade Guide*.
- The network interface names used for the private interconnects on the new node must be the same as that of the existing nodes in the cluster.

Complete the following preparatory steps on the new node before you add it to an existing SFCFSHA cluster.

To prepare the new node

1 Navigate to the folder that contains the installer program. Verify that the new node meets installation requirements.Verify that the new node meets installation requirements.

```
# ./installer -precheck
```

2 Install Veritas InfoScale Enterprise RPMs only without configuration on the new system. Make sure all the VRTS RPMs available on the existing nodes are also available on the new node.

```
# ./installer
```

Do not configure SFCFSHA when prompted.

```
Would you like to configure InfoScale Enterprise after installation?
 [y,n,q] (n) n
```

Adding a node to a cluster using the Veritas InfoScale installer

You can add a node to a cluster using the `-addnode` option with the Veritas InfoScale installer.

The Veritas InfoScale installer performs the following tasks:

- Verifies that the node and the existing cluster meet communication requirements.
- Verifies the products and RPMs installed but not configured on the new node.
- Discovers the network interfaces on the new node and checks the interface settings.
- Creates the following files on the new node:
  `/etc/llttab`
  `/etc/VRTSvcs/conf/sysname`
- Updates and copies the following files to the new node from the existing node:
  `/etc/llthosts`
  `/etc/gabtab`
  `/etc/VRTSvcs/conf/config/main.cf`
- Copies the following files from the existing cluster to the new node:
  `/etc/vxfenmode`
  `/etc/vxfendg`
  `/etc/vx/.uuids/clusuuid`
  `/etc/sysconfig/llt`
  `/etc/sysconfig/gab`
  `/etc/sysconfig/vxfen`
- Configures disk-based or server-based fencing depending on the fencing mode in use on the existing cluster.
- Adds the new node to the CVM, ClusterService service groups in the VCS configuration.

  Note: For other service groups configured under VCS, update the configuration for the new node manually.

- Starts SFCFSHA processes and configures CVM and CFS on the new node.

At the end of the process, the new node joins the SFCFSHA cluster.

Prerequisite

Enable the required LLT ports on the firewall.

See "Enabling LLT ports in firewall" on page 373.

Note: If you have configured server-based fencing on the existing cluster, make sure that the CP server does not contain entries for the new node. If the CP server already contains entries for the new node, remove these entries before adding the node to the cluster, otherwise the process may fail with an error.

See "Removing the node configuration from the CP server" on page 290.

To add the node to an existing cluster using the installer

1. Log in as the root user on one of the nodes of the existing cluster.
2. Run the Veritas InfoScale installer with the -addnode option.

   ```
   # cd /opt/VRTS/install

   # ./installer -addnode
   ```

   The installer displays the copyright message and the location where it stores the temporary installation logs.
3. Enter the name of a node in the existing SFCFSHA cluster.

   The installer uses the node information to identify the existing cluster.

   ```
   Enter the name of any one node of the InfoScale ENTERPRISE cluster
   where you would like to add one or more new nodes: sys1
   ```
4. Review and confirm the cluster information.
5. Enter the name of the systems that you want to add as new nodes to the cluster.

   ```
   Enter the system names separated by spaces
   to add to the cluster: sys5
   ```

   Confirm if the installer prompts if you want to add the node to the cluster.

   The installer checks the installed products and RPMs on the nodes and discovers the network interfaces.

6 Enter the name of the network interface that you want to configure as the first private heartbeat link.

```
Enter the NIC for the first private heartbeat
link on sys5: [b,q,?] eth1

Enter the NIC for the second private heartbeat
link on sys5: [b,q,?] eth2
```

Note: At least two private heartbeat links must be configured for high availability of the cluster.

7 Depending on the number of LLT links configured in the existing cluster, configure additional private heartbeat links for the new node.

The installer verifies the network interface settings and displays the information.

8 Review and confirm the information.

9 If you have configured SMTP, SNMP, or the global cluster option in the existing cluster, you are prompted for the NIC information for the new node.

```
Enter the NIC for VCS to use on sys5: eth3
```

10 The installer prompts you with an option to mount the shared volumes on the new node. Select **y** to mount them.

When completed, the installer confirms the volumes are mounted. The installer indicates the location of the log file, summary file, and response file with details of the actions performed.

11 If the existing cluster uses server-based fencing, the installer will configure server-based fencing on the new nodes.

The installer then starts all the required processes and joins the new node to cluster.

The installer indicates the location of the log file, summary file, and response file with details of the actions performed.

If you have enabled security on the cluster, the installer displays the following message:

```
Since the cluster is in secure mode, check the main.cf whether
you need to modify the usergroup that you would like to grant
read access. If needed, use the following commands to modify:

# haconf -makerw

# hauser -addpriv <user group> GuestGroup

# haconf -dump -makero
```

12 Confirm that the new node has joined the SFCFSHA cluster using `lltstat -n` and `gabconfig -a` commands.

Adding the node to a cluster manually

Perform this procedure after you install Veritas InfoScale Enterprise only if you plan to add the node to the cluster manually.

Table 19-2 Procedures for adding a node to a cluster manually

Step	Description
Start the Veritas Volume Manager (VxVM) on the new node.	See "Starting Veritas Volume Manager (VxVM) on the new node" on page 267.
Configure the cluster processes on the new node.	See "Configuring cluster processes on the new node" on page 268.

Table 19-2 Procedures for adding a node to a cluster manually *(continued)*

Step	Description
Configure fencing for the new node to match the fencing configuration on the existing cluster. If the existing cluster is configured to use server-based I/O fencing, configure server-based I/O fencing on the new node.	See "Starting fencing on the new node" on page 273.
Start VCS.	See "To start VCS on the new node" on page 274.
Configure CVM and CFS.	See "Configuring Cluster Volume Manager (CVM) and Cluster File System (CFS) on the new node" on page 274.
If the ClusterService group is configured on the existing cluster, add the node to the group.	See "Configuring the ClusterService group for the new node" on page 275.

Starting Veritas Volume Manager (VxVM) on the new node

Veritas Volume Manager (VxVM) uses license keys to control access. As you run the `vxinstall` utility, answer **n** to prompts about licensing. You installed the appropriate license when you ran the `installer` program.

To start VxVM on the new node

1. To start VxVM on the new node, use the vxinstall utility:

   ```
   # vxinstall
   ```

2. Enter **n** when prompted to set up a system wide disk group for the system.

   The installation completes.

3. Verify that the daemons are up and running. Enter the command:

   ```
   # vxdisk list
   ```

   Make sure the output displays the shared disks without errors.

Configuring cluster processes on the new node

Perform the steps in the following procedure to configure cluster processes on the new node.

1. For Red Hat Linux, modify the file /etc/sysctl.conf on the new system to set the shared memory and other parameter required by your application; refer to the your application documentation for details. The value of the shared memory parameter is put to effect when the system restarts.

   Do not apply for SUSE Linux.

2. Edit the /etc/llthosts file on the existing nodes. Using vi or another text editor, add the line for the new node to the file. The file resembles:

   ```
   0 sys1
   1 sys2
   2 sys5
   ```

3. Copy the `/etc/llthosts` file from one of the existing systems over to the new system. The `/etc/llthosts` file must be identical on all nodes in the cluster.

4. Create an `/etc/llttab` file on the new system. For example:

   ```
   set-node sys5
   set-cluster 101

   link eth1 eth-[MACID for eth1] - ether - -
   link eth2 eth-[MACID for eth2] - ether - -
   ```

   Except for the first line that refers to the node, the file resembles the /etc/llttab files on the existing nodes. The second line, the cluster ID, must be the same as in the existing nodes.

5. Use vi or another text editor to create the file `/etc/gabtab` on the new node. This file must contain a line that resembles the following example:

   ```
   /sbin/gabconfig -c -nN
   ```

   Where N represents the number of systems in the cluster including the new node. For a three-system cluster, N would equal 3.

6. Edit the /etc/gabtab file on each of the existing systems, changing the content to match the file on the new system.

7 Use vi or another text editor to create the file `/etc/VRTSvcs/conf/sysname` on the new node. This file must contain the name of the new node added to the cluster.

For example:

sys5

8 Create the Unique Universal Identifier file `/etc/vx/.uuids/clusuuid` on the new node:

```
# /opt/VRTSvcs/bin/uuidconfig.pl -rsh -clus -copy \
-from_sys sys1 -to_sys sys5
```

9 Start the LLT, GAB, and ODM drivers on the new node:

For systemd environments with supported Linux distributions::

```
# systemctl start llt
# systemctl start gab
# systemctl restart vxodm
```

For other supported Linux distributions:

```
# /etc/init.d/llt start

# /etc/init.d/gab start

# /etc/init.d/vxodm restart
```

10 On the new node, verify that the GAB port memberships:

```
# gabconfig -a
GAB Port Memberships
===============================================================
Port a gen df204 membership 012
```

Setting up the node to run in secure mode

You must follow this procedure only if you are adding a node to a cluster that is running in secure mode. If you are adding a node to a cluster that is not running in a secure mode, proceed with configuring LLT and GAB.

Table 19-3 uses the following information for the following command examples.

Table 19-3 The command examples definitions

Name	Fully-qualified host name (FQHN)	Function
sys5	sys5.nodes.example.com	The new node that you are adding to the cluster.

Configuring the authentication broker on node sys5

To configure the authentication broker on node sys5

1 Extract the embedded authentication files and copy them to temporary directory:

```
# mkdir -p /var/VRTSvcs/vcsauth/bkup

# cd /tmp; gunzip -c /opt/VRTSvcs/bin/VxAT.tar.gz | tar xvf -
```

2 Edit the setup file manually:

```
# cat /etc/vx/.uuids/clusuuid 2>&1
```

The output is a string denoting the UUID. This UUID (without { and }) is used as the ClusterName for the setup file.

```
{UUID}

# cat /tmp/eat_setup 2>&1
```

The file content must resemble the following example:

```
AcceptorMode=IP_ONLY

BrokerExeName=vcsauthserver

ClusterName=UUID

DataDir=/var/VRTSvcs/vcsauth/data/VCSAUTHSERVER

DestDir=/opt/VRTSvcs/bin/vcsauth/vcsauthserver

FipsMode=0

IPPort=14149

RootBrokerName=vcsroot_uuid

SetToRBPlusABorNot=0

SetupPDRs=1

SourceDir=/tmp/VxAT/version
```

3 Set up the embedded authentication file:

```
# cd /tmp/VxAT/version/bin/edition_number; \
./broker_setup.sh/tmp/eat_setup

# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/vssregctl -s -f
/var/VRTSvcs/vcsauth/data/VCSAUTHSERVER/root/.VRTSat/profile \
/VRTSatlocal.conf -b 'Security\Authentication \
\Authentication Broker' -k UpdatedDebugLogFileName \
-v /var/VRTSvcs/log/vcsauthserver.log -t string
```

4 Copy the broker credentials from one node in the cluster to sys5 by copying the entire `bkup` directory.

The `bkup` directory content resembles the following example:

```
# cd /var/VRTSvcs/vcsauth/bkup/

# ls

CMDSERVER  HAD  VCS_SERVICES  WAC
```

5 Import the VCS_SERVICES domain.

```
# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/atutil import -z \
/var/VRTSvcs/vcsauth/data/VCSAUTHSERVER -f /var/VRTSvcs/vcsauth/bkup \
/VCS_SERVICES -p password
```

6 Import the credentials for HAD, CMDSERVER, and WAC.

```
# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/atutil import -z \
/var/VRTSvcs/vcsauth/data/VCS_SERVICES -f /var/VRTSvcs/vcsauth/bkup \
/HAD -p password
```

7 Start the vcsauthserver process on sys5.

```
# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/vcsauthserver.sh
```

8 Perform the following tasks:

```
# mkdir /var/VRTSvcs/vcsauth/data/CLIENT

# mkdir /var/VRTSvcs/vcsauth/data/TRUST

# export EAT_DATA_DIR='/var/VRTSvcs/vcsauth/data/TRUST'

# /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/vssat setuptrust -b \
localhost:14149 -s high
```

9 Create the `/etc/VRTSvcs/conf/config/.secure` file:

```
# touch /etc/VRTSvcs/conf/config/.secure
```

Starting fencing on the new node

Perform the following steps to start fencing on the new node.

To start fencing on the new node

1 For disk-based fencing on at least one node, copy the following files from one of the nodes in the existing cluster to the new node:

```
/etc/sysconfig/vxfen
/etc/vxfendg
/etc/vxfenmode
```

See "Configuring server-based fencing on the new node" on page 278.

2 Start fencing on the new node:

3 On the new node, verify that the GAB port memberships are a, b, and d:

```
# gabconfig -a

GAB Port Memberships
===============================================================
Port a gen    df204 membership 012
Port b gen    df20d membership 012
Port d gen    df20a membership 012
```

After adding the new node

Start VCS on the new node.

To start VCS on the new node

1. Start VCS on the new node:

```
# hastart
```

   VCS brings the CVM group online.

2. Verify that the CVM group is online:

```
# hagrp -state
```

Configuring Cluster Volume Manager (CVM) and Cluster File System (CFS) on the new node

Modify the existing cluster configuration to configure Cluster Volume Manager (CVM) and Cluster File System (CFS) for the new node.

To configure CVM and CFS on the new node

1. Make a backup copy of the main.cf file on the existing node, if not backed up in previous procedures. For example:

```
# cd /etc/VRTSvcs/conf/config
# cp main.cf main.cf.2node
```

2. On one of the nodes in the existing cluster, set the cluster configuration to read-write mode:

```
# haconf -makerw
```

3. Add the new node to the VCS configuration, if not already added:

```
# hasys -add sys5
```

4. To enable the existing cluster to recognize the new node, run the following commands on one of the existing nodes:

```
# hagrp -modify cvm SystemList -add sys5 2
# hagrp -modify cvm AutoStartList -add sys5
# hares -modify cvm_clus CVMNodeId -add sys5 2
# haconf -dump -makero
# /etc/vx/bin/vxclustadm -m vcs reinit
# /etc/vx/bin/vxclustadm nidmap
```

5 On the remaining nodes of the existing cluster, run the following commands:

```
# /etc/vx/bin/vxclustadm -m vcs reinit
# /etc/vx/bin/vxclustadm nidmap
```

6 Copy the configuration files from one of the nodes in the existing cluster to the new node:

```
# rcp /etc/VRTSvcs/conf/config/main.cf \
sys5:/etc/VRTSvcs/conf/config/main.cf
# rcp /etc/VRTSvcs/conf/config/CFSTypes.cf \
sys5:/etc/VRTSvcs/conf/config/CFSTypes.cf
# rcp /etc/VRTSvcs/conf/config/CVMTypes.cf \
sys5:/etc/VRTSvcs/conf/config/CVMTypes.cf
```

7 The `/etc/vx/tunefstab` file sets non-default tunables for local-mounted and cluster-mounted file systems.

If you have configured a `/etc/vx/tunefstab` file to tune cluster-mounted file systems on any of the existing cluster nodes, you may want the new node to adopt some or all of the same tunables.

To adopt some or all tunables, review the contents of the file, and copy either the file, or the portions desired, into the `/etc/vx/tunefstab` file on the new cluster node.

Configuring the ClusterService group for the new node

If the ClusterService group is configured on the existing cluster, add the node to the group by performing the steps in the following procedure on one of the nodes in the existing cluster.

To configure the ClusterService group for the new node

1 On an existing node, for example sys1, write-enable the configuration:

```
# haconf -makerw
```

2 Add the node sys5 to the existing ClusterService group.

```
# hagrp -modify ClusterService SystemList -add sys5 2

# hagrp -modify ClusterService AutoStartList -add sys5
```

3 Modify the IP address and NIC resource in the existing group for the new node.

```
# hares -modify gcoip Device eth0 -sys sys5

# hares -modify gconic Device eth0 -sys sys5
```

4 Save the configuration by running the following command from any node.

```
# haconf -dump -makero
```

Adding a node using response files

Typically, you can use the response file that the installer generates on one system to add nodes to an existing cluster.

To add nodes using response files

1 Make sure the systems where you want to add nodes meet the requirements.

2 Make sure all the tasks required for preparing to add a node to an existing SFCFSHA cluster are completed.

3 Copy the response file to one of the systems where you want to add nodes.

See "Sample response file for adding a node to a SFCFSHA cluster" on page 277.

4 Edit the values of the response file variables as necessary.

See "Response file variables to add a node to a SFCFSHA cluster" on page 277.

5 Mount the product disc and navigate to the folder that contains the installation program.

6 Start adding nodes from the system to which you copied the response file. For example:

```
# ./installer -responsefile /tmp/response_file
```

Where /tmp/*response_file* is the response file's full path name.

Depending on the fencing configuration in the existing cluster, the installer configures fencing on the new node. The installer then starts all the required processes and joins the new node to cluster. The installer indicates the location of the log file and summary file with details of the actions performed.

Response file variables to add a node to a SFCFSHA cluster

Table 19-4 lists the response file variables that you can define to add a node to an SFCFSHA cluster.

Table 19-4 Response file variables for adding a node to an SFCFSHA cluster

Variable	Description
$CFG{opt}{addnode}	Adds a node to an existing cluster. List or scalar: scalar Optional or required: required
$CFG{newnodes}	Specifies the new nodes to be added to the cluster. List or scalar: list Optional or required: required

Sample response file for adding a node to a SFCFSHA cluster

The following example shows a response file for adding a node to a SFCFSHA cluster.

```
our %CFG;

$CFG{clustersystems}=[ qw(sys1) ];
$CFG{newnodes}=[ qw(sys5) ];
$CFG{opt}{addnode}=1;
$CFG{opt}{configure}=1;
$CFG{opt}{vr}=1;

$CFG{prod}=" ENTERPRISE80";

$CFG{systems}=[ qw(sys1 sys5) ];
$CFG{vcs_allowcomms}=1;
$CFG{vcs_clusterid}=101;
$CFG{vcs_clustername}="clus1";
$CFG{vcs_lltlink1}{sys5}="eth1";
$CFG{vcs_lltlink2}{sys5}="eth2";

1;
```

Configuring server-based fencing on the new node

This section describes the procedures to configure server-based fencing on a new node.

To configure server-based fencing on the new node

1. Log in to each CP server as the root user.
2. Update each CP server configuration with the new node information:

```
# cpsadm -s cps1.example.com \
-a add_node -c clus1 -h sys5 -n2

Node 2 (sys5) successfully added
```

3. Verify that the new node is added to the CP server configuration:

```
# cpsadm -s cps1.example.com -a list_nodes
```

   The new node must be listed in the output.

4. Copy the certificates to the new node from the peer nodes.

See "Generating the client key and certificates manually on the client nodes " on page 149.

Adding the new node to the vxfen service group

Perform the steps in the following procedure to add the new node to the vxfen service group.

To add the new node to the vxfen group using the CLI

1. On one of the nodes in the existing SFCFSHA cluster, set the cluster configuration to read-write mode:

```
# haconf -makerw
```

2. Add the node sys5 to the existing vxfen group.

```
# hagrp -modify vxfen SystemList -add sys5 2
```

3. Save the configuration by running the following command from any node in the SFCFSHA cluster:

```
# haconf -dump -makero
```

Adding nodes to a cluster that is using authentication for SFDB tools

To add a node to a cluster that is using authentication for SFDB tools, perform the following steps as the root user

1. Export authentication data from a node in the cluster that has already been authorized, by using the `-o export_broker_config` option of the `sfae_auth_op` command.

   Use the `-f` option to provide a file name in which the exported data is to be stored.

   ```
   # /opt/VRTS/bin/sfae_auth_op \
   -o export_broker_config -f exported-data
   ```

2. Copy the exported file to the new node by using any available copy mechanism such as `scp` or `rcp`.

3. Import the authentication data on the new node by using the `-o import_broker_config` option of the `sfae_auth_op` command.

   Use the `-f` option to provide the name of the file copied in Step 2.

   ```
   # /opt/VRTS/bin/sfae_auth_op \
    -o import_broker_config -f exported-data
   Setting up AT
   Importing broker configuration
   Starting SFAE AT broker
   ```

4. Stop the `vxdbd` daemon on the new node.

   ```
   # /opt/VRTS/bin/sfae_config disable
   vxdbd has been disabled and the daemon has been stopped.
   ```

5 Enable authentication by setting the `AUTHENTICATION` key to `yes` in the `/etc/vx/vxdbed/admin.properties` configuration file.

If `/etc/vx/vxdbed/admin.properties` does not exist, then use `cp /opt/VRTSdbed/bin/admin.properties.example /etc/vx/vxdbed/admin.properties`

6 Start the `vxdbd` daemon.

```
# /opt/VRTS/bin/sfae_config enable
vxdbd has been enabled and the daemon has been started.
It will start automatically on reboot.
```

The new node is now authenticated to interact with the cluster to run SFDB commands.

Updating the Storage Foundation for Databases (SFDB) repository after adding a node

If you are using Database Storage Checkpoints, Database FlashSnap, or SmartTier for Oracle in your configuration, update the SFDB repository to enable access for the new node after it is added to the cluster.

To update the SFDB repository after adding a node

1 Copy the `/var/vx/vxdba/rep_loc` file from one of the nodes in the cluster to the new node.

2 If the `/var/vx/vxdba/auth/user-authorizations` file exists on the existing cluster nodes, copy it to the new node.

If the `/var/vx/vxdba/auth/user-authorizations` file does not exist on any of the existing cluster nodes, no action is required.

This completes the addition of the new node to the SFDB repository.

