# Storage Foundation for Oracle® RAC 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 Configuring SF Oracle RAC ..... 17

Chapter 1 Preparing to configure SF Oracle RAC ..... 18

About this document ..... 18
I/O fencing requirements ..... 18
Coordinator disk requirements for I/O fencing ..... 19
CP server requirements ..... 19
About planning to configure I/O fencing ..... 22
Typical SF Oracle RAC cluster configuration with disk-based I/O fencing ..... 24
Typical SF ORACLE RAC cluster configuration with server-based I/O fencing ..... 25
Recommended CP server configurations ..... 25
Preparing to configure server-based fencing for SF Oracle RAC ..... 27

Chapter 2 Configuring SF Oracle RAC using the script-based installer ..... 28

Configuring the SF Oracle RAC components using the script-based installer ..... 28
Configuring the SF Oracle RAC cluster ..... 30
Creation of SF Oracle RAC configuration files ..... 50
Stopping and starting SF Oracle RAC processes ..... 50
Setting up disk-based I/O fencing using installer ..... 51
Initializing disks as VxVM disks ..... 51
Identifying disks to use as coordinator disks ..... 52
Checking shared disks for I/O fencing ..... 52
Configuring disk-based I/O fencing using installer ..... 56
Refreshing keys or registrations on the existing coordination points for disk-based fencing using the installer ..... 58
Setting up server-based I/O fencing using installer ..... 60
Refreshing keys or registrations on the existing coordination points for server-based fencing using the installer ..... 65

Setting the order of existing coordination points for server-based fencing using the installer ..... 67
Configuring SFDB ..... 70

Chapter 3 Performing an automated SF Oracle RAC configuration ..... 72

Configuring SF Oracle RAC using response files ..... 72
Response file variables to configure SF Oracle RAC ..... 73
Sample response file for configuring SF Oracle RAC ..... 85
Configuring I/O fencing using response files ..... 86
Response file variables to configure disk-based I/O fencing ..... 86
Sample response file for configuring disk-based I/O fencing ..... 89
Configuring CP server using response files ..... 89
Response file variables to configure CP server ..... 90
Sample response file for configuring the CP server on SFHA cluster ..... 92
Response file variables to configure server-based I/O fencing ..... 93
Sample response file for configuring server-based I/O fencing ..... 95

Section 2 Post-installation and configuration tasks ..... 97

Chapter 4 Verifying the installation ..... 98

Performing a postcheck on a node ..... 98
Verifying SF Oracle RAC installation using VCS configuration file ..... 98
Verifying LLT, GAB, and cluster operation ..... 99
Verifying LLT ..... 99
Verifying GAB ..... 101
Verifying the cluster ..... 103
Verifying the cluster nodes ..... 104

Chapter 5 Performing additional post-installation and configuration tasks ..... 107

Setting the environment variables for SF Oracle RAC ..... 107
About enabling LDAP authentication for clusters that run in secure mode ..... 108
About configuring authentication for SFDB tools ..... 108
Configuring vxdbd for SFDB tools authentication ..... 109
Configuring Volume Replicator ..... 109

Running SORT Data Collector to collect configuration information ..... 110

Section 3 Upgrade of SF Oracle RAC ..... 111

Chapter 6 Planning to upgrade SF Oracle RAC ..... 112

About the upgrade ..... 112
Supported upgrade paths ..... 114
Creating backups ..... 118
Considerations for upgrading SF Oracle RAC to 8.0 on systems configured with an Oracle resource ..... 119
Using Install Bundles to simultaneously install or upgrade full releases (base, maintenance, rolling patch), and individual patches ..... 119

Chapter 7 Performing a full upgrade of SF Oracle RAC using the product installer ..... 122

About full upgrades ..... 122
Preparing to perform a full upgrade to SF Oracle RAC 8.0 ..... 123
Upgrading to SF Oracle RAC 8.0 ..... 125
Upgrading SF Oracle RAC using the Veritas installation program ..... 127
Upgrading SFDB ..... 129

Chapter 8 Performing an automated full upgrade of SF Oracle RAC using response files ..... 130

Upgrading SF Oracle RAC using a response file ..... 130
Response file variables to upgrade SF Oracle RAC ..... 131
Sample response file for upgrading SF Oracle RAC ..... 134

Chapter 9 Performing a phased upgrade of SF Oracle RAC ..... 136

About phased upgrade ..... 136
Performing a phased upgrade of SF Oracle RAC from version 6.2.1 and later release ..... 137
Step 1: Performing pre-upgrade tasks on the first half of the cluster ..... 138
Step 2: Upgrading the first half of the cluster ..... 141
Step 3: Performing pre-upgrade tasks on the second half of the cluster ..... 142

Step 4: Performing post-upgrade tasks on the first half of the cluster ..... 144
Step 5: Upgrading the second half of the cluster ..... 145
Step 6: Performing post-upgrade tasks on the second half of the cluster ..... 146

Chapter 10 Performing a rolling upgrade of SF Oracle RAC ..... 148

About rolling upgrade ..... 148
Preparing to perform a rolling upgrade to SF Oracle RAC 8.0 ..... 151
Performing a rolling upgrade of SF Oracle RAC using the product installer ..... 153
Performing a rolling upgrade of SF Oracle RAC from 7.4.2 or later to 8.0 ..... 158
Performing rolling upgrades in CVR environments ..... 160

Chapter 11 Upgrading Volume Replicator ..... 163

Upgrading Volume Replicator ..... 163
Upgrading VVR without disrupting replication ..... 163
Upgrading VVR sites for InfoScale 7.3.1 or earlier ..... 164
Upgrading VVR sites with InfoScale 7.4 or later ..... 165
Post-upgrade tasks for VVR sites ..... 166

Chapter 12 Performing post-upgrade tasks ..... 167

Relinking Oracle RAC libraries with the SF Oracle RAC libraries ..... 167
Re-joining the backup boot disk group into the current disk group ..... 169
Reverting to the backup boot disk group after an unsuccessful upgrade ..... 169
Setting or changing the product license level ..... 170
Resetting DAS disk names to include host name in FSS environments ..... 170
Upgrading disk layout versions ..... 171
CVM master node needs to assume the logowner role for VCS managed VVR resources ..... 172
Switching on Quotas ..... 173
Upgrading the disk group version ..... 173

Section 4 Installation of Oracle RAC ..... 175

Chapter 13 Before installing Oracle RAC ..... 176

Important preinstallation information for Oracle RAC ..... 176
About preparing to install Oracle RAC ..... 176
Preparing to install Oracle RAC using the SF Oracle RAC installer or manually ..... 177
Identifying the public virtual IP addresses for use by Oracle ..... 178
Setting the kernel parameters ..... 179
Verifying that RPMs and patches required by Oracle are installed ..... 179
Verifying the user `nobody` exists ..... 180
Launching the SF Oracle RAC installer ..... 180
Creating users and groups for Oracle RAC ..... 180
Creating storage for OCR and voting disk ..... 184
Configuring private IP addresses for Oracle RAC ..... 200
Verifying that multicast is functional on all private network interfaces ..... 204
Creating Oracle Clusterware/Grid Infrastructure and Oracle database home directories manually ..... 204
Setting up user equivalence ..... 210

Chapter 14 Installing Oracle RAC ..... 211

About installing Oracle RAC ..... 211
Installing the Oracle Clusterware/Grid Infrastructure software ..... 212
Installing Oracle Clusterware/Grid Infrastructure using the Oracle Universal Installer ..... 213
Configuring LLT links in the GPnP profile ..... 214
Installing the Oracle RAC database software ..... 215
Installing the Oracle RAC database using the SF Oracle RAC script-based installer ..... 216
Installing the Oracle RAC database using the Oracle Universal Installer ..... 219
Verifying the Oracle Clusterware/Grid Infrastructure and database installation ..... 219

Chapter 15 Performing an automated Oracle RAC installation ..... 221

About installing Oracle RAC using response files ..... 221
Before you install ..... 221

Installing Oracle RAC ..... 223
Response file variable definitions for Oracle RAC ..... 224
Sample response file for installing Oracle RAC ..... 243

Chapter 16 Performing Oracle RAC post-installation tasks ..... 246

Adding Oracle RAC patches or patchsets ..... 246
Configuring the CSSD resource ..... 247
Configuring the CSSD resource using the SF Oracle RAC script-based installer ..... 248
Configuring the CSSD resource manually ..... 250
Preventing automatic startup of Oracle Clusterware/Grid Infrastructure ..... 251
Relinking the SF Oracle RAC libraries with Oracle RAC ..... 252
Relinking the SF Oracle RAC libraries with Oracle RAC using the SF Oracle RAC script-based installer ..... 252
Relinking SF Oracle RAC libraries with Oracle RAC manually ..... 253
Updating the CSS misscount setting ..... 255
Creating the Oracle RAC database ..... 256
Configuring VCS service groups for Oracle RAC ..... 256
Supported types of database management ..... 257
Sample service group configurations ..... 257
Configuring VCS service groups manually for traditional Oracle databases ..... 262
Configuring VCS service groups manually for container Oracle databases ..... 266
Managing database restart after failure ..... 271
Location of VCS log files ..... 272
Preventing automatic database startup ..... 272
Removing existing PrivNIC or MultiPrivNIC resources ..... 273
Removing permissions for communication ..... 273

Chapter 17 Upgrading Oracle RAC ..... 274

Preparing to upgrade Oracle RAC ..... 274
Upgrading Oracle RAC binaries ..... 275
Migrating the Oracle RAC database ..... 276
Performing post-upgrade tasks ..... 276

Section 5 Adding and removing nodes ..... 277

Chapter 18 Adding a node to SF Oracle RAC clusters ..... 278

About adding a node to a cluster ..... 278
Before adding a node to a cluster ..... 279
Adding a node to a cluster using the Veritas InfoScale installer ..... 282
Preparing the new nodes for installing Oracle RAC using the installer ..... 286
Adding the node to a cluster manually ..... 293
Starting Veritas Volume Manager (VxVM) on the new node ..... 294
Configuring cluster processes on the new node ..... 295
Setting up the node to run in secure mode ..... 297
Starting fencing on the new node ..... 300
After adding the new node ..... 300
Configuring server-based fencing on the new node ..... 301
Configuring Cluster Volume Manager (CVM) and Cluster File System (CFS) on the new node ..... 302
Configuring the ClusterService group for the new node ..... 304
Preparing the new node manually for installing Oracle RAC ..... 305
Adding a node to the cluster using the SF Oracle RAC response file ..... 311
Response file variables to add a node to an SF Oracle RAC cluster ..... 313
Sample response file for adding a node to an SF Oracle RAC cluster ..... 314
Configuring private IP addresses for Oracle RAC on the new node ..... 315
Using link aggregation or NIC bonding ..... 315
Using Oracle High Availability IP (HAIP) ..... 316
Adding the new node to Oracle RAC ..... 316
Adding nodes to a cluster that is using authentication for SFDB tools ..... 317
Updating the Storage Foundation for Databases (SFDB) repository after adding a node ..... 318
Sample configuration file for adding a node to the cluster ..... 319

Chapter 19 Removing a node from SF Oracle RAC clusters ..... 324

About removing a node from a cluster ..... 324
Removing a node from a cluster ..... 325
Modifying the VCS configuration files on existing nodes ..... 327

Modifying the Cluster Volume Manager (CVM) configuration on the existing nodes to remove references to the deleted node ..... 329
Removing the node configuration from the CP server ..... 329
Removing security credentials from the leaving node ..... 330
Updating the Storage Foundation for Databases (SFDB) repository after removing a node ..... 331
Sample configuration file for removing a node from the cluster ..... 331

Section 6 Configuration of disaster recovery environments ..... 335

Chapter 20 Configuring disaster recovery environments ..... 336

Disaster recovery options for SF Oracle RAC ..... 336
Hardware requirements for campus cluster ..... 337
Supported replication technologies for global clusters ..... 337
About setting up a campus cluster for disaster recovery ..... 338
About setting up a global cluster environment for SF Oracle RAC ..... 340
About configuring a parallel global cluster using Volume Replicator (VVR) for replication ..... 341

Section 7 Installation reference ..... 343

Appendix A Installation scripts ..... 344

Installation script options ..... 344
About using the postcheck option ..... 349

Appendix B Tunable files for installation ..... 353

About setting tunable parameters using the installer or a response file ..... 353
Setting tunables for an installation, configuration, or upgrade ..... 354
Setting tunables with no other installer-related operations ..... 355
Setting tunables with an un-integrated response file ..... 356
Preparing the tunables file ..... 357
Setting parameters for the tunables file ..... 358
Tunables value parameter definitions ..... 358

Appendix C Sample installation and configuration values ..... 366

About the installation and configuration worksheets ..... 366
SF Oracle RAC worksheet ..... 367

Cluster Server component information ..... 369
I/O fencing information ..... 370
SF Oracle RAC add user information ..... 371
Global cluster information ..... 372
Oracle RAC worksheet ..... 373
Replicated cluster using VVR worksheet ..... 380
Replicated cluster using SRDF worksheet ..... 381
Required installation information for Oracle Clusterware/Grid Infrastructure ..... 382
Required installation information for Oracle database ..... 385

Appendix D Configuration files ..... 387

About VCS configuration file ..... 387
About the LLT and GAB configuration files ..... 388
About I/O fencing configuration files ..... 392
Sample configuration files ..... 395
sfrac02_main.cf file ..... 396
sfrac03_main.cf file ..... 397
sfrac04_main.cf file ..... 398
sfrac05_main.cf file ..... 400
sfrac06_main.cf file ..... 401
sfrac07_main.cf and sfrac08_main.cf files ..... 402
sfrac09_main.cf and sfrac10_main.cf files ..... 404
sfrac11_main.cf file ..... 407
sfrac12_main.cf and sfrac13_main.cf files ..... 408
sfrac14_main.cf file ..... 411
sfrac15_main.cf and sfrac16_main.cf files ..... 412
sfrac17_main.cf file ..... 413
Sample configuration files for CP server ..... 414

Appendix E Configuring the secure shell or the remote shell for communications ..... 421

About configuring secure shell or remote shell communication modes before installing products ..... 421
Manually configuring passwordless ssh ..... 422
Setting up ssh and rsh connection using the installer -comsetup command ..... 425
Setting up ssh and rsh connection using the pwdutil.pl utility ..... 427
Restarting the ssh session ..... 430
Enabling rsh for Linux ..... 430

Appendix F Automatic Storage Management ..... 432

About ASM in SF Oracle RAC environments ..... 432
ASM configuration with SF Oracle RAC ..... 433
Configuring ASM in SF Oracle RAC environments ..... 434
Creating database storage for ASM ..... 435
Creating ASM disk groups and instances ..... 435
Verifying the ASM setup ..... 436
Configuring VCS service groups for database instances on ASM ..... 437
Configuring Flex ASM in SF Oracle RAC environments ..... 439
Creating storage for Flex ASM ..... 440
Verifying the Flex ASM setup ..... 440
Configuring VCS service groups for database instances on ASM for Flex ASM cluster ..... 441

Appendix G Creating a test database ..... 445

About creating a test database ..... 445
Creating a database for Oracle ..... 445
Creating the database storage on CFS ..... 446

Appendix H High availability agent information ..... 448

About agents ..... 448
VCS agents included within SF Oracle RAC ..... 449
VCS agents for Oracle included within SF Oracle RAC ..... 449
CVMCluster agent ..... 450
Entry points for CVMCluster agent ..... 450
Attribute definition for CVMCluster agent ..... 451
CVMCluster agent type definition ..... 452
CVMCluster agent sample configuration ..... 452
CVMVxconfigd agent ..... 453
Entry points for CVMVxconfigd agent ..... 453
Attribute definition for CVMVxconfigd agent ..... 454
CVMVxconfigd agent type definition ..... 455
CVMVxconfigd agent sample configuration ..... 456
CVMVolDg agent ..... 456
Entry points for CVMVolDg agent ..... 456
Attribute definition for CVMVolDg agent ..... 457
CVMVolDg agent type definition ..... 459
CVMVolDg agent sample configuration ..... 459
CFSMount agent ..... 460
Entry points for CFSMount agent ..... 460

Attribute definition for CFSMount agent ..... 461
CFSMount agent type definition ..... 463
CFSMount agent sample configuration ..... 464
CFSfsckd agent ..... 464
Entry points for CFSfsckd agent ..... 464
Attribute definition for CFSfsckd agent ..... 465
CFSfsckd agent type definition ..... 466
CFSfsckd agent sample configuration ..... 467
CSSD agent ..... 467
Functions of the CSSD agent ..... 467
Attributes of the CSSD agent ..... 469
States of the CSSD agent ..... 470
Disabling monitoring of Oracle Grid Infrastructure processes temporarily ..... 471
Sample service group configurations with the CSSD agent ..... 472
Type definition of the CSSD resource ..... 472
Sample configuration of the CSSD resource ..... 473
VCS agents for Oracle ..... 473
Oracle agent functions ..... 473
Resource type definition for the Oracle agent ..... 481
Netlsnr agent functions ..... 492
Resource type definition for the Netlsnr agent ..... 493
ASMDG agent functions ..... 498
Resource type definition for the ASMDG agent ..... 499
CRSResource agent ..... 501
Functions of the CRSResource agent ..... 502
States of the CRSResource agent ..... 502
Attributes of the CRSResource agent ..... 502
VCS service group dependencies with the CRSResource agent ..... 503
Resource type definition for the CRSResource agent ..... 510
Sample configuration for the CRSResource agent ..... 510

Appendix I SF Oracle RAC deployment scenarios ..... 512

Configuration diagrams for setting up server-based I/O fencing ..... 512
Two unique client clusters served by 3 CP servers ..... 512
Client cluster served by highly available CPS and 2 SCSI-3 disks ..... 513
Two node campus cluster served by remote CP server and 2 SCSI-3 disks ..... 515
Multiple client clusters served by highly available CP server and 2 SCSI-3 disks ..... 517

Appendix J Configuring LLT over UDP ..... 519

Using the UDP layer for LLT ..... 519
When to use LLT over UDP ..... 519
Manually configuring LLT over UDP using IPv4 ..... 519
Broadcast address in the /etc/llttab file ..... 520
The link command in the /etc/llttab file ..... 521
The set-addr command in the /etc/llttab file ..... 521
Selecting UDP ports ..... 522
Configuring the netmask for LLT ..... 522
Configuring the broadcast address for LLT ..... 523
Sample configuration: direct-attached links ..... 523
Sample configuration: links crossing IP routers ..... 525
Using the UDP layer of IPv6 for LLT ..... 526
When to use LLT over UDP ..... 526
Manually configuring LLT over UDP using IPv6 ..... 527
The link command in the /etc/llttab file ..... 527
The set-addr command in the /etc/llttab file ..... 528
Selecting UDP ports ..... 528
Sample configuration: direct-attached links ..... 529
Sample configuration: links crossing IP routers ..... 530
About configuring LLT over UDP multiport ..... 532
Manually configuring LLT over UDP multiport ..... 532
Enabling LLT ports in firewall ..... 534
Disabling the UDP multiport feature ..... 535

Appendix K Using LLT over RDMA ..... 536

Using LLT over RDMA ..... 536
About RDMA over RoCE or InfiniBand networks in a clustering environment ..... 536
How LLT supports RDMA capability for faster interconnects between applications ..... 537
Using LLT over RDMA: supported use cases ..... 538
Configuring LLT over RDMA ..... 538
Choosing supported hardware for LLT over RDMA ..... 539
Installing RDMA, InfiniBand or Ethernet drivers and utilities ..... 540
Configuring RDMA over an Ethernet network ..... 541
Configuring RDMA over an InfiniBand network ..... 543
Tuning system performance ..... 547
Manually configuring LLT over RDMA ..... 548
LLT over RDMA sample /etc/llttab ..... 553
Verifying LLT configuration ..... 553
Troubleshooting LLT over RDMA ..... 554

IP addresses associated to the RDMA NICs do not automatically plumb on node restart ..... 554
Ping test fails for the IP addresses configured over InfiniBand interfaces ..... 555
After a node restart, by default the Mellanox card with Virtual Protocol Interconnect (VPI) gets configured in InfiniBand mode ..... 555
The LLT module fails to start ..... 555

Section 1

# Configuring SF Oracle RAC


- Chapter 1. Preparing to configure SF Oracle RAC
- Chapter 2. Configuring SF Oracle RAC using the script-based installer
- Chapter 3. Performing an automated SF Oracle RAC configuration

Chapter 1

# Preparing to configure SF Oracle RAC

This chapter includes the following topics:


- About this document
- I/O fencing requirements
- About planning to configure I/O fencing


## About this document

This document provides instructions for performing the following tasks:

- Configuring SF Oracle RAC
- Installing and configuring Oracle RAC
- Adding or removing nodes from a cluster
- Uninstalling SF Oracle RAC

---

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

---

## 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 19.

- CP servers
  See "CP server requirements" on page 19.