For information on using SFDB tools features:

See *Veritas InfoScale™ Storage and Availability Management for Oracle Databases*

See *Veritas InfoScale™ Storage and Availability Management for DB2 Databases*

Sample configuration file for adding a node to the cluster

You may use this sample file as reference information to understand the configuration changes that take place when you add a node to a cluster.

```
include "OracleASMTypes.cf"
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"
include "Db2udbTypes.cf"
include "OracleTypes.cf"
include "SybaseTypes.cf"

cluster sys1_230 (
        ClusterAddress = "10.198.89.19"
        SecureClus = 1
        DefaultGuestAccess = 1
        UseFence = SCSI3
        HacliUserLevel = COMMANDROOT
        )

system sys1 (
        )

system sys2 (
        )

group ClusterService (
        SystemList = { sys1 = 0, sys2 = 1 }
        AutoStartList = { sys1, sys2 }
        OnlineRetryLimit = 3
        OnlineRetryInterval = 120
        )

        Application wac (
                StartProgram = "/opt/VRTSvcs/bin/wacstart -secure"
                StopProgram = "/opt/VRTSvcs/bin/wacstop"
                MonitorProcesses = { "/opt/VRTSvcs/bin/wac -secure" }
                RestartLimit = 3
                )

        IP gcoip (
```

```
                Device = en0
                Address = "10.198.89.19"
                NetMask = "255.255.248.0"
                )

        NIC gconic (
                Device = en0
                NetworkHosts = { "10.198.88.1" }
                )

        gcoip requires gconic
        wac requires gcoip


        // resource dependency tree
        //
        //      group ClusterService
        //      {
        //      Application wac
        //          {
        //          IP gcoip
        //              {
        //              NIC gconic
        //              }
        //          }
        //      }


group cpi_share_dg_sys1_cpi_cvm_vol_29870_sg (
        SystemList = { sys1 = 0, sys2 = 1 }
        AutoFailOver = 0
        Parallel = 1
        AutoStartList = { sys1, sys2 }
        )

        CFSMount cfsmount1 (
                Critical = 0
                MountPoint = "/cpi_auto/cpi_mnt_29870"
                BlockDevice = "/dev/vx/dsk/cpi_share_dg_sys1/cpi_cvm_vol_
    29870"
                MountOpt @sys1 = rw
                MountOpt @sys2 = rw
                NodeList = { sys1, sys2 }
```

```
        )

    CVMVolDg cvmvoldg1 (
        Critical = 0
        CVMDiskGroup = cpi_share_dg_sys1
        CVMVolume = { cpi_cvm_vol_29870 }
        CVMActivation @sys1 = sw
        CVMActivation @sys2 = sw
        )

    requires group cvm online local firm
    cfsmount1 requires cvmvoldg1


    // resource dependency tree
    //
    //      group cpi_share_dg_sys1_cpi_cvm_vol_29870_sg
    //      {
    //      CFSMount cfsmount1
    //          {
    //          CVMVolDg cvmvoldg1
    //          }
    //      }


group cvm (
    SystemList = { sys1 = 0, sys2 = 1 }
    AutoFailOver = 0
    Parallel = 1
    AutoStartList = { sys1, sys2 }
    )

    CFSfsckd vxfsckd (
        ActivationMode @sys1 = { cpi_share_dg_sys1 = sw }
        ActivationMode @sys2 = { cpi_share_dg_sys1 = sw }
        )

    CVMCluster cvm_clus (
        CVMClustName = sys1_230
        CVMNodeId = { sys1 = 0, sys2 = 1 }
        CVMTransport = gab
        CVMTimeout = 200
        )
```

```
        CVMVxconfigd cvm_vxconfigd (
                Critical = 0
                CVMVxconfigdArgs = { syslog }
                )

        cvm_clus requires cvm_vxconfigd
        vxfsckd requires cvm_clus


        // resource dependency tree
        //
        //      group cvm
        //      {
        //      CFSfsckd vxfsckd
        //          {
        //          CVMCluster cvm_clus
        //              {
        //              CVMVxconfigd cvm_vxconfigd
        //              }
        //          }
        //      }

group vxfen (
        SystemList = { sys1 = 0, sys2 = 1 }
        AutoFailOver = 0
        Parallel = 1
        )

        CoordPoint coordpoint
```

Chapter 20

Removing a node from SFCFSHA clusters

This chapter includes the following topics:

- About removing a node from a cluster
- Removing a node from a cluster
- Modifying the VCS configuration files on existing nodes
- Modifying the Cluster Volume Manager (CVM) configuration on the existing nodes to remove references to the deleted node
- Removing the node configuration from the CP server
- Removing security credentials from the leaving node
- Updating the Storage Foundation for Databases (SFDB) repository after removing a node
- Sample configuration file for removing a node from the cluster

About removing a node from a cluster

You can remove one or more nodes from an SFCFSHA cluster. The following table provides a summary of the tasks required to remove a node to an existing SFCFSHA cluster.

Table 20-1 Tasks for removing a node from a cluster

Step	Description
Prepare to remove the node: ■ Back up the configuration file. ■ Check the status of the nodes and the service groups. ■ Take the service groups offline and removing the database instances.	See "Removing a node from a cluster" on page 286.
Remove the node from the cluster.	See "Removing a node from a cluster" on page 286.
Modify the cluster configuration on remaining nodes. ■ Edit the /etc/llthosts file. ■ Edit the /etc/gabtab file. ■ Modify the VCS configuration to remove the node. ■ Modify the CVM configuration to remove the node.	See "Modifying the VCS configuration files on existing nodes" on page 287. See "Modifying the Cluster Volume Manager (CVM) configuration on the existing nodes to remove references to the deleted node" on page 290.
If the existing cluster is configured to use server-based I/O fencing, remove the node configuration from the Coordination Point (CP) server.	See "Removing the node configuration from the CP server" on page 290.
For a cluster that is running in a secure mode, remove the security credentials from the leaving node.	See "Removing security credentials from the leaving node " on page 291.
Updating the Storage Foundation for Databases (SFDB) repository after removing a node	See "Updating the Storage Foundation for Databases (SFDB) repository after removing a node" on page 291.

The Veritas product installer does not support removing a node. You must remove a node manually. The example procedures describe how to remove a node from a cluster with three nodes.

Removing a node from a cluster

Perform the following steps to remove a node from a cluster. The procedure can be done from any node remaining in the cluster or from a remote host.

To prepare to remove a node from a cluster

1 Take your application service groups offline if they are under Cluster Server (VCS) control on the node you want to remove.

```
# hagrp -offline app_group -sys sys5
```

2 Stop the applications that use Veritas File System (VxFS) or Cluster File System (CFS) mount points and are not configured under VCS. Use native application commands to stop the applications.

To remove a node from a cluster

1 Unmount the VxFS/CFS file systems that are not configured under VCS.

```
# umount mount_point
```

2 Stop VCS on the node:

```
# hastop -local
```

3 Stop SFCFSHA on the node using the Veritas InfoScale Enterprise installer.

```
# cd /opt/VRTS/install

# ./installer -stop sys5
```

The installer stops all SFCFSHA processes.

4 Modify the VCS configuration files on the existing nodes to remove references to the deleted node.

See "Modifying the VCS configuration files on existing nodes" on page 287.

5 Modify the Cluster Volume Manager (CVM) configuration on the existing nodes to remove references to the deleted node.

See "Modifying the Cluster Volume Manager (CVM) configuration on the existing nodes to remove references to the deleted node" on page 290.

Modifying the VCS configuration files on existing nodes

Modify the configuration files on the remaining nodes of the cluster to remove references to the deleted nodes.

Tasks for modifying the cluster configuration files:

- Edit the /etc/llthosts file

- Edit the /etc/gabtab file
- Modify the VCS configuration to remove the node

For an example main.cf:

See "Sample configuration file for removing a node from the cluster" on page 292.

To edit the /etc/llthosts file

- On each of the existing nodes, edit the `/etc/llthosts` file to remove lines that contain references to the removed nodes.

  For example, if sys5 is the node removed from the cluster, remove the line "2 sys5" from the file:

  ```
  0 sys1
  1 sys2
  2 sys5
  ```

  Change to:

  ```
  0 sys1
  1 sys2
  ```

To edit the /etc/gabtab file

- Modify the following command in the `/etc/gabtab` file to reflect the number of systems after the node is removed:

  ```
  /sbin/gabconfig -c -nN
  ```

  where N is the number of remaining nodes in the cluster.

  For example, with two nodes remaining, the file resembles:

  ```
  /sbin/gabconfig -c -n2
  ```

Modify the VCS configuration file main.cf to remove all references to the deleted node.

Use one of the following methods to modify the configuration:

- Edit the `/etc/VRTSvcs/conf/config/main.cf` file
  This method requires application down time.
- Use the command line interface
  This method allows the applications to remain online on all remaining nodes.

The following procedure uses the command line interface and modifies the sample VCS configuration to remove references to the deleted node. Run the steps in the procedure from one of the existing nodes in the cluster. The procedure allows you

to change the VCS configuration while applications remain online on the remaining nodes.

To modify the cluster configuration using the command line interface (CLI)

1. Back up the `/etc/VRTSvcs/conf/config/main.cf` file.

   ```
   # cd /etc/VRTSvcs/conf/config

   # cp main.cf main.cf.3node.bak
   ```

2. Change the cluster configuration to read-write mode:

   ```
   # haconf -makerw
   ```

3. Remove the node from the AutoStartList attribute of the service group by specifying the remaining nodes in the desired order:

   ```
   # hagrp -modify cvm AutoStartList sys1 sys2
   ```

4. Remove the node from the SystemList attribute of the service group:

   ```
   # hagrp -modify cvm SystemList -delete sys5
   ```

   If the system is part of the SystemList of a parent group, it must be deleted from the parent group first.

5. Remove the node from the CVMNodeId attribute of the service group:

   ```
   # hares -modify cvm_clus CVMNodeId -delete sys5
   ```

6. If you have the other service groups (such as the database service group or the ClusterService group) that have the removed node in their configuration, perform step 4 and step 5 for each of them.

7. Remove the deleted node from the NodeList attribute of all CFS mount resources:

   ```
   # hares -modify CFSMount NodeList -delete sys5
   ```

8. Remove the deleted node from the system list of any other service groups that exist on the cluster. For example, to delete the node sys5:

   ```
   # hagrp -modify appgrp SystemList -delete sys5
   ```

9. Remove the deleted node from the cluster system list:

   ```
   # hasys -delete sys5
   ```

10 Save the new configuration to disk:

```
# haconf -dump -makero
```

11 Verify that the node is removed from the VCS configuration.

```
# grep -i sys5 /etc/VRTSvcs/conf/config/main.cf
```

If the node is not removed, use the VCS commands as described in this procedure to remove the node.

Modifying the Cluster Volume Manager (CVM) configuration on the existing nodes to remove references to the deleted node

To modify the CVM configuration on the existing nodes to remove references to the deleted node

- On the remaining nodes of the existing cluster, run the following commands:

```
# /etc/vx/bin/vxclustadm -m vcs reinit
# /etc/vx/bin/vxclustadm nidmap
```

Removing the node configuration from the CP server

After removing a node from a SFCFSHA cluster, perform the steps in the following procedure to remove that node's configuration from the CP server.

Note: The `cpsadm` command is used to perform the steps in this procedure. For detailed information about the `cpsadm` command, see the *Storage Foundation Cluster File System High Availability Administrator's Guide*.

To remove the node configuration from the CP server

1. Log into the CP server as the root user.
2. View the list of VCS users on the CP server.

   If the CP server is configured to use HTTPS-based communication, run the following command:

   ```
   # cpsadm -s cp_server -a list_users
   ```

   Where *cp_server* is the virtual IP/ virtual hostname of the CP server.
3. Remove the node entry from the CP server:

   ```
   # cpsadm -s cp_server -a rm_node  -h sys5 -c clus1 -n 2
   ```

4. View the list of nodes on the CP server to ensure that the node entry was removed:

   ```
   # cpsadm -s cp_server -a list_nodes
   ```

Removing security credentials from the leaving node

If the leaving node is part of a cluster that is running in a secure mode, you must remove the security credentials from node sys5. Perform the following steps.

To remove the security credentials

1. Stop the AT process.

   ```
   # /opt/VRTSvcs/bin/vcsauth/vcsauthserver/bin/vcsauthserver.sh \
   stop
   ```

2. Remove the credentials.

   ```
   # rm -rf /var/VRTSvcs/vcsauth/data/
   ```

Updating the Storage Foundation for Databases (SFDB) repository after removing a node

After removing a node from a cluster, you do not need to perform any steps to update the SFDB repository.

For information on removing the SFDB repository after removing the product:

Sample configuration file for removing a node from the cluster

You may use this sample file as reference information to understand the configuration changes involved when you remove a node from a cluster.

The existing sample configuration before removing the node *sys3* is as follows:

- The existing cluster*clus1*comprises three nodes *sys1*, *sys2*, and *sys3* and hosts a single database.
- The database is stored on CFS.
- The database is managed by a VCS database agent.
  The agent starts, stops, and monitors the database.

Note: The following sample file shows in **bold** the configuration information that is removed when the node *sys3* is removed from the cluster.

```
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"

cluster clus1 (
        UserNames = { admin = bopHo }
        Administrators = { admin }
        UseFence = SCSI3
        )

system sys1 (
        )
system sys2 (
        )
system sys3 (
       )
```

Note: In the following group *app_grp*, the *sys3* node must be removed.

```
group app_grp (
        SystemList = { sys1 = 0, sys2 = 1, sys3 = 2 }
        AutoFailOver = 0
        Parallel = 1
        AutoStartList = { sys1, sys2, sys3 }
        )
```

Note: In the following application resource, the *sys3* node information must be removed.

```
App app1 (
          Critical = 0
          Sid @sys1 = vrts1
          Sid @sys2 = vrts2
          Sid @sys3 = vrts3
          )

CFSMount appdata_mnt (
         Critical = 0
         MountPoint = "/oradata"
         BlockDevice = "/dev/vx/dsk/appdatadg/appdatavol"
         )

CVMVolDg appdata_voldg (
         Critical = 0
         CVMDiskGroup = appdatadg
         CVMVolume = { appdatavol }
         CVMActivation = sw
         )

requires group cvm online local firm
app1 requires appdata_mnt
appdata_mnt requires appdata_voldg
```

Note: In the following CVM and CVMCluster resources, the *sys3* node information must be removed.

```
group cvm (
        SystemList = { sys1 = 0, sys2 = 1, sys3 =2}
        AutoFailOver = 0
        Parallel = 1
        AutoStartList = { sys1, sys2, sys3 }
        )

        CFSfsckd vxfsckd (
                )
```

```
 CVMCluster cvm_clus (
       CVMClustName = clus1
       CVMNodeId = { sys1 = 0, sys2 = 1, sys3 =2 }
       CVMTransport = gab
       CVMTimeout = 200
       )


      CVMVxconfigd cvm_vxconfigd (
              Critical = 0
              CVMVxconfigdArgs = { syslog }
              )


vxfsckd requires cvm_clus
cvm_clus requires cvm_vxconfigd
```

Section 7

Configuration and Upgrade reference

- Appendix A. Installation scripts
- Appendix B. Configuration files
- Appendix C. Configuring the secure shell or the remote shell for communications
- Appendix D. High availability agent information
- Appendix E. Sample SFCFSHA cluster setup diagrams for CP server-based I/O fencing
- Appendix F. Configuring LLT over UDP
- Appendix G. Using LLT over RDMA

Appendix A

Installation scripts

This appendix includes the following topics:

- Installation script options
- About using the postcheck option

Installation script options

Table A-1 shows command line options for the installation script. For an initial install or upgrade, options are not usually required. The installation script options apply to all Veritas InfoScale product scripts, except where otherwise noted.

Table A-1 Available command line options

Command Line Option	Function
-addnode	Adds a node to a high availability cluster.
-allpkgs	Displays all RPMs required for the specified product. The RPMs are listed in correct installation order. The output can be used to create scripts for command line installs, or for installations over a network.
-comcleanup	The `-comcleanup` option removes the secure shell or remote shell configuration added by installer on the systems. The option is only required when installation routines that performed auto-configuration of the shell are abruptly terminated.
-comsetup	The `-comsetup` option is used to set up the ssh or rsh communication between systems without requests for passwords or passphrases.

Table A-1 Available command line options *(continued)*

Command Line Option	Function
-configcps	The `-configcps` option is used to configure CP server on a running system or cluster.
-configure	Configures the product after installation.
-disable_dmp_native_support	Disables Dynamic Multi-pathing support for the native LVM volume groups and ZFS pools during upgrade. Retaining Dynamic Multi-pathing support for the native LVM volume groups and ZFS pools during upgrade increases RPM upgrade time depending on the number of LUNs and native LVM volume groups and ZFS pools configured on the system.
-fencing	Configures I/O fencing in a running cluster.
-fips	The `-fips` option is used to enable or disable security with fips mode on a running VCS cluster. It could only be used together with `-security` or `-securityonenode` option.
–hostfile *full_path_to_file*	Specifies the location of a file that contains a list of hostnames on which to install.
-install	Used to install products on system
-online_upgrade	Used to perform online upgrade. Using this option, the installer upgrades the whole cluster and also supports customer's application zero down time during the upgrade procedure. Now this option only supports VCS and ApplicationHA.
-patch_path	Defines the path of a patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed .
-patch2_path	Defines the path of a second patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed.

Table A-1 Available command line options *(continued)*

Command Line Option	Function
-patch3_path	Defines the path of a third patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed.
-patch4_path	Defines the path of a fourth patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed.
-patch5_path	Defines the path of a fifth patch level release to be integrated with a base or a maintenance level release in order for multiple releases to be simultaneously installed.
–keyfile *ssh_key_file*	Specifies a key file for secure shell (SSH) installs. This option passes `-I ssh_key_file` to every SSH invocation.
–kickstart *dir_path*	Produces a kickstart configuration file for installing with Linux RHEL Kickstart. The file contains the list of required RPMs in the correct order for installing, in a format that can be used for Kickstart installations. The *dir_path* indicates the path to the directory in which to create the file.
-license	Registers or updates product licenses on the specified systems.
–logpath *log_path*	Specifies a directory other than `/opt/VRTS/install/logs` as the location where installer log files, summary files, and response files are saved.
-noipc	Disables the installer from making outbound networking calls to Veritas Services and Operations Readiness Tool (SORT) in order to automatically obtain patch and release information updates.
-nolic	Allows installation of product RPMs without entering a license key. Licensed features cannot be configured, started, or used when this option is specified.

Table A-1 Available command line options *(continued)*

Command Line Option	Function
-pkgtable	Displays product's RPMs in correct installation order by group.
–postcheck	Checks for different HA and file system-related processes, the availability of different ports, and the availability of cluster-related service groups.
-precheck	Performs a preinstallation check to determine if systems meet all installation requirements. Veritas recommends doing a precheck before installing a product.
-prod	Specifies the product for operations.
-component	Specifies the component for operations.
-redirect	Displays progress details without showing the progress bar.
-require	Specifies an installer patch file.
–responsefile *response_file*	Automates installation and configuration by using system and configuration information stored in a specified file instead of prompting for information. The *response_file* must be a full path name. You must edit the response file to use it for subsequent installations. Variable field definitions are defined within the file.
-rolling_upgrade	Starts a rolling upgrade. Using this option, the installer detects the rolling upgrade status on cluster systems automatically without the need to specify rolling upgrade phase 1 or phase 2 explicitly.
-rollingupgrade_phase1	The `-rollingupgrade_phase1` option is used to perform rolling upgrade Phase-I. In the phase, the product kernel RPMs get upgraded to the latest version.
-rollingupgrade_phase2	The `-rollingupgrade_phase2` option is used to perform rolling upgrade Phase-II. In the phase, VCS and other agent RPMs upgrade to the latest version. Product kernel drivers are rolling-upgraded to the latest protocol version.

Table A-1 Available command line options *(continued)*

Command Line Option	Function
-rsh	Specify this option when you want to use RSH and RCP for communication between systems instead of the default SSH and SCP. See "About configuring secure shell or remote shell communication modes before installing products" on page 318.
-security	The -security option is used to convert a running VCS cluster between secure and non-secure modes of operation.
-securityonenode	The `-securityonenode` option is used to configure a secure cluster node by node.
-securitytrust	The `-securitytrust` option is used to setup trust with another broker.
–serial	Specifies that the installation script performs install, uninstall, start, and stop operations on each system in a serial fashion. If this option is not specified, these operations are performed simultaneously on all systems.
-settunables	Specify this option when you want to set tunable parameters after you install and configure a product. You may need to restart processes of the product for the tunable parameter values to take effect. You must use this option together with the `-tunablesfile` option.
-start	Starts the daemons and processes for the specified product.
-stop	Stops the daemons and processes for the specified product.
-timeout	The `-timeout` option is used to specify the number of seconds that the script should wait for each command to complete before timing out. Setting the `-timeout` option overrides the default value of 1200 seconds. Setting the `-timeout` option to 0 prevents the script from timing out. The `-timeout` option does not work with the `-serial option`

Table A-1 Available command line options *(continued)*

Command Line Option	Function
–tmppath *tmp_path*	Specifies a directory other than `/opt/VRTStmp` as the working directory for the installation scripts. This destination is where initial logging is performed and where RPMs are copied on remote systems before installation.
-tunables	Lists all supported tunables and create a tunables file template.
-tunables_file *tunables_file*	Specify this option when you specify a tunables file. The tunables file should include tunable parameters.
-uninstall	This option is used to uninstall the products from systems
-upgrade	Specifies that an existing version of the product exists and you plan to upgrade it.
-version	Checks and reports the installed products and their versions. Identifies the installed and missing RPMs and patches where applicable for the product. Provides a summary that includes the count of the installed and any missing RPMs and patches where applicable. Lists the installed patches, patches, and available updates for the installed product if an Internet connection is available.
-yumgroupxml	The `-yumgroupxml` option is used to generate a yum group definition XML file. The `createrepo` command can use the file on Redhat Linux to create a yum group for automated installation of all RPMs for a product. An available location to store the XML file should be specified as a complete path. The `-yumgroupxml` option is supported on Redhat Linux and supported RHEL compatible distributions only.

About using the postcheck option

You can use the installer's post-check to determine installation-related problems and to aid in troubleshooting.

Note: This command option requires downtime for the node.