---

**Note:** Irrespective of whether you use coordinator disks or CP server for I/O fencing, ensure that the shared storage supports SCSI-3 persistent reservations.

---

## 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

SF Oracle RAC 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:** SF Oracle RAC requires at least 3 coordination points for I/O fencing.

---

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

**Table 1-1** 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:<br>■ 550 MB in the /opt directory (additionally, the language pack requires another 15 MB)<br>■ 300 MB in /usr<br>■ 20 MB in /var<br>■ 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 SF Oracle RAC clusters (application clusters). |

Table 1-2 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 1-2** 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:<br>■ Linux:<br>  ■ RHEL 7.9, 8.2, 8.4<br>  ■ SLES 12, 15<br>Review other details such as supported operating system levels and architecture for the supported operating systems.<br>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 for Oracle RAC Administrator's Guide*.

# About planning to configure I/O fencing

After you configure SF Oracle RAC 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 or server-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.

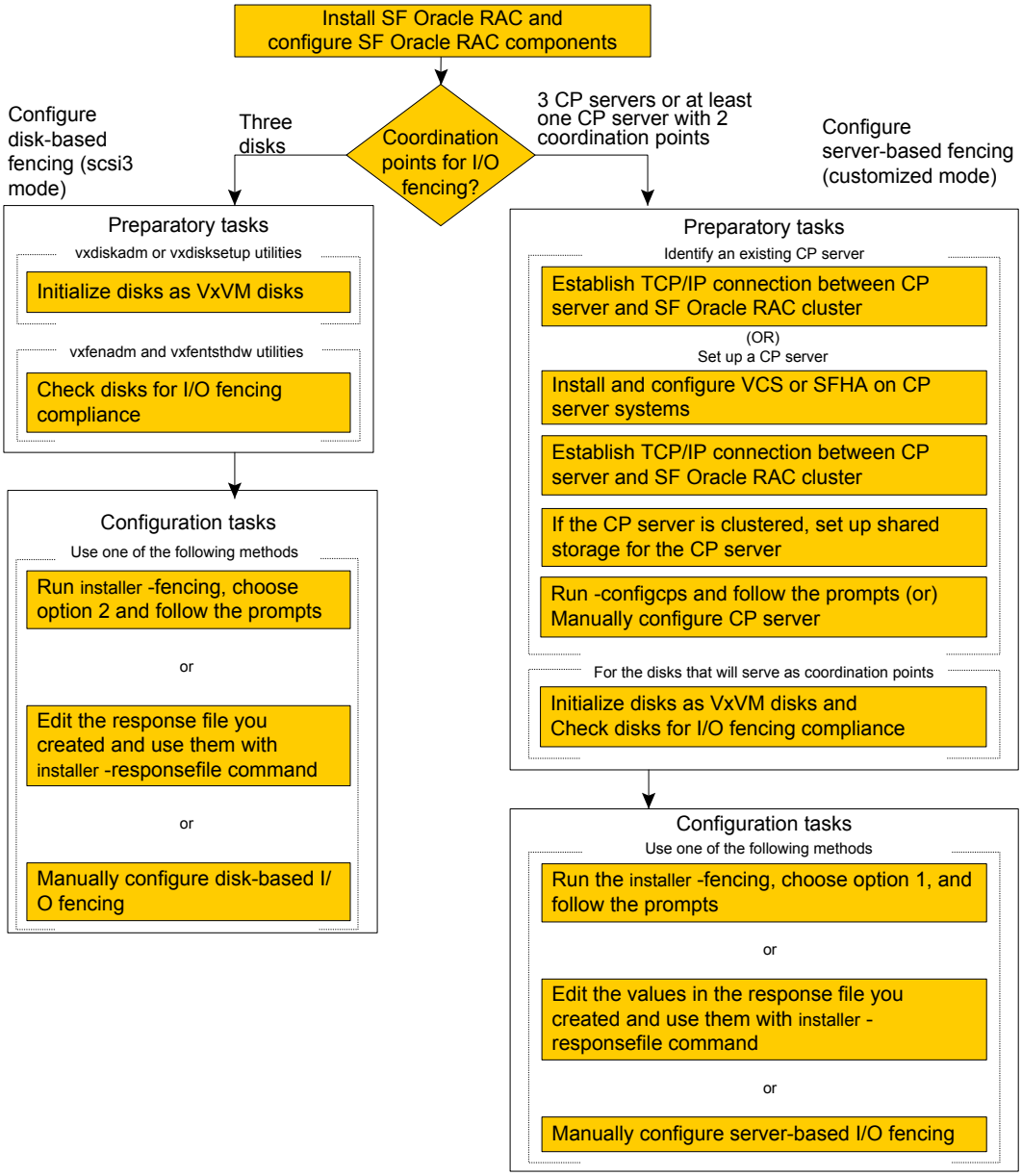
---

**Note:** Irrespective of whether you use coordinator disks or CP servers, SF Oracle RAC requires at least 3 coordination points.

---

Figure 1-1 illustrates a high-level flowchart to configure I/O fencing for the SF Oracle RAC cluster.

**Figure 1-1** Workflow to configure 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 51.<br>See “Setting up server-based I/O fencing using installer” on page 60. |
| Using response files | See “Response file variables to configure disk-based I/O fencing” on page 86.<br>See “Response file variables to configure server-based I/O fencing” on page 93.<br>See “Configuring I/O fencing using response files” on page 86. |
| Manually editing configuration files | |

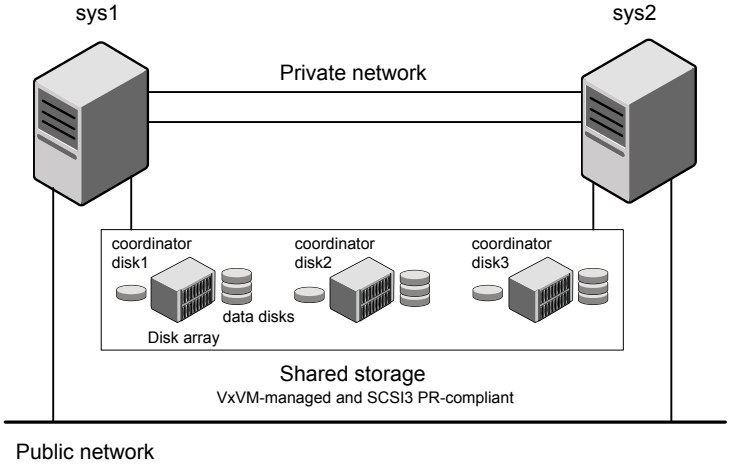
You can also migrate from one I/O fencing configuration to another.

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

## Typical SF Oracle RAC cluster configuration with disk-based I/O fencing

Figure 1-2 displays a typical VCS configuration with two nodes and shared storage. The configuration uses three coordinator disks for I/O fencing.

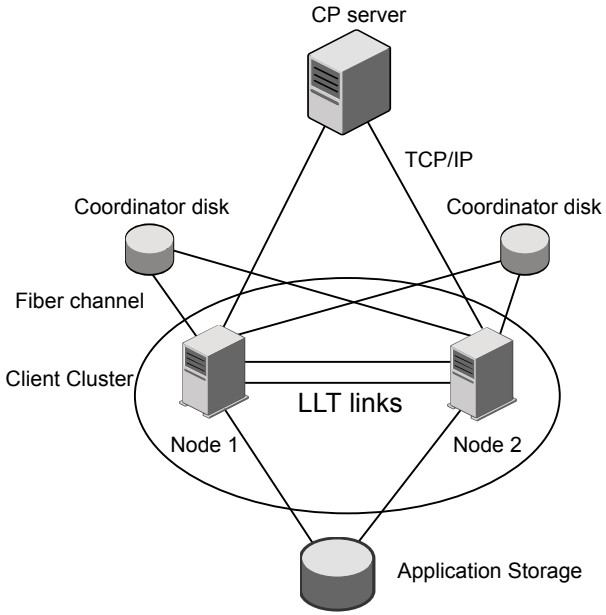
**Figure 1-2** Typical SF Oracle RAC cluster configuration with disk-based I/O fencing



## Typical SF ORACLE RAC cluster configuration with server-based I/O fencing

Figure 1-3 displays a configuration using a SF Oracle RAC cluster (with two nodes), a single CP server, and two coordinator disks. The nodes within the SF Oracle RAC cluster are connected to and communicate with each other using LLT links.

**Figure 1-3** CP server, SF Oracle RAC 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 1-4 on page 26.
- Multiple application clusters use a single CP server and single or multiple pairs of coordinator disks (two) as their coordination points
  See Figure 1-5 on page 27.

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 1-4 displays a configuration using three CP servers that are connected to multiple application clusters.

**Figure 1-4** Three CP servers connecting to multiple application clusters

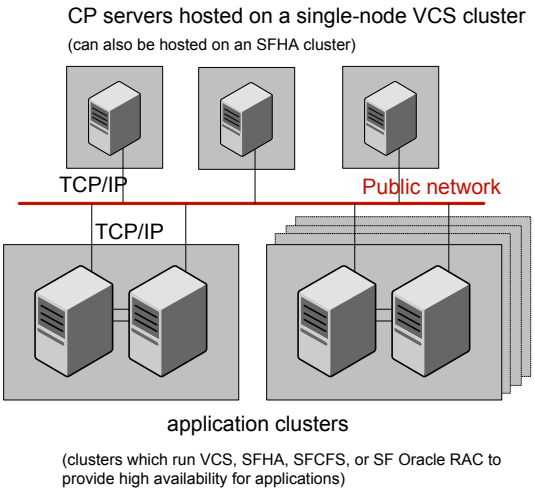
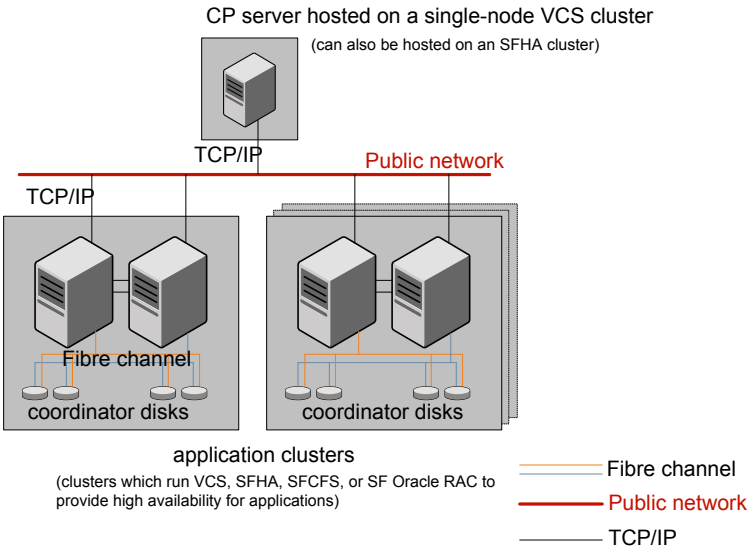



Figure 1-5 displays a configuration using a single CP server that is connected to multiple application clusters with each application cluster also using two coordinator disks.

**Figure 1-5** Single CP server with two coordinator disks for each application cluster




See "Configuration diagrams for setting up server-based I/O fencing" on page 512.

# Preparing to configure server-based fencing for SF Oracle RAC

If you plan to use server-based I/O fencing, you need to configure the coordination point (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.

For instructions on configuring the CP server, see the *Cluster Server Installation Guide*.

Chapter 2

# Configuring SF Oracle RAC using the script-based installer

This chapter includes the following topics:


- Configuring the SF Oracle RAC components using the script-based installer
- Setting up disk-based I/O fencing using installer
- Setting up server-based I/O fencing using installer
- Configuring SFDB


## Configuring the SF Oracle RAC components using the script-based installer

Make sure that you have performed the necessary pre-configuration tasks if you want to configure the cluster in secure mode.

Start the `installer` program if you quit the installer after installation.

By default, the communication between the systems is selected as SSH. If SSH is used for communication between systems, the SSH commands execute without prompting for passwords or confirmations.

At the end of the configuration, the VCS, CVM, and CFS components are configured to provide a cluster-aware environment.

---

**Note:** If you want to reconfigure SF Oracle RAC, 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. For resources that are not configured under VCS, use native application commands to stop the resources.

---

If you encounter issues during the configuration, see the *Storage Foundation for Oracle RAC Administrator's Guide*, Chapter "Performance and troubleshooting" for information on resolving the issue.

**To configure the SF Oracle RAC components**

1. Log in as the superuser on any of the nodes in the cluster.
2. Start the configuration program.

   ```
   # ./installer -configure \
   sys1 sys2
   ```

   Choose **Storage Foundation for Oracle RAC** to configure SF Oracle RAC.

   The installer displays the copyright message and specifies the directory where the logs are created.
3. Select the option **Configure SF Oracle RAC sub-components**.

   ```
   1)  Configure SF Oracle RAC sub-components
   2)  Prepare to Install Oracle
   3)  Install Oracle Clusterware/Grid Infrastructure and Database
   4)  Post Oracle Installation Tasks
   5)  Exit SF Oracle RAC Configuration
   Choose option: [1-5,q] (1)
   ```

4. If you had quit the installer in the process of an active configuration, the installer discovers that installer process and provides the option of resuming the configuration or starting a new configuration. Provide a suitable response.

   ```
   The installer has discovered an existing installer process.
   The process exited while performing configure of
   SF Oracle RAC on sys1.
   Do you want to resume this process? [y,n,q,?] (y) n
   ```

5. Configure Cluster Server to set up the SF Oracle RAC cluster.

   See "Configuring the SF Oracle RAC cluster" on page 30.
6. Add VCS users.

   See "Adding VCS users" on page 45.

7 Configure SMTP email notification.

See "Configuring SMTP email notification" on page 46.

8 Configure SNMP trap notification.

See "Configuring SNMP trap notification" on page 48.

9 Configure global clusters, if you chose to enable GCO during the installation.

See "Configuring global clusters" on page 49.

10 Stop the SF Oracle RAC resources.

See "Stopping and starting SF Oracle RAC processes" on page 50.

# Configuring the SF Oracle RAC cluster

You must configure Cluster Server to set up the SF Oracle RAC cluster.

You can configure a basic cluster or an advanced cluster. A basic cluster configuration requires the cluster name and ID and the private heartbeat links for LLT. The remaining configuration options presented by the installer are optional and may be used if you plan to configure an advanced cluster.

Refer to the *Cluster Server Installation Guide* for more information.

## 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 519.

See "Using LLT over RDMA: supported use cases " on page 538.

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)
```

**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.

A basic cluster is now configured. The remaining configuration settings are optional.

---

**Note:** You can proceed through the subsequent screens by just accepting the default value **n**.

---

## 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.

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

6 Confirm the default netmask or enter another one:

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

7 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)
```

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

See "Setting up trust relationships for your SF Oracle RAC cluster" on page 39.

See "Configuring a secure cluster node by node" on page 41.

## Configuring SF Oracle RAC in secure mode

Configuring SF Oracle RAC in secure mode ensures that all the communication between the systems is encrypted and users are verified against security credentials.

SF Oracle RAC user names and passwords are not used when a cluster is running in secure mode.

**To configure SF Oracle RAC in secure mode**

1. To install and configure SF Oracle RAC 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.

## Setting up trust relationships for your SF Oracle RAC cluster

If you need to use an external authentication broker for authenticating VCS users, you must set up a trust relationship between VCS and the broker. For example, if Veritas InfoScale Operations Manager is your external authentication broker, the trust relationship ensures that VCS accepts the credentials that VOM issues.

Perform the following steps to set up a trust relationship between your SF Oracle RAC cluster and a broker.

**To set up a trust relationship**

1. Ensure that you are logged in as superuser on one of the nodes in the cluster.
2. Enter the following command:

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

   The installer specifies the location of the log files. It then lists the cluster information such as cluster name, cluster ID, node names, and service groups.
3. When the installer prompts you for the broker information, specify the IP address, port number, and the data directory for which you want to establish trust relationship with the broker.

   ```
   Input the broker name of IP address: 15.193.97.204
   ```

   ```
   Input the broker port: (14545)
   ```

   Specify a port number on which broker is running or press Enter to accept the default port.

   ```
   Input the data directory to setup trust with: (/var/VRTSvcs/
   vcsauth/data/HAD)
   ```

   Specify a valid data directory or press Enter to accept the default directory.
4. The installer performs one of the following actions:
   - If you specified a valid directory, the installer prompts for a confirmation.

     ```
     Are you sure that you want to setup trust for the VCS cluster
     with the broker 15.193.97.204 and port 14545? [y,n,q] y
     ```

     The installer sets up trust relationship with the broker for all nodes in the cluster and displays a confirmation.

     ```
     Setup trust with broker 15.193.97.204 on cluster node1
     ........Done
     ```

     ```
     Setup trust with broker 15.193.97.204 on cluster node2
     ........Done
     ```

     The installer specifies the location of the log files, summary file, and response file and exits.
   - If you entered incorrect details for broker IP address, port number, or directory name, the installer displays an error. It specifies the location of the log files, summary file, and response file and exits.

## 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 2-1 lists the tasks that you must perform to configure a secure cluster.

**Table 2-1** Configuring a secure cluster node by node

| Task | Reference |
|---|---|
| Configure security on one node | See "Configuring the first node" on page 41. |
| Configure security on the remaining nodes | See "Configuring the remaining nodes" on page 42. |
| Complete the manual configuration steps | See "Completing the secure cluster configuration" on page 43. |

### 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 48.
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 SF Oracle RAC based on the configuration details you provided.

See "Configuring global clusters" on page 49.

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

If you had installed a valid HA/DR license, the installer provides you an option to configure this cluster as global cluster. If not, the installer proceeds with other set of questions for CVM and CFS.

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 SF Oracle RAC global clusters.