When you use the `postcheck` option, it can help you troubleshoot the following VCS-related issues:

- The heartbeat link does not exist.
- The heartbeat link cannot communicate.
- The heartbeat link is a part of a bonded or aggregated NIC.
- A duplicated cluster ID exists (if LLT is not running at the check time).
- The VRTSllt pkg version is not consistent on the nodes.
- The llt-linkinstall value is incorrect.
- The /etc/llthosts and /etc/lltab configuration is incorrect.
- the `/etc/gabtab` file is incorrect.
- The incorrect GAB linkinstall value exists.
- The VRTSgab pkg version is not consistent on the nodes.
- The `main.cf` file or the `types.cf` file is invalid.
- The `/etc/VRTSvcs/conf/sysname` file is not consistent with the hostname.
- The cluster UUID does not exist.
- The `uuidconfig.pl` file is missing.
- The VRTSvcs pkg version is not consistent on the nodes.
- The `/etc/vxfenmode` file is missing or incorrect.
- The `/etc/vxfendg file` is invalid.
- The vxfen link-install value is incorrect.
- The VRTSvxfen pkg version is not consistent.

The `postcheck` option can help you troubleshoot the following SFHA or SFCFSHA issues:

- Volume Manager cannot start because the `/etc/vx/reconfig.d/state.d/install-db` file has not been removed.
- Volume Manager cannot start because the `volboot` file is not loaded.
- Volume Manager cannot start because no license exists.
- Cluster Volume Manager cannot start because the CVM configuration is incorrect in the `main.cf` file. For example, the Autostartlist value is missing on the nodes.

- Cluster Volume Manager cannot come online because the node ID in the `/etc/llthosts` file is not consistent.
- Cluster Volume Manager cannot come online because Vxfen is not started.
- Cluster Volume Manager cannot start because gab is not configured.
- Cluster Volume Manager cannot come online because of a CVM protocol mismatch.
- Cluster Volume Manager group name has changed from "cvm", which causes CVM to go offline.

You can use the installer's post-check option to perform the following checks:

General checks for all products:

- All the required RPMs are installed.
- The versions of the required RPMs are correct.
- There are no verification issues for the required RPMs.

Checks for Volume Manager (VM):

- Lists the daemons which are not running (`vxattachd`, `vxconfigbackupd`, `vxesd`, `vxrelocd` ...).
- Lists the disks which are not in 'online' or 'online shared' state (`vxdisk list`).
- Lists the diskgroups which are not in 'enabled' state (`vxdg list`).
- Lists the volumes which are not in 'enabled' state (`vxprint -g <dgname>`).
- Lists the volumes which are in 'Unstartable' state (`vxinfo -g <dgname>`).
- Lists the volumes which are not configured in `/etc/fstab` .

Checks for File System (FS):

- Lists the VxFS kernel modules which are not loaded (`vxfs/fdd/vxportal`.).
- Whether all VxFS file systems present in `/etc/fstab` file are mounted.
- Whether all VxFS file systems present in `/etc/fstab` are in disk layout 12 or higher.
- Whether all mounted VxFS file systems are in disk layout 12 or higher.

Checks for Cluster File System:

- Whether FS and ODM are running at the latest protocol level.
- Whether all mounted CFS file systems are managed by VCS.
- Whether cvm service group is online.

Appendix B

Configuration files

This appendix includes the following topics:

- About the LLT and GAB configuration files
- About the AMF configuration files
- About I/O fencing configuration files
- Sample configuration files for CP server

About the LLT and GAB configuration files

Low Latency Transport (LLT) and Group Membership and Atomic Broadcast (GAB) are VCS communication services. LLT requires /etc/llthosts and /etc/lltab files. GAB requires /etc/gabtab file.

Table B-1 lists the LLT configuration files and the information that these files contain.

Table B-1 LLT configuration files

File	Description
/etc/sysconfig/llt	This file stores the start and stop environment variables for LLT: ■ LLT_START—Defines the startup behavior for the LLT module after a system reboot. Valid values include: 1—Indicates that LLT is enabled to start up. 0—Indicates that LLT is disabled to start up. ■ LLT_STOP—Defines the shutdown behavior for the LLT module during a system shutdown. Valid values include: 1—Indicates that LLT is enabled to shut down. 0—Indicates that LLT is disabled to shut down. The installer sets the value of these variables to 1 at the end of SFCFSHA configuration. If you manually configured VCS, make sure you set the values of these environment variables to 1. Assign the buffer pool memory for RDMA operations: ■ LLT_BUFPOOL_MAXMEM—Maximum assigned memory that LLT can use for the LLT buffer pool. This buffer pool is used to allocate memory for RDMA operations and packet allocation, which are delivered to the LLT clients. The default value is calculated based on the total system memory, the minimum value is 1GB, and the maximum value is 10GB. You must specify the value in GB. Define the number of RDMA queue pairs per LLT link ■ LLT_RDMA_QPS — Maximum number of RDMA queue pairs per LLT link. You can change the value of this parameter in the `/etc/sysconfig/llt` file. The default value is 4. However, you can extend this value to maximum 8 queue pairs per LLT link. Enable or disable the adaptive window feature: ■ For performance reason, the adaptive window feature (LLT_ENABLE_AWINDOW) is enabled by default for 5 (cfs) and port 24(cvm). You can disable the adaptive window feature by manually changing the value of the LLT_ENABLE_AWINDOW parameter to zero. To enable adaptive window for ports other than 5 and 24, add the port numbers in LLT_AW_PORT_LIST separated by comma. For example: LLT_AW_PORT_LIST=: '"5,24,0,1,5,14"'. If you want to disable the adaptive window feature for any of the ports, remove that specific port from this parameter. Example: LLT_AW_PORT_LIST='"5"' Configure LLT over TCP ■ When you configure LLT over the Transmission Control Protocol (TCP) layer for clusters using wide-area networks and routers, you can use the LLT_TCP_CONNS parameter to control the number of TCP connections that can be established between peer nodes. The default value of this parameter is 8 connections. However, you can extend this value to maximum 64 connections.

Table B-1 LLT configuration files *(continued)*

File	Description

/etc/llthosts	The file `llthosts` is a database that contains one entry per system. This file links the LLT system ID (in the first column) with the LLT host name. This file must be identical on each node in the cluster. A mismatch of the contents of the file can cause indeterminate behavior in the cluster.
	
	For example, the file /etc/llthosts contains the entries that resemble:
	
	`0 sys1`
	`1 sys2`

/etc/lltab	The file `lltab` contains the information that is derived during installation and used by the utility `lltconfig(1M)`. After installation, this file lists the LLT network links that correspond to the specific system.
	
	For example, the file /etc/lltab contains the entries that resemble:
	
	`set-node sys1`
	`set-cluster 2`
	`link eth1 eth1 - ether - -`
	`link eth2 eth2 - ether - -`
	
	If you use aggregated interfaces, then the file contains the aggregated interface name instead of the eth-*MAC_address*.
	
	`set-node sys1`
	`set-cluster 2`
	`link eth1 eth-00:04:23:AC:12:C4 - ether - -`
	`link eth2 eth-00:04:23:AC:12:C5 - ether - -`
	
	The first line identifies the system. The second line identifies the cluster (that is, the cluster ID you entered during installation). The next two lines begin with the `link` command. These lines identify the two network cards that the LLT protocol uses.
	
	If you configured a low priority link under LLT, the file also includes a "link-lowpri" line.
	
	Refer to the `lltab(4)` manual page for details about how the LLT configuration may be modified. The manual page describes the ordering of the directives in the `lltab` file.

Table B-2 lists the GAB configuration files and the information that these files contain.

Table B-2 GAB configuration files

File	Description
/etc/sysconfig/gab	This file stores the start and stop environment variables for GAB: ■ GAB_START—Defines the startup behavior for the GAB module after a system reboot. Valid values include: 1—Indicates that GAB is enabled to start up. 0—Indicates that GAB is disabled to start up. ■ GAB_STOP—Defines the shutdown behavior for the GAB module during a system shutdown. Valid values include: 1—Indicates that GAB is enabled to shut down. 0—Indicates that GAB is disabled to shut down. The installer sets the value of these variables to 1 at the end of Storage Foundation Cluster File System High Availability configuration.
/etc/gabtab	After you install SFCFSHA, the file /etc/gabtab contains a `gabconfig(1)` command that configures the GAB driver for use. The file /etc/gabtab contains a line that resembles: `/sbin/gabconfig -c -nN` The `-c` option configures the driver for use. The `-nN` specifies that the cluster is not formed until at least `N` nodes are ready to form the cluster. Veritas recommends that you set N to be the total number of nodes in the cluster. **Note:** Veritas does not recommend the use of the `-c -x` option for `/sbin/gabconfig`. Using `-c -x` can lead to a split-brain condition. Use the `-c` option for `/sbin/gabconfig` to avoid a split-brain condition.

About the AMF configuration files

Asynchronous Monitoring Framework (AMF) kernel driver provides asynchronous event notifications to the VCS agents that are enabled for intelligent resource monitoring.

Table B-3 lists the AMF configuration files.

Table B-3 AMF configuration files

File	Description
/etc/sysconfig/amf	This file stores the start and stop environment variables for AMF: ■ AMF_START—Defines the startup behavior for the AMF module after a system reboot or when AMF is attempted to start using the init script. Valid values include: 1—Indicates that AMF is enabled to start up. (default) 0—Indicates that AMF is disabled to start up. ■ AMF_STOP—Defines the shutdown behavior for the AMF module during a system shutdown or when AMF is attempted to stop using the init script. Valid values include: 1—Indicates that AMF is enabled to shut down. (default) 0—Indicates that AMF is disabled to shut down.
/etc/amftab	After you install VCS, the file `/etc/amftab` contains a `amfconfig`(1) command that configures the AMF driver for use. The AMF init script uses this `/etc/amftab` file to configure the AMF driver. The `/etc/amftab` file contains the following line by default: `/opt/VRTSamf/bin/amfconfig -c`

About I/O fencing configuration files

Table B-4 lists the I/O fencing configuration files.

Table B-4 I/O fencing configuration files

File	Description
/etc/sysconfig/vxfen	This file stores the start and stop environment variables for I/O fencing: ■ VXFEN_START—Defines the startup behavior for the I/O fencing module after a system reboot. Valid values include: 1—Indicates that I/O fencing is enabled to start up. 0—Indicates that I/O fencing is disabled to start up. ■ VXFEN_STOP—Defines the shutdown behavior for the I/O fencing module during a system shutdown. Valid values include: 1—Indicates that I/O fencing is enabled to shut down. 0—Indicates that I/O fencing is disabled to shut down. The installer sets the value of these variables to 1 at the end of Storage Foundation Cluster File System High Availability configuration.

Table B-4 I/O fencing configuration files *(continued)*

File	Description
/etc/vxfendg	This file includes the coordinator disk group information. This file is not applicable for server-based fencing and majority-based fencing.

Table B-4 I/O fencing configuration files *(continued)*

File	Description
/etc/vxfenmode	This file contains the following parameters: ■ vxfen_mode ■ scsi3—For disk-based fencing. ■ customized—For server-based fencing. ■ disabled—To run the I/O fencing driver but not do any fencing operations. ■ majority— For fencing without the use of coordination points. ■ vxfen_mechanism This parameter is applicable only for server-based fencing. Set the value as cps. ■ scsi3_disk_policy ■ dmp—Configure the vxfen module to use DMP devices The disk policy is dmp by default. If you use iSCSI devices, you must set the disk policy as dmp. **Note:** You must use the same SCSI-3 disk policy on all the nodes. ■ List of coordination points This list is required only for server-based fencing configuration. Coordination points in server-based fencing can include coordinator disks, CP servers, or both. If you use coordinator disks, you must create a coordinator disk group containing the individual coordinator disks. Refer to the sample file /etc/vxfen.d/vxfenmode_cps for more information on how to specify the coordination points and multiple IP addresses for each CP server. ■ single_cp This parameter is applicable for server-based fencing which uses a single highly available CP server as its coordination point. Also applicable for when you use a coordinator disk group with single disk. ■ autoseed_gab_timeout This parameter enables GAB automatic seeding of the cluster even when some cluster nodes are unavailable. This feature is applicable for I/O fencing in SCSI3 and customized mode. 0—Turns the GAB auto-seed feature on. Any value greater than 0 indicates the number of seconds that GAB must delay before it automatically seeds the cluster. -1—Turns the GAB auto-seed feature off. This setting is the default. ■ detect_false_pesb 0—Disables stale key detection. 1—Enables stale key detection to determine whether a preexisting split brain is a true condition or a false alarm. Default: 0 **Note:** This parameter is considered only when `vxfen_mode=customized`.

Table B-4 I/O fencing configuration files *(continued)*

File	Description

/etc/vxfentab	When I/O fencing starts, the vxfen startup script creates this /etc/vxfentab file on each node. The startup script uses the contents of the /etc/vxfendg and /etc/vxfenmode files. Any time a system is rebooted, the fencing driver reinitializes the vxfentab file with the current list of all the coordinator points.
	Note: The /etc/vxfentab file is a generated file; do not modify this file.
	For disk-based I/O fencing, the /etc/vxfentab file on each node contains a list of all paths to each coordinator disk along with its unique disk identifier. A space separates the path and the unique disk identifier. An example of the /etc/vxfentab file in a disk-based fencing configuration on one node resembles as follows:
	■ DMP disk:
	`/dev/vx/rdmp/sdx3 HITACHI%5F1724-100%20%20FAStT%5FDISKS%5F6 00A0B8000215A5D000006804E795D0A3`
	`/dev/vx/rdmp/sdy3 HITACHI%5F1724-100%20%20FAStT%5FDISKS%5F6 00A0B8000215A5D000006814E795D0B3`
	`/dev/vx/rdmp/sdz3 HITACHI%5F1724-100%20%20FAStT%5FDISKS%5F6 00A0B8000215A5D000006824E795D0C3`
	For server-based fencing, the /etc/vxfentab file also includes the security settings information.
	For server-based fencing with single CP server, the /etc/vxfentab file also includes the single_cp settings information.
	This file is not applicable for majority-based fencing.

Sample configuration files for CP server

The `/etc/vxcps.conf` file determines the configuration of the coordination point server (CP server.)

See "Sample CP server configuration (/etc/vxcps.conf) file output" on page 317.

The following are example main.cf files for a CP server that is hosted on a single node, and a CP server that is hosted on an SFHA cluster.

- The main.cf file for a CP server that is hosted on a single node:
  See "Sample main.cf file for CP server hosted on a single node that runs VCS" on page 312.
- The main.cf file for a CP server that is hosted on an SFHA cluster:
  See "Sample main.cf file for CP server hosted on a two-node SFHA cluster" on page 314.

The example main.cf files use IPv4 addresses.

Sample main.cf file for CP server hosted on a single node that runs VCS

The following is an example of a single CP server node main.cf.

For this CP server single node main.cf, note the following values:

- Cluster name: cps1
- Node name: cps1

```
include "types.cf"
include "/opt/VRTScps/bin/Quorum/QuorumTypes.cf"

// cluster name:  cps1
// CP server: cps1

cluster cps1 (
     UserNames = { admin = bMNfMHmJNiNNlVNhMK, haris = fopKojNvpHouNn,
             "cps1.example.com@root@vx" = aj,
             "root@cps1.example.com" = hq }
     Administrators = { admin, haris,
             "cps1.example.com@root@vx",
             "root@cps1.example.com" }
     SecureClus = 1
     HacliUserLevel = COMMANDROOT
     )

system cps1 (
     )

group CPSSG (
      SystemList = { cps1 = 0 }
      AutoStartList = { cps1 }
      )

      IP cpsvip1 (
            Critical = 0
            Device @cps1 = eth0
            Address = "10.209.3.1"
            NetMask = "255.255.252.0"
            )
```

```
        IP cpsvip2 (
              Critical = 0
              Device @cps1 = eth1
              Address = "10.209.3.2"
              NetMask = "255.255.252.0"
              )

        NIC cpsnic1 (
             Critical = 0
             Device @cps1 = eth0
             PingOptimize = 0
             NetworkHosts @cps1 = { "10.209.3.10" }
             )

        NIC cpsnic2 (
             Critical = 0
             Device @cps1 = eth1
             PingOptimize = 0
             )

        Process vxcpserv (
             PathName = "/opt/VRTScps/bin/vxcpserv"
             ConfInterval = 30
             RestartLimit = 3
             )

        Quorum quorum (
              QuorumResources = { cpsvip1, cpsvip2 }
              )

    cpsvip1 requires cpsnic1
    cpsvip2 requires cpsnic2
    vxcpserv requires quorum


    // resource dependency tree
    //
    // group CPSSG
    // {
    // IP cpsvip1
    //     {
    //     NIC cpsnic1
    //     }
```

```
// IP cpsvip2
//     {
//     NIC cpsnic2
//     }
// Process vxcpserv
//     {
//     Quorum quorum
//     }
// }
```

Sample main.cf file for CP server hosted on a two-node SFHA cluster

The following is an example of a main.cf, where the CP server is hosted on an SFHA cluster.

For this CP server hosted on an SFHA cluster main.cf, note the following values:

- Cluster name: cps1
- Nodes in the cluster: cps1, cps2

```
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"
include "/opt/VRTScps/bin/Quorum/QuorumTypes.cf"


// cluster: cps1
// CP servers:
// cps1
// cps2

cluster cps1 (
     UserNames = { admin = ajkCjeJgkFkkIskEjh,
            "cps1.example.com@root@vx" = JK,
            "cps2.example.com@root@vx" = dl }
     Administrators = { admin, "cps1.example.com@root@vx",
            "cps2.example.com@root@vx" }
     SecureClus = 1
     )

system cps1 (
     )

system cps2 (
```

```
    )

group CPSSG (
      SystemList = { cps1 = 0, cps2 = 1 }
      AutoStartList = { cps1, cps2 } )

      DiskGroup cpsdg (
            DiskGroup = cps_dg
            )

      IP cpsvip1 (
            Critical = 0
            Device @cps1 = eth0
            Device @cps2 = eth0
            Address = "10.209.81.88"
            NetMask = "255.255.252.0"
            )

      IP cpsvip2 (
            Critical = 0
            Device @cps1 = eth1
            Device @cps2 = eth1
            Address = "10.209.81.89"
            NetMask = "255.255.252.0"
            )

      Mount cpsmount (
            MountPoint = "/etc/VRTScps/db"
            BlockDevice = "/dev/vx/dsk/cps_dg/cps_volume"
            FSType = vxfs
            FsckOpt = "-y"
            )

      NIC cpsnic1 (
           Critical = 0
           Device @cps1 = eth0
           Device @cps2 = eth0
           PingOptimize = 0
           NetworkHosts @cps1 = { "10.209.81.10 }
           )

      NIC cpsnic2 (
           Critical = 0
```

```
            Device @cps1 = eth1
            Device @cps2 = eth1
            PingOptimize = 0
            )

      Process vxcpserv (
            PathName = "/opt/VRTScps/bin/vxcpserv"
            )

      Quorum quorum (
            QuorumResources = { cpsvip1, cpsvip2 }
            )

      Volume cpsvol (
            Volume = cps_volume
            DiskGroup = cps_dg
            )

cpsmount requires cpsvol
cpsvip1 requires cpsnic1
cpsvip2 requires cpsnic2
cpsvol requires cpsdg
vxcpserv requires cpsmount
vxcpserv requires quorum


// resource dependency tree
//
// group CPSSG
// {
// IP cpsvip1
//     {
//     NIC cpsnic1
//     }
// IP cpsvip2
//     {
//     NIC cpsnic2
//     }
// Process vxcpserv
//     {
//     Quorum quorum
//     Mount cpsmount
//         {
```

```
//          Volume cpsvol
//              {
//              DiskGroup cpsdg
//              }
//          }
//      }
// }
```

Sample CP server configuration (/etc/vxcps.conf) file output

The following is an example of a coordination point server (CP server) configuration file `/etc/vxcps.conf` output.

```
##  The vxcps.conf file determines the
## configuration for Veritas CP Server.
cps_name=cps1
vip=[10.209.81.88]
vip=[10.209.81.89]:56789
vip_https=[10.209.81.88]:55443
vip_https=[10.209.81.89]
port=14250
port_https=443
security=1
db=/etc/VRTScps/db
ssl_conf_file=/etc/vxcps_ssl.properties
```

Appendix C

Configuring the secure shell or the remote shell for communications

This appendix includes the following topics:

- About configuring secure shell or remote shell communication modes before installing products
- Manually configuring passwordless ssh
- Setting up ssh and rsh connection using the installer -comsetup command
- Setting up ssh and rsh connection using the pwdutil.pl utility
- Restarting the ssh session
- Enabling rsh for Linux

About configuring secure shell or remote shell communication modes before installing products

Establishing communication between nodes is required to install Veritas InfoScale software from a remote system, or to install and configure a system. The system from which the installer is run must have permissions to run `rsh` (remote shell) or `ssh` (secure shell) utilities. You need to run the installer with superuser privileges on the systems where you plan to install the Veritas InfoScale software.

You can install products to remote systems using either secure shell (ssh) or remote shell (rsh). Veritas recommends that you use ssh as it is more secure than rsh.

You can set up ssh and rsh connections in many ways.

- You can manually set up the ssh and rsh connection with UNIX shell commands.
- You can run the `installer -comsetup` command to interactively set up ssh and rsh connection.
- You can run the password utility, `pwdutil.pl`.

This section contains an example of how to set up ssh password free communication. The example sets up ssh between a source system (sys1) that contains the installation directories, and a target system (sys2). This procedure also applies to multiple target systems.

Note: The product installer supports establishing passwordless communication.

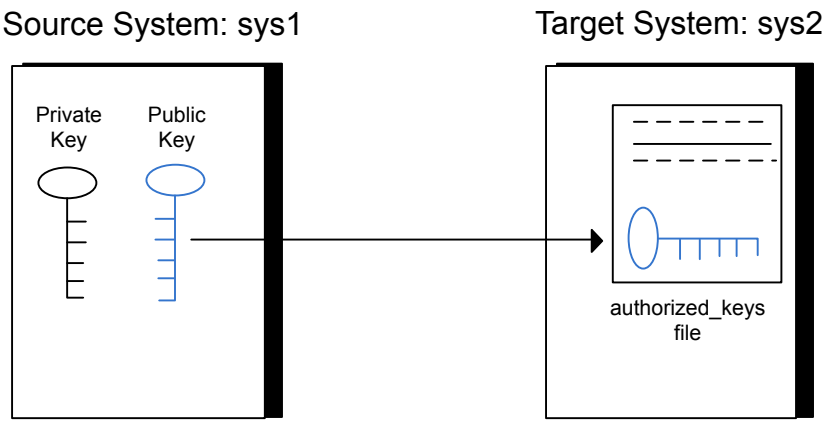
Manually configuring passwordless ssh

The ssh program enables you to log into and execute commands on a remote system. ssh enables encrypted communications and an authentication process between two untrusted hosts over an insecure network.

In this procedure, you first create a DSA key pair. From the key pair, you append the public key from the source system to the authorized_keys file on the target systems.

Figure C-1 illustrates this procedure.

Figure C-1 Creating the DSA key pair and appending it to target systems



Read the ssh documentation and online manual pages before enabling ssh. Contact your operating system support provider for issues regarding ssh configuration.

Visit the Openssh website that is located at: http://www.openssh.com/ to access online manuals and other resources.

To create the DSA key pair

1. On the source system (sys1), log in as root, and navigate to the root directory.

   ```
   sys1 # cd /root
   ```

2. To generate a DSA key pair on the source system, type the following command:

   ```
   sys1 # ssh-keygen -t dsa
   ```

   System output similar to the following is displayed:

   ```
   Generating public/private dsa key pair.
   Enter file in which to save the key (/root/.ssh/id_dsa):
   ```

3. Press Enter to accept the default location of `/root/.ssh/id_dsa`.
4. When the program asks you to enter the passphrase, press the Enter key twice.

   ```
   Enter passphrase (empty for no passphrase):
   ```

   Do not enter a passphrase. Press Enter.

   ```
   Enter same passphrase again:
   ```

   Press Enter again.
5. Output similar to the following lines appears.

   ```
   Your identification has been saved in /root/.ssh/id_dsa.
   Your public key has been saved in /root/.ssh/id_dsa.pub.
   The key fingerprint is:
   1f:00:e0:c2:9b:4e:29:b4:0b:6e:08:f8:50:de:48:d2 root@sys1
   ```

To append the public key from the source system to the authorized_keys file on the target system, using secure file transfer

1 From the source system (sys1), move the public key to a temporary file on the target system (sys2).

Use the secure file transfer program.

In this example, the file name `id_dsa.pub` in the root directory is the name for the temporary file for the public key.

Use the following command for secure file transfer:

```
sys1 # sftp sys2
```

If the secure file transfer is set up for the first time on this system, output similar to the following lines is displayed:

```
Connecting to sys2 ...
The authenticity of host 'sys2 (10.182.00.00)'
can't be established. DSA key fingerprint is
fb:6f:9f:61:91:9d:44:6b:87:86:ef:68:a6:fd:88:7d.
Are you sure you want to continue connecting (yes/no)?
```

2 Enter `yes`.

Output similar to the following is displayed:

```
Warning: Permanently added 'sys2,10.182.00.00'
(DSA) to the list of known hosts.
root@sys2 password:
```

3 Enter the root password of sys2.

4 At the `sftp` prompt, type the following command:

```
sftp> put /root/.ssh/id_dsa.pub
```

The following output is displayed:

```
Uploading /root/.ssh/id_dsa.pub to /root/id_dsa.pub
```

5 To quit the SFTP session, type the following command:

```
sftp> quit
```

6 Add the `id_dsa.pub` keys to the `authorized_keys` file on the target system. To begin the `ssh` session on the target system (sys2 in this example), type the following command on sys1:

```
sys1 # ssh sys2
```

Enter the root password of sys2 at the prompt:

```
password:
```

Type the following commands on sys2:

```
sys2 # cat /root/id_dsa.pub >> /root/.ssh/authorized_keys
sys2 # rm  /root/id_dsa.pub
```

7 Run the following commands on the source installation system. If your ssh session has expired or terminated, you can also run these commands to renew the session. These commands bring the private key into the shell environment and make the key globally available to the user `root`:

```
sys1 # exec /usr/bin/ssh-agent $SHELL
sys1 # ssh-add

  Identity added: /root/.ssh/id_dsa
```

This shell-specific step is valid only while the shell is active. You must execute the procedure again if you close the shell during the session.

To verify that you can connect to a target system

1 On the source system (sys1), enter the following command:

```
sys1 # ssh -l root sys2 uname -a
```

where sys2 is the name of the target system.

2 The command should execute from the source system (sys1) to the target system (sys2) without the system requesting a passphrase or password.

3 Repeat this procedure for each target system.

Setting up ssh and rsh connection using the installer -comsetup command

You can interactively set up the ssh and rsh connections using the `installer -comsetup` command.

Enter the following:

```
# ./installer -comsetup

Input the name of the systems to set up communication:
Enter the <platform> system names separated by spaces:
[q,?] sys2
Set up communication for the system sys2:

  Checking communication on sys2 ..................... Failed

CPI ERROR V-9-20-1303 ssh permission was denied on sys2. rsh
permission was denied on sys2. Either ssh or rsh is required
to be set up and ensure that it is working properly between the local
node and sys2 for communication

Either ssh or rsh needs to be set up between the local system and
sys2 for communication

Would you like the installer to setup ssh or rsh communication
automatically between the systems?
Superuser passwords for the systems will be asked. [y,n,q,?] (y) y

Enter the superuser password for system sys2:

     1)  Setup ssh between the systems
     2)  Setup rsh between the systems
     b)  Back to previous menu

Select the communication method [1-2,b,q,?] (1) 1

Setting up communication between systems. Please wait.
Re-verifying systems.

 Checking communication on sys2 ..................... Done

Successfully set up communication for the system sys2
```

Setting up ssh and rsh connection using the pwdutil.pl utility

The password utility, `pwdutil.pl`, is bundled under the `scripts` directory. The users can run the utility in their script to set up the ssh and rsh connection automatically.

```
# ./pwdutil.pl -h
Usage:

Command syntax with simple format:

    pwdutil.pl check|configure|unconfigure ssh|rsh <hostname|IP addr>
    [<user>] [<password>] [<port>]

Command syntax with advanced format:

    pwdutil.pl [--action|-a 'check|configure|unconfigure']
               [--type|-t 'ssh|rsh']
               [--user|-u  '<user>']
               [--password|-p '<password>']
               [--port|-P '<port>']
               [--hostfile|-f '<hostfile>']
               [--keyfile|-k '<keyfile>']
               [-debug|-d]
               <host_URI>

    pwdutil.pl -h | -?
```

Table C-1 Options with pwdutil.pl utility

Option	Usage
--action|-a 'check|configure|unconfigure'	Specifies action type, default is 'check'.
--type|-t 'ssh|rsh'	Specifies connection type, default is 'ssh'.
--user|-u '<user>'	Specifies user id, default is the local user id.
--password|-p '<password>'	Specifies user password, default is the user id.
--port|-P '<port>'	Specifies port number for ssh connection, default is 22

Table C-1 Options with pwdutil.pl utility *(continued)*

Option	Usage
--keyfile|-k '<keyfile>'	Specifies the private key file.
--hostfile|-f '<hostfile>'	Specifies the file which list the hosts.
-debug	Prints debug information.
-h|-?	Prints help messages.
<host_URI>	Can be in the following formats: <hostname> <user>:<password>@<hostname> <user>:<password>@<hostname>: <port>

You can check, configure, and unconfigure ssh or rsh using the `pwdutil.pl`utility. For example:

- To check ssh connection for only one host:

  ```
  pwdutil.pl check ssh hostname
  ```

- To configure ssh for only one host:

  ```
  pwdutil.pl configure ssh hostname user password
  ```

- To unconfigure rsh for only one host:

  ```
  pwdutil.pl unconfigure rsh hostname
  ```

- To configure ssh for multiple hosts with same user ID and password:

  ```
  pwdutil.pl -a configure -t ssh -u user -p password hostname1
  hostname2 hostname3
  ```

- To configure ssh or rsh for different hosts with different user ID and password:

  ```
  pwdutil.pl -a configure -t ssh user1:password1@hostname1
  user2:password2@hostname2
  ```

- To check or configure ssh or rsh for multiple hosts with one configuration file:

  ```
  pwdutil.pl -a configure -t ssh --hostfile /tmp/sshrsh_hostfile
  ```

- To keep the host configuration file secret, you can use the 3rd party utility to encrypt and decrypt the host file with password.
  For example:

```
### run openssl to encrypt the host file in base64 format
# openssl aes-256-cbc -a -salt -in /hostfile -out /hostfile.enc
enter aes-256-cbc encryption password: <password>
Verifying - enter aes-256-cbc encryption password: <password>

### remove the original plain text file
# rm /hostfile

### run openssl to decrypt the encrypted host file
# pwdutil.pl -a configure -t ssh `openssl aes-256-cbc -d -a
-in /hostfile.enc`
enter aes-256-cbc decryption password: <password>
```

- To use the ssh authentication keys which are not under the default $HOME/.ssh directory, you can use --keyfile option to specify the ssh keys. For example:

```
### create a directory to host the key pairs:
# mkdir /keystore

### generate private and public key pair under the directory:
# ssh-keygen -t rsa -f /keystore/id_rsa

### setup ssh connection with the new generated key pair under
the directory:
# pwdutil.pl -a configure -t ssh --keyfile /keystore/id_rsa
user:password@hostname
```

You can see the contents of the configuration file by using the following command:

```
# cat /tmp/sshrsh_hostfile
user1:password1@hostname1
user2:password2@hostname2
user3:password3@hostname3
user4:password4@hostname4

# all default: check ssh connection with local user
hostname5
The following exit values are returned:

0     Successful completion.
```

```
1     Command syntax error.
2     Ssh or rsh binaries do not exist.
3     Ssh or rsh service is down on the remote machine.
4     Ssh or rsh command execution is denied due to password is required.
5     Invalid password is provided.
255   Other unknown error.
```

Restarting the ssh session

After you complete this procedure, ssh can be restarted in any of the following scenarios:

- After a terminal session is closed
- After a new terminal session is opened
- After a system is restarted
- After too much time has elapsed, to refresh ssh

To restart ssh

1. On the source installation system (sys1), bring the private key into the shell environment.

   ```
   sys1 # exec /usr/bin/ssh-agent $SHELL
   ```

2. Make the key globally available for the user `root`

   ```
   sys1 # ssh-add
   ```

Enabling rsh for Linux

The following section describes how to enable remote shell.

Veritas recommends configuring a secure shell environment for Veritas InfoScale product installations.

See “Manually configuring passwordless ssh” on page 319.

See the operating system documentation for more information on configuring remote shell.

To enable rsh for RHEL

- Run the following commands to enable rsh passwordless connection:

```
# systemctl start rsh.socket
# systemctl start rlogin.socket
# systemctl enable rsh.socket
# systemctl enable rlogin.socket
# echo rsh >> /etc/securetty
# echo rlogin >> /etc/securetty
#echo "+ +" >> /root/.rhosts
```

To disable rsh for RHEL

- Run the following commands to disable rsh passwordless connection:

```
# systemctl stop rsh.socket
# systemctl stop rlogin.socket
# systemctl disable rsh.socket
# systemctl disable rlogin.socket
```

Appendix D

High availability agent information

This appendix includes the following topics:

- About agents
- Enabling and disabling intelligent resource monitoring for agents manually
- CVMCluster agent
- CVMVxconfigd agent
- CVMVolDg agent
- CFSMount agent
- CFSfsckd agent

About agents

An agent is defined as a process that starts, stops, and monitors all configured resources of a type, and reports their status to Cluster Server (VCS). Agents have both entry points and attributes. Entry points are also known as agent functions and are referred to as "agent functions" throughout the document.

Attributes contain data about the agent. An attribute has a definition and a value. You change attribute values to configure resources, which are defined as the individual components that work together to provide application services to the public network. For example, a resource may be a physical component such as a disk or a network interface card, a software component such as Oracle or a web server, or a configuration component such as an IP address or mounted file system.

Attributes are either optional or required, although sometimes the attributes that are optional in one configuration may be required in other configurations. Many optional attributes have predefined or default values, which you should change as required. A variety of internal use only attributes also exist. Do not modify these attributes—modifying them can lead to significant problems for your clusters. Attributes have type and dimension. Some attribute values can accept numbers, others can accept alphanumeric values or groups of alphanumeric values, while others are simple boolean on/off values.

The entry points and attributes for each SFCFSHA agent are described in this appendix.

VCS agents included within SFCFSHA

SFCFSHA includes the following VCS agents:

- CVMCluster agent
- CVMVxconfigd agent
- CVMVolDg agent
- CFSMount agent
- CFSfsckd
- Coordination Point agent

An SFCFSHA installation automatically configures the CVMCluster resource and the CVMVxconfigd resource.

You must configure the CVMVolDg agent for each shared disk group. If the database uses cluster file systems, configure the CFSMount agent for each volume in the disk group.

Use the information in this appendix about the entry points and attributes of the listed agents to make necessary configuration changes. For information on how to modify the VCS configuration:

See the *Cluster Server Administrator's Guide*.

Enabling and disabling intelligent resource monitoring for agents manually

Review the following procedures to enable or disable intelligent resource monitoring manually. The intelligent resource monitoring feature is enabled by default. The IMF resource type attribute determines whether an IMF-aware agent must perform intelligent resource monitoring.

To enable intelligent resource monitoring

1 Make the VCS configuration writable.

```
# haconf -makerw
```

2 Run the following command to enable intelligent resource monitoring.

- To enable intelligent monitoring of offline resources:

```
# hatype -modify resource_type IMF -update Mode 1
```

- To enable intelligent monitoring of online resources:

```
# hatype -modify resource_type IMF -update Mode 2
```

- To enable intelligent monitoring of both online and offline resources:

```
# hatype -modify resource_type IMF -update Mode 3
```

3 If required, change the values of the MonitorFreq key and the RegisterRetryLimit key of the IMF attribute.

Review the agent-specific recommendations in the attribute definition tables to set these attribute key values.

See "Attribute definition for CVMVxconfigd agent" on page 337.

See "Attribute definition for CFSMount agent" on page 344.

See "Attribute definition for CFSfsckd agent" on page 348.

4 Save the VCS configuration.

```
# haconf -dump -makero
```

5 Make sure that the AMF kernel driver is configured on all nodes in the cluster.

```
/opt/VRTSamf/bin/amf status
```

If the AMF kernel driver is configured, the output resembles:

```
AMF: Module loaded and configured
```

Configure the AMF driver if the command output returns that the AMF driver is not loaded or not configured.

See "Administering the AMF kernel driver" on page 333.

6 Restart the agent. Run the following commands on each node.

```
# haagent -stop agent_name -force -sys sys_name
# haagent -start agent_name -sys sys_name
```

To disable intelligent resource monitoring

1 Make the VCS configuration writable.

```
# haconf -makerw
```

2 To disable intelligent resource monitoring for all the resources of a certain type, run the following command:

```
# hatype -modify resource_type IMF -update Mode 0
```

3 To disable intelligent resource monitoring for a specific resource, run the following command:

```
# hares -override resource_name IMF
# hares -modify resource_name IMF -update Mode 0
```

4 Save the VCS configuration.

```
# haconf -dump -makero
```

Note: VCS provides haimfconfig script to enable or disable the IMF functionality for agents. You can use the script with VCS in running or stopped state. Use the script to enable or disable IMF for the IMF-aware bundled agents, enterprise agents, and custom agents.

Administering the AMF kernel driver

Review the following procedures to start, stop, or unload the AMF kernel driver.

To start the AMF kernel driver

1. Set the value of the AMF_START variable to 1 in the following file, if the value is not already 1:

   ```
   # /etc/sysconfig/amf
   ```

2. Start the AMF kernel driver. Run the following command:

   ```
   # systemctl start amf
   ```

To stop the AMF kernel driver

1. Set the value of the AMF_STOP variable to 1 in the following file, if the value is not already 1:

   ```
   # /etc/sysconfig/amf
   ```

2. Stop the AMF kernel driver. Run the following command:

   ```
   # systemctl stop amf
   ```

To unload the AMF kernel driver

1. If agent downtime is not a concern, use the following steps to unload the AMF kernel driver:
   - Stop the agents that are registered with the AMF kernel driver.
     The `amfstat` command output lists the agents that are registered with AMF under the Registered Reapers section.
     See the `amfstat` manual page.
   - Stop the AMF kernel driver.
     See "To stop the AMF kernel driver" on page 333.
   - Start the agents.
2. If you want minimum downtime of the agents, use the following steps to unload the AMF kernel driver:
   - Run the following command to disable the AMF driver even if agents are still registered with it.

     ```
     # amfconfig -Uof
     ```

   - Stop the AMF kernel driver.
     See "To stop the AMF kernel driver" on page 333.

CVMCluster agent

The CVMCluster agent controls system membership on the cluster port that is associated with Veritas Volume Manager (VxVM).

The CVMCluster agent performs the following functions:

- Joins a node to the CVM cluster port.
- Removes a node from the CVM cluster port.
- Monitors the node's cluster membership state.

Entry points for CVMCluster agent

Table D-1 describes the entry points used by the CVMCluster agent.

Table D-1 CVMCluster agent entry points

Entry Point	Description
Online	Joins a node to the CVM cluster port. Enables the Volume Manager cluster functionality by automatically importing the shared disk groups.
Offline	Removes a node from the CVM cluster port.
Monitor	Monitors the node’s CVM cluster membership state.

Attribute definition for CVMCluster agent

Table D-2 describes the user-modifiable attributes of the CVMCluster resource type.

Table D-2 CVMCluster agent attributes

Attribute	Description
CVMClustName	Name of the cluster. - Type and dimension: string-scalar
CVMNodeAddr	List of host names and IP addresses. - Type and dimension: string-association
CVMNodeId	Associative list. The first part names the system; the second part contains the LLT ID number for the system. - Type and dimension: string-association

Table D-2 CVMCluster agent attributes *(continued)*

Attribute	Description
CVMTransport	Specifies the cluster messaging mechanism. ■ Type and dimension: string-scalar ■ Default = gab **Note:** Do not change this value.
PortConfigd	The port number that is used by CVM for vxconfigd-level communication. ■ Type and dimension: integer-scalar
PortKmsgd	The port number that is used by CVM for kernel-level communication. ■ Type and dimension: integer-scalar
CVMTimeout	Timeout in seconds used for CVM cluster reconfiguration. ■ Type and dimension: integer-scalar ■ Default = 200
CVMNodePreference	The preference value that determines which nodes in the CVM cluster are the most appropriate candidates to run the master role. The preference values are in the range from -2147483648 to 2147483647. If you do not specify custom preferences, CVM gives preference to the node with the maximum visibility to the storage to become the CVM master node. ■ Type and dimension: string ■ Default = 0
CVMDGSubClust	Enables or disables the application isolation feature. The values for the attribute are as follows: ■ 1 The application isolation feature is enabled when the cluster starts. The shared disk groups will not be auto-imported on all nodes in the cluster as in the traditional CVM environment. ■ 0 If the attribute is set to 0, the feature is disabled when the cluster starts and all the shared disk groups are auto-imported on all nodes in the cluster.

CVMCluster agent type definition

The following type definition is included in the file, `CVMTypes.cf`:

```
type CVMCluster (
        static keylist RegList = { CVMNodePreference, CVMDGSubClust }
        static int NumThreads = 1
        static int OnlineRetryLimit = 2
        static int OnlineTimeout = 400
        static str ArgList[] = { CVMTransport, CVMClustName, CVMNodeAddr, CVMNodeId,
                                 PortConfigd, PortKmsgd, CVMTimeout, CVMDGSubClust }
        str CVMClustName
        str CVMNodeAddr{}
        str CVMNodeId{}
        str CVMTransport
        str CVMNodePreference
        int PortConfigd
        int PortKmsgd
        int CVMTimeout
        boolean CVMDGSubClust = 0
```

Note: The attributes `CVMNodeAddr`, `PortConfigd`, and `PortKmsgd` are not used in an SFCFSHA environment. GAB, the required cluster communication messaging mechanism, does not use them.

CVMCluster agent sample configuration

The following is an example definition for the CVMCluster service group:

```
CVMCluster cvm_clus (
        Critical = 0
        CVMClustName = clus1
        CVMNodeId = { sys1 = 0, sys2 = 1 }
        CVMTransport = gab
        CVMTimeout = 200
         )
```

CVMVxconfigd agent

The CVMVxconfigd agent starts and monitors the vxconfigd daemon. The vxconfigd daemon maintains disk and disk group configurations, communicates configuration changes to the kernel, and modifies the configuration information that is stored on disks. CVMVxconfigd must be present in the CVM service group.

The CVMVxconfigd agent is an OnOnly agent; the agent starts the resource when the cluster starts up and VCS restarts the resource when necessary. The Operations attribute specifies these default aspects of startup.

Veritas recommends starting the vxconfigd daemon with the `syslog` option, which enables logging of debug messages. Note that the SFCFSHA installation configures the `syslog` option for the CVMVxconfigd agent.

This agent is IMF-aware and uses asynchronous monitoring framework (AMF) kernel driver for IMF notification. For more information about the Intelligent Monitoring Framework (IMF) and intelligent resource monitoring, refer to the *Cluster Server Administrator's Guide*.

Entry points for CVMVxconfigd agent

Table D-3 describes the entry points for the CVMVxconfigd agent.