See the appropriate *Veritas InfoScale* installation guide for instructions to set up SF Oracle RAC global clusters.

---

**Note:** If you installed a HA/DR license to set up 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.
4. Verify and confirm the configuration of the global cluster.

```
Global Cluster Option configuration verification:

NIC: eth0
IP: 192.168.1.16
Netmask: 255.255.240.0

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

For complete details for configuring your disaster recovery environment once clusters are installed and configured:

See the *Veritas InfoScale™ Disaster Recovery Implementation Guide*.

# Creation of SF Oracle RAC configuration files

The program consolidates all the information gathered in the preceding configuration tasks and creates configuration files.

If you chose to configure the cluster in secure mode, the installer also configures the Veritas Product Authentication Service, which creates an Authentication Broker with root and authentication mode.

Review the output as the configuration program starts VCS, creates VCS configuration files, and copies the files to each node.

# Stopping and starting SF Oracle RAC processes

The installer stops and starts SF Oracle RAC processes and configures the SF Oracle RAC agents.

**Note:** Do not opt to start SF Oracle RAC now if you want to configure private heartbeats to use aggregated interfaces that the installer has not discovered or to use aggregated interfaces on nodes that run SLES.

**To stop SF Oracle RAC processes**

1 Enter **y** to stop SF Oracle RAC processes.

```
Do you want to stop SF Oracle RAC processes now? [y,n,q,?] (y)
```

2 Review the output as the installer stops and starts the SF Oracle RAC processes.

Note that SF Oracle RAC configuration program starts I/O fencing feature in disabled mode. SF Oracle RAC requires you to configure and enable I/O fencing feature.

# Setting up disk-based I/O fencing using installer

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

## 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.

# Identifying disks to use as coordinator disks

Make sure you initialized disks as VxVM disks.

See "Initializing disks as VxVM disks" on page 51.

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 52.

# Checking shared disks for I/O fencing

Make sure that the shared storage you set up while preparing to configure SF Oracle RAC 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 for Oracle RAC 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 53.
- Verifying that nodes have access to the same disk
  See "Verifying that the nodes have access to the same disk" on page 53.
- Testing the shared disks for SCSI-3
  See "Testing the disks using vxfentsthdw utility" on page 54.

## 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 for Oracle RAC 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 421.

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.

# Configuring disk-based I/O fencing using installer

---

**Note:** The installer stops and starts SF Oracle RAC to complete I/O fencing configuration. Make sure to unfreeze any frozen VCS service groups in the cluster for the installer to successfully stop SF Oracle RAC.

---

**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 SF Oracle RAC 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 fencing in disabled mode
   Select the fencing mechanism to be configured in this
   Application Cluster [1-3,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 52.

**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 SF Oracle RAC 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)
```

## 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 SF Oracle RAC 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 SF Oracle RAC 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

See " About planning to configure I/O fencing" on page 22.

See "Recommended CP server configurations" on page 25.

This section covers the following example procedures:

| | |
|---|---|
| Mix of CP servers and coordinator disks | See "To configure server-based fencing for the SF Oracle RAC cluster (one CP server and two coordinator disks)" on page 60. |

**To configure server-based fencing for the SF Oracle RAC 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 SF Oracle RAC cluster. The SF Oracle RAC cluster is also referred to as the application cluster or the client cluster.
- The coordination disks are verified for SCSI3-PR compliance.
  See "Checking shared disks for I/O fencing" on page 52.

**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 SF Oracle RAC 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 SF Oracle RAC (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 SF Oracle RAC (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 392.

**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 SF Oracle RAC (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
```

## 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 SF Oracle RAC 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 SF Oracle RAC 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}
```

# 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 3

# Performing an automated SF Oracle RAC configuration

This chapter includes the following topics:


- Configuring SF Oracle RAC using response files
- Response file variables to configure SF Oracle RAC
- Sample response file for configuring SF Oracle RAC
- 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 CP server
- Sample response file for configuring the CP server on SFHA cluster
- Response file variables to configure server-based I/O fencing
- Sample response file for configuring server-based I/O fencing


## Configuring SF Oracle RAC using response files

Typically, you can use the response file that the installer generates after you perform SF Oracle RAC configuration on one cluster to configure SF Oracle RAC on other clusters.

**To configure SF Oracle RAC using response files**

1. Make sure the Veritas InfoScale Availability or Enterprise RPMs are installed on the systems where you want to configure SF Oracle RAC.
2. Copy the response file to one of the cluster systems where you want to configure SF Oracle RAC.
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 SF Oracle RAC" on page 73.
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.
5. Configure I/O fencing.

   For instructions on configuring I/O fencing using a response file, see the chapter *Configuring I/O fencing using a response file* in this document.
6. Complete the SF Oracle RAC post-installation tasks.

   For instructions, see the chapter *Performing post-installation and configuration tasks* in this document.

# Response file variables to configure SF Oracle RAC

Table 3-1 lists the response file variables that you can define to configure SF Oracle RAC.

**Table 3-1** Response file variables specific to configuring SF Oracle RAC

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{opt}{configure} | Scalar | Performs the configuration if the RPMs are already installed.<br>(Required)<br>Set the value to 1 to configure SF Oracle RAC. |

**Table 3-1** Response file variables specific to configuring SF Oracle RAC *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{config_sfrac_subcomponents} | Scalar | Set the variable to 1 to configure the SF Oracle RAC components.<br>(Required)<br>**Note:** You must set the *CFG{opt}{configure}* variable to 1. |
| CFG{accepteula} | Scalar | Specifies whether you agree with EULA.pdf on the media.<br>(Required) |
| CFG{activecomponent} | List | Defines the component to be configured.<br>The value is SFRAC80 for SF Oracle RAC.<br>(Required) |
| CFG{keys}{keyless}<br>CFG{keys}{license} | List | `CFG{keys}{keyless}` gives a list of keyless keys to be registered on the system.<br>`CFG{keys}{license}` gives a list of user defined keys to be registered on the system.<br>(Optional) |
| CFG{systems} | List | List of systems on which the product is to be configured.<br>(Required) |
| CFG{prod} | Scalar | Defines the product for operations.<br>The value is ENTERPRISE80 for Veritas InfoScale Enterprise.<br>(Required) |
| CFG{opt}{keyfile} | Scalar | Defines the location of an ssh keyfile that is used to communicate with all remote systems.<br>(Optional) |

**Table 3-1** Response file variables specific to configuring SF Oracle RAC *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{opt}{rsh} | Scalar | Defines that *rsh* must be used instead of ssh as the communication method between systems.<br>(Optional) |
| CFG{opt}{logpath} | Scalar | Mentions the location where the log files are to be copied. The default location is /opt/VRTS/install/logs.<br>**Note:** The installer copies the response files and summary files also to the specified *logpath* location.<br>(Optional) |
| CFG{uploadlogs} | Scalar | Defines a Boolean value 0 or 1.<br>The value 1 indicates that the installation logs are uploaded to the Veritas website.<br>The value 0 indicates that the installation logs are not uploaded to the Veritas website.<br>(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 3-2 lists the response file variables that specify the required information to configure a basic SF Oracle RAC cluster.

**Table 3-2** Response file variables specific to configuring a basic SF Oracle RAC cluster

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{donotreconfigurevcs} | Scalar | Defines if you need to re-configure VCS.<br>(Optional) |
| CFG{donotreconfigurefencing} | Scalar | Defines if you need to re-configure fencing.<br>(Optional) |
| CFG{vcs_clusterid} | Scalar | An integer between 0 and 65535 that uniquely identifies the cluster.<br>(Required) |
| CFG{vcs_clustername} | Scalar | Defines the name of the cluster.<br>(Required) |
| 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).<br>(Required) |

Table 3-3 lists the response file variables that specify the required information to configure LLT over Ethernet.

**Table 3-3** Response file variables specific to configuring private LLT over Ethernet

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{vcs_lltlink#}<br>{"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.<br>You must enclose the system name within double quotes.<br>(Required) |

**Table 3-3** Response file variables specific to configuring private LLT over Ethernet *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{vcs_lltlinklowpri#}<br>{"system"} | Scalar | Defines a low priority heartbeat link. Typically, lltlinklowpri is used on a public network link to provide an additional layer of communication.<br>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.<br>You must enclose the system name within double quotes.<br>(Optional) |

Table 3-4 lists the response file variables that specify the required information to configure LLT over UDP.

**Table 3-4** 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.<br>(Required) |
| CFG{vcs_udplink<n>_address}<br>{<sys1>} | Scalar | Stores the IP address (IPv4 or IPv6) that the heartbeat link uses on node1.<br>You can have four heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective heartbeat links.<br>(Required) |

**Table 3-4** Response file variables specific to configuring LLT over UDP *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| CFG<br>{vcs_udplinklowpri<n>_address}<br>{<sys1>} | Scalar | Stores the IP address (IPv4 or IPv6) that the low priority heartbeat link uses on node1.<br>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.<br>(Required) |
| CFG{vcs_udplink<n>_port}<br>{<sys1>} | Scalar | Stores the UDP port (16-bit integer value) that the heartbeat link uses on node1.<br>You can have four heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective heartbeat links.<br>(Required) |
| CFG{vcs_udplinklowpri<n>_port}<br>{<sys1>} | Scalar | Stores the UDP port (16-bit integer value) that the low priority heartbeat link uses on node1.<br>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.<br>(Required) |
| CFG{vcs_udplink<n>_netmask}<br>{<sys1>} | Scalar | Stores the netmask (prefix for IPv6) that the heartbeat link uses on node1.<br>You can have four heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective heartbeat links.<br>(Required) |

**Table 3-4** Response file variables specific to configuring LLT over UDP *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| CFG {vcs_udplinklowpri<n>_netmask}<br>{<sys1>} | Scalar | Stores the netmask (prefix for IPv6) that the low priority heartbeat link uses on node1.<br>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.<br>(Required) |

Table 3-5 lists the response file variables that specify the required information to configure LLT over RDMA.

**Table 3-5** Response file variables specific to configuring LLT over RDMA

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{lltoverrdma}=1 | Scalar | Indicates whether to configure heartbeat link using LLT over RDMA.<br>(Required) |
| CFG{vcs_rdmalink<n>_address}<br>{<sys1>} | Scalar | Stores the IP address (IPv4 or IPv6) that the heartbeat link uses on node1.<br>You can have four heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective heartbeat links.<br>(Required) |

**Table 3-5** Response file variables specific to configuring LLT over RDMA *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| CFG<br>{vcs_rdmalinklowpri<n>_address}<br>{<sys1>} | Scalar | Stores the IP address (IPv4 or IPv6) that the low priority heartbeat link uses on node1.<br>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.<br>(Required) |
| CFG{vcs_rdmalink<n>_port}<br>{<sys1>} | Scalar | Stores the RDMA port (16-bit integer value) that the heartbeat link uses on node1.<br>You can have four heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective heartbeat links.<br>(Required) |
| CFG{vcs_rdmalinklowpri<n>_port}<br>{<sys1>} | Scalar | Stores the RDMA port (16-bit integer value) that the low priority heartbeat link uses on node1.<br>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.<br>(Required) |
| CFG{vcs_rdmalink<n>_netmask}<br>{<sys1>} | Scalar | Stores the netmask (prefix for IPv6) that the heartbeat link uses on node1.<br>You can have four heartbeat links and <n> for this response file variable can take values 1 to 4 for the respective heartbeat links.<br>(Required) |

**Table 3-5** Response file variables specific to configuring LLT over RDMA *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| CFG {vcs_rdmalinklowpri<n>_netmask}<br>{<sys1>} | Scalar | Stores the netmask (prefix for IPv6) that the low priority heartbeat link uses on node1.<br>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.<br>(Required) |

Table 3-6 lists the response file variables that specify the required information to configure virtual IP for SF Oracle RAC cluster.

**Table 3-6** Response file variables specific to configuring virtual IP for SF Oracle RAC cluster

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{vcs_csgnic}<br>{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.<br>(Optional) |
| CFG{vcs_csgvip} | Scalar | Defines the virtual IP address for the cluster.<br>(Optional) |
| CFG{vcs_csgnetmask} | Scalar | Defines the Netmask of the virtual IP address for the cluster.<br>(Optional) |

Table 3-7 lists the response file variables that specify the required information to configure the SF Oracle RAC cluster in secure mode.

**Table 3-7** Response file variables specific to configuring SF Oracle RAC 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.<br>■ 1—Configure the first node<br>■ 2—Configure the other node |
| CFG{secusrgrps} | List | Defines the user groups which get read access to the cluster.<br>List or scalar: list<br>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.<br>(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.<br>Optional or required: optional |
| CFG{vcs_eat_security_fips} | Scalar | Specifies that the enabled security is FIPS compliant. |

Table 3-8 lists the response file variables that specify the required information to configure VCS users.

**Table 3-8** Response file variables specific to configuring VCS users

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{vcs_userenpw} | List | List of encoded passwords for VCS users<br>The value in the list can be "Administrators Operators Guests"<br>**Note:** The order of the values for the vcs_userenpw list must match the order of the values in the vcs_username list.<br>(Optional) |
| CFG{vcs_username} | List | List of names of VCS users<br>(Optional) |
| CFG{vcs_userpriv} | List | List of privileges for VCS users<br>**Note:** The order of the values for the vcs_userpriv list must match the order of the values in the vcs_username list.<br>(Optional) |

Table 3-9 lists the response file variables that specify the required information to configure VCS notifications using SMTP.

**Table 3-9** 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.<br>(Optional) |
| CFG{vcs_smtprecp} | List | List of full email addresses (example: user@example.com) of SMTP recipients.<br>(Optional) |

**Table 3-9** 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.<br>(Optional) |

Table 3-10 lists the response file variables that specify the required information to configure VCS notifications using SNMP.

**Table 3-10** 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).<br>(Optional) |
| CFG{vcs_snmpcons} | List | List of SNMP console system names<br>(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.<br>(Optional) |

Table 3-11 lists the response file variables that specify the required information to configure SF Oracle RAC global clusters.

**Table 3-11** Response file variables specific to configuring SF Oracle RAC global clusters

| Variable | List or Scalar | Description |
|---|---|---|
| CFG{vcs_gconic}<br>{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.<br>(Optional) |
| CFG{vcs_gcovip} | Scalar | Defines the virtual IP address to that the Global Cluster Option uses.<br>(Optional) |
| CFG{vcs_gconetmask} | Scalar | Defines the Netmask of the virtual IP address that the Global Cluster Option uses.<br>(Optional) |

# Sample response file for configuring SF Oracle RAC

The following sample response file configures SF Oracle RAC on two nodes, sys1 and sys2.

```
our %CFG;

$CFG{accepteula}=1;
$CFG{opt}{install}=1;
$CFG{prod}="ENTERPRISE80";
$CFG{activecomponent}="SFRAC80";
$CFG{systems}=[ qw(sys1 sys2) ];
$CFG{opt}{configure}=1;
$CFG{config_sfrac_subcomponents} = 1;
$CFG{vcs_allowcomms}=1;
$CFG{vcs_clusterid}=101;
$CFG{vcs_clustername}="clus1";
$CFG{vcs_lltlink1}{sys1}="eth1";
$CFG{vcs_lltlink1}{sys2}="eth1";
$CFG{vcs_lltlink2}{sys1}="eth2";
$CFG{vcs_lltlink2}{sys2}="eth2";
```

```
$CFG{vcs_username}=[ qw(admin) ];
$CFG{vcs_userpriv}=[ qw(Administrators) ];
$CFG{vcs_userenpw}=[ qw(gpqIpkPmqLqqOyqKpn) ];
$CFG{uploadlogs}=0;

1;
```

# 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 SF Oracle RAC.

**To configure I/O fencing using response files**

1. Make sure that SF Oracle RAC 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 22.
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 89.

   See "Sample response file for configuring server-based I/O fencing" on page 95.
4. Edit the values of the response file variables as necessary.

   See "Response file variables to configure disk-based I/O fencing" on page 86.

   See "Response file variables to configure server-based I/O fencing" on page 93.
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 3-12 lists the response file variables that specify the required information to configure disk-based I/O fencing for SF Oracle RAC.

**Table 3-12** 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.<br>(Required) |
| CFG{fencing_option} | Scalar | Specifies the I/O fencing configuration mode.<br>■ 1—Coordination Point Server-based I/O fencing<br>■ 2—Coordinator disk-based I/O fencing<br>■ 3—Disabled-based I/O fencing<br>■ 4—Online fencing migration<br>■ 5—Refresh keys/registrations on the existing coordination points<br>■ 6—Change the order of existing coordination points |
| CFG{fencing_dgname} | Scalar | Specifies the disk group for I/O fencing.<br>(Optional)<br>**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.<br>(Optional)<br>**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 3-12** 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.<br>**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.<br>Enter "0" if you do not want to configure the Coordination Point agent using the installer.<br>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.<br>**Note:** This field is obsolete if the **fencing_config_cpagent** field is given a value of '0'. |

**Table 3-12** 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 86.

```
# 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 3-13 describes the response file variables to configure CP server.

**Table 3-13** 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 3-13** 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 SFHA cluster

Review the response file variables and their definitions.

See Table 3-13 on page 90.

```
#
# 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 3-14 lists the fields in the response file that are relevant for server-based customized I/O fencing.

**Table 3-14** 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.<br>Enter "0" if you do not want to configure the Coordination Point agent using the installer.<br>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.<br>**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. |

**Table 3-14** Coordination point server (CP server) based fencing response file definitions *(continued)*

| Response file field | Definition |
|---|---|
| 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).<br>Enter either a "1" or "0".<br>Entering a "1" indicates reuse, and entering a "0" indicates do not reuse.<br>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. |
| 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. |

**Table 3-14** Coordination point server (CP server) based fencing response file definitions *(continued)*

| Response file field | Definition |
|---|---|
| CFG{fencing_option} | Specifies the I/O fencing configuration mode.<br>■ 1—Coordination Point Server-based I/O fencing<br>■ 2—Coordinator disk-based I/O fencing<br>■ 3—Disabled-based I/O fencing<br>■ 4—Online fencing migration<br>■ 5—Refresh keys/registrations on the existing coordination points<br>■ 6—Change the order of existing coordination points |
| 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
emc_clariion0_12) ];
$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;
```

Section 2

# Post-installation and configuration tasks


- Chapter 4. Verifying the installation
- Chapter 5. Performing additional post-installation and configuration tasks

Chapter 4

# Verifying the installation

This chapter includes the following topics:


- Performing a postcheck on a node
- Verifying SF Oracle RAC installation using VCS configuration file
- Verifying LLT, GAB, and cluster operation


## Performing a postcheck on a node

The installer's `postcheck` command can help you to determine installation-related problems and provide troubleshooting information.

See "About using the postcheck option" on page 349.

**To run the postcheck command on a node**

1. Run the installer with the `-postcheck` option.

   ```
   # ./installer -postcheck system_name
   ```

2. Review the output for installation-related information.

## Verifying SF Oracle RAC installation using VCS configuration file

The configuration file, main.cf, is created on each node at /etc/VRTSvcs/conf/config/. Review the main.cf configuration file after the SF Oracle RAC installation and before the Oracle installation.

Verify the following information in the main.cf file:

- The cluster definition within the main.cf includes the cluster information that was provided during the configuration. The cluster information includes the cluster name, cluster address, and the names of cluster users and administrators.
- The UseFence = SCSI3 attribute is present in the file.
- If you configured the cluster in secure mode, the “SecureClus = 1” cluster attribute is set.

For more information on the configuration file:

See “About VCS configuration file ” on page 387.

# Verifying LLT, GAB, and cluster operation

Verify the operation of LLT, GAB, and the cluster using the VCS commands.

**To verify LLT, GAB, and cluster operation**

1. Log in to any node in the cluster as superuser.
2. Make sure that the PATH environment variable is set to run the VCS commands.
3. Verify LLT operation.

   See “Verifying LLT” on page 99.
4. Verify GAB operation.

   See “Verifying GAB” on page 101.
5. Verify the cluster operation.

   See “Verifying the cluster” on page 103.

## Verifying LLT

Use the `lltstat` command to verify that links are active for LLT. If LLT is configured correctly, this command shows all the nodes in the cluster. The command also returns information about the links for LLT for the node on which you typed the command.

Refer to the `lltstat(1M)` manual page for more information.

**To verify LLT**

1. Log in as superuser on the node sys1.
2. Run the `lltstat` command on the node sys1 to view the status of LLT.

   ```
   lltstat -n
   ```

The output on sys1 resembles:

```
LLT node information:
    Node             State           Links
    *0 sys1          OPEN           2
     1 sys2          OPEN           2
```

Each node has two links and each node is in the OPEN state. The asterisk (*) denotes the node on which you typed the command.

If LLT does not operate, the command does not return any LLT links information: If only one network is connected, the command returns the following LLT statistics information:

```
LLT node information:
    Node                 State    Links
   * 0 sys1            OPEN        2
     1 sys2            OPEN        2
     2 sys5            OPEN        1
```

3 Log in as superuser on the node sys2.

4 Run the `lltstat` command on the node sys2 to view the status of LLT.