Table D-3 CVMVxconfigd entry points

Entry Point	Description
Online	Starts the `vxconfigd` daemon
Offline	N/A
Monitor	Monitors whether `vxconfigd` daemon is running
imf_init	Initializes the agent to interface with the AMF kernel module. This function runs when the agent starts up.
imf_getnotification	Gets notification about the `vxconfigd` process state. This function runs after the agent initializes with the AMF kernel module. This function continuously waits for notification. If the `vxconfigd` process fails, the function initiates a traditional CVMVxconfigd monitor entry point.
imf_register	Registers or unregisters the `vxconfigd` process id (pid) with the AMF kernel module. This function runs after the resource goes into steady online state.

Attribute definition for CVMVxconfigd agent

Table D-4 describes the modifiable attributes of the CVMVxconfigd resource type.

Table D-4 CVMVxconfigd agent attribute

Attribute	Description
CVMVxconfigdArgs	List of the arguments that are sent to the `online` entry point. Veritas recommends always specifying the `syslog` option. ■ Type and dimension: keylist
IMF	This resource-type level attribute determines whether the CVMVxconfigd agent must perform intelligent resource monitoring. You can also override the value of this attribute at resource-level. This attribute includes the following keys: ■ Mode: Define this attribute to enable or disable intelligent resource monitoring. Valid values are as follows: ■ 0—Does not perform intelligent resource monitoring ■ 2—Performs intelligent resource monitoring for online resources and performs poll-based monitoring for offline resources Default: 0 ■ MonitorFreq: This key value specifies the frequency at which the agent invokes the monitor agent function. The value of this key is an integer. Default: 1 You can set this key to a non-zero value for cases where the agent requires to perform both poll-based and intelligent resource monitoring. If the value is 0, the agent does not perform poll-based process check monitoring. After the resource registers with the AMF kernel driver, the agent calls the monitor agent function as follows: ■ After every (MonitorFreq x MonitorInterval) number of seconds for online resources ■ After every (MonitorFreq x OfflineMonitorInterval) number of seconds for offline resources ■ RegisterRetryLimit: If you enable intelligent resource monitoring, the agent invokes the imf_register agent function to register the resource with the AMF kernel driver. The value of the RegisterRetyLimit key determines the number of times the agent must retry registration for a resource. If the agent cannot register the resource within the limit that is specified, then intelligent monitoring is disabled until the resource state changes or the value of the Mode key changes. Default: 3. ■ Type and dimension: integer-association For more details of IMF attribute for the agent type, refer to the *Cluster Server Administrator's Guide*.

CVMVxconfigd agent type definition

The following type definition is included in the CVMTypes.cf file:

```
type CVMVxconfigd (
        static int IMF{} = { Mode=2, MonitorFreq=1, RegisterRetryLimit=3 }
        static int FaultOnMonitorTimeouts = 2
        static int RestartLimit = 5
        static str ArgList[] = { CVMVxconfigdArgs }
        static str Operations = OnOnly
        keylist CVMVxconfigdArgs
)
```

CVMVxconfigd agent sample configuration

The following is an example definition for the `CVMVxconfigd` resource in the CVM service group:

```
CVMVxconfigd cvm_vxconfigd (
        Critical = 0
        CVMVxconfigdArgs = { syslog }
        )
```

CVMVolDg agent

The CVMVolDg agent manages the CVM disk groups and CVM volumes and volume sets within the disk groups by performing the following functions:

- Imports the shared disk group from the CVM master node
- Starts the volumes and volume sets in the disk group
- Monitors the disk group, volumes, and volume sets
- Optionally, deports the disk group when the dependent applications are taken offline. The agent deports the disk group only if the appropriate attribute is set.

Configure the CVMVolDg agent for each disk group used by a Oracle service group. A disk group must be configured to only one Oracle service group.If cluster file systems are used for the database, configure the CFSMount agent for each volume or volume set in the disk group.

Entry points for CVMVolDg agent

Table D-5 describes the entry points used by the CVMVolDg agent.

Table D-5 CVMVolDg agent entry points

Entry Point	Description
Online	Imports the shared disk group from the CVM master node, if the disk group is not already imported. Starts all volumes and volume sets in the shared disk group specified by the CVMVolume attribute. Sets the disk group activation mode to shared-write if the value of the CVMActivation attribute is sw. You can set the activation mode on both slave and master systems.
Offline	Removes the temporary files created by the online entry point. If the `CVMDeportOnOffline` attribute is set to 1 and if the shared disk group does not contain open volumes on any node in the cluster, the disk group is deported from the CVM master node.
Monitor	Determines whether the disk group, the volumes, and the volume sets are online. The agent takes a volume set offline if the file system metadata volume of a volume set is discovered to be offline in a monitor cycle. **Note:** If the CFSMount resource goes offline and the file system on the volume set is unmounted, the agent retains the online state of the volume set even if the file system metadata volume in the volume set is offline. This is because the CVMVolDg agent is unable to determine whether or not the volumes that are offline are metadata volumes.
Clean	Removes the temporary files created by the online entry point.

Attribute definition for CVMVolDg agent

Table D-6 describes the user-modifiable attributes of the CVMVolDg resource type.

Table D-6 CVMVolDg agent attributes

Attribute	Description
CVMDiskGroup (required)	Shared disk group name. ■ Type and dimension: string-scalar

Table D-6 CVMVolDg agent attributes *(continued)*

Attribute	Description
CVMVolume (required)	Name of shared volumes or volume sets. This list is used to check that the volumes or volume sets are in the correct state before allowing the resource to come online, and that the volumes remain in an enabled state. ■ Type and dimension: string-keylist
CVMActivation (required)	Activation mode for the disk group. ■ Type and dimension: string-scalar ■ Default = `sw (shared-write)` This is a localized attribute.
CVMVolumeIoTest(optional)	List of volumes and volume sets that will be periodically polled to test availability. The polling is in the form of 4 KB reads every monitor cycle to a maximum of 10 of the volumes or volume sets in the list. For volume sets, reads are done on a maximum of 10 component volumes in each volume set. ■ Type and dimension: string-keylist
CVMDeportOnOffline (optional)	Indicates whether or not the shared disk group must be deported when the last online CVMVolDg resource for a disk group is taken offline. The value 1 indicates that the agent will deport the shared disk group from the CVM master node, if not already deported, when the last online CVMVolDg resource for the disk group is taken offline. The value 0 indicates that the agent will not deport the shared disk group when the CVMVolDg resource is taken offline. ■ Type and dimension: integer-scalar ■ Default = 0 **Note:** If multiple CVMVolDg resources are configured for a shared disk group, set the value of the attribute to either 1 or 0 for all of the resources. The CVM disk group is deported based on the order in which the CVMVolDg resources are taken offline. If the CVMVolDg resources in the disk group contain a mixed setting of 1 and 0 for the `CVMDeportOnOffline` attribute, the disk group is deported only if the attribute value is 1 for the last CVMVolDg resource taken offline. If the attribute value is 0 for the last CVMVolDg resource taken offline, the disk group is not deported. The deport operation fails if the shared disk group contains open volumes.

Table D-6 CVMVolDg agent attributes *(continued)*

Attribute	Description
ClearClone	Indicates whether or not the disk group is imported with the `-c` option. The `-c` option clears the clone and the udid_mismatch flags from the disks of the disk groups and also updates the UDID when you import the disk group. ■ Type and dimension: integer-scalar ■ Default = 0
NodeList	List of nodes that belong to a disk group sub-cluster. This attribute is required only if you want to configure application isolation capability in your environment.

CVMVolDg agent type definition

The CVMTypes.cf file includes the CVMVolDg type definition:

```
type CVMVolDg (
        static keylist RegList = { CVMActivation, CVMVolume, NodeList }
        static int OnlineRetryLimit = 2
        static int OnlineTimeout = 400
        static int OnlineWaitLimit = 10
        static str ArgList[] = { CVMDiskGroup, CVMVolume, CVMActivation, CVMV
CVMDGAction, CVMDeportOnOffline, State, CVMDeactivateOnOffline, ClearClone, N
        str CVMDiskGroup
        str CVMDGAction
        keylist CVMVolume
        str CVMActivation
        keylist CVMVolumeIoTest
        int CVMDeportOnOffline
        int CVMDeactivateOnOffline
        int ClearClone
        temp int voldg_stat
        keylist NodeList
)
```

CVMVolDg agent sample configuration

Each Oracle service group requires a CVMVolDg resource type to be defined. The following is a sample configuration:

```
CVMVolDg cvmvoldg1 (
 Critical = 0
```

```
 CVMDiskgroup = testdg
 CVMVolume = { vol1, vol2, mvol1, mvol2, snapvol, vset1 }
 CVMVolumeIoTest = { snapvol, vset1 }
 CVMActivation @sys1 = sw
 CVMActivation @sys2 = sw
 CVMDeportOnOffline = 1
)
```

CFSMount agent

The CFSMount agent brings online, takes offline, and monitors a cluster file system mount point.

The agent executable is located in /opt/VRTSvcs/bin/CFSMount/CFSMountAgent.

The CFSMount type definition is described in the /etc/VRTSvcs/conf/config/CFSTypes.cf file.

This agent is IMF-aware and uses asynchronous monitoring framework (AMF) kernel driver for IMF notification. For more information about the Intelligent Monitoring Framework (IMF) and intelligent resource monitoring, refer to the *Cluster Server Administrator's Guide*.

Entry points for CFSMount agent

Table D-7 provides the entry points for the CFSMount agent.

Table D-7 CFSMount agent entry points

Entry Point	Description
Online	Mounts a block device in cluster mode.
Offline	Unmounts the file system, forcing unmount if necessary, and sets primary to secondary if necessary.
Monitor	Determines if the file system is mounted. Checks mount status using the `fsclustadm` command.
Clean	Generates a null operation for a cluster file system mount.
imf_init	Initializes the agent to interface with the AMF kernel driver, which is the IMF notification module for the agent. This function runs when the agent starts up.

Table D-7 CFSMount agent entry points *(continued)*

Entry Point	Description
imf_getnotification	Gets notification about resource state changes. This function runs after the agent initializes with the AMF kernel module. This function continuously waits for notification and takes action on the resource upon notification.
imf_register	Registers or unregisters resource entities with the AMF kernel module. This function runs for each resource after the resource goes into steady state (online or offline).

Attribute definition for CFSMount agent

Table D-8 lists user-modifiable attributes of the CFSMount Agent resource type.

Table D-8 CFSMount Agent attributes

Attribute	Description
MountPoint	Directory for the mount point. ■ Type and dimension: string-scalar
BlockDevice	Block device for the mount point. ■ Type and dimension: string-scalar
NodeList	List of nodes on which to mount. If NodeList is NULL, the agent uses the service group system list. ■ Type and dimension: string-keylist

Table D-8 CFSMount Agent attributes *(continued)*

Attribute	Description
IMF	Resource-type level attribute that determines whether the CFSMount agent must perform intelligent resource monitoring. You can also override the value of this attribute at resource-level. This attribute includes the following keys: ■ Mode: Define this attribute to enable or disable intelligent resource monitoring. Valid values are as follows: ■ 0—Does not perform intelligent resource monitoring ■ 1—Performs intelligent resource monitoring for offline resources and performs poll-based monitoring for online resources ■ 2—Performs intelligent resource monitoring for online resources and performs poll-based monitoring for offline resources ■ 3—Performs intelligent resource monitoring for both online and for offline resources Default: 0 ■ MonitorFreq: This key value specifies the frequency at which the agent invokes the monitor agent function. The value of this key is an integer. Default: 1 You can set this key to a non-zero value for cases where the agent requires to perform both poll-based and intelligent resource monitoring. If the value is 0, the agent does not perform poll-based process check monitoring. After the resource registers with the AMF kernel driver, the agent calls the monitor agent function as follows: ■ After every (MonitorFreq x MonitorInterval) number of seconds for online resources ■ After every (MonitorFreq x OfflineMonitorInterval) number of seconds for offline resources ■ RegisterRetryLimit: If you enable intelligent resource monitoring, the agent invokes the imf_register agent function to register the resource with the AMF kernel driver. The value of the RegisterRetyLimit key determines the number of times the agent must retry registration for a resource. If the agent cannot register the resource within the limit that is specified, then intelligent monitoring is disabled until the resource state changes or the value of the Mode key changes. Default: 3. ■ Type and dimension: integer-association See "Enabling and disabling intelligent resource monitoring for agents manually" on page 330.

Table D-8 CFSMount Agent attributes *(continued)*

Attribute	Description

MountOpt
(optional)
	Options for the mount command. To create a valid MountOpt attribute string:
	■ Use the VxFS type-specific options only.
	■ Do not use the -o flag to specify the VxFS-specific options.
	■ Do not use the `-t` vxfs file system type option.
	■ Be aware the cluster option is not required.
	■ Specify options in comma-separated list:
	`ro`
	`ro,cluster`
	`blkclear,mincache=closesync`
	■ Type and dimension: string-scalar

Policy (optional)	List of nodes to assume the primaryship of the cluster file system if the primary fails. If set to NULL or if none of the hosts specified in the list is active when the primary fails, a node is randomly selected from the set of active nodes to assume primaryship. ■ Type and dimension: string-scalar

CFSMount agent type definition

The `CFSTypes.cf` file includes the CFSMount agent type definition:

```
type CFSMount (
        static int IMF{} = { Mode=3, MonitorFreq=1, RegisterRetryLimit=3 }
        static keylist RegList = { MountOpt, Policy, NodeList, ForceOff, SetPrimary }
        static keylist SupportedActions = { primary }
        static int FaultOnMonitorTimeouts = 1
        static int OnlineWaitLimit = 1
        static str ArgList[] = { MountPoint, BlockDevice, MountOpt, Primary, AMFMountType }
        str MountPoint
        str MountType
        str BlockDevice
        str MountOpt
        keylist NodeList
        keylist Policy
        temp str Primary
        str SetPrimary
        temp str RemountRes
        temp str AMFMountType
```

```
        str ForceOff
)
```

CFSMount agent sample configuration

Each Oracle service group requires a CFSMount resource type to be defined:

```
CFSMount ora_mount (
        MountPoint = "/oradata"
        BlockDevice = "/dev/vx/dsk/oradatadg/oradatavol1"
        Primary = sys2;
        )
```

To see CFSMount defined in a more extensive example:

CFSfsckd agent

The CFSfsckd agent starts, stops, and monitors the `vxfsckd` process. The CFSfsckd agent executable is `/opt/VRTSvcs/bin/CFSfsckd/CFSfsckdAgent`. The type definition is in the `/etc/VRTSvcs/conf/config/CFSTypes.cf` file. The configuration is added to the `main.cf` file after running the `cfscluster config` command.

This agent is IMF-aware and uses asynchronous monitoring framework (AMF) kernel driver for IMF notification. For more information about the Intelligent Monitoring Framework (IMF) and intelligent resource monitoring, refer to the *Cluster Server Administrator's Guide*.

Entry points for CFSfsckd agent

Table D-9 describes the CFSfsckd agent entry points.

Table D-9 CFSfsckd agent entry points

Entry Points	Description
Online	Starts the vxfsckd process.
Offline	Kills the vxfsckd process.
Monitor	Checks whether the vxfsckd process is running.
Clean	A null operation for a cluster file system mount.
imf_init	Initializes the agent to interface with the AMF kernel driver, which is the IMF notification module for the agent. This function runs when the agent starts up.

Table D-9 CFSfsckd agent entry points *(continued)*

Entry Points	Description
imf_getnotification	Gets notification about resource state changes. This function runs after the agent initializes with the AMF kernel module. This function continuously waits for notification and takes action on the resource upon notification.
imf_register	Registers or unregisters resource entities with the AMF kernel module. This function runs for each resource after the resource goes into steady state (online or offline).

Attribute definition for CFSfsckd agent

Table D-10 lists user-modifiable attributes of the CFSfsckd Agent resource type.

Table D-10 CFSfsckd Agent attributes

Attribute	Description
IMF	Resource-type level attribute that determines whether the CFSfsckd agent must perform intelligent resource monitoring. You can also override the value of this attribute at resource-level. This attribute includes the following keys: ■ Mode: Define this attribute to enable or disable intelligent resource monitoring. Valid values are as follows: ■ 0—Does not perform intelligent resource monitoring ■ 1—Performs intelligent resource monitoring for offline resources and performs poll-based monitoring for online resources ■ 2—Performs intelligent resource monitoring for online resources and performs poll-based monitoring for offline resources ■ 3—Performs intelligent resource monitoring for both online and for offline resources Default: 0 ■ MonitorFreq: This key value specifies the frequency at which the agent invokes the monitor agent function. The value of this key is an integer. Default: 1 You can set this key to a non-zero value for cases where the agent requires to perform both poll-based and intelligent resource monitoring. If the value is 0, the agent does not perform poll-based process check monitoring. After the resource registers with the AMF kernel driver, the agent calls the monitor agent function as follows: ■ After every (MonitorFreq x MonitorInterval) number of seconds for online resources ■ After every (MonitorFreq x OfflineMonitorInterval) number of seconds for offline resources ■ RegisterRetryLimit: If you enable intelligent resource monitoring, the agent invokes the imf_register agent function to register the resource with the AMF kernel driver. The value of the RegisterRetyLimit key determines the number of times the agent must retry registration for a resource. If the agent cannot register the resource within the limit that is specified, then intelligent monitoring is disabled until the resource state changes or the value of the Mode key changes. Default: 3. ■ Type and dimension: integer-association See "Enabling and disabling intelligent resource monitoring for agents manually" on page 330.

CFSfsckd agent type definition

The CFSfsckd type definition:

```
type CFSfsckd (
        static int IMF{} = { Mode=3, MonitorFreq=1, RegisterRetryLimit=3 }
        static int RestartLimit = 1
```

```
        str ActivationMode{}
)
```

CFSfsckd agent sample configuration

This is a sample of CFSfsckd configuration:

```
CFSfsckd vxfsckd (
)
```

Appendix E

Sample SFCFSHA cluster setup diagrams for CP server-based I/O fencing

This appendix includes the following topics:

- Configuration diagrams for setting up server-based I/O fencing

Configuration diagrams for setting up server-based I/O fencing

The following CP server configuration diagrams can be used as guides when setting up CP server within your configuration:

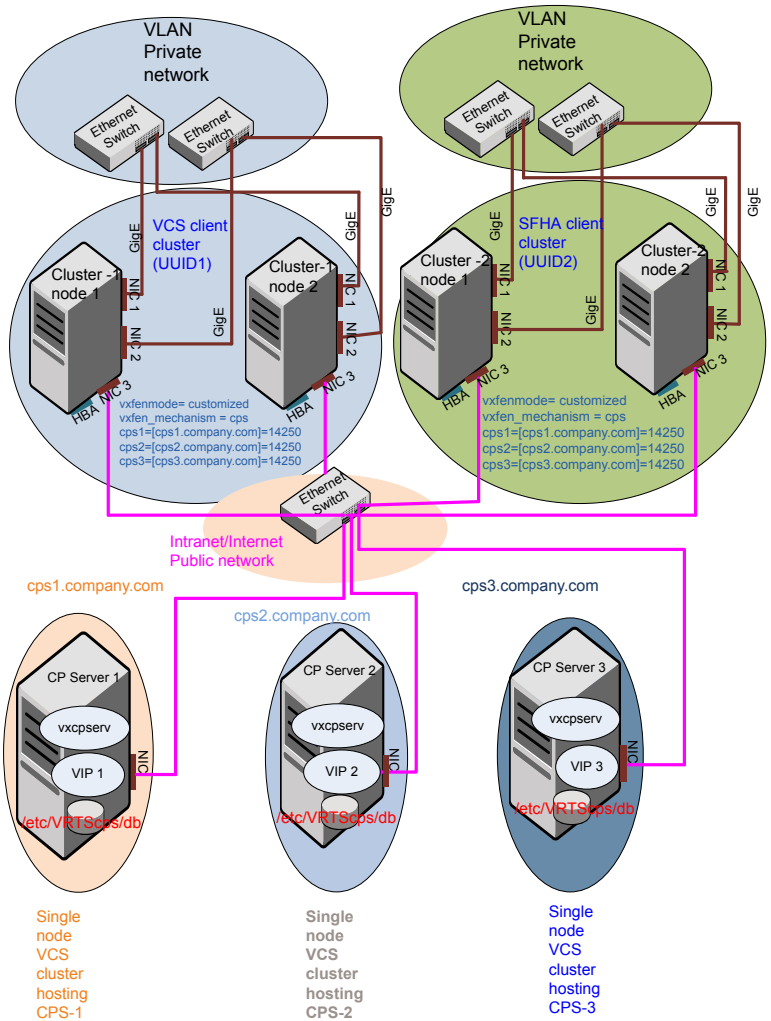
- Two unique client clusters that are served by 3 CP servers:
  See Figure E-1 on page 352.
- Client cluster that is served by highly available CP server and 2 SCSI-3 disks:
- Two node campus cluster that is served be remote CP server and 2 SCSI-3 disks:
- Multiple client clusters that are served by highly available CP server and 2 SCSI-3 disks:

Two unique client clusters served by 3 CP servers

Figure E-1 displays a configuration where two unique client clusters are being served by 3 CP servers (coordination points). Each client cluster has its own unique user ID (UUID1 and UUID2).

In the `vxfenmode` file on the client nodes, vxfenmode is set to `customized` with vxfen mechanism set to `cps`.

Figure E-1 Two unique client clusters served by 3 CP servers



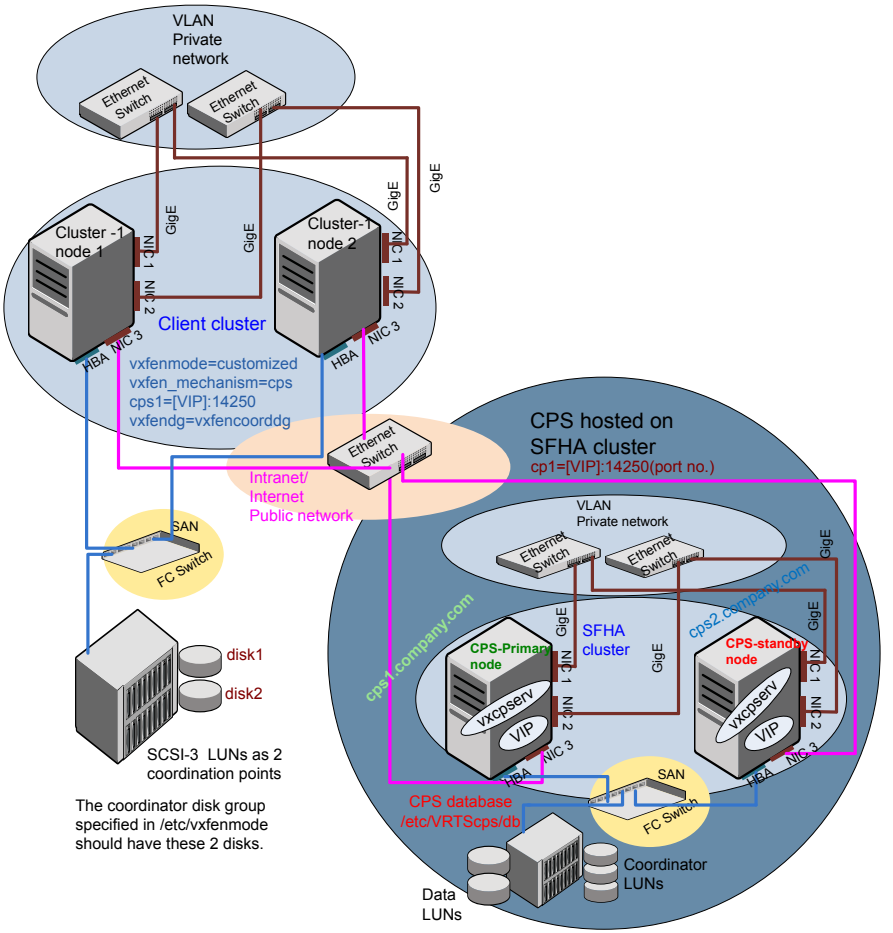
Client cluster served by highly available CPS and 2 SCSI-3 disks

Figure E-2 displays a configuration where a client cluster is served by one highly available CP server and 2 local SCSI-3 LUNs (disks).

In the `vxfenmode` file on the client nodes, `vxfenmode` is set to customized with vxfen mechanism set to `cps`.

The two SCSI-3 disks are part of the disk group vxfencoorddg. The third coordination point is a CP server hosted on an SFHA cluster, with its own shared database and coordinator disks.

Figure E-2 Client cluster served by highly available CP server and 2 SCSI-3 disks



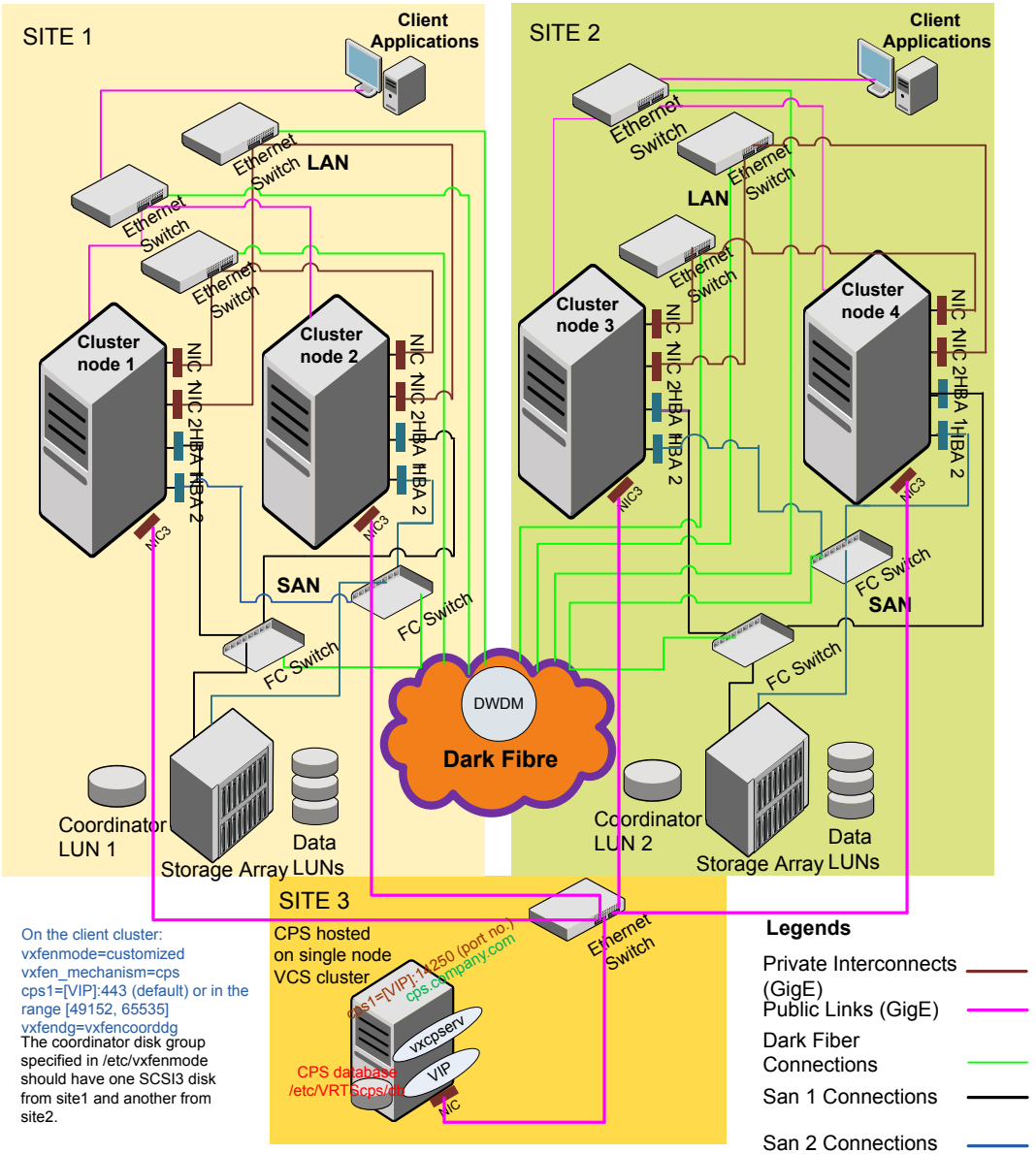
Two node campus cluster served by remote CP server and 2 SCSI-3 disks

Figure E-3 displays a configuration where a two node campus cluster is being served by one remote CP server and 2 local SCSI-3 LUN (disks).

In the `vxfenmode` file on the client nodes, vxfenmode is set to `customized` with vxfen mechanism set to `cps`.

The two SCSI-3 disks (one from each site) are part of disk group vxfencoorddg. The third coordination point is a CP server on a single node VCS cluster.

Figure E-3 Two node campus cluster served by remote CP server and 2 SCSI-3



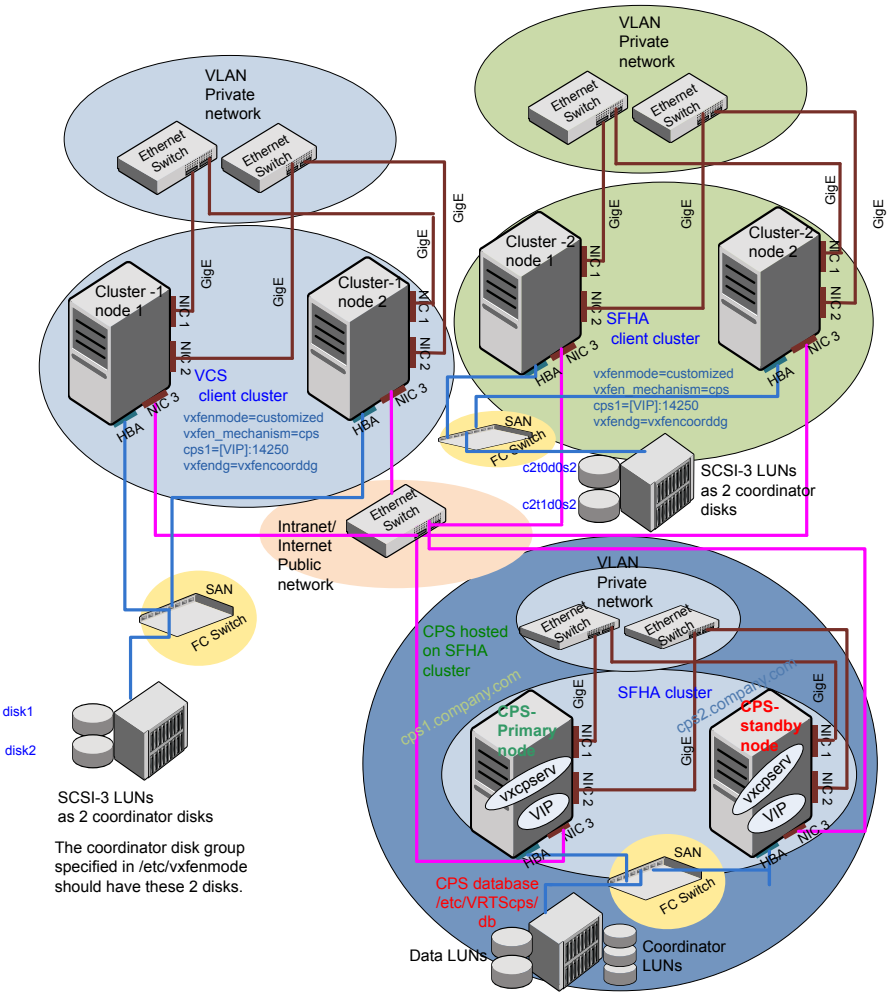
Multiple client clusters served by highly available CP server and 2 SCSI-3 disks

Figure E-4 displays a configuration where multiple client clusters are being served by one highly available CP server and 2 local SCSI-3 LUNS (disks).

In the `vxfenmode` file on the client nodes, vxfenmode is set to `customized` with vxfen mechanism set to `cps`.

The two SCSI-3 disks are are part of the disk group vxfencoorddg. The third coordination point is a CP server, hosted on an SFHA cluster, with its own shared database and coordinator disks.

Figure E-4 Multiple client clusters served by highly available CP server and 2 SCSI-3 disks



Appendix F

Configuring LLT over UDP

This appendix includes the following topics:

- Using the UDP layer for LLT
- Manually configuring LLT over UDP using IPv4
- Using the UDP layer of IPv6 for LLT
- Manually configuring LLT over UDP using IPv6
- About configuring LLT over UDP multiport

Using the UDP layer for LLT

SFCFSHA provides the option of using LLT over the UDP (User Datagram Protocol) layer for clusters using wide-area networks and routers. UDP makes LLT packets routable and thus able to span longer distances more economically.

When to use LLT over UDP

Use LLT over UDP in the following situations:

- LLT must be used over WANs
- When hardware, such as blade servers, do not support LLT over Ethernet

LLT over UDP is slower than LLT over Ethernet. Use LLT over UDP only when the hardware configuration makes it necessary.

Manually configuring LLT over UDP using IPv4

The following checklist is to configure LLT over UDP:

- Make sure that the LLT private links are on separate subnets. Set the broadcast address in /etc/lltab explicitly depending on the subnet for each link.
  See "Broadcast address in the /etc/lltab file" on page 359.
- Make sure that each NIC has an IP address that is configured before configuring LLT.
- Make sure the IP addresses in the /etc/lltab files are consistent with the IP addresses of the network interfaces.
- Make sure that each link has a unique not well-known UDP port.
  See "Selecting UDP ports" on page 361.
- Set the broadcast address correctly for direct-attached (non-routed) links.
  See "Sample configuration: direct-attached links" on page 362.
- For the links that cross an IP router, disable broadcast features and specify the IP address of each link manually in the /etc/lltab file.
  See "Sample configuration: links crossing IP routers" on page 364.

Broadcast address in the /etc/lltab file

The broadcast address is set explicitly for each link in the following example.

- Display the content of the /etc/lltab file on the first node sys1:

```
sys1 # cat /etc/llttab

set-node sys1
set-cluster 1
link link1 udp - udp  50000  -  192.168.9.1 192.168.9.255
link link2 udp - udp  50001  -  192.168.10.1 192.168.10.255
```

  Verify the subnet mask using the ifconfig command to ensure that the two links are on separate subnets.

- Display the content of the /etc/lltab file on the second node sys2:

```
sys2 # cat /etc/llttab

set-node sys2
set-cluster 1
link link1 udp - udp  50000  -  192.168.9.2 192.168.9.255
link link2 udp - udp  50001  -  192.168.10.2 192.168.10.255
```

  Verify the subnet mask using the ifconfig command to ensure that the two links are on separate subnets.

The link command in the /etc/lltab file

Review the link command information in this section for the /etc/lltab file. See the following information for sample configurations:

- See "Sample configuration: direct-attached links" on page 362.
- See "Sample configuration: links crossing IP routers" on page 364.

Table F-1 describes the fields of the link command that are shown in the /etc/lltab file examples. Note that some of the fields differ from the command for standard LLT links.

Table F-1 Field description for link command in /etc/lltab

Field	Description
tag-name	A unique string that is used as a tag by LLT; for example link1, link2,....
device	The device path of the UDP protocol; for example udp. A place holder string. On other unix platforms like Solaris or HP, this entry points to a device file (for example, /dev/udp). Linux does not have devices for protocols. So this field is ignored.
node-range	Nodes using the link. "-" indicates all cluster nodes are to be configured for this link.
link-type	Type of link; must be "udp" for LLT over UDP.
udp-port	Unique UDP port in the range of 49152-65535 for the link. See "Selecting UDP ports" on page 361.
MTU	"-" is the default, which has a value of 8192. The value may be increased or decreased depending on the configuration. Use the `lltstat -l` command to display the current value.
IP address	IP address of the link on the local node.
bcast-address	■ For clusters with enabled broadcasts, specify the value of the subnet broadcast address. ■ "-" is the default for clusters spanning routers.

The set-addr command in the /etc/lltab file

The `set-addr` command in the /etc/lltab file is required when the broadcast feature of LLT is disabled, such as when LLT must cross IP routers.

See "Sample configuration: links crossing IP routers" on page 364.

Table F-2 describes the fields of the set-addr command.

Table F-2 Field description for set-addr command in /etc/llttab

Field	Description
node-id	The node ID of the peer node; for example, 0.
link tag-name	The string that LLT uses to identify the link; for example link1, link2,....
address	IP address assigned to the link for the peer node.

Selecting UDP ports

When you select a UDP port, select an available 16-bit integer from the range that follows:

- Use available ports in the private range 49152 to 65535
- Do not use the following ports:
  - Ports from the range of well-known ports, 0 to 1023
  - Ports from the range of registered ports, 1024 to 49151

To check which ports are defined as defaults for a node, examine the file /etc/services. You should also use the `netstat` command to list the UDP ports currently in use. For example:

```
# netstat -au | more
Active Internet connections (servers and established)
Proto Recv-Q Send-Q Local Address Foreign Address         State
udp        0      0 *:32768                 *:*
udp        0      0 *:956                   *:*
udp        0      0 *:tftp                  *:*
udp        0      0 *:sunrpc                *:*
udp        0      0 *:ipp                   *:*
```

Look in the UDP section of the output; the UDP ports that are listed under Local Address are already in use. If a port is listed in the /etc/services file, its associated name is displayed rather than the port number in the output.

Configuring the netmask for LLT

For nodes on different subnets, set the netmask so that the nodes can access the subnets in use. Run the following command and answer the prompt to set the netmask:

```
# ifconfig interface_name netmask netmask
```

For example:

- For the first network interface on the node sys1:

  ```
  IP address=192.168.9.1, Broadcast address=192.168.9.255,
  Netmask=255.255.255.0
  ```

  For the first network interface on the node sys2:

  ```
  IP address=192.168.9.2, Broadcast address=192.168.9.255,
  Netmask=255.255.255.0
  ```

- For the second network interface on the node sys1:

  ```
  IP address=192.168.10.1, Broadcast address=192.168.10.255,
  Netmask=255.255.255.0
  ```

  For the second network interface on the node sys2:

  ```
  IP address=192.168.10.2, Broadcast address=192.168.10.255,
  Netmask=255.255.255.0
  ```

Configuring the broadcast address for LLT

For nodes on different subnets, set the broadcast address in /etc/llttab depending on the subnet that the links are on.

An example of a typical /etc/llttab file when nodes are on different subnets. Note the explicitly set broadcast address for each link.

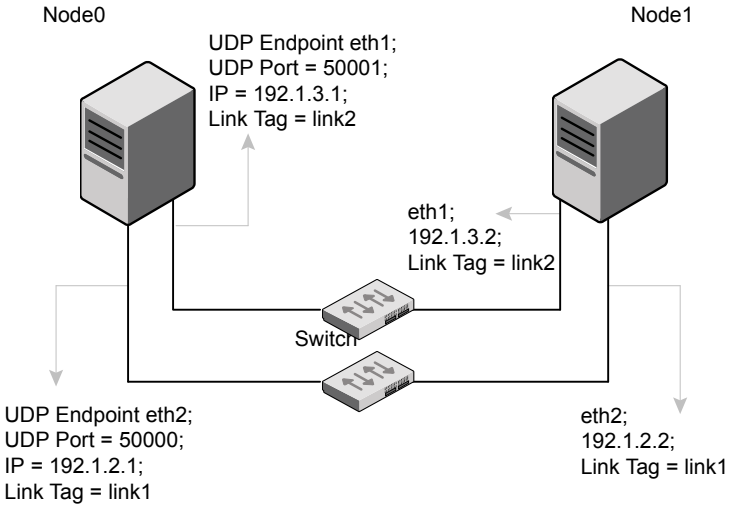
```
# cat /etc/llttab
set-node nodexyz
set-cluster 100

link link1 udp - udp 50000 - 192.168.30.1 192.168.30.255
link link2 udp - udp 50001 - 192.168.31.1 192.168.31.255
```

Sample configuration: direct-attached links

Figure F-1 depicts a typical configuration of direct-attached links employing LLT over UDP.

Figure F-1 A typical configuration of direct-attached links that use LLT over UDP



The configuration that the /etc/llttab file for Node 0 represents has directly attached crossover links. It might also have the links that are connected through a hub or switch. These links do not cross routers.

LLT sends broadcast requests to peer nodes to discover their addresses. So the addresses of peer nodes do not need to be specified in the /etc/llttab file using the `set-addr` command. For direct attached links, you do need to set the broadcast address of the links in the /etc/llttab file. Verify that the IP addresses and broadcast addresses are set correctly by using the `ifconfig -a` command.

```
set-node Node0
set-cluster 1
#configure Links
#link tag-name device node-range link-type udp port MTU \
IP-address bcast-address
link link1 udp - udp 50000 - 192.1.2.1 192.1.2.255
link link2 udp - udp 50001 - 192.1.3.1 192.1.3.255
```

The file for Node 1 resembles:

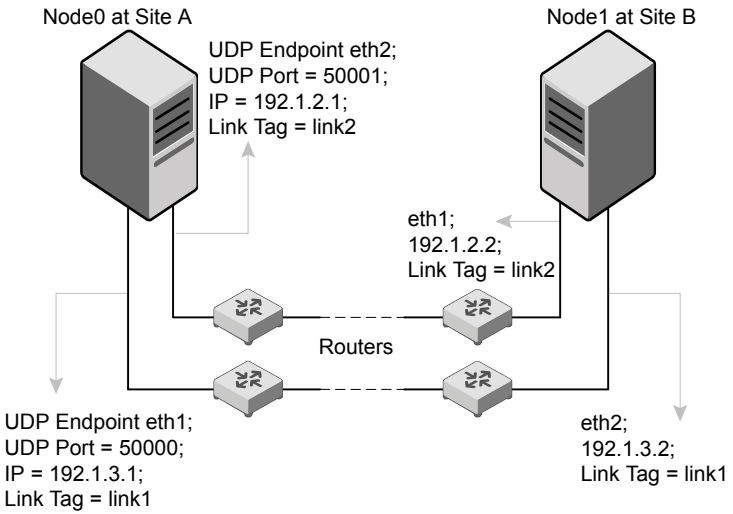
```
set-node Node1
set-cluster 1
#configure Links
#link tag-name device node-range link-type udp port MTU \
IP-address bcast-address
```

```
link link1 udp - udp 50000 - 192.1.2.2 192.1.2.255
link link2 udp - udp 50001 - 192.1.3.2 192.1.3.255
```

Sample configuration: links crossing IP routers

Figure F-2 depicts a typical configuration of links crossing an IP router employing LLT over UDP. The illustration shows two nodes of a four-node cluster.

Figure F-2 A typical configuration of links crossing an IP router



The configuration that the following `/etc/lltab` file represents for Node 1 has links crossing IP routers. Notice that IP addresses are shown for each link on each peer node. In this configuration broadcasts are disabled. Hence, the broadcast address does not need to be set in the `link` command of the `/etc/lltab` file.

```
set-node Node1
set-cluster 1

link link1 udp - udp 50000 - 192.1.3.1 -
link link2 udp - udp 50001 - 192.1.4.1 -

#set address of each link for all peer nodes in the cluster
#format: set-addr node-id link tag-name address
set-addr        0 link1 192.1.1.1
set-addr        0 link2 192.1.2.1
set-addr        2 link1 192.1.5.2
set-addr        2 link2 192.1.6.2
```

```
set-addr        3 link1 192.1.7.3
set-addr        3 link2 192.1.8.3

#disable LLT broadcasts
set-bcasthb     0
set-arp         0
```

The /etc/llttab file on Node 0 resembles:

```
set-node Node0
set-cluster 1

link link1 udp - udp 50000 - 192.1.1.1 -
link link2 udp - udp 50001 - 192.1.2.1 -

#set address of each link for all peer nodes in the cluster
#format: set-addr node-id link tag-name address
set-addr        1 link1 192.1.3.1
set-addr        1 link2 192.1.4.1
set-addr        2 link1 192.1.5.2
set-addr        2 link2 192.1.6.2
set-addr        3 link1 192.1.7.3
set-addr        3 link2 192.1.8.3

#disable LLT broadcasts
set-bcasthb     0
set-arp         0
```

Using the UDP layer of IPv6 for LLT

Storage Foundation Cluster File System High Availability 7.4.3 provides the option of using LLT over the UDP (User Datagram Protocol) layer for clusters using wide-area networks and routers. UDP makes LLT packets routable and thus able to span longer distances more economically.