```
lltstat -n
```

The output on sys2 resembles:

```
LLT node information:
    Node             State           Links
     0 sys1          OPEN           2
    *1 sys2          OPEN           2
```

5 To view additional information about LLT, run the `lltstat -nvv` command on each node.

For example, run the following command on the node sys1 in a two-node cluster:

```
lltstat -nvv active
```

The output on sys1 resembles:

```
Node          State      Link    Status      Address
*0 sys1      OPEN
                         eth1 UP     08:00:20:93:0E:34
                         eth2 UP     08:00:20:93:0E:38
```

```
1 sys2       OPEN
                          eth1 UP     08:00:20:8F:D1:F2
                          eth2 DOWN
```

The command reports the status on the two active nodes in the cluster, sys1 and sys2.

For each correctly configured node, the information must show the following:

- A state of OPEN
- A status for each link of UP
- An address for each link

However, the output in the example shows different details for the node sys2. The private network connection is possibly broken or the information in the /etc/lltab file may be incorrect.

**6** To obtain information about the ports open for LLT, type `lltstat -p` on any node.

For example, type `lltstat -p` on the node sys1 in a two-node cluster:

```
lltstat -p
```

The output resembles:

```
LLT port information:
  Port  Usage       Cookie
  0     gab          0x0
        opens:       0 2 3 4 5 6 7 8 9 10 11 ... 60 61 62 63
        connects:    0 1
  7     gab          0x7
        opens:       0 2 3 4 5 6 7 8 9 10 11 ... 60 61 62 63
        connects:    0 1
  31    gab          0x1F
        opens:       0 2 3 4 5 6 7 8 9 10 11 ... 60 61 62 63
        connects:    0 1
```

# Verifying GAB

Verify the GAB operation using the `gabconfig -a` command. This command returns the GAB port membership information. The output displays the nodes that have membership with the modules you installed and configured. You can use GAB port membership as a method of determining if a specific component of the SF Oracle RAC stack communicates with its peers.

Table 4-1 lists the different ports that the software configures for different functions.

**Table 4-1** GAB port description

| Port | Function |
|---|---|
| a | GAB |
| b | I/O fencing |
| d | Oracle Disk Manager (ODM) |
| f | Cluster File System (CFS) |
| h | Cluster Server (VCS: High Availability Daemon) |
| m | Cluster Volume Manager (CVM)<br>CVM uses port m for SmartIO VxVM cache coherency using Group Lock Manager (GLM). |
| o | VCSMM driver |
| u | Cluster Volume Manager (CVM)<br>(to ship commands from slave node to master node)<br>Port u in the `gabconfig` output is visible with CVM protocol version >= 100. Run the `vxdctl protocolversion` command to check the protocol version. |
| v | Cluster Volume Manager (CVM) |
| w | vxconfigd (module for CVM) |
| y | Cluster Volume Manager (CVM) I/O shipping |

For more information on GAB, refer to the *Cluster Server Administrator's Guide*.

**To verify GAB**

- To verify the GAB operation, type the following command on each node:

  ```
  # /sbin/gabconfig -a
  ```

  For example, the command returns the following output:

  ```
  GAB Port Memberships
  ===============================================================
  Port a gen   58f501 membership 01
  Port b gen   58f502 membership 01
  Port d gen   58f50e membership 01
  Port f gen   58f511 membership 01
  Port h gen   58f505 membership 01
  Port m gen   58f507 membership 01
  Port o gen   58f513 membership 01
  Port u gen   58f50d membership 01
  Port v gen   58f509 membership 01
  Port w gen   58f50b membership 01
  Port y gen   58f508 membership 01
  ```

# Verifying the cluster

Verify the status of the cluster using the `hastatus` command. This command returns the system state and the group state.

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

Refer to the *Cluster Server Administrator's Guide* for a description of system states and the transitions between them.

**To verify the cluster**

1 To verify the status of the cluster, type the following command:

```
# hastatus -summary
```

The output resembles:

```
-- SYSTEM STATE
-- System              State                Frozen

A  sys1               RUNNING                0
A  sys2               RUNNING                0

-- GROUP STATE
-- Group           System      Probed  AutoDisabled   State

B  cvm             sys1         Y       N              ONLINE
B  cvm             sys2         Y       N              ONLINE
```

2 Review the command output for the following information:

- The system state
  If the value of the system state is RUNNING, the cluster is successfully started.
- The cvm group state
  In the sample output, the group state lists the cvm group, which is ONLINE on both the nodes sys1 and sys2.

# Verifying the cluster nodes

Verify the information of the cluster systems using the `hasys -display` command. The information for each node in the output should be similar.

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

Refer to the *Cluster Server Administrator's Guide* for information about the system attributes for VCS.

**To verify the cluster nodes**

◆ On one of the nodes, type the `hasys -display` command:

```
# hasys -display
```

The example in the following procedure is for SPARC and it shows the output when the command is run on the node sys1. The list continues with similar information for sys2 (not shown) and any other nodes in the cluster.

```
#System  Attribute            Value

sys1     AgentsStopped        0

sys1     AvailableCapacity    CPU 7.84

                              Mem 2600.00

                              Swap 4095.00

sys1     CPUThresholdLevel    Critical 90 Warning 80 Note 70
                              Info 60

sys1     CPUUsage             0

sys1     CPUUsageMonitoring   Enabled 0 ActionThreshold 0
                              ActionTimeLimit 0 Action  NONE
                              NotifyThreshold 0 NotifyTimeLimit 0

sys1     Capacity             CPU 8.00 Mem 3448.00 Swap 4095.00

sys1     ConfigBlockCount     293

sys1     ConfigCheckSum       37283

sys1     ConfigDiskState      CURRENT

sys1     ConfigFile           /etc/VRTSvcs/conf/config

sys1     ConfigInfoCnt        0

sys1     ConfigModDate        Wed 23 Mar 2016 12:36:17 PM IST

sys1     ConnectorState       Down

sys1     CurrentLimits

sys1     DiskHbStatus

sys1     DynamicLoad

sys1     EngineRestarted      0

sys1     EngineVersion        8.0.0.0000

sys1     FencingWeight        0

sys1     Frozen               0

sys1     GUIIPAddr

sys1     HostUtilization      CPU 2 Mem 24 Swap 0
```

```
sys1    LLTNodeId            0

sys1    LicenseType          PERMANENT_SITE

sys1    Limits

sys1    LinkHbStatus         eth1 UP eth2 UP

sys1    LoadTimeCounter      0

sys1    LoadTimeThreshold    600

sys1    LoadWarningLevel     80

sys1    MemThresholdLevel    Critical 90 Warning 80
                             Note 70 Info 60

sys1    MeterRecord          AvailableGC 2 ForecastGC 0

sys1    NoAutoDisable        0

sys1    NodeId               0

sys1    OnGrpCnt             1

sys1    PhysicalServer

sys1    ShutdownTimeout      600

sys1    SourceFile           ./main.cf

sys1    SwapThresholdLevel   Critical 90 Warning 80 Note 70
                             Info 60

sys1    Site

sys1    SysName              sys1

sys1    HostAvailableForecast

sys1    ReservedCapacity

sys1    ServerAvailableCapacity

sys1    ServerAvailableForecast

sys1    ServerCapacity

sys1    ServerReservedCapacity
```

Chapter 5

# Performing additional post-installation and configuration tasks

This chapter includes the following topics:


- Setting the environment variables for SF Oracle RAC
- About enabling LDAP authentication for clusters that run in secure mode
- About configuring authentication for SFDB tools
- Configuring Volume Replicator
- Running SORT Data Collector to collect configuration information


## Setting the environment variables for SF Oracle RAC

Set the MANPATH variable in the .profile file (or other appropriate shell setup file for your system) to enable viewing of manual pages.

Based on the shell you use, type one of the following:

For sh, ksh, or bash

```
# export MANPATH=$MANPATH:\
/opt/VRTS/man
```

For csh

```
# setenv MANPATH /usr/share/man:\
/opt/VRTS/man
```

For csh
```
# setenv MANPATH $MANPATH:/opt/VRTS/man
```

Some terminal programs may display garbage characters when you view the man pages. Set the environment variable LC_ALL=C to resolve this issue.

Set the PATH environment variable in the .profile file (or other appropriate shell setup file for your system) on each system to include installation and other commands.

Based on the shell you use, type one of the following:

For sh, ksh, or bash
```
# PATH=/usr/sbin:/sbin:/usr/bin:\
/opt/VRTS/bin\
$PATH; export PATH
```

# 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.

See the *Cluster Server Configuration and Upgrade Guide* for more information.

# 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 109. |
| 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 317. |

## 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.

# Configuring Volume Replicator

Perform this step only if you have not already configured VVR during the installation.

By default, the installer installs the required VVR configuration files irrespective of whether or not you choose to enable VVR. To configure VVR manually in SF Oracle RAC, simply start VVR using the `vxstart_vvr` command . The command starts the VVR daemons and configures the ports. You may change the default settings at any time.

For instructions on changing the default settings, see the *Volume Replicator Administrator's Guide*.

**To configure VVR**

1. Log into each node in the cluster as the root user.
2. Start VVR:

```
# vxstart_vvr start
VxVM VVR INFO V-5-2-3935 Using following ports:
heartbeat: 4145
vradmind: 8199
vxrsyncd: 8989
data: Anonymous-Ports
To change, see vrport(1M) command
VxVM VVR  V-5-2-5942 Starting Communication daemon: [OK]
```

# Running SORT Data Collector to collect configuration information

SORT Data Collector now supersedes the VRTSexplorer utility. Run the Data Collector with the `VxExplorer` option to gather information about the system.

Visit the SORT Website and download the UNIX Data Collector appropriate for your operating system.

https://sort.veritas.com/land

For more information:

https://sort.veritas.com/public/help/wwhelp/wwhimpl/js/html/wwhelp.htm

Section 3

# Upgrade of SF Oracle RAC


- Chapter 6. Planning to upgrade SF Oracle RAC
- Chapter 7. Performing a full upgrade of SF Oracle RAC using the product installer
- Chapter 8. Performing an automated full upgrade of SF Oracle RAC using response files
- Chapter 9. Performing a phased upgrade of SF Oracle RAC
- Chapter 10. Performing a rolling upgrade of SF Oracle RAC
- Chapter 11. Upgrading Volume Replicator
- Chapter 12. Performing post-upgrade tasks

Chapter 6

# Planning to upgrade SF Oracle RAC

This chapter includes the following topics:


- About the upgrade
- Supported upgrade paths
- Creating backups
- Considerations for upgrading SF Oracle RAC to 8.0 on systems configured with an Oracle resource
- 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
- Phased Upgrade
- Rolling Upgrade

Table 6-1 describes the product mapping after an upgrade.

**Table 6-1** Veritas InfoScale product mapping after upgrade

| Product (6.2.x and earlier) | Product (7.0 and later) | Component (7.0 and later) |
|---|---|---|
| SF Oracle RAC | Veritas InfoScale Enterprise | SF Oracle RAC |

**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 SF Oracle RAC installation, the installer upgrades and restores the configuration to SF Oracle RAC. 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*.

SF Oracle RAC software must be at the same version across all nodes in an SF Oracle RAC cluster after the upgrade, that is 8.0.

Review the following information before you upgrade:

- If you are running an Oracle RAC version that is not supported in this release, first upgrade Oracle RAC to the supported version before upgrading SF Oracle RAC. For instructions, see the product documentation of the version in use.

Table 6-2 lists the supported upgrade paths for upgrades on RHEL and Oracle Linux.

**Table 6-2** 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<br>RHEL 7 GA, Update 1, 2, 3, 4, 5, 6, 7, 8, 9<br>Oracle Linux 6 Update 4, 5, 6, 7<br>Oracle Linux 7 GA, Update 2, 3 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 9<br>Oracle Linux 8 Update 2, 4 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |

**Table 6-2** Supported upgrade paths on RHEL and Oracle Linux *(continued)*

| From product version | From OS version | To OS version | To product version | To Component |
|---|---|---|---|---|
| 7.1 | RHEL 6 update 4, 5, 6, 7, 8, 9<br>RHEL 7 Update 1, 2, 3, 4<br>Oracle Linux 6 Update 6, 7, 8<br>Oracle Linux 7 Update 1, 2 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 9<br>Oracle Linux 8 Update 2, 4 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.2 | RHEL 6 update 6, 7, 8, 9<br>RHEL 7 Update 1, 2, 3, 4<br>Oracle Linux 6 Update 4, 5, 6, 7, 8, 9<br>Oracle Linux 7 Update 1, 2, 3, 4 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 9<br>Oracle Linux 8 Update 2, 4 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.3 | RHEL 6 update 7, 8, 9<br>RHEL 7 Update 1,2, 3, 4<br>Oracle Linux 6 Update 7, 8, 9<br>Oracle Linux 7 Update 1, 2, 3 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 9<br>Oracle Linux 8 Update 2, 4 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |

**Table 6-2** Supported upgrade paths on RHEL and Oracle Linux *(continued)*

| From product version | From OS version | To OS version | To product version | To Component |
|---|---|---|---|---|
| 7.3.1 | RHEL 6 update 7, 8, 9<br>RHEL 7 Update 3, 4, 5<br>Oracle Linux 6 Update 7, 8, 9<br>Oracle Linux 7 Update 3, 4 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 9<br>Oracle Linux 8 Update 2, 4 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.4 | RHEL 6 Update 7, 8, 9, 10<br>RHEL 7 Update 3, 4,5<br>Oracle Linux 6 Update 7, 8, 9<br>Oracle Linux 7 Update 3, 4 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 9<br>Oracle Linux 8 Update 2, 4 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.4.1 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 4, 5, 6, 7, 8, 9<br>Oracle Linux 8 Update 1, 2 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 9<br>Oracle Linux 8 Update 2, 4 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.4.2 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 7<br>Oracle Linux 8 Update 1 | RHEL 7 Update 9<br>RHEL 8 Update 2, 4<br>Oracle Linux 7 Update 9<br>Oracle Linux 8 Update 2, 4 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |

Table 6-3 lists the supported upgrade paths for upgrades on SLES.

**Table 6-3** 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<br>SLES 12 GA, SP1, SP2, SP3, SP4 | SLES 12 SP5<br>SLES 15 SP2 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.1, 7.2 | SLES 11 SP3, SP4<br>SLES 12 GA, SP1, SP2 | SLES 12 SP5<br>SLES 15 SP2 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.3 | SLES 11 SP3, SP4<br>SLES 12 SP1, SP2 | SLES 12 SP5<br>SLES 15 SP2 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.3.1 | SLES 11 SP3, SP4<br>SLES 12 SP2, SP4, SP5 | SLES 12 SP5<br>SLES 15 SP2 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.4 | SLES 11 SP3, SP4<br>SLES 12 SP2, SP3 | SLES 12 SP5<br>SLES 15 SP2 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.4.1 | SLES 11 SP3, SP4<br>SLES 12 SP2, SP3, SP4, SP5<br>SLES 15 GA, SP1, SP2 | SLES 12 SP5<br>SLES 15 SP2 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |
| 7.4.2 | SLES 12 SP 4, SP5<br>SLES 15 SP1, SP2 | SLES 12 SP5<br>SLES 15 SP2 | Veritas InfoScale Enterprise 8.0 | SF Oracle RAC |

# 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.

# Considerations for upgrading SF Oracle RAC to 8.0 on systems configured with an Oracle resource

If you plan to upgrade SF Oracle RAC 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*.

# 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 6-4** 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) |

**Table 6-4** Release Levels *(continued)*

| Level | Content | Form factor | Applies to | Release types | Download location |
|---|---|---|---|---|---|
| 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 7

# Performing a full upgrade of SF Oracle RAC using the product installer

This chapter includes the following topics:


- About full upgrades
- Preparing to perform a full upgrade to SF Oracle RAC 8.0
- Upgrading to SF Oracle RAC 8.0


## About full upgrades

A full upgrade involves upgrading all the nodes in the cluster at the same time. All components are upgraded during the process. The cluster remains unavailable for the duration of the upgrade.

---

**Note:** You can not roll back the upgrade to a previous version after you upgrade to version 8.0.

---

You can perform the upgrade using Veritas script-based installation program, `installer`.

You can also perform a full upgrade using a response file. You can create a response file by using the response file template or by customizing a response file that is generated by the script-based installer.

# Preparing to perform a full upgrade to SF Oracle RAC 8.0

Perform the preparatory steps in this section if you are performing a full upgrade of the cluster. Before you upgrade, make sure that your systems meet the hardware and software requirements for this release.

**To prepare to upgrade SF Oracle RAC**

1. Log in as superuser to one of the nodes in the cluster.
2. Back up the following configuration files on each node: `main.cf, types.cf, CVMTypes.cf, CFSTypes.cf, OracleTypes.cf, OracleASMTypes.cf, PrivNIC.cf, MultiPrivNIC.cf, CRSResource.cf, /etc/llttab, /etc/llthosts, /etc/gabtab, /etc/vxfentab, /etc/vxfendg, /etc/vxfenmode`

   For example:

   ```
   # cp /etc/VRTSvcs/conf/config/main.cf \
   /etc/VRTSvcs/conf/config/main.cf.save
   ```

3. 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.

4. If you plan to continue using the Storage Foundation for Databases (SFDB) tools, you must prepare to migrate the SFDB repository database before upgrading to SF Oracle RAC 8.0.
5. If a cache area is online, you must take the cache area offline before upgrading the VxVM RPM. Use the following command to take the cache area offline:

   ```
   # sfcache offline cachename
   ```

6 From any node in the cluster, stop all applications that use VxFS or VxVM disk groups, whether local or CFS.

If the applications are under VCS control:

```
# hagrp -offline grp_name -any
```

If the applications are not under VCS control:

Use native application commands to stop the application.

7 From any node in the cluster, stop all Oracle RAC resources.

- If the database instances are managed by VCS, take the corresponding VCS service groups offline. As superuser, enter:

```
# hagrp -offline group_name -any
```

- If the database instances are not managed by VCS, then run the following on one node:

```
$ srvctl stop database -db db_name
```

8 - If the Oracle database is managed by VCS, set the AutoStart value to 0 to prevent the database service group from starting automatically when VCS starts:

```
# haconf -makerw
# hagrp -modify oracle_group AutoStart 0
# haconf -dump -makero
```

- If the Oracle database is not managed by VCS, change the management policy for the database to manual:

```
$ srvctl modify database -db db-name -policy manual
```

9 Stop VCS on all nodes:

```
# hastop -all
```

**10** Unmount the VxFS file system on each node, which is not under VCS control.

```
# mount |grep vxfs
```

```
# fuser -m /mount_point
```

```
# umount /mount_point
```

Make sure that no processes are running which make use of mounted shared file system or shared volumes.

```
# fuser -cu /mount_point
```

**11** If you plan to upgrade the operating system, stop all ports.

```
# ./installer -stop
```

```
# /opt/VRTS/install/installer -stop sys1 sys2
```

# Upgrading to SF Oracle RAC 8.0

For details on installing and upgrading Veritas InfoScale using the installer with the -yum option, refer to the *Veritas InfoScale Installation guide*.

**To upgrade SF Oracle RAC and operating system**

**1** If you want to upgrade the operating system, perform the following steps:

- On each node, rename the `/etc/llttab` file to prevent LLT from starting automatically when the node starts:

  ```
  # mv /etc/llttab /etc/llttab.save
  ```

- Upgrade the operating system on all nodes in the cluster.
  For instructions, see the operating system documentation.

  **Note:** If reboot is required, use `shutdown -r now` command to reboot the nodes.

- After the system restarts, restore the `/etc/llttab` file to its original name:

```
# mv /etc/llttab.save /etc/llttab
```

2 Upgrade to SF Oracle RAC 8.0.

See "Upgrading SF Oracle RAC using the Veritas installation program" on page 127.

You can also perform a silent upgrade:

See "Upgrading SF Oracle RAC using a response file" on page 130.

3 Manually mount the VxFS and CFS file systems that are not managed by VCS.

4 Relink the SF Oracle RAC libraries with Oracle.

See "Relinking Oracle RAC libraries with the SF Oracle RAC libraries" on page 167.

5 Bring the Oracle database service group online.

- If the Oracle database is managed by VCS:

```
# hagrp -online Oracle_group -any
```

- If the Oracle database is not managed by VCS:

```
# srvctl start database -db db_name
```

6 Start all applications that are not managed by VCS. Use native application commands to start the applications.

7
- If the Oracle database is managed by VCS, reset the AutoStart value to 1 to enable VCS to bring the database service group online automatically when VCS starts:

```
# haconf -makerw
# hagrp -modify oradb_grpname AutoStart 1
# haconf -dump -makero
```

- If the Oracle database is not managed by VCS, change the management policy for the database to automatic:

```
$ srvctl modify database -db db_name -policy AUTOMATIC
```

8 Complete other post-upgrade steps.

For instructions, see the chapter *Performing post-upgrade tasks* in this document.

- See "Re-joining the backup boot disk group into the current disk group" on page 169.
- See "Reverting to the backup boot disk group after an unsuccessful upgrade" on page 169.
- See "Setting or changing the product license level" on page 170.
- See "Upgrading disk layout versions" on page 171.
- See "Upgrading the disk group version" on page 173.

**9** If you want to upgrade CP server systems that use VCS or SFHA to 8.0, make sure that you upgrade VCS or SFHA on the CP server systems. Then, upgrade all application clusters to version 8.0.

For instructions to upgrade VCS or SFHA on the CP server systems, see the VCS or SFHA installation guide.

## Upgrading SF Oracle RAC using the Veritas installation program

Use the `installer` Veritas installation programs to upgrade SF Oracle RAC.

---

**Note:** Root Disk Encapsulation (RDE) is not supported on Linux from 7.3.1 onwards.

---

The installer performs the following tasks to upgrade SF Oracle RAC:

- Verifies the compatibility of the systems before the upgrade.
- Stops the SF Oracle RAC processes before the upgrade.
- Uninstalls SF Oracle RAC.
- Installs the Veritas InfoScale Enterprise 8.0 RPMs on the nodes.
- Starts SF Sybase CE 8.0 on all the nodes.
- Displays the location of the log files, summary file, and response file.

**To upgrade to SF Oracle RAC 8.0 using the installer program**

1. Start the installation program:

   ```
   # ./installer sys1 sys2
   ```

2. From the menu, choose **Upgrade a product > Full upgrade**.

   The installer verifies the systems for compatibility.

   ---

   **Note:** If `had` is stopped before upgrade, the installer displays the following warning:

   VCS is not running before upgrade. Please make sure all the configurations are valid before upgrade.

   If the configuration files are valid, you may ignore the message.

   ---

   Review the messages displayed and make sure that you meet the requirements before proceeding with the upgrade.

3. Press **Enter** to continue with the upgrade.

   Enter y to agree to the End User License Agreement (EULA).

   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 book disk group before the upgrade proceeds. If you want to split the boot disk group to create a backup, answer y.

4. Enter the name of the backup boot disk group when prompted. Press **Enter** to accept the default.

   You are prompted to start the split operation.

5. Enter **y** to continue with the split operation.

   The split operation can take some time to complete.

   ---

   **Note:** Verify the boot device from which the system is set to boot. Make sure that the system is set to start from the boot device with the required version of SF Oracle RAC.

   ---

6 Enter **y** to stop the SF Oracle RAC processes.

```
Do you want to stop SF Oracle RAC processes now? [y,n,q,?] (y)
```

The installer stops the processes and uninstalls SF Oracle RAC. After the uninstallation, the installer installs Veritas InfoScale SF Oracle RAC 8.0 and starts SF Oracle RAC 8.0 on all the nodes.

7 Complete the remaining tasks to finish the upgrade:

See "Upgrading to SF Oracle RAC 8.0" on page 125.

# 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 8

# Performing an automated full upgrade of SF Oracle RAC using response files

This chapter includes the following topics:


- Upgrading SF Oracle RAC using a response file
- Response file variables to upgrade SF Oracle RAC
- Sample response file for upgrading SF Oracle RAC


## Upgrading SF Oracle RAC using a response file

Perform the steps in the following procedure to upgrade to SF Oracle RAC 8.0 using a response file.

**To upgrade SF Oracle RAC using a response file**

1. Upgrade the operating system, if required.

   For instructions, see the operating system documentation.

2. Create a response file using one of the available options.

   **Note:** Make sure that you replace the host names in the response file with the names of the systems that you plan to upgrade.

   For information on various options available for creating a response file:

   For response file variable definitions:

   See "Response file variables to upgrade SF Oracle RAC" on page 131.

   For a sample response file:

   See "Sample response file for upgrading SF Oracle RAC" on page 134.

3. Navigate to the directory on the installation media that contains the installation program.

4. Start the installation:

   ```
   # ./installer -responsefile /tmp/response_file
   ```

   Where `/tmp/response_file` is the full path name of the response file.

5. Complete the post-upgrade steps.

   See the chapter "Performing post-upgrade tasks" in this document.

# Response file variables to upgrade SF Oracle RAC

Table 8-1 lists the response file variables that you can define to configure SF Oracle RAC.

**Table 8-1** Response file variables for upgrading SF Oracle RAC

| Variable | Description |
|---|---|
| CFG{accepteula} | Specifies whether you agree with the EULA.pdf file on the media.<br>List or scalar: scalar<br>Optional or required: required |

**Table 8-1** Response file variables for upgrading SF Oracle RAC *(continued)*

| Variable | Description |
|---|---|
| CFG{systems} | List of systems on which the product is to be installed or uninstalled.<br>List or scalar: list<br>Optional or required: required |
| CFG{keys}{keyless}<br>CFG{keys}{licensefile} | `CFG{keys}{keyless}` gives a list of keyless keys to be registered on the system.<br>`CFG{keys}{licensefile}` gives the absolute file path to the permanent license key to be registered on the system.<br>List or scalar: list<br>Optional or required: required |
| CFG{opt}{keyfile} | Defines the location of an ssh keyfile that is used to communicate with all remote systems.<br>List or scalar: scalar<br>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.<br>List or scalar: scalar<br>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.<br>List or scalar: scalar<br>Optional or required: optional |
| CFG{mirrordgname}{system} | If the root dg is encapsulated and you select split mirror is selected:<br>Splits the target disk group name for a system.<br>List or scalar: scalar<br>Optional or required: optional |

**Table 8-1** Response file variables for upgrading SF Oracle RAC *(continued)*

| Variable | Description |
|---|---|
| CFG{splitmirror}{system} | If the root dg is encapsulated and you select split mirror is selected:<br>Indicates the system where you want a split mirror backup disk group created.<br>List or scalar: scalar<br>Optional or required: optional |
| 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.<br>List or scalar: scalar<br>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 .<br>List or scalar: scalar<br>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.<br>List or scalar: scalar<br>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.<br>List or scalar: scalar<br>Optional or required: optional |

**Table 8-1** Response file variables for upgrading SF Oracle RAC *(continued)*

| Variable | Description |
|---|---|
| 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.<br>List or scalar: scalar<br>Optional or required: optional |
| 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.<br>List or scalar: scalar<br>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.<br>List or scalar: scalar<br>Optional or required: optional |
| CFG{secusrgrps} | Defines the usergroup names that are granted read access to the cluster.<br>List or scalar: scalar<br>Optional or required: optional |

# Sample response file for upgrading SF Oracle RAC

The following sample response file upgrades SF Oracle RAC to version 8.0 on two nodes, sys1 and sys2.

```
our %CFG;

$CFG{accepteula}=1;
$CFG{keys}{keyless}=[ qw(ENTERPRISE) ];
$CFG{prod}="ENTERPRISE80";
$CFG{crs_home}="/u01/app/grid/product/11.2.0/gridhome";
$CFG{db_home}="/u02/app/oracle/product/11.2.0/dbhome_1";
```

```
$CFG{opt}{gco}=1;
$CFG{opt}{noipc}=1;
$CFG{opt}{upgrade}=1;
$CFG{opt}{vr}=1;
$CFG{opt}{vvr}=1;
$CFG{oracle_group}="oinstall";
$CFG{oracle_user}="oracle";
$CFG{relink_oracle_database}=1;
$CFG{systems}=[ qw(sys1 sys2) ];
$CFG{vcs_allowcomms}=1;

1;
```

Chapter 9

# Performing a phased upgrade of SF Oracle RAC

This chapter includes the following topics:


- About phased upgrade
- Performing a phased upgrade of SF Oracle RAC from version 6.2.1 and later release


## About phased upgrade

The phased upgrade methodology involves upgrading half of the nodes in the cluster at a time.

For supported upgrade paths:

See "Supported upgrade paths" on page 114.

---

**Note:** You can perform a phased upgrade only if the Oracle RAC binaries are present on the local file system.

---

**Caution:** There is a potential for dependency problems between product components that no longer match when upgrading part of a cluster at a time. Follow the phased upgrade procedures carefully to avoid these problems.

---

**Note:** There will be some downtime involved. Review the procedures and carefully plan your downtime before proceeding with any steps. The sample procedures assume that Oracle RAC binaries are installed on local file systems for each node in the cluster.

Set up remote passwordless communication between the first sub-cluster nodes and the second sub-cluster nodes. For instructions, see the appendix *Configuring the secure shell or the remote shell for communications* in this document.

The examples in the procedures assume a four-node SF Oracle RAC cluster with the nodes *sys1* and *sys2* constituting the first half of the cluster and the nodes *sys3* and *sys4* constituting the second half of the cluster.

# Performing a phased upgrade of SF Oracle RAC from version 6.2.1 and later release

Table 9-1 illustrates the phased upgrade process. Each column describes the steps to be performed on the corresponding subcluster and the status of the subcluster when operations are performed on the other subcluster.

**Table 9-1** Summary of phased upgrade

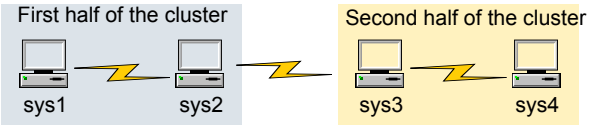
| First half of the cluster | Second half of the cluster |
|---|---|
| SF Oracle RAC cluster before the upgrade:   | |
| **STEP 1**: Perform the following pre-upgrade steps:<br>■ Switch failover applications.<br>■ Stop all parallel applications.<br>See "Step 1: Performing pre-upgrade tasks on the first half of the cluster" on page 138.<br>**STEP 2**: Upgrade SF Oracle RAC.<br>See "Step 2: Upgrading the first half of the cluster" on page 141. | The second half of the cluster is up and running. |

**Table 9-1** Summary of phased upgrade *(continued)*

| First half of the cluster | Second half of the cluster |
|---|---|
| The first half of the cluster is not running.<br> <br> | **STEP 3**: Perform the following pre-upgrade steps:<br>■ Stop all parallel and failover applications.<br>■ Stop SF Oracle RAC.<br>See "Step 3: Performing pre-upgrade tasks on the second half of the cluster" on page 142.<br>**The downtime starts now.** |
| **STEP 4**: Perform the following post-upgrade steps:<br>■ Start SF Oracle RAC.<br>■ Start all applications.<br>See "Step 4: Performing post-upgrade tasks on the first half of the cluster" on page 144.<br>**The downtime ends here.** | The second half of the cluster is not running.<br> <br> |
| The first half of the cluster is up and running.<br> <br> | **STEP 5**: Upgrade SF Oracle RAC.<br>See "Step 5: Upgrading the second half of the cluster" on page 145.<br>**STEP 6**: Perform the following post-upgrade steps:<br>■ Start SF Oracle RAC.<br>■ Start all applications.<br>See "Step 6: Performing post-upgrade tasks on the second half of the cluster" on page 146. |
| The phased upgrade is complete and both the first and the second half of the cluster are running.<br> <br> | |

# Step 1: Performing pre-upgrade tasks on the first half of the cluster

Perform the following pre-upgrade steps on the first half of the cluster.

### To perform the pre-upgrade tasks on the first half of the cluster

1. Back up the following configuration files: `main.cf, types.cf, CVMTypes.cf, CFSTypes.cf, OracleTypes.cf, OracleASMTypes.cf, , PrivNIC.cf, MultiPrivNIC.cf, /etc/llttab, /etc/llthosts, /etc/gabtab, /etc/vxfentab, /etc/vxfendg, /etc/vxfenmode`

   For example:

   ```
   # cp /etc/VRTSvcs/conf/config/main.cf \
   /etc/VRTSvcs/conf/config/main.cf.save
   # cp /etc/VRTSvcs/conf/config/types.cf \
   /etc/VRTSvcs/conf/config/types.cf.save
   # cp /etc/VRTSvcs/conf/config/OracleTypes.cf \
   /etc/VRTSvcs/conf/config/OracleTypes.cf.save
   # cp /etc/VRTSvcs/conf/config/PrivNIC.cf \
   /etc/VRTSvcs/conf/config/PrivNIC.cf.save
   # cp /etc/VRTSvcs/conf/config/MultiPrivNIC.cf \
   /etc/VRTSvcs/conf/config/MultiPrivNIC.cf.save
   ```

   The installer verifies that recent backups of configuration files in the VxVM private region are saved in `/etc/vx/cbr/bk`.

   If not, the following warning message is displayed: `Warning: Backup /etc/vx/cbr/bk directory.`

2. Stop all applications that are not configured under VCS but dependent on Oracle RAC or resources controlled by VCS. Use native application commands to stop the application.

3. Stop the applications configured under VCS. Stop the Oracle database:

   - If the Oracle RAC instance is managed by VCS:

     ```
     # hagrp -offline oracle_group -sys sys1
     # hagrp -offline oracle_group -sys sys2
     ```

   - If the Oracle RAC instance is not managed by VCS, log in as the Oracle user on one of the nodes in the first half of the cluster and shut down the instances:

     ```
     $ srvctl stop instance -db db_name \
     -node node_name
     ```

4. - If the Oracle database is managed by VCS, set the AutoStart value to 0 to prevent the database service group from starting automatically when VCS starts:

```
# haconf -makerw
# hagrp -modify oracle_group AutoStart 0
# haconf -dump -makero
```

- If the Oracle database is not managed by VCS, change the management policy for the database to manual:

```
$ srvctl modify database -db db_name -policy manual
```

5 Unmount the CFS file systems that are not managed by VCS.

- Make sure that no processes are running which make use of mounted shared file system. To verify that no processes use the CFS mount point:

```
# mount | grep vxfs | grep cluster
```

```
# fuser -cu /mount_point
```

- Unmount the CFS file system:

```
# umount /mount_point
```

6 Stop the parallel service groups and switch over failover service groups on each of the nodes in the first half of the cluster:

```
# hastop -local -evacuate
```

7 Unmount the VxFS file systems that are not managed by VCS.

- Make sure that no processes are running which make use of mounted shared file system. To verify that no processes use the VxFS mount point:

```
# mount | grep vxfs
```

```
# fuser -cu /mount_point
```

- Unmount the VxFS file system:

```
# umount /mount_point
```

8 Verify that no VxVM volumes (other than VxVM boot volumes) remain open. Stop any open volumes that are not managed by VCS.

```
# vxvol -g diskgroup stopall
# vxprint -Aht -e v_open
```

9 If a cache area is online, you must take the cache area offline before upgrading the VxVM RPM. On the nodes in the first subcluster, use the following command to take the cache area offline:

```
# sfcache offline cachename
```

10 If you plan to upgrade the operating system, stop all ports.

For versions up to 7.0:

```
# /opt/VRTS/install/installsfrac<version> -stop sys1 sys2
```

Where <*version*> is the specific release version.

For 7.0 and later versions:

```
# /opt/VRTS/install/installer -stop sys1 sys2
```

## Step 2: Upgrading the first half of the cluster

Perform the following steps to upgrade the first half of the cluster.

**To upgrade the first half of the cluster**

1 If you plan to upgrade the operating system, rename the `/etc/llttab` file to prevent LLT from starting automatically when the node starts:

```
# mv /etc/llttab /etc/llttab.save
```

2 Upgrade the operating system, if required.

For instructions, see the operating system documentation.

3 If you upgraded the operating system, restart the nodes if required:

```
# shutdown -r now
```

4 After you finish upgrading the operating system, change the LLT configuration to its original configuration.

```
# mv /etc/llttab.save /etc/llttab
```

5 Make sure that you can run secure shell or remote shell from the node where you launched the installer to the nodes in the second subcluster without requests for a password.

6 Upgrade SF Oracle RAC using the product installer. Navigate to the product directory on the installation media. When you invoke the installer, select the **Full Upgrade** option. The installer automatically detects the phased upgrade though you select the Full Upgrade option.