When to use LLT over UDP

Use LLT over UDP in the following situations:

- LLT must be used over WANs
- When hardware, such as blade servers, do not support LLT over Ethernet

Manually configuring LLT over UDP using IPv6

The following checklist is to configure LLT over UDP:

- For UDP6, the multicast address is set to "-".
- Make sure that each NIC has an IPv6 address that is configured before configuring LLT.
- Make sure the IPv6 addresses in the /etc/llttab files are consistent with the IPv6 addresses of the network interfaces.
- Make sure that each link has a unique not well-known UDP port.
  See "Selecting UDP ports" on page 367.
- For the links that cross an IP router, disable multicast features and specify the IPv6 address of each link manually in the /etc/llttab file.
  See "Sample configuration: links crossing IP routers" on page 369.

The link command in the /etc/llttab file

Review the link command information in this section for the /etc/llttab file. See the following information for sample configurations:

- See "Sample configuration: direct-attached links" on page 368.
- See "Sample configuration: links crossing IP routers" on page 369.

Note that some of the fields in Table F-3 differ from the command for standard LLT links.

Table F-3 describes the fields of the link command that are shown in the /etc/llttab file examples.

Table F-3 Field description for link command in /etc/llttab

Field	Description
tag-name	A unique string that is used as a tag by LLT; for example link1, link2,....
device	The device name of the UDP protocol; for example udp6.
node-range	Nodes using the link. "-" indicates all cluster nodes are to be configured for this link.
link-type	Type of link; must be "udp6" for LLT over UDP.
udp-port	Unique UDP port in the range of 49152-65535 for the link. See "Selecting UDP ports" on page 367.

Table F-3 Field description for link command in /etc/lltab *(continued)*

Field	Description
MTU	"-" is the default, which has a value of 8192. The value may be increased or decreased depending on the configuration. Use the `lltstat -l` command to display the current value.
IPv6 address	IPv6 address of the link on the local node.
mcast-address	"-" is the default for clusters spanning routers.

The set-addr command in the /etc/lltab file

The `set-addr` command in the /etc/lltab file is required when the multicast feature of LLT is disabled, such as when LLT must cross IP routers.

See "Sample configuration: links crossing IP routers" on page 369.

Table F-4 describes the fields of the set-addr command.

Table F-4 Field description for set-addr command in /etc/lltab

Field	Description
node-id	The ID of the peer node; for example, 0.
link tag-name	The string that LLT uses to identify the link; for example link1, link2,....
address	IPv6 address assigned to the link for the peer node.

Selecting UDP ports

When you select a UDP port, select an available 16-bit integer from the range that follows:

- Use available ports in the private range 49152 to 65535
- Do not use the following ports:
  - Ports from the range of well-known ports, 0 to 1023
  - Ports from the range of registered ports, 1024 to 49151

To check which ports are defined as defaults for a node, examine the file /etc/services. You should also use the `netstat` command to list the UDP ports currently in use. For example:

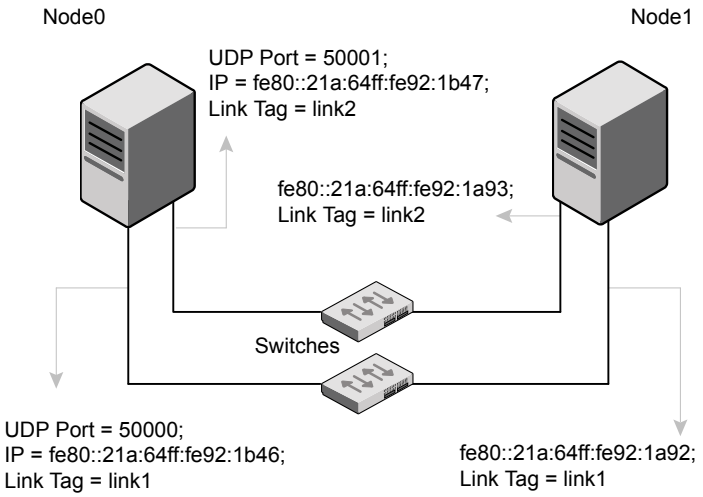
```
# netstat -au | more
Active Internet connections (servers and established)
Proto Recv-Q Send-Q Local Address Foreign Address         State
udp        0      0 *:32768                 *:*
udp        0      0 *:956                   *:*
udp        0      0 *:tftp                  *:*
udp        0      0 *:sunrpc                *:*
udp        0      0 *:ipp                   *:*
```

Look in the UDP section of the output; the UDP ports that are listed under Local Address are already in use. If a port is listed in the /etc/services file, its associated name is displayed rather than the port number in the output.

Sample configuration: direct-attached links

Figure F-3 depicts a typical configuration of direct-attached links employing LLT over UDP.

Figure F-3 A typical configuration of direct-attached links that use LLT over UDP



The configuration that the /etc/lltab file for Node 0 represents has directly attached crossover links. It might also have the links that are connected through a hub or switch. These links do not cross routers.

LLT uses IPv6 multicast requests for peer node address discovery. So the addresses of peer nodes do not need to be specified in the /etc/lltab file using the `set-addr`

command. Use the `ifconfig -a` command to verify that the IPv6 address is set correctly.

```
set-node Node0
set-cluster 1
#configure Links
#link tag-name device node-range link-type udp port MTU \
IP-address mcast-address
link link1 udp6 - udp6 50000 - fe80::21a:64ff:fe92:1b46 -
link link1 udp6 - udp6 50001 - fe80::21a:64ff:fe92:1b47 -
```

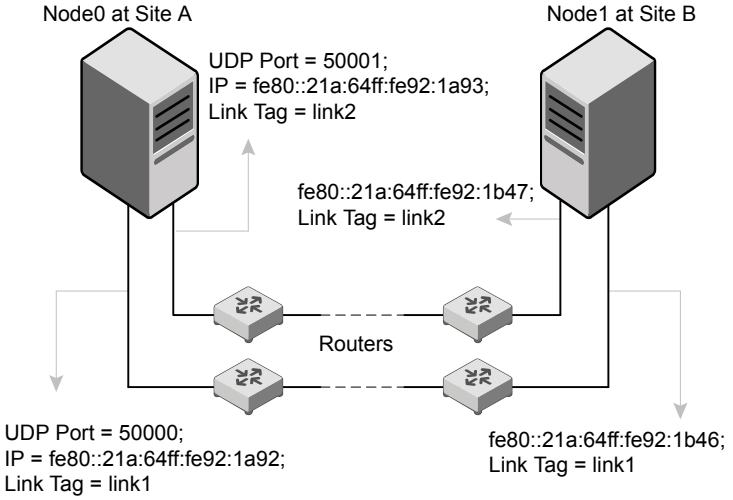
The file for Node 1 resembles:

```
set-node Node1
set-cluster 1
#configure Links
#link tag-name device node-range link-type udp port MTU \
IP-address mcast-address
link link1 udp6 - udp6 50000 - fe80::21a:64ff:fe92:1a92 -
link link1 udp6 - udp6 50001 - fe80::21a:64ff:fe92:1a93 -
```

Sample configuration: links crossing IP routers

Figure F-4 depicts a typical configuration of links crossing an IP router employing LLT over UDP. The illustration shows two nodes of a four-node cluster.

Figure F-4 A typical configuration of links crossing an IP router



The configuration that the following `/etc/llttab` file represents for Node 1 has links crossing IP routers. Notice that IPv6 addresses are shown for each link on each peer node. In this configuration multicasts are disabled.

```
set-node Node1
set-cluster 1

link link1 udp6 - udp6 50000 - fe80::21a:64ff:fe92:1a92 -
link link1 udp6 - udp6 50001 - fe80::21a:64ff:fe92:1a93 -

#set address of each link for all peer nodes in the cluster
#format: set-addr node-id link tag-name address
set-addr 0 link1 fe80::21a:64ff:fe92:1b46
set-addr 0 link2 fe80::21a:64ff:fe92:1b47
set-addr 2 link1 fe80::21a:64ff:fe92:1d70
set-addr 2 link2 fe80::21a:64ff:fe92:1d71
set-addr 3 link1 fe80::209:6bff:fe1b:1c94
set-addr 3 link2 fe80::209:6bff:fe1b:1c95

#disable LLT multicasts
set-bcasthb     0
set-arp         0
```

The /etc/llttab file on Node 0 resembles:

```
set-node Node0
set-cluster 1

link link1 udp6 - udp6 50000 - fe80::21a:64ff:fe92:1b46 -
link link2 udp6 - udp6 50001 - fe80::21a:64ff:fe92:1b47 -

#set address of each link for all peer nodes in the cluster
#format: set-addr node-id link tag-name address
set-addr 1 link1 fe80::21a:64ff:fe92:1a92
set-addr 1 link2 fe80::21a:64ff:fe92:1a93
set-addr 2 link1 fe80::21a:64ff:fe92:1d70
set-addr 2 link2 fe80::21a:64ff:fe92:1d71
set-addr 3 link1 fe80::209:6bff:fe1b:1c94
set-addr 3 link2 fe80::209:6bff:fe1b:1c95

#disable LLT multicasts
set-bcasthb     0
set-arp         0
```

About configuring LLT over UDP multiport

LLT uses UDP sockets for communication among the cluster nodes and creates one UDP socket per LLT link. In a Flexible Storage Sharing (FSS) environment, data can be read from and written to remote disks. In such a case, one socket per LLT link may not be enough for large read-write operations. More sockets are needed to achieve parallelism and throughput to meet the needs of the high data-generating applications.

Configuring LLT over UDP multiport enables you to create additional sockets per link. These sockets are reserved only for I/O shipping.

Note: For the multiport feature to work, LLT requires at least six consecutive network port numbers to be configured.

Manually configuring LLT over UDP multiport

Perform the following steps to configure LLT over UDP multiport.

Preparing for configuration

1. Set the maximum transmission unit (MTU) of all the high-priority links to the highest value (9000) to ensure optimal performance. You may also choose to use the default MTU (1500).

   Ensure that the network path MTU is the same as the MTU of the NIC.

   To change the MTU size of a NIC permanently:

   a. Edit the `/etc/sysconfig/network-scripts/ifcfg-eth0` file

      ```
      # vi /etc/sysconfig/network-scripts/ifcfg-eth0
      ```

   b. Add MTU settings at the end of the file:

      MTU=9000

   c. Save and close the file and restart networking

      ```
      # service network restart
      ```

2. Enable the LLT ports in firewall. See "Enabling LLT ports in firewall" on page 373.

Configuring LLT over UDP Multiport

1 Use the following shell script to increase the size of network buffers which consequently increases the send and receive TCP/IP buffers. This script also tunes the Rx /Tx queue size to max and enables the receive side scaling (RSS) functionality of the NIC.

```
#---------------------------------------
set -x

for card in `cat /etc/llttab | grep -v "lowpri" | grep -w "link" |
                awk '{print $2}'`;
do
    echo -e "Changeing buffers of $card"
    ethtool -G $card rx 4096
    ethtool -G $card rx-jumbo 4096
    ethtool -G $card tx 4096
    ethtool -N $card rx-flow-hash udp4 sdfn ethtool -N
                      $card rx-flow-hash tcp4 sdfn
    sysctl -w net.ipv4.conf.${card}.arp_ignore=1
done

sysctl -w net.core.rmem_max=1600000000
sysctl -w net.core.wmem_max=1600000000
sysctl -w net.core.netdev_max_backlog=250000
sysctl -w net.core.rmem_default=4194304
sysctl -w net.core.wmem_default=4194304
sysctl -w net.core.optmem_max=4194304
sysctl -w net.ipv4.udp_rmem_min=819200
sysctl -w net.ipv4.udp_wmem_min=819200
sysctl -w net.core.netdev_budget=600
set +x
#-------------------------------------------
```

2 Install Veritas InfoScale using the installer and select UDP as LLT protocol

```
# ./installer
```

The installer automatically enables the UDP Multiport feature and creates four additional sockets for each LLT link.

3 Verify that UDP multiport links are enabled

```
# lltstat -nvvr configured
```

Enabling LLT ports in firewall

You can use any firewall tool to enable the network ports.

While enabling ports make sure that:

- No other application is using the LLT consumable network ports (50000 to 50006).
- These ports are enabled in security groups in case of InfoScale installations in any cloud environment.

By default, LLT uses 50000 to 50001 port range for clustering and 50002 to 50006 for I/O shipping sockets.

Enabling ports using iptables

Ingress table:

```
iptables -A INPUT -p udp --dport 50000:50006 -j ACCEPT
```

Egress table:

```
iptables -A OUTPUT -p udp --sport 50000:50006 -j ACCEPT
```

Enable ports in the etc/llttab file

```
link eth1 udp - udp 50000 - 192.168.10.1 -
link eth2 udp - udp 50001 - 192.168.11.1 -
```

You can also use the following tunables to choose the number of sockets per link. By default 4 sockets are created for each link

Tunable	Description
set-udpports	Changes the port range to be used for I/O shipping if you do not want to use port range 50002 and onwards. Usage: set-udpports <initial_port_number> Example: set-udpports 60000 In this case, LLT uses the port 50000 and 50001 for clustering and 60000 and the subsequent port numbers for I/O shipping.
set-udpthreads	Specifies how many threads per socket needs to be created. Usage: set-udpthreads <number of threads per socket> Example: set-udpthreads 3

Tunable	Description
set-udpsockets	Specifies how many sockets per link needs to be created. Usage: set-udpsockets <number of sockets per link> Example: set-udpsockets 6

Disabling the UDP multiport feature

The UDP multiport feature is enabled by default. You can manually disable it by setting the value of the **LLT_UDP_MULTIPORT** variable to **zero**. This variable is defined in the `/etc/sysconfig/llt` file.

That is, in the `/etc/sysconfig/llt` file, set **LLT_UDP_MULTIPORT= 0**

Appendix G

Using LLT over RDMA

This appendix includes the following topics:

- Using LLT over RDMA
- About RDMA over RoCE or InfiniBand networks in a clustering environment
- How LLT supports RDMA capability for faster interconnects between applications
- Using LLT over RDMA: supported use cases
- Configuring LLT over RDMA
- Troubleshooting LLT over RDMA

Using LLT over RDMA

This section describes how LLT works with RDMA, lists the hardware requirements for RDMA, and the procedure to configure LLT over RDMA.

About RDMA over RoCE or InfiniBand networks in a clustering environment

Remote direct memory access (RDMA) is a direct memory access capability that allows server to server data movement directly between application memories with minimal CPU involvement. Data transfer using RDMA needs RDMA-enabled network cards and switches. Networks designed with RDMA over Converged Ethernet (RoCE) and InfiniBand architecture support RDMA capability. RDMA provides fast interconnect between user-space applications or file systems between nodes over these networks. In a clustering environment, RDMA capability allows applications on separate nodes to transfer data at a faster rate with low latency and less CPU usage.

See “How LLT supports RDMA capability for faster interconnects between applications” on page 376.

How LLT supports RDMA capability for faster interconnects between applications

LLT and GAB support fast interconnect between applications using RDMA technology over InfiniBand and Ethernet media (RoCE). To leverage the RDMA capabilities of the hardware and also support the existing LLT functionalities, LLT maintains two channels (RDMA and non-RDMA) for each of the configured RDMA links. Both RDMA and non-RDMA channels are capable of transferring data between the nodes and LLT provides separate APIs to their clients, such as, CFS, CVM, to use these channels. The RDMA channel provides faster data transfer by leveraging the RDMA capabilities of the hardware. The RDMA channel is mainly used for data-transfer when the client is capable to use this channel. The non-RDMA channel is created over the UDP layer and LLT uses this channel mainly for sending and receiving heartbeats. Based on the health of the non-RDMA channel, GAB decides cluster membership for the cluster. The connection management of the RDMA channel is separate from the non-RDMA channel, but the connect and disconnect operations for the RDMA channel are triggered based on the status of the non-RDMA channel

If the non-RDMA channel is up but due to some issues in RDMA layer the RDMA channel is down, in such cases the data-transfer happens over the non-RDMA channel with a lesser performance until the RDMA channel is fixed. The system logs displays the message when the RDMA channel is up or down.

LLT uses the Open Fabrics Enterprise Distribution (OFED) layer and the drivers installed by the operating system to communicate with the hardware. LLT over RDMA allows applications running on one node to directly access the memory of an application running on another node that are connected over an RDMA-enabled network. In contrast, on nodes connected over a non-RDMA network, applications cannot directly read or write to an application running on another node. LLT clients such as, CFS and CVM, have to create intermediate copies of data before completing the read or write operation on the application, which increases the latency period and affects performance in some cases.

LLT over an RDMA network enables applications to read or write to applications on another node over the network without the need to create intermediate copies. This leads to low latency, higher throughput, and minimized CPU host usage thus improving application performance. Cluster volume manager and Cluster File Systems, which are clients of LLT and GAB, can use LLT over RDMA capability for specific use cases.

See “Using LLT over RDMA: supported use cases ” on page 377.

Using LLT over RDMA: supported use cases

You can configure the LLT over RDMA capability for the following use cases:

- Storage Foundation Smart IO feature on flash storage devices: The Smart IO feature provides file system caching on flash devices for increased application performance by reducing IO bottlenecks. It also reduces IO loads on storage controllers as the Smart IO feature meets most of the application IO needs. As the IO requirements from the storage array are much lesser, you require lesser number of servers to maintain the same IO throughput.
- Storage Foundation IO shipping feature: The IO shipping feature in Storage Foundation Cluster File System HA (SFCFSHA) provides the ability to ship IO data between applications on peer nodes without service interruption even if the IO path on one of the nodes in the cluster goes down.
- Storage Foundation Flexible Storage Sharing feature : The Flexible Storage Sharing feature in cluster volume manager allows network shared storage to co-exist with physically shared storage. It provides server administrators the ability to provision clusters for Storage Foundation Cluster File System HA (SFCFSHA) and Storage Foundation for Oracle RAC (SFRAC) or SFCFSHA applications without requiring physical shared storage.

Both Cluster File System (CFS) and Cluster Volume Manager (CVM) are clients of LLT and GAB. These clients use LLT as the transport protocol for data transfer between applications on nodes. Using LLT data transfer over an RDMA network boosts performance of file system data transfer and IO transfer between nodes.

To enable RDMA capability for faster application data transfer between nodes, you must install RDMA-capable network interface cards, RDMA-supported network switches, configure the operating system for RDMA, and configure LLT.

Ensure that you select RDMA-supported hardware and configure LLT to use RDMA functionality.

See “Choosing supported hardware for LLT over RDMA” on page 378.

See “Configuring LLT over RDMA” on page 377.

Configuring LLT over RDMA

This section describes the required hardware and configuration needed for LLT to support RDMA capability. The high-level steps to configure LLT over RDMA are as follows:

Table G-1 lists the high-level steps to configure LLT over RDMA.

Step	Action	Description
Choose supported hardware	Choose RDMA capable network interface cards (NICs), network switches, and cables.	See "Choosing supported hardware for LLT over RDMA" on page 378.
Check the supported operating system	Verify that it is a supported Linux operating system.	All supported Linux flavors
Install RDMA, InfiniBand or Ethernet drivers and utilities	Install the packages to access the RDMA, InfiniBand or Ethernet drivers and utilities.	See "Installing RDMA, InfiniBand or Ethernet drivers and utilities" on page 379.
Configure RDMA over an Ethernet network	Load RDMA and Ethernet drivers.	See "Configuring RDMA over an Ethernet network" on page 380.
Configuring RDMA over an InfiniBand network	Load RDMA and InfiniBand drivers.	See "Configuring RDMA over an InfiniBand network" on page 382.
Tune system performance	Tune CPU frequency and boot parameters for systems.	See "Tuning system performance" on page 386.
Configure LLT manually	Configure LLT to use RDMA capability. Alternatively, you can use the installer to automatically configure LLT to use RDMA.	See "Manually configuring LLT over RDMA" on page 388.
Verify LLT configuration	Run LLT commands to test the LLT over RDMA configuration.	See "Verifying LLT configuration" on page 392.

Choosing supported hardware for LLT over RDMA

To configure LLT over RDMA you need to use the hardware that is RDMA enabled.

Table G-2

Hardware	Supported types	Reference
Network card	Mellanox-based Host Channel Adapters (HCAs) (VPI, ConnectX, ConnectX-2 and 3)	For detailed installation information, refer to the hardware vendor documentation.

Table G-2 *(continued)*

Hardware	Supported types	Reference
Network switch	Mellanox, InfiniBand switches Ethernet switches must be Data Center Bridging (DCB) capable	For detailed installation information, refer to the hardware vendor documentation.
Cables	Copper and Optical Cables, InfiniBand cables	For detailed installation information, refer to the hardware vendor documentation.

Warning: When you install the Mellanox NIC for using RDMA capability, do not install Mellanox drivers that come with the hardware. LLT uses the Mellanox drivers that are installed by default with the Linux operating system. LLT might not be configurable if you install Mellanox drivers provided with the hardware.

Installing RDMA, InfiniBand or Ethernet drivers and utilities

Install the following RPMs to get access to the required RDMA, InfiniBand or Ethernet drivers and utilities. Note that the rpm version of the RPMs may differ for each of the supported Linux flavors.

Veritas does not support any external Mellanox OFED packages. The supported packages are listed in this section.

Veritas recommends that you use the Yellowdog Updater Modified (yum) package management utility to install RPMs on RHEL systems and use Zypper, a command line package manager, on SUSE systems.