```
# cd product image folder
```

Using the product installer:

```
# ./installer -upgrade sys1 sys2
```

At this point, you may see the following installer prompt:

```
1)  Relink Oracle Database Binary
2)  Exit SF Oracle RAC Configuration
```

It is recommended that you choose to exit the configuration.

**Note:** After you complete the upgrade of the first half of the cluster, no GAB ports will be shown in the output when you run the `gabconfig -a` command.

## Step 3: Performing pre-upgrade tasks on the second half of the cluster

Perform the following pre-upgrade steps on the second half of the cluster.

**To perform the pre-upgrade tasks on the second half of the cluster**

1 Stop all applications that are not configured under VCS but dependent on Oracle RAC or resources controlled by VCS. Use native application commands to stop the application.

**Note:** The downtime starts now.

2 Stop all applications that are configured under VCS. Stop the Oracle database:

- If the Oracle RAC instance is managed by VCS:

```
# hagrp -offline oracle_group -sys sys3
# hagrp -offline oracle_group -sys sys4
```

- If the Oracle RAC instance is not managed by VCS, log in as the Oracle user on one of the nodes in the second half of the cluster and shut down the instances:

```
$ srvctl stop instance -db db_name \
-node node_name
```

3 Unmount the CFS file systems that are not managed by VCS.

- Make sure that no processes are running which make use of mounted shared file system. To verify that no processes use the VxFS or CFS mount point:

```
# mount | grep vxfs | grep cluster
```

```
# fuser -cu /mount_point
```

- Unmount the non-system VxFS file system:

```
# umount /mount_point
```

4 Stop VCS on each of the nodes in the second half of the cluster:

```
# hastop -local
```

5 Unmount the VxFS file systems that are not managed by VCS.

- Make sure that no processes are running which make use of mounted shared file system. To verify that no processes use the VxFS or CFS mount point:

```
# mount | grep vxfs
```

```
# fuser -cu /mount_point
```

- Unmount the non-system VxFS file system:

```
# umount /mount_point
```

6 Verify that no VxVM volumes (other than VxVM boot volumes) remain open. Stop any open volumes that are not managed by VCS.

```
# vxvol -g diskgroup stopall
# vxprint -Aht -e v_open
```

7 If a cache area is online, you must take the cache area offline before upgrading the VxVM RPM. On the nodes in the second subcluster, use the following command to take the cache area offline:

```
# sfcache offline cachename
```

8 Stop all ports.

For 6.0.1x and later versions:

```
# /opt/VRTS/install/installsfrac<version> -stop sys1 sys2
```

Where <*version*> is the specific release version.

For 7.0 and later versions:

```
# /opt/VRTS/install/installer -stop sys1 sys2
```

## Step 4: Performing post-upgrade tasks on the first half of the cluster

Perform the following post-upgrade steps on the first half of the cluster.

**To perform the post-upgrade tasks on the first half of the cluster**

1 On any one node on the first half of the cluster, force GAB to form a cluster.

To avoid an issue with starting LLT, run: `# systemctl restart veki`.

```
# systemctl start llt
# systemctl start gab
# gabconfig -x
```

2 On the first half of the cluster, start SF Oracle RAC:

```
# cd /opt/VRTS/install
# ./installer -start sys1 sys2
```

3 On the first half of the cluster, manually mount the VxFS or CFS file systems that are not managed by VCS.

4 Relink the SF Oracle RAC libraries with Oracle.

See "Relinking Oracle RAC libraries with the SF Oracle RAC libraries" on page 167.

5 Bring the Oracle database service group online.

If the Oracle database is managed by VCS:

```
# hagrp -online oracle_group -sys node_name
```

If the Oracle database is not managed by VCS:

```
$ srvctl start instance -db db_name \
-node node_name
```

**Note:** The downtime ends here.

6 On the first half of the cluster, start all applications that are not managed by VCS. Use native application commands to start the applications.

## Step 5: Upgrading the second half of the cluster

Perform the following steps to upgrade the second half of the cluster.

**To upgrade the second half of the cluster**

1 If you plan to upgrade the operating system, rename the `/etc/llttab` file to prevent LLT from starting automatically when the node starts:

```
# mv /etc/llttab /etc/llttab.save
```

2 Upgrade the operating system, if required.

For instructions, see the operating system documentation.

3 
```
# shutdown -r now
```

If you upgraded the operating system, restart the nodes:

4 After you finish upgrading the operating system, change the LLT configuration to its original configuration.

```
# mv /etc/llttab.save /etc/llttab
```

5 Make sure that you can run secure shell or remote shell from the node where you launched the installer to the nodes in the second subcluster without requests for a password.

6 On the second half of the cluster, upgrade SF Oracle RAC using the product installer. Navigate to the product directory on the installation media.

```
# cd /dvd_mount/
```

When you invoke the installer, select the **Full Upgrade** option. The installer automatically detects the phased upgrade though you select the Full Upgrade option.

Using the product installer:

```
# ./installer -upgrade sys3 sys4
```

## Step 6: Performing post-upgrade tasks on the second half of the cluster

Perform the following post-upgrade steps on the second half of the cluster.

**To perform the post-upgrade tasks on the second half of the cluster**

1 Run the following command to avoid an issue with starting LLT:

```
# systemctl restart veki
```

2 On the second half of the cluster, start SF Oracle RAC:

```
# cd /opt/VRTS/install
# ./installer -start sys3 sys4
```

3 Manually mount the VxFS and CFS file systems that are not managed by VCS.

4 Relink the SF Oracle RAC libraries with Oracle.

See "Relinking Oracle RAC libraries with the SF Oracle RAC libraries" on page 167.

5 Upgrade VxVM disk group version.

6 Upgrade disk layout version.

See "Upgrading disk layout versions" on page 171.

7 Bring the Oracle database service group online.

If the Oracle database is managed by VCS:

```
# hagrp -online oracle_group -sys sys3
# hagrp -online oracle_group -sys sys4
```

If the Oracle database is not managed by VCS:

```
$ srvctl start instance -db db_name \
-node node_name
```

8
- If the Oracle database is managed by VCS, reset the AutoStart value to 1 to enable VCS to bring the database service group online automatically when VCS starts:

  ```
  # haconf -makerw
  # hagrp -modify oracle_group AutoStart 1
  # haconf -dump -makero
  ```

- If the Oracle database is not managed by VCS, change the management policy for the database to automatic:

  ```
  $ srvctl modify database -db db_name -policy AUTOMATIC
  ```

9 Start all applications that are not managed by VCS. Use native application commands to start the applications.

10 Set or change the product license level, if required.

See "Setting or changing the product license level" on page 170.

11 Migrate the SFDB repository database.

12 If you want to upgrade application clusters that use CP server-based fencing to 8.0, make sure that you upgrade VCS or SFHA on the CP server systems. Then, upgrade all application clusters to version 8.0.

For instructions to upgrade VCS or SFHA on the CP server systems, see the VCS or SFHA installation guide.

Chapter 10

# Performing a rolling upgrade of SF Oracle RAC

This chapter includes the following topics:


- About rolling upgrade
- Preparing to perform a rolling upgrade to SF Oracle RAC 8.0
- Performing a rolling upgrade of SF Oracle RAC using the product installer
- Performing a rolling upgrade of SF Oracle RAC 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.

If you plan to upgrade the operating system, first complete the operating system upgrade on all nodes before you upgrade the product.

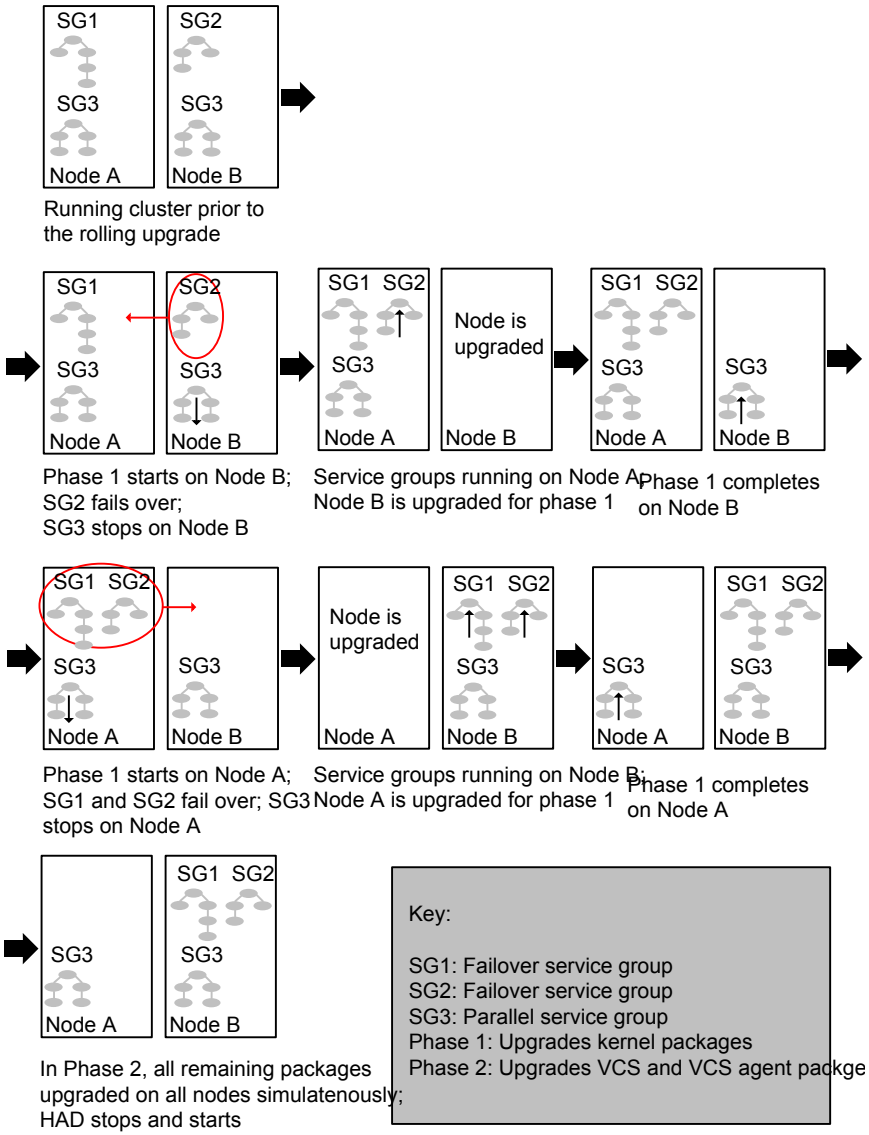
---

The following is an overview of the flow for a rolling upgrade:

1. The installer performs prechecks on the cluster.
2. 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.

**Figure 10-1** Example of the installer performing a rolling upgrade




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.

# Preparing to perform a rolling upgrade to SF Oracle RAC 8.0

Perform the preparatory steps in this section if you are performing a rolling upgrade of the cluster. Before you upgrade, make sure that your systems meet the hardware and software requirements for this release.

---

**Note:** If you plan to upgrade the operating system, ensure that you upgrade all nodes before you start the rolling upgrade of SF Oracle RAC.

---

**To prepare to upgrade SF Oracle RAC**

Perform the steps on the first subcluster.

1. Log in as superuser to one of the nodes in the subcluster.
2. Back up the following configuration files on your system: `main.cf, types.cf, CVMTypes.cf, CFSTypes.cf, OracleTypes.cf, OracleASMTypes.cf, , PrivNIC.cf, MultiPrivNIC.cf, CRSResource.cf, /etc/llttab, /etc/llthosts, /etc/gabtab, /etc/vxfentab, /etc/vxfendg, /etc/vxfenmode`

   For example:

   ```
   # cp /etc/VRTSvcs/conf/config/main.cf \
   /etc/VRTSvcs/conf/config/main.cf.save
   ```

3. 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:** Back up the `/etc/vx/cbr/bk` directory.

   ---

4 Stop the applications that use VxFS or VxVM disk groups on each node, whether local or CFS.

If the applications are under VCS control:

```
# hagrp -offline grp_name -sys node_name
```

If the applications are not under VCS control:

Use native application commands to stop the application.

5 Switch over all failover service groups to the nodes in the other subcluster:

```
# hagrp -switch grp_name -to sys_name
```

6 Stop the Oracle RAC resources on each node.

- If the database instances are managed by VCS, take the corresponding VCS service groups offline. As superuser, enter:

```
# hagrp -offline grp_name -sys node_name
```

- If the database instances are not managed by VCS, then run the following on one node:

```
$ srvctl stop instance -db db_name \
-node node_name
```

7 - If the Oracle database is managed by VCS, set the AutoStart value to 0 to prevent the database service group from starting automatically when VCS starts. Failing to perform this step results in the database attempting to come online after the upgrade; the attempt fails due to the presence of old libraries on the system.

```
# haconf -makerw
# hagrp -modify oracle_group AutoStart 0
# haconf -dump -makero
```

- If the Oracle database is not managed by VCS, change the management policy for the database to manual:

```
$ srvctl modify database -db db_name -policy manual
```

8 Unmount all the CFS file system which is not under VCS control.

```
# mount |grep vxfs | grep cluster

# fuser -m /mount_point

# umount /mount_point
```

Make sure that no processes are running which make use of mounted shared file system or shared volumes.

```
# fuser -cu /mount_point
```

9 If a cache area is online, you must take the cache area offline before upgrading the VxVM RPM. Use the following command to take the cache area offline:

```
# sfcache offline cachename
```

10 Take all the parallel VCS service groups offline on each of the nodes in the current subcluster:

```
# hagrp -offline grp_name -sys sys_name
```

11 If you plan to continue using the Storage Foundation for Databases (SFDB) tools, you must prepare to migrate the SFDB repository database before upgrading to SF Oracle RAC 8.0.

# Performing a rolling upgrade of SF Oracle RAC 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.

**To perform a rolling upgrade**

1 If you want to upgrade the operating system, perform the following steps:

- Change to the `/opt/VRTS/install` directory on the node where you want to upgrade the operating system:

```
# cd /opt/VRTS/install
```

- Stop SF Oracle RAC:

  ```
  # ./installer -stop
  ```

- Upgrade the operating system. For instructions, see the operating system documentation.
- Restart the nodes, if required:

  ```
  # shutdown -r now
  ```

2. Phase 1 of rolling upgrade begins on the first subcluster. Complete the preparatory steps on the first subcluster.
3. Log in as superuser and mount the SF Oracle RAC 8.0 installation media.
4. From root, start the installer.

   ```
   # ./installer
   ```

5. From the menu, select `Upgrade a Product` and from the sub menu, select `Rolling Upgrade`.
6. 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**.
7. The installer checks system communications, release compatibility, version information, and lists the cluster name, ID, and cluster nodes. Type **y** to continue.
8. 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.
9. 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.
10. Review the end-user license agreement, and type **y** if you agree to its terms.
11. 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.
12. 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.

---

**13** 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.

**14** 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.

**15** 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.

---

**Note:** The Oracle service group is offline as the `AutoStart` attribute is set to 0 to prevent the service group from starting automatically. The service group is started later in the process.

---

**16** Reboot the nodes on the subcluster being upgraded.

If some processes fail to start, you may need to reboot the nodes and manually check the cluster's status.

If the installer reboots the nodes, restart the installer.

**17** Manually mount the VxFS and CFS file systems that VCS does not manage.

**18** Relink the SF Oracle RAC libraries with Oracle by choosing the option **Relink Oracle Database Binary** from the program menu.

See "Relinking Oracle RAC libraries with the SF Oracle RAC libraries" on page 167.

**19** If the boot disk is encapsulated, the installer strongly recommends a restart of the nodes. Restart the nodes as prompted by the installer.

**20** Bring the Oracle database service group online.

- If VCS manages the Oracle database:

```
# hagrp -online oracle_group -sys node_name
```

- If VCS does not manage the Oracle database:

```
$ srvctl start instance -db db_name \
-node node_name
```

**21** Start all applications that are not managed by VCS. Use native application commands to start the applications.

**22** Complete the preparatory steps on the nodes that you have not yet upgraded.

**23** 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 8 through step 21.

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.

This completes phase 1 of the upgrade.

**24**
- If VCS manages the Oracle database, reset the `AutoStart` value to 1 to enable VCS to bring the database service group online when VCS starts:

```
# haconf -makerw
# hagrp -modify oracle_group AutoStart 1
# haconf -dump -makero
```

- If VCS does not manage the Oracle database, change the management policy for the database to automatic:

```
$ srvctl modify database -db db_name -policy AUTOMATIC
```

**25** Migrate the SFDB repository database.

**26** Phase 2 of the upgrade begins here. This phase includes downtime for the VCS engine (HAD), which does not include application downtime. Type **y** to continue.

**27** The installer determines the remaining RPMs to upgrade. Press **Enter** to continue.

**28** 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]

**29** 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.

**30** If you have network connection to the Internet, the installer checks for updates.

If updates are discovered, you can apply them now.

**31** 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 SF Oracle RAC 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 11

# 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.

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.

## 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 12

# Performing post-upgrade tasks

This chapter includes the following topics:


- Relinking Oracle RAC libraries with the SF Oracle RAC libraries
- Re-joining the backup boot disk group into the current disk group
- Reverting to the backup boot disk group after an unsuccessful upgrade
- Setting or changing the product license level
- Resetting DAS disk names to include host name in FSS environments
- Upgrading disk layout versions
- CVM master node needs to assume the logowner role for VCS managed VVR resources
- Switching on Quotas
- Upgrading the disk group version


## Relinking Oracle RAC libraries with the SF Oracle RAC libraries

You must relink the Oracle RAC libraries with the SF Oracle RAC libraries after upgrading SF Oracle RAC.

### To relink Oracle RAC using the installer

1. Run the installer:

```
# cd /opt/VRTS/install
# ./installer -configure sys1 sys2
```

2. Select the component **Storage Foundation for Oracle RAC (SF Oracle RAC)**.

3. Select the option **Post Oracle Installation Tasks**.

```
1)  Configure SF Oracle RAC sub-components
2)  Prepare to Install Oracle
3)  Install Oracle Clusterware/Grid Infrastructure and Database
4)  Post Oracle Installation Tasks
5)  Exit SF Oracle RAC Configuration
Choose option: [1-5,q] (1) 4
```

4. Select the option **Relink Oracle Database Binary**.

```
1)  Configure CSSD agent
2)  Relink Oracle Database Binary
3)  Exit SF Oracle RAC Configuration
b)  Back to previous menu
Choose option: [1-3,b,q] (1) 2
```

5. Provide the following information:

```
Enter Oracle UNIX user name: [b] (oracle)
Enter Oracle UNIX group name: [b] (oinstall)
Enter absolute path of Oracle Clusterware/Grid Infrastructure
Home directory: [b]
Enter absolute path of Oracle Database Home directory: [b]
```

The installer detects the Oracle version.

6. Enter **y** to proceed with relinking.

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

# 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.

# Setting or changing the product license level

SIf 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. 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.

Once you have obtained a valid license key file you can run the `vxlicinstupgrade` command to add it to your system. The license key file must not be saved in the root directory (`/`) or the default license directory on the local host (`/etc/vx/licesnes/lic`). You can save the license key file inside any other directory on the local host.

---

**Note:** You can also use the `vxkeyless` command to update the license keys to the keyless license model. Within 60 days of choosing this option, you must install a valid license key file corresponding to the entitled license level.

---

For more information and instructions, see the chapter *Licensing SF Oracle RAC*.

# 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.

# 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 *Storage Foundation for Oracle RAC Administrator's Guide*.

# 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.

# 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
```

# Upgrading the disk group version

After you upgrade from previous versions to 8.0, you have to upgrade the disk group version manually.

To upgrade disk group version, you have to first upgrade the cluster protocol version using the `vxdctl upgrade` command.

```
# vxdctl list
Volboot file
version: 3/1
seqno:   0.1
cluster protocol version: 140
hostid:  sys1
hostguid:  {fca678ac-e0ef-11e2-b22c-5e26fd3b6f13}
#
# vxdctl upgrade
#

# vxdctl list

Volboot file
version: 3/1
seqno:   0.2
cluster protocol version: 260
```

```
hostid:  sys1
hostguid:  {fca678ac-e0ef-11e2-b22c-5e26fd3b6f13}
```

Verify if the cluster protocol version shows 260 and disk group version is upgraded to 310.

```
# vxdctl list |grep version

version:   260
#
# vxdg upgrade dg_name
#
# vxdg list dg_name |grep version

version:   310
```

Section 4

# Installation of Oracle RAC


- Chapter 13. Before installing Oracle RAC
- Chapter 14. Installing Oracle RAC
- Chapter 15. Performing an automated Oracle RAC installation
- Chapter 16. Performing Oracle RAC post-installation tasks
- Chapter 17. Upgrading Oracle RAC

Chapter 13

# Before installing Oracle RAC

This chapter includes the following topics:


- Important preinstallation information for Oracle RAC
- About preparing to install Oracle RAC
- Preparing to install Oracle RAC using the SF Oracle RAC installer or manually


## Important preinstallation information for Oracle RAC

Before you install Oracle RAC, ensure that you meet the following requirements:

- Review the Oracle documentation for additional requirements pertaining to your version of Oracle.
- Keep the Oracle worksheets handy with the values appropriate for your installation setup.
  See "Oracle RAC worksheet" on page 373.

---

**Note:** The manual procedures use variables, which are indicated by *italicized* text. Ensure that you replace these variables with actual values.

---

## About preparing to install Oracle RAC

Use one of the following ways to perform the pre-installation tasks:

| | |
|---|---|
| SF Oracle RAC installer | The SF Oracle RAC installer provides a menu-driven command line interface to step you through the pre-installation tasks.<br>**Note:** Some of the pre-installation steps are not supported by the SF Oracle RAC installer and must be done manually as described in the manual procedures. |
| Manual | You need to perform the pre-installation tasks manually as described in the manual procedures. |
| Response file | You can pre-configure the systems for Oracle RAC installation using an SF Oracle RAC response file. The SF Oracle RAC response file in tandem with the Oracle RAC response files simplify the process of automating and standardizing Oracle RAC installations.<br>**Note:** You can use the response file to automate only those pre-configuration tasks that are supported by the SF Oracle RAC installer.<br>For instructions, see the chapter "Installing Oracle RAC using a response file" in this document. |

The examples in this chapter assume a two-node cluster comprising the nodes sys1 and sys2.

Before installing Oracle RAC, review the Oracle installation manuals and the appropriate Oracle support Web sites. Some of the pre-installation tasks, wherever indicated in the document, must be done in accordance with the instructions in the Oracle installation manuals. The instructions for these tasks are not provided in this document.

---

**Note:** The instructions in this chapter use variables and sample values whereever required. Replace these variables and sample values with values that conform to your installation requirements.

Before you start the preparatory tasks, you may want to update the sample worksheet with the correct installation values and keep them handy during the process.

---

# Preparing to install Oracle RAC using the SF Oracle RAC installer or manually

This section provides instructions for both manual as well as SF Oracle RAC installer-based procedures.

**To prepare to install Oracle RAC**

1. Identify the public virtual IP addresses for use by Oracle.

   See "Identifying the public virtual IP addresses for use by Oracle" on page 178.

2. Set the kernel parameters.

   See "Setting the kernel parameters" on page 179.

3. Verify that RPMs and patches required by Oracle are installed.

   See "Verifying that RPMs and patches required by Oracle are installed" on page 179.

4. Verify that the user `nobody` exists.

   See "Verifying the user `nobody` exists" on page 180.

5. Launch the SF Oracle RAC installer.

   See "Launching the SF Oracle RAC installer" on page 180.

6. Create Oracle user and groups.

   See "Creating users and groups for Oracle RAC" on page 180.

7. Create the storage for OCR and voting disk.

   See "Creating storage for OCR and voting disk " on page 184.

8. Configure the private network for Oracle RAC.

   See "Configuring private IP addresses for Oracle RAC" on page 200.

9. Verify that multicast is functional on all private network interfaces.

   See "Verifying that multicast is functional on all private network interfaces" on page 204.

10. Create the Oracle Clusterware/Grid Infrastructure and Oracle database home directories manually.

    See "Creating Oracle Clusterware/Grid Infrastructure and Oracle database home directories manually" on page 204.

11. Set up Oracle user equivalence on all nodes.

    See "Setting up user equivalence" on page 210.

## Identifying the public virtual IP addresses for use by Oracle

Identify separate public virtual IP addresses for each node in the cluster. Oracle requires one public virtual IP address for the Oracle listener process on each node. Public virtual IP addresses are used by client applications to connect to the Oracle database. Oracle Clusterware/Grid Infrastructure manages the virtual IP addresses.

The IP address and the corresponding host name should be registered in the domain name service (DNS). Alternatively, an entry for the virtual IP address and virtual public name can be placed in the `/etc/hosts` file as shown in the following example:

```
10.182.79.239 sys1-vip
10.182.79.240 sys2-vip
```

The `/etc/hosts` file on each node of the cluster should have these entries.

Oracle recommends that the public node name for the virtual IP address be in the following format *hostname-vip*. For example, sys1-vip.

---

**Note:** The public node name (in other words, the alias for the virtual IP address) for the nodes must be different from the host's current fully qualified domain name (FQDN).

---

For Oracle RAC 11g Release 2 and later versions: Additionally, you need a Single Client Access Name (SCAN) registered in Enterprise DNS that resolves to three IP addresses (recommended) using a round robin algorithm or at least one IP address. The IP addresses must be on the same subnet as your public network in the cluster. SCAN provides a single name for clients to access an Oracle database running in a cluster.

---

**Note:** The virtual IP addresses that are used for SCAN IP resolution must be on the same subnet. Oracle RAC does not support their configuration on different subnets.

---

## Setting the kernel parameters

Set the kernel parameter values to meet Oracle RAC deployment requirements. The Oracle Universal Installer (OUI) verifies the settings at the time of installation to ensure that they satisfy the minimum requirements. The Oracle RAC installation fails if the kernel parameters are not configured properly. The settings can also be tuned to optimize system performance.

For instructions and guidelines, see the Oracle Metalink document: 169706.1

## Verifying that RPMs and patches required by Oracle are installed

Oracle may require certain RPMs and patches to be installed before the installation of Oracle RAC.

For the list of RPMs and patches and related instructions, see the Oracle documentation.

## Verifying the user `nobody` exists

To verify the user "nobody" exists on each system in the cluster:

```
# id nobody
uid=65001(nobody) gid=65001(nobody) groups=65001(nobody)
```

If the user does not exist, create the user:

```
# groupadd -g 65001 nobody
# useradd -g 65001 -u 65001 nobody
```

---

**Note:** Make sure that the user ID and group ID are the same across the nodes in your cluster.

---

## Launching the SF Oracle RAC installer

You can use the installation program to complete the Oracle RAC preparatory tasks.

**To launch the SF Oracle RAC script-based installer**

1. Log in as the root user on any one node and start the installation program:

   ```
   # cd /opt/VRTS/install
   ```

   ```
   # ./installer -configure sys1 sys2
   ```

2. Review the installer instructions and press **Enter** to proceed. From the configuration program menu, select **Prepare to Install Oracle**.
3. Select the pre-installation task you want to perform.
   - Enter **1** to select the option **Create Oracle Users and Groups**:
     See "Creating users and groups for Oracle RAC" on page 180.
   - Enter **2** to select the option **Create Storage for OCR and voting disk**:
     See "Creating storage for OCR and voting disk " on page 184.
   - Enter **3** to select the option **Oracle Network Configuration**:

## Creating users and groups for Oracle RAC

Create the following groups and users for Oracle RAC 11g Release 2 and later versions:

- Oracle Inventory group
- dba group

- Oracle grid user
- Oracle user

Create additional users and groups as required by Oracle. Before creating Oracle users and groups, see the Oracle documentation.

You must assign Oracle Inventory as the primary group and dba as the secondary group. Oracle requires secondary groups for identifying operating system accounts that have database administrative (SYSDBA) privileges and for those accounts that have limited sets of database administrative (SYSOPER) privileges. Create the groups on all systems and assign the Oracle user to these groups.

Use one of the following ways to create the Oracle user and groups:

| | |
|---|---|
| Using the SF Oracle RAC script-based installer | See "Creating the Oracle user and groups using the SF Oracle RAC script-based installer" on page 181. |
| Manual | See "Creating the Oracle user and groups manually" on page 183. |

## Creating the Oracle user and groups using the SF Oracle RAC script-based installer

This procedure provides instructions for creating the Oracle user and groups using the SF Oracle RAC installer.

**To create the Oracle user and groups on all nodes in the cluster**

1. From the SF Oracle RAC installer menu, enter **1** to select the option **Create Oracle Users and Groups**.

2. Provide the following information for creating the Oracle user and groups: user name, group name, user ID, group ID, and the full path of the Oracle user home directory.

   The user ID and group ID must not be in use on any node in the cluster. The installer suggests unused values, which you may use or change as required. The configuration program assigns the same values on all the nodes.

   **Note:** If you are configuring GCO, then the user IDs and group IDs of all nodes on both the primary and secondary clusters must be the same. While configuring the user ID and group ID values on the secondary site, make sure that they are identical to the values used at the primary site.

```
Enter Oracle UNIX user name: [b] grid
Enter Oracle user's ID (numerical):  [b] (1168)
Enter Oracle UNIX group name: [b] (oinstall)
Enter Oracle group's ID (numerical):  [b] (1000)
Enter absolute path of Oracle user's Home directory: [b] /home/grid
```

3. Enter the information for the secondary group.

```
Do you want to create a secondary group
for Oracle user? [y,n,q,b,?] (n) y
Enter Oracle UNIX group name:  [b] (dba, oper etc.)
Enter Oracle group's ID (numerical): [b] (1996)

    Creating secondary group dba for
    Oracle user oracle on sys1 ..... Done
    Creating secondary group dba for
    Oracle user oracle on sys2 ..... Done
Do you want to create another secondary group
for Oracle user? [y,n,q,b,?] (n)
```

4 Create a password for the oracle user on each node.

```
# passwd oracle
```

For Oracle RAC 12c Release 2 and later versions: Create a password for the grid user on each node.

```
# passwd grid
```

5 Set up passwordless SSH as the oracle/grid user to install Oracle RAC binaries.

For instructions, see the appendix *Configuring the secure shell or the remote shell for communications* in this document.

## Creating the Oracle user and groups manually

Depending on the Oracle RAC version, create the necessary Oracle groups and users. Be sure to assign the same group ID, user ID, and home directory for the user on each system.

---

**Note:** When you create the user and group, make sure that you specify a user and group ID that is not in use.

---

**To create the operating system Oracle user and group on each system**

1 Create the primary and secondary group on each system.

Primary group:

```
# groupadd -g grp_id grp_name
```

Secondary group:

```
# groupadd -g grp_id_sec grp_name_sec
```

2 Create the Oracle user and the user home directory on each system:

```
# useradd -g grp_name -u user_id \
-G grp_name_sec -m -d user_home user_name
```

For Oracle RAC 12c Release 2 and later versions: Repeat the steps to create the grid user.

3 Create a password for the oracle user on each node.

```
# passwd oracle
```

For Oracle RAC 12c Release 2 and later versions: Create a password for the grid user on each node.

```
# passwd grid
```

4 Set up passwordless SSH as the oracle/grid user to install Oracle RAC binaries.

For instructions, see the appendix *Configuring the secure shell or the remote shell for communications* in this document.

## Creating storage for OCR and voting disk

Create appropriate storage for Oracle Cluster Registry (OCR) and the voting disk depending on the version of Oracle RAC.

---

**Note:** SF Oracle RAC supports both normal and external redundancy for OCR and voting disk storage. If you plan to use normal redundancy for OCR and voting disk, ensure that you follow the manual procedures for the same.

---

| | |
|---|---|
| Oracle RAC 12c R2 | The OCR and the voting disk resides only on Oracle ASM disk group.<br>**Note:** To prevent OCR data from failures, you can use CVM raw volumes to create ASM disk groups. |
| Other supported Oracle RAC versions | The OCR and the voting disk can reside on ASM or in directories in a cluster file system.<br>**Note:** You can use CVM raw volumes to create ASM disk groups. |

Use one of the following ways to create the storage:

| | |
|---|---|
| SF Oracle RAC script-based installer | See "Creating storage for OCR and voting disk using the SF Oracle RAC script-based installer" on page 185. |
| Manual | See "Creating storage for OCR and voting disk manually" on page 190. |

You need to create CVM volumes or a CFS mount point for database file storage later in the installation process:

See "Creating the Oracle RAC database" on page 256.

## Creating storage for OCR and voting disk using the SF Oracle RAC script-based installer

Oracle RAC provides the following levels of redundancy for OCR and voting disk:

- Normal
- High
- External

Different redundancy levels require different storage configurations.

SF Oracle RAC supports all of the redundancy levels with the following storage configuration:

| | |
|---|---|
| Normal | Requires multiple OCR and voting disk locations. |
| High | Requires multiple OCR and voting disk locations. |
| External | Requires mirroring of volumes. The CVM volumes are mirrored by SF Oracle RAC. |

The InfoScale Enterprise installer enables you to create OCR and voting disk storage on a clustered file system. The file system may reside on shared storage or on Direct Attached Storage (DAS) or internal disk drives shared over the network. After creating the storage, the installer adds the storage configuration to VCS for high availability.

However, for Oracle 12cR2, the OCR and voting disks are stored only on Oracle ASM. Hence, create Oracle ASM on raw CVM volumes during Oracle Grid installation. You can then use the Oracle ASM disk group to manage Oracle OCR and Voting files.

The installer performs the following tasks:

- Creates CVM volumes for OCR and voting disk (two-way mirrored or unmirrored)
- Creates the OCR and voting disk volumes and sets the ownership
- Starts the volumes
- Creates the mount point and mounts it on all the nodes
- Sets the ownership for the CFS mount point
- Adds the CFSMount and CVMVolDg resources to the VCS configuration in the cvm group so that the resources are brought online automatically when the nodes start

- Brings the CFSMount and CVMVolDg resources online

**To create storage for OCR and voting disk using the SF Oracle RAC installer**

1. From the SF Oracle RAC installer menu, enter **2** to select the option **Create Storage for OCR and Voting disk**.

   The following menu displays:

   ```
   1)  Create Oracle Users and Groups
   2)  Create Storage for OCR and Voting disk
   3)  Oracle Network Configuration
   4)  Exit SF Oracle RAC Configuration
   b)  Back to previous menu
   Choose option: [1-4,b,q] (1) 2
   ```

2. Enter **y** to create the storage.

   ```
   Do you want the installer to assist you in creating disk groups,
   volumes and file systems for Oracle? [y,n,q] (n) y
   ```

   If you want to create the storage manually, enter **n**. The installer displays instructions for creating the storage manually. You may skip the remaining steps.

**3** Select an appropriate option for the disk group.

```
1)  Create a disk group (Shared Storage)
2)  Use an existing disk group (Shared Storage)
3)  Create a disk group (FSS)
4)  Use an existing disk group (FSS)
b)  Back to previous menu
Choose option: [1-4,b,q]
```

If you want to mirror the CVM volumes, click **Yes** at the following prompt:

```
Do you want to enable mirroring? [y,n,q] (y) y
```

Based on the options you choose for a disk group:

| | |
|---|---|
| Create a disk group<br>(Shared Storage) | The installer displays the list of existing shared disks that do not belong to any disk group. Specify the disk (by entering the serial numbers displayed next to the disk name) that you want to use to create the disk group. If you chose to mirror the CVM volumes, select at least two disks.<br>Enter the name of the disk group.<br>`Enter the disk group name:  [b] (ocrvotedg)` |
| Use an existing disk group<br>(Shared Storage) | If you choose to use an existing disk group, the installer displays the list of existing non-FSS disk groups. Select a disk group. If you chose to mirror the CVM volumes, select an existing disk group that contains at least two disks for mirroring. |

| | |
|---|---|
| Create a disk group<br>(FSS) | The installer displays the list of existing local disks that do not belong to any disk group. Specify the disk (by entering the serial numbers displayed next to the disk name) that you want to use to create the disk group. Mirroring is enabled by default. Select at least two disks from different nodes. If you select only one disk or or if you select all disks from one node, the installer will prompt to choose disks again<br><br>`The following disks are not part of any disk group`<br>`1) emc_clariion0_103 auto:cdsdisk -`<br>`- online sys1`<br>`2) emc_clariion0_79 auto:cdsdisk -`<br>`- online exported sys2`<br>`b) Back to previous menu`<br><br>If no eligible disks are found, the installer displays the message:<br><br>`No eligible disk found`<br><br>Enter the name of the disk group.<br><br>`Enter the disk group name: [b] (ocrvotedg)` |
| Use an existing disk group<br>(FSS) | If you choose to use an existing disk group, the installer displays the list of usable FSS disk groups. Disk groups that contain disks from only one node are not listed.<br><br>To view the list of disks that are contained in the disk group, run the following command:<br><br># **vxdisk -g *disk_group* list** |

**4** Review and confirm the configuration information displayed:

```
CVM Master node: sys1
Selected disks (including mirroring):
   1. Disk_2
   2. Disk_3
Disk group name: ocrvotedg
Is this information correct? [y,n,q] (y)
```

The installer initializes the disk groups.

5 Specify whether you want to create separate file systems for OCR and voting disk:

```
Do you want to create separate filesystems for ocr and vote?
[y,n,q] (y)
```

6 Enter the name and size of the volume on which you want to store OCR and voting disk information.

- If you have chosen to create a shared file system for OCR and voting disk, enter the following information:

```
Enter the volume name for OCR and Voting disk:
[b] (ocrvotevol)
Enter the volume size for OCR and Voting disk in MB
(Refer to Oracle documentation to set it properly): [b] (920)
```

- If you have chosen to create separate file systems for OCR and voting disk, enter the following information:

```
Enter the volume name for OCR: [b] (ocrvol)
Enter the volume size for OCR in MB
(Refer to Oracle documentation to set it properly): [b] (600)
Enter the volume name for Vote: [b] (votevol)
Enter the volume size for Vote: [b] (320)
```

---

**Note:** For Oracle 12cR2, create Oracle ASM disk group on the raw CVM volumes where you plan to store OCR and voting files.

---

7 Enter the Oracle UNIX user name.

```
Enter Oracle UNIX user name:  [b]
```

8 Enter the Oracle UNIX group name.

```
Enter Oracle UNIX group name: [b] (oinstall)
```

9 Press **Return** to continue.

10 Review and confirm the configuration information. The installer creates and starts the volumes on all nodes in the cluster.