Note: Install the OpenSM package only if you configure an InfiniBand network. All other packages are required with both InfiniBand and Ethernet networks.

Table G-3 lists the drivers and utilities required for RDMA, InfiniBand or Ethernet network.

Packages	RHEL	SUSE
Userland device drivers for RDMA operations	▪ `libmthca` ▪ `libmlx4` ▪ `rdma` ▪ `librdmacm-utils`	▪ `libmthca-rdmav2` ▪ `libmlx4-rdmav2` ▪ `ofed` ▪ `librdmacm`

Table G-3 lists the drivers and utilities required for RDMA, InfiniBand or Ethernet network. *(continued)*

Packages	RHEL	SUSE
OpenSM related package (InfiniBand only)	■ `opensm` ■ `opensm-libs` ■ `libibumad`	■ `opensm` ■ `libibumad3`
InfiniBand troubleshooting and performance tests	■ `Ibutils` ■ `infiniband-diags` ■ `Perftest`	■ `Ibutils` ■ `infiniband-diags`
libibverbs packages for userland InfiniBand operations	■ `libibverbs-devel` ■ `libibverbs-utils`	■ `libibverbs`

Configuring RDMA over an Ethernet network

Configure the RDMA and Ethernet drivers so that LLT can use the RDMA capable hardware.

See "Enable RDMA over Converged Ethernet (RoCE)" on page 380.

See "Configuring RDMA and Ethernet drivers" on page 381.

See "Configuring IP addresses over Ethernet Interfaces" on page 381.

Enable RDMA over Converged Ethernet (RoCE)

The following steps are applicable only on a system installed with RHEL Linux or supported RHEL-compatible distributions. On SUSE Linux, the RDMA is enabled by default.

1. Make sure that the SFHA stack is stopped and the LLT and GAB modules are not loaded.

2. Skip this step if you are on a RHEL 7 or supported RHEL-compatible distributions. Alternatively, create or modify the `/etc/modprobe.d/mlx4.conf` configuration file and add the value `options mlx4_core hpn=1` to the file. This enables RDMA over Converged Ethernet (RoCE) in Mellanox drivers (installed by default with the operating system).

3 Verify whether the Mellanox drivers are loaded.

```
# lsmod | grep mlx4_en

# lsmod | grep mlx4_core
```

4 Unload the Mellanox drivers if the drivers are loaded.

```
# rmmod mlx4_ib

# rmmod mlx4_en

# rmmod mlx4_core
```

Configuring RDMA and Ethernet drivers

Load the Mellanox drivers that are installed by default with the operating system and enable the RDMA service.

1 (RHEL and supported RHEL-compatible distributions only) Load the Mellanox drivers.

```
# modprobe mlx4_core

# modprobe mlx4_ib

# modprobe mlx4_en
```

2 Enable RDMA service on the Linux operating system.

On RHEL Linux: `# chkconfig --level 235 rdma on`

On SUSE Linux: `# chkconfig --level 235 openibd on`

Configuring IP addresses over Ethernet Interfaces

Perform the following steps to configure IP addresses over the network interfaces which you plan to configure under LLT. These interfaces must not be aggregated interfaces.

1. Configure IP addresses using Linux `ifconfig` command. Make sure that the IP address for each link must be from a different subnet.

   Typical private IP addresses that you can use are:

   Node0:

   link0: 192.168.1.1

   link1: 192.168.2.1

   Node1:

   link0: 192.168.1.2

   link1: 192.168.2.2

2. Run IP ping test between nodes to ensure that there is network level connectivity between nodes.
3. Configure IP addresses to start automatically after the system restarts or reboots by creating a new configuration file or by modifying the existing file.
   - On RHEL or supported RHEL-compatible distributions, modify the `/etc/sysconfig/network-scripts/` directory by modifying the `ifcfg-eth` (Ethernet) configuration file.
   - On SUSE, modify the `/etc/sysconfig/network/` by modifying the `ifcfg-eth` (Ethernet) configuration file.
     For example, for an Ethernet interface eth0, create the `ifcfg-eth0` file with values for the following parameters.

```
DEVICE=eth0
BOOTPROTO=static
IPADDR=192.168.27.1
NETMASK=255.255.255.0
NETWORK=192.168.27.0
BROADCAST=192.168.27.255
NM_CONTROLLED=no # This line ensures IPs are plumbed correctly after bootup
ONBOOT=yes
STARTMODE='auto' # This line is only for SUSE
```

Configuring RDMA over an InfiniBand network

While configuring RDMA over an InfiniBand network, you need to configure the InfiniBand drivers, configure the OpenSM service, and configure IP addresses for the InfiniBand interfaces.

See “Configuring RDMA and InfiniBand drivers” on page 383.

See “ Configuring the OpenSM service” on page 384.

See “Configuring IP addresses over InfiniBand Interfaces” on page 385.

Configuring RDMA and InfiniBand drivers

Configure the RDMA and InfiniBand drivers so that LLT can use the RDMA capable hardware.

1. Ensure that the following RDMA and InfiniBand drivers are loaded. Use the `lsmod` command to verify whether a driver is loaded.

   The InfiniBand interfaces are not visible by default until you load the InfiniBand drivers. This procedure is only required for initial configuration.

   ```
   # modprobe rdma_cm

   # modprobe rdma_ucm

   # modprobe mlx4_en

   # modprobe mlx4_ib

   # modprobe ib_mthca

   # modprobe ib_ipoib

   # modprobe ib_umad
   ```

2. Load the drivers at boot time by appending the configuration file on the operating system.

   On RHEL and SUSE Linux, append the `/etc/rdma/rdma.conf` and `/etc/infiniband/openib.conf` files respectively with the following values:

   ```
   ONBOOT=yes

   RDMA_UCM_LOAD=yes

   MTHCA_LOAD=yes

   IPOIB_LOAD=yes

   SDP_LOAD=yes

   MLX4_LOAD=yes

   MLX4_EN_LOAD=yes
   ```

3. Enable RDMA service on the Linux operating system.

   On RHEL Linux:

   ```
   # chkconfig --level 235 rdma on
   ```

   On SUSE Linux:

   ```
   # chkconfig --level 235 openibd on
   ```

Configuring the OpenSM service

OpenSM is an InfiniBand compliant Subnet Manager and Subnet Administrator, which is required to initialize the InfiniBand hardware. In the default mode, OpenSM

scans the IB fabric, initializes the hardware, and checks the fabric occasionally for changes.

For InfiniBand network, make sure to configure subnet manager if you have not already configured the service.

1. Modify the OpenSM configuration file if you plan to configure multiple links under LLT.

   On RHEL, update the `/etc/sysconfig/opensm` file.

2. Start OpenSM.

   For systemd environments with supported Linux distributions, run:

   ```
   # systemctl start opensm
   ```

   For other supported Linux distributions, run:

   ```
   # /etc/init.d/opensm start
   ```

3. Enable Linux service to start OpenSM automatically after restart.

   For systemd environments with supported RHEL distributions, run:

   ```
   # systemctl enable opensm
   ```

   For systemd environments with supported SLES distributions, run:

   ```
   # systemctl enable opensmd
   ```

Configuring IP addresses over InfiniBand Interfaces

Perform the following steps to configure IP addresses over the network interfaces which you plan to configure under LLT. These interfaces must not be aggregated interfaces.

1. Configure IP addresses using Linux `ifconfig` command. Make sure that the IP address for each link must be from a different subnet.

   Typical private IP addresses that you can use are: **192.168.12.1, 192.168.12.2, 192.168.12.3** and so on.

2. Run the InfiniBand ping test between nodes to ensure that there is InfiniBand level connectivity between nodes.

   - On one node, start the ibping server.

     ```
     # ibping -S
     ```

   - On the node, get the GUID of an InfiniBand interface that you need to ping from another node.

     ```
     # ibstat
     ```

```
CA 'mlx4_0'
Number of ports: 2
 --
Port 1:
State: Active
---
Port GUID: 0x0002c90300a02af1
Link layer: InfiniBand
```

- Ping the peer node by using its GUID.

  `# ibping -G 0x0002c90300a02af1`

  Where, *0x0002c90300a02af1* is the GUID of the server.

3 Configure IP addresses automatically after restart by creating a new configuration file or by modifying the existing file.

- On RHEL, modify the `/etc/sysconfig/network-scripts/` directory by modifying the `ifcfg-ibX` (InfiniBand) configuration file.
- On SUSE, modify the `/etc/sysconfig/network/` by modifying the `ifcfg-ibX` (InfiniBand) configuration file.

```
For example, for an Infiniband interface ib0,
create ifcfg-ib0 file with values for the following parameters.

DEVICE=ib0
BOOTPROTO=static
IPADDR=192.168.27.1
NETMASK=255.255.255.0
NETWORK=192.168.27.0
BROADCAST=192.168.27.255
NM_CONTROLLED=no # This line ensures IPs are plumbed correctly
after bootup and the Network manager does not interfere
with the interfaces
ONBOOT=yes
STARTMODE='auto' # This line is only for SUSE
```

Tuning system performance

Run IP ping test to ensure that systems are tuned for the best performance. The latency should be less than 30us, if not then your system may need tuning. However, the latency may vary based on your system configuration.

To tune your system, perform the following steps. For additional tuning, follow the performance tuning guide from Mellanox.

Performance Tuning Guidelines for Mellanox Network Adapters

Tuning the CPU frequency

To tune the CPU frequency of a system, perform the following steps:

1. Verify whether the CPU frequency is already tuned.

```
# cat /proc/cpuinfo | grep Hz


model name      : Intel(R) Xeon(R) CPU E5-2643 0 @ 3.30GHz
cpu MHz         : 3300.179
```

2. If the CPU frquency displayed by the `cpu MHz` and `model name` attribute is the same, then the CPU frequency is already tuned. You can skip the next steps.

   If the CPU frequency displayed by the `cpu Mhz` and `model name` attribute is not the same, then follow the next steps to tune the frequency.

3. Go to system console and restart the system.
4. Press F11 to enter into BIOS settings.
5. Go to BIOS menu > Launch System setup > BIOS settings > System Profile Settings > System Profile > Max performance.

   The menu options might vary with system type.

Tuning the boot parameter settings

To tune the boot parameter settings, perform the following steps.

1. In the `/boot/grub/grub.conf` file or any other boot loader configuration file, ensure that the value of the `intel_iommu` is set to **off**.
2. Append the `/boot/grub/grub.conf` file or any other boot loader configuration file with the following parameters if they are not listed in the configuration file.

   `intel_idle.max_cstate=0 processor.max_cstate=1`
3. Restart the system.

On RHEL 7 and supported RHEL-compatible distributions:

1. In the `/etc/default/grub` file, append the GRUB_CMDLINE_LINUX variable with intel_idle.max_cstate=0 processor.max_cstate=1
2. After `/etc/default/grub` is modified, run the following command:

   ```
   grub2-mkconfig -o /boot/grub2/grub.cfg
   ```

3. Restart the system.

Manually configuring LLT over RDMA

You can automatically configure LLT to use RDMA using the installer. To manually configure LLT over RDMA follow the steps that are given in this section.

The following checklist is to configure LLT over RDMA:

- Make sure that the LLT private links are on separate subnets. Set the broadcast address in `/etc/lltab` explicitly depending on the subnet for each link.
  See "Broadcast address in the `/etc/lltab` file" on page 388.
- Make sure that each RDMA enabled NIC (RNIC) over an InfiniBand or Ethernet network has an IP address that is configured before configuring LLT.
- Make sure that the IP addresses in the `/etc/lltab` files are consistent with the IP addresses of the network interfaces (InfiniBand or Ethernet network interfaces).
- Make sure that each link has a unique and a private IP range for the UDP port.
  See "Selecting UDP ports" on page 390.
- See the sample comfiguration for direct-attached (non-routed) links.
  See "Sample configuration: direct-attached links" on page 391.

Broadcast address in the `/etc/lltab` file

The broadcast address is set explicitly for each link in the following example.

- Display the content of the `/etc/lltab` file on the first node sys1:

  ```
  sys1 # cat /etc/lltab

  set-node sys1
  set-cluster 1
  link link1 udp - rdma  50000  -  192.168.9.1 192.168.9.255
  link link2 udp - rdma  50001  -  192.168.10.1 192.168.10.255
  ```

Verify the subnet mask using the ifconfig command to ensure that the two links are on separate subnets.

- Display the content of the `/etc/llttab` file on the second node sys2:

```
sys2 # cat /etc/llttab

set-node sys2
set-cluster 1
link link1 udp - rdma  50000  -  192.168.9.2 192.168.9.255
link link2 udp - rdma  50001  -  192.168.10.2 192.168.10.255
```

Verify the subnet mask using the ifconfig command to ensure that the two links are on separate subnets.

The link command in the `/etc/llttab` file

Review the link command information in this section for the `/etc/llttab` file. See the following information for sample configurations:

-

Table G-4 describes the fields of the link command that are shown in the `/etc/llttab` file examples. Note that some of the fields differ from the command for standard LLT links.

Table G-4 Field description for link command in `/etc/llttab`

Field	Description
tag-name	A unique string that is used as a tag by LLT; for example link1, link2,....
device	The device path of the UDP protocol; for example udp. A place holder string. Linux does not have devices for protocols. So this field is ignored.
node-range	Nodes using the link. "-" indicates all cluster nodes are to be configured for this link.
link-type	Type of link; must be "rdma" for LLT over RDMA.
udp-port	Unique UDP port in the range of 49152-65535 for the link. See “Selecting UDP ports” on page 361.
MTU	"-" is the default, which has a value of 8192. Do not change this default value for the RDMA links.

Table G-4 Field description for link command in `/etc/llttab` *(continued)*

Field	Description
IP address	IP address of the link on the local node.
bcast-address	Specify the value of the subnet broadcast address.

Selecting UDP ports

When you select a UDP port, select an available 16-bit integer from the range that follows:

- Use available ports in the private range 49152 to 65535
- Do not use the following ports:
  - Ports from the range of well-known ports, 0 to 1023
  - Ports from the range of registered ports, 1024 to 49151

To check which ports are defined as defaults for a node, examine the file /etc/services. You should also use the `netstat` command to list the UDP ports currently in use. For example:

```
# netstat -au | more
Active Internet connections (servers and established)
Proto Recv-Q Send-Q Local Address Foreign Address         State
udp        0      0 *:32768                 *:*
udp        0      0 *:956                   *:*
udp        0      0 *:tftp                  *:*
udp        0      0 *:sunrpc                *:*
udp        0      0 *:ipp                   *:*
```

Look in the UDP section of the output; the UDP ports that are listed under Local Address are already in use. If a port is listed in the /etc/services file, its associated name is displayed rather than the port number in the output.

Configuring the netmask for LLT

For nodes on different subnets, set the netmask so that the nodes can access the subnets in use. Run the following command and answer the prompt to set the netmask:

```
# ifconfig interface_name netmask netmask
```

For example:

- For the first network interface on the node sys1:

```
IP address=192.168.9.1, Broadcast address=192.168.9.255,
Netmask=255.255.255.0
```

For the first network interface on the node sys2:

```
IP address=192.168.9.2, Broadcast address=192.168.9.255,
Netmask=255.255.255.0
```

- For the second network interface on the node sys1:

```
IP address=192.168.10.1, Broadcast address=192.168.10.255,
Netmask=255.255.255.0
```

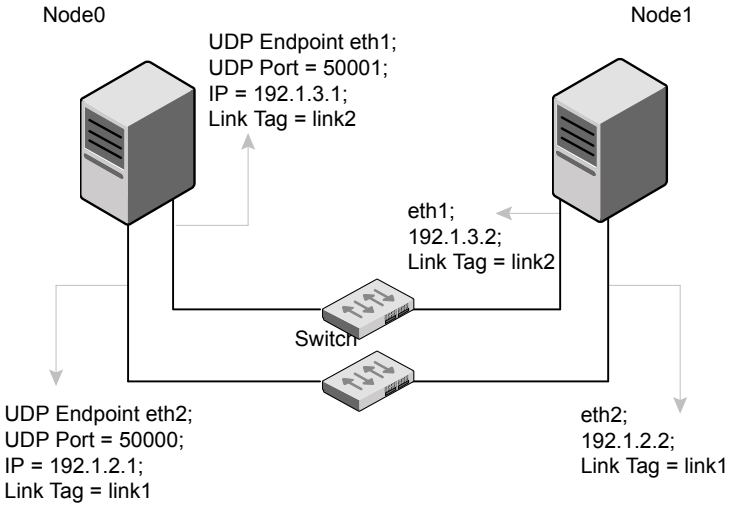
For the second network interface on the node sys2:

```
IP address=192.168.10.2, Broadcast address=192.168.10.255,
Netmask=255.255.255.0
```

Sample configuration: direct-attached links

Figure G-1 depicts a typical configuration of direct-attached links employing LLT over UDP.

Figure G-1 A typical configuration of direct-attached links that uses LLT over RDMA



The configuration that the /etc/lltab file for Node 0 represents has directly attached crossover links. It might also have the links that are connected through a hub or switch. These links do not cross routers.

LLT sends broadcasts to peer nodes to discover their addresses. So the addresses of peer nodes do not need to be specified in the /etc/lltab file using the `set-addr` command. For direct attached links, you do need to set the broadcast address of the links in the /etc/lltab file. Verify that the IP addresses and broadcast addresses are set correctly by using the `ifconfig -a` command.

```
set-node Node0
set-cluster 1
#configure Links
#link tag-name device node-range  link-type udp port MTU IP-addressbast-address
link link1 udp - rdma 50000 - 192.1.2.1 192.1.2.255
link link2 udp - rdma 50001 - 192.1.3.1 192.1.3.255
```

The file for Node 1 resembles:

```
set-node Node1
set-cluster 1
#configure Links
#link tag-name device node-range link-type udp port MTU IP-address bast-address
link link1 udp - rdma 50000 - 192.1.2.2 192.1.2.255
link link2 udp - rdma 50001 - 192.1.3.2 192.1.3.255
```

LLT over RDMA sample /etc/lltab

The following is a sample of LLT over RDMA in the etc/lltab file.

```
set-node sys1
set-cluster clus1
link eth1 udp - rdma 50000 - 192.168.10.1 - 192.168.10.255
link eth2 udp - rdma 50001 - 192.168.11.1 - 192.168.11.255
link-lowpri eth0 udp - rdma 50004 - 10.200.58.205 - 10.200.58.255
```

Verifying LLT configuration

After starting LLT, GAB and other component, run the following commands to verify the LLT configuration.

1 Run the `lltstat -l` command to view the RDMA link configuration. View the link-type configured to rdma for the RDMA links.

```
# lltstat -l


LLT link information:
link 0  link0 on rdma hipri
mtu 8192, sap 0x2345, broadcast 192.168.27.255, addrlen 4
txpkts 171  txbytes 10492
rxpkts 105  rxbytes 5124
latehb 0  badcksum 0  errors 0
```

2 Run the `lltstat -nvv -r` command to view the RDMA and non-RDMA channel connection state.

LLT internally configures each RDMA link in two modes (RDMA and non-RDMA) to allow both RDMA and non-RDMA traffic to use the same link. The GAB membership-related traffic goes over the non-RDMA channel while node to node data-transfer goes over high-speed RDMA channel for better performance.

```
# lltstat -rnvv active


LLT node information:
Node          State Link  Status TxRDMA RxRDMA Address
* 0 thorpc365 OPEN  link0   UP    UP     UP    192.168.27.1
                    link1   UP    UP     UP    192.168.28.1
                    link2   UP    N/A    N/A   00:15:17:97:91:2E

1 thorpc366   OPEN  link0   UP    UP     UP    192.168.27.2
                    link1   UP    UP     UP    192.168.28.2
                    link2   UP    N/A    N/A   00:15:17:97:A1:7C
```

Troubleshooting LLT over RDMA

This section lists the issues and their resolutions.

IP addresses associated to the RDMA NICs do not automatically plumb on node restart

If IP addresses do not plumb automatically, you might experience LLT failure.

Resolution: Assign unique IP addresses to RNICs and assign the same in the configuration script. For example, on an ethernet network, the `ifcfg-eth` script must be modified with the unique IP address of the RNIC.

See "Configuring IP addresses over InfiniBand Interfaces" on page 385.

Ping test fails for the IP addresses configured over InfiniBand interfaces

Resolution: Check the physical configuration and configure OpenSM. If you configured multiple links, then make sure that you have configured OpenSM to monitor multiple links in the configuration file. On RHEL and supported RHEL-compatible distributions, configure the `/etc/sysconfig/opensm` file.

See " Configuring the OpenSM service" on page 384.

After a node restart, by default the Mellanox card with Virtual Protocol Interconnect (VPI) gets configured in InfiniBand mode

After restart, you might expect the Mellanox VPI RNIC to get configured in the Ethernet mode. By default, the card gets configured in the InfiniBand mode.

Resolution: Update the Mellanox configuration file. On RHEL and supported RHEL-compatible distributions, configure the `/etc/rdma/mlx4.conf` file.

The LLT module fails to start

For Linux distributions:

```
# systemctl start llt
```

When you try to start LLT, it may fail to start and you may see the following message:

```
Starting LLT:
LLT: loading module...
LLT:Error loading LLT dependency rdma_cm.
Make sure module rdma_cm is available on the system.
```

Description: Check the system log at /var/log/messages. If the log file lists the following error, the issue may be because the IPv6 module is not available on the system. In addition, the LLT module has indirect dependency on the IPv6 module.

```
ib_addr: Unknown symbol ipv6_dev_get_saddr
ib_addr: Unknown symbol ip6_route_output
ib_addr: Unknown symbol ipv6_chk_addr
```

Resolution: Load the IPv6 module. If you do not want to configure the IPv6 module on the node, then configure the IPv6 module to start in the disabled mode.

To start IPv6 in the disabled mode:

- In the `/etc/modprobe.d/` directory, create a file `ipv6.conf` and add the following line to the file

  ```
  options ipv6 disable=1
  ```

The LLT module starts up without any issues once the file loads the IPv6 module in the disabled mode.