11 Enter the CFS mount point for OCR and voting disk information.

- If you have chosen to create a shared file system for OCR and voting disk, enter the following information:

```
Enter the mount point location for CFS
(common for all the nodes) [b] (/ocrvote)
```

- If you have chosen to create separate file systems for OCR and voting disk, enter the following information:

```
Enter the mount point location for OCR storage
(common for all the nodes): [b] (/ocr)
Enter the mount point location for Vote storage
(common for all the nodes): [b] (/vote)
```

The installer creates the CFS mount points and sets the ownership. Press **Return** to continue.

**12** Verify that the corresponding CVMVolDg and CFSMount resources are online on all nodes in the cluster:

---

**Note:** It takes a few minutes for the CVMVolDg resource to come online.

---

```
# hares -state ocrvotemnt_resname

# hares -state ocrvotevol_resname
```

## Creating storage for OCR and voting disk manually

You can create the storage for OCR and voting disk on a Veritas clustered file system. Whether you create volumes or file system directories, you can add them to the VCS configuration to make them highly available.

The following procedures assume external redundancy for OCR and voting disk.

**To create the storage for OCR and voting disks on CFS**

1. Create a shared VxVM disk group:

```
# vxdg -s init ocrvote_dgname
     disk_name2
     disk_name3
```

2. From the CVM master, create a mirrored volume (for example `ocrvotevol`) for OCR and voting disk:

```
# vxassist -g ocrvote_dgname make ocrvote_volname 640M nmirrors=2
```

3. From the CVM master, create a file system with the volume (`ocrvotevol`).

```
# mkfs -t vxfs /dev/vx/rdsk/ocrvote_dgname/ocrvote_volname
```

4. On each system, create a directory (for example, `/ocrvote`) on which to mount the file system containing OCR and voting disk.

```
# mkdir /ocrvote_mnt
```

5. On each system, mount the file system containing OCR and voting disk:

```
# mount -t vxfs -o cluster /dev/vx/dsk/ocrvote_dgname/\
ocrvote_volname
     ocrvote_mnt
```

6. From any system, change permissions on the file system containing OCR and voting disk.

   For example:

```
# chown -R grid:oinstall ocrvote_mnt
```

7. Add the storage resources to the VCS configuration to make them highly available.

   See "Adding the storage resources to the VCS configuration" on page 191.

## Adding the storage resources to the VCS configuration

You need to add the CVMVolDg and CFSMount resources to the VCS configuration.

---

**Note:** Set the attribute "Critical" to "0" for all the resources in the cvm service group. This ensures that critical CVM and CFS resources are always online.

---

**To add the storage resources created on CFS to the VCS configuration**

1. Change the permission on the VCS configuration file to read-write mode:

```
# haconf -makerw
```

2. Configure the CVM volumes under VCS:

```
# hares -add ocrvotevol_resname CVMVolDg cvm_grpname

# hares -modify ocrvotevol_resname Critical 0

# hares -modify ocrvotevol_resname CVMDiskGroup ocrvote_dgname

# hares -modify ocrvotevol_resname CVMVolume -add ocrvote_volname

# hares -modify ocrvotevol_resname CVMActivation sw
```

3. Set up the file system under VCS:

```
# hares -add ocrvotemnt_resname CFSMount cvm_grpname

# hares -modify ocrvotemnt_resname Critical 0

# hares -modify ocrvotemnt_resname MountPoint ocrvote_mnt

# hares -modify ocrvotemnt_resname BlockDevice  \
/dev/vx/dsk/ocrvote_dgname/ocrvote_volname
```

4. Link the parent and child resources:

```
# hares -link ocrvotevol_resname cvm_clus

# hares -link ocrvotemnt_resname ocrvotevol_resname

# hares -link ocrvotemnt_resname vxfsckd
```

5. Enable the resources:

```
# hares -modify ocrvotevol_resname Enabled 1

# hares -modify ocrvotemnt_resname Enabled 1

# haconf -dump -makero
```

**6** Verify the configuration of the CVMVolDg and CFSMount resources in the main.cf file.

For example:

```
CFSMount ocrvote_mnt_ocrvotedg (
                Critical = 0
                MountPoint = "/ocrvote"
                BlockDevice = "/dev/vx/dsk/ocrvotedg/ocrvotevol"
                )
CVMVolDg ocrvote_voldg_ocrvotedg (
               Critical = 0
               CVMDiskGroup = ocrvotedg
               CVMVolume = { ocrvotevol }
               CVMActivation = sw
               )
ocrvote_mnt_ocrvotedg requires ocrvote_voldg_ocrvotedg
ocrvote_mnt_ocrvotedg requires vxfsckd
ocrvote_voldg_ocrvotedg requires cvm_clus
```

**7** Bring the CFSMount and CVMVolDg resources online on all systems in the cluster:

```
# hares -online ocrvotevol_resname -sys node_name
```

```
# hares -online ocrvotemnt_resname -sys node_name
```

Verify that the resources are online on all systems in the cluster:

```
# hares -state ocrvotevol_resname
```

```
# hares -state ocrvotemnt_resname
```

## Configuring atleast resource dependency for OCR and voting disks

If you select **Normal Redundancy** for OCR and voting disks during the installation of Oracle RAC, ensure that you configure atleast resource dependency for high availability of the OCR and voting disk.

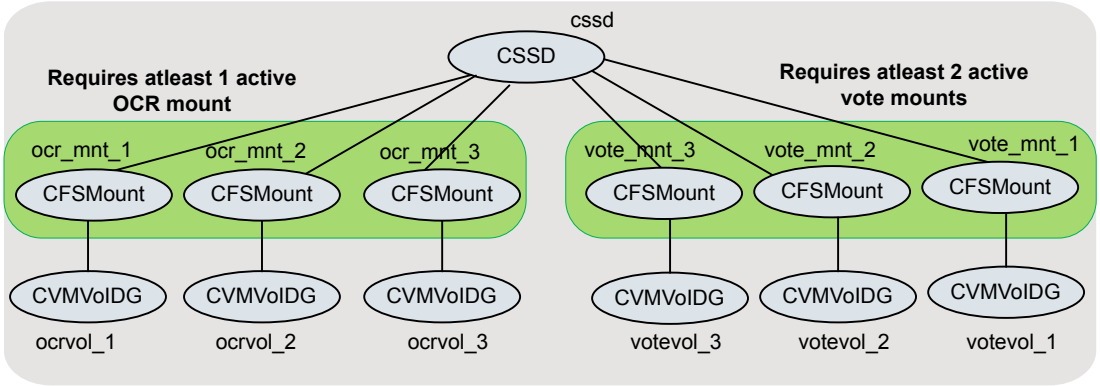
Figure 13-1 illustrates the atleast resource dependency for OCR and voting disks.

---

**Note:** See the Oracle documentation to determine the minimum OCR (x) and voting disks (y) required by Oracle Clusterware/Grid.

---

**Figure 13-1** Atleast resource dependency for OCR and voting disks




Perform the steps in this procedure to configure atleast resource dependency for OCR and voting disks.

**To configure atleast resource dependency for OCR and voting disks**

1 Create the OCR and voting disk disk groups:

```
# vxdg -s init ocrdg_1_name disk_name1
# vxdg -s init ocrdg_2_name disk_name2
# vxdg -s init ocrdg_3_name disk_name3
# vxdg -s init votedg_1_name disk_name4
# vxdg -s init votedg_2_name disk_name5
# vxdg -s init votedg_3_name disk_name6
```

2 Create volumes for OCR and voting disk:

**Note:** The size will vary depending on the Oracle RAC version you use.

```
# vxassist -g ocrdg_1_name make ocrvol_1_name 640M
# vxassist -g ocrdg_2_name make ocrvol_2_name 640M
# vxassist -g ocrdg_3_name make ocrvol_3_name 640M
# vxassist -g votedg_1_name make votevol_1_name 640M
# vxassist -g votedg_2_name make votevol_2_name 640M
# vxassist -g votedg_3_name make votevol_3_name 640M
```

**3** From the CVM master, create a file system with the volumes:

```
# mkfs -t vxfs /dev/vx/rdsk/ocrdg_1_name/ocrvol_1_name
# mkfs -t vxfs /dev/vx/rdsk/ocrdg_2_name/ocrvol_2_name
# mkfs -t vxfs /dev/vx/rdsk/ocrdg_3_name/ocrvol_3_name
# mkfs -t vxfs /dev/vx/rdsk/votedg_1_name/votevol_1_name
# mkfs -t vxfs /dev/vx/rdsk/votedg_2_name/votevol_2_name
# mkfs -t vxfs /dev/vx/rdsk/votedg_3_name/votevol_3_name
```

**4** On each system, create a directory on which to mount the file system containing the OCR and voting disks:

```
# mkdir /ocr_mnt_1
# mkdir /ocr_mnt_2
# mkdir /ocr_mnt_3
# mkdir /vote_mnt_1
# mkdir /vote_mnt_2
# mkdir /vote_mnt_3
```

**5** On each system, mount the file system containing OCR and voting disk:

```
# mount -t vxfs -o cluster /dev/vx/dsk/ocrdg_1_name/\
ocrvol_1_name /ocr_mnt_1
# mount -t vxfs -o cluster /dev/vx/dsk/ocrdg_2_name/\
ocrvol_2_name /ocr_mnt_2
# mount -t vxfs -o cluster /dev/vx/dsk/ocrdg_3_name/\
ocrvol_3_name /ocr_mnt_3
# mount -t vxfs -o cluster /dev/vx/dsk/votedg_1_name/\
votevol_1_name /vote_mnt_1
# mount -t vxfs -o cluster /dev/vx/dsk/votedg_2_name/\
votevol_2_name /vote_mnt_2
# mount -t vxfs -o cluster /dev/vx/dsk/votedg_3_name/\
votevol_3_name /vote_mnt_3
```

6 From any system, change the permissions on the file system containing the OCR and voting disk:

```
# chown -R grid:oinstall /ocr_mnt_1
# chown -R grid:oinstall /ocr_mnt_2
# chown -R grid:oinstall /ocr_mnt_3
# chown -R grid:oinstall /vote_mnt_1
# chown -R grid:oinstall /vote_mnt_2
# chown -R grid:oinstall /vote_mnt_3
```

7 Add the storage resources to the VCS configuration to make them highly available.

See "To add the storage resources to the VCS configuration" on page 197.

### To add the storage resources to the VCS configuration

1 Change the permission on the VCS configuration file to read-write mode:

```
# haconf -makerw
```

2 Configure the CVM volumes under VCS:

```
# hares -add ocrvol_1_resname CVMVolDg cvm_grpname
# hares -modify ocrvol_1_resname Critical 0
# hares -modify ocrvol_1_resname CVMDiskGroup ocrdg_1_name
# hares -modify ocrvol_1_resname CVMVolume -add ocrvol_1_name
# hares -modify ocrvol_1_resname CVMActivation sw

# hares -add ocrvol_2_resname CVMVolDg cvm_grpname
# hares -modify ocrvol_2_resname Critical 0
# hares -modify ocrvol_2_resname CVMDiskGroup ocrdg_2_name
# hares -modify ocrvol_2_resname CVMVolume -add ocrvol_2_name
# hares -modify ocrvol_2_resname CVMActivation sw

# hares -add ocrvol_3_resname CVMVolDg cvm_grpname
# hares -modify ocrvol_3_resname Critical 0
# hares -modify ocrvol_3_resname CVMDiskGroup ocrdg_3_name
# hares -modify ocrvol_3_resname CVMVolume -add ocrvol_3_name
# hares -modify ocrvol_3_resname CVMActivation sw

# hares -add votevol_1_resname CVMVolDg cvm_grpname
# hares -modify votevol_1_resname Critical 0
# hares -modify votevol_1_resname CVMDiskGroup votedg_1_name
# hares -modify votevol_1_resname CVMVolume -add votevol_1_name
# hares -modify votevol_1_resname CVMActivation sw

# hares -add votevol_2_resname CVMVolDg cvm_grpname
# hares -modify votevol_2_resname Critical 0
# hares -modify votevol_2_resname CVMDiskGroup votedg_2_name
# hares -modify votevol_2_resname CVMVolume -add votevol_2_name
# hares -modify votevol_2_resname CVMActivation sw

# hares -add votevol_3_resname CVMVolDg cvm_grpname
# hares -modify votevol_3_resname Critical 0
# hares -modify votevol_3_resname CVMDiskGroup votedg_3_name
# hares -modify votevol_3_resname CVMVolume -add votevol_3_name
# hares -modify votevol_3_resname CVMActivation sw
```

**3** Set up the file system under VCS:

```
# hares -add ocrmnt_1_resname CFSMount cvm_grpname
# hares -modify ocrmnt_1_resname Critical 0
# hares -modify ocrmnt_1_resname MountPoint /ocr_mnt_1
# hares -modify ocrmnt_1_resname BlockDevice \
/dev/vx/dsk/ocrdg_1_name/ocrvol_1_name

# hares -add ocrmnt_2_resname CFSMount cvm_grpname
# hares -modify ocrmnt_2_resname Critical 0
# hares -modify ocrmnt_2_resname MountPoint /ocr_mnt_2
# hares -modify ocrmnt_2_resname BlockDevice \
/dev/vx/dsk/ocrdg_2_name/ocrvol_2_name

# hares -add ocrmnt_3_resname CFSMount cvm_grpname
# hares -modify ocrmnt_3_resname Critical 0
# hares -modify ocrmnt_3_resname MountPoint /ocr_mnt_3
# hares -modify ocrmnt_3_resname BlockDevice \
/dev/vx/dsk/ocrdg_3_name/ocrvol_3_name

# hares -add votemnt_1_resname CFSMount cvm_grpname
# hares -modify votemnt_1_resname Critical 0
# hares -modify votemnt_1_resname MountPoint /vote_mnt_1
# hares -modify votemnt_1_resname BlockDevice \
/dev/vx/dsk/votedg_1_name/votevol_1_name

# hares -add votemnt_2_resname CFSMount cvm_grpname
# hares -modify votemnt_2_resname Critical 0
# hares -modify votemnt_2_resname MountPoint /vote_mnt_2
# hares -modify votemnt_2_resname BlockDevice \
/dev/vx/dsk/votedg_2_name/votevol_2_name

# hares -add votemnt_3_resname CFSMount cvm_grpname
# hares -modify votemnt_3_resname Critical 0
# hares -modify votemnt_3_resname MountPoint /vote_mnt_3
# hares -modify votemnt_3_resname BlockDevice \
/dev/vx/dsk/votedg_3_name/votevol_3_name
```

**4** Link the parent and child resources:

```
# hares -link ocrvol_1_resname cvm_clus
# hares -link ocrmnt_1_resname ocrvol_1_resname
# hares -link ocrmnt_1_resname vxfsckd

# hares -link ocrvol_2_resname cvm_clus
# hares -link ocrmnt_2_resname ocrvol_2_resname
# hares -link ocrmnt_2_resname vxfsckd

# hares -link ocrvol_3_resname cvm_clus
# hares -link ocrmnt_3_resname ocrvol_3_resname
# hares -link ocrmnt_3_resname vxfsckd

# hares -link votevol_1_resname cvm_clus
# hares -link votemnt_1_resname votevol_1_resname
# hares -link votemnt_1_resname vxfsckd

# hares -link votevol_2_resname cvm_clus
# hares -link votemnt_2_resname votevol_2_resname
# hares -link votemnt_2_resname vxfsckd

# hares -link votevol_3_resname cvm_clus
# hares -link votemnt_3_resname votevol_3_resname
# hares -link votemnt_3_resname vxfsckd
```

**5** Enable the resources:

```
# hares -modify ocrvol_1_resname Enabled 1
# hares -modify ocrvol_2_resname Enabled 1
# hares -modify ocrvol_3_resname Enabled 1

# hares -modify votevol_1_resname Enabled 1
# hares -modify votevol_2_resname Enabled 1
# hares -modify votevol_3_resname Enabled 1

# hares -modify ocrmnt_1_resname Enabled 1
# hares -modify ocrmnt_2_resname Enabled 1
# hares -modify ocrmnt_3_resname Enabled 1

# hares -modify votemnt_1_resname Enabled 1
# hares -modify votemnt_2_resname Enabled 1
# hares -modify votemnt_3_resname Enabled 1
```

# Configuring private IP addresses for Oracle RAC

Private IP addresses are required by Oracle RAC to provide communication between the cluster nodes. Depending on your private network configuration, you may need one or more IP addresses. Veritas recommends the use of link aggregation, NIC bonding, or Oracle High Availability IP (HAIP) to provide redundant, dedicated private communication.

InfoScale supports SF Oracle RAC installer-based configuration of HAIP.

Use one of the following options to configure the private network for Oracle RAC.

**Table 13-1** Options for configuring the private network for Oracle RAC

| Option | Description |
|---|---|
| Link aggregation or NIC bonding | See "Using link aggregation or NIC bonding" on page 200. |
| Oracle High Availability IP (HAIP) | See "Using Oracle High Availability IP (HAIP)" on page 201. |

## Using link aggregation or NIC bonding

To configure link aggregation or NIC bonding, see the operating system documentation.

---

**Note:** Ensure that a link configured under an aggregated link or NIC bond is not configured as a separate LLT link.

---

When LLT is configured over a bonded interface, perform one of the following steps to prevent GAB from reporting jeopardy membership:

- Configure an additional network interface under LLT in addition to the bonded NIC.
- Add the following line in the `/etc/llttab` file:

  ```
  set-dbg-minlinks <n>
  ```

  Where <*n*> is the number of links configured under an aggregated link.
  Restart LLT.

  ```
  # systemctl start llt
  ```

## Using Oracle High Availability IP (HAIP)

Starting with Oracle RAC 11.2.0.2, Oracle introduced the High Availability IP (HAIP) feature for supporting IP address failover. The purpose of HAIP is to perform load balancing across all active interconnect interfaces and fail over existing non-responsive interfaces to available interfaces. HAIP has the ability to activate a maximum of four private interconnect connections. These private network adapters can be configured during the installation of Oracle Grid Infrastructure or after the installation using the `oifcfg` utility.

Oracle Grid Infrastructure currently creates an alias IP (known as virtual private IP) on a private network interface using the 169.254.*.* subnet for HAIP. However, if that subnet range is already in use, Oracle Grid Infrastructure will not attempt to use it. Therefore, to enable HAIP using the Oracle Grid Infrastructure installer, configure all private interfaces with static private IP addresses before starting the Grid installation. During the installation of Oracle Grid Infrastructure, select these interfaces as 'private' interfaces. It is recommended to use all the LLT links as cluster interconnects.

You can configure HAIP in one of the following ways:

| | |
|---|---|
| SF Oracle RAC script-based installer | See "Configuring static IP addresses for HAIP using the SF Oracle RAC script-based installer" on page 201. |
| Manual | See "Configuring static IP addresses for HAIP manually" on page 203. |

### Configuring static IP addresses for HAIP using the SF Oracle RAC script-based installer

The high availability IP address is managed by Oracle Grid Infrastructure. SF Oracle RAC supports the configuration of multiple network interfaces for HAIP. One static IP address is required for each network interface. All nodes must use a common network interface for the private interconnects in the cluster.

Review the following information before you configure HAIP:

- Do not use the public network interface for Oracle HAIP configuration.
- All LLT links are used for configuring HAIP.
- The same network interfaces must be used for private interconnect on each node of the cluster.
- The private IP addresses must be added to /etc/hosts on all nodes before installing Oracle Clusterware. If you choose the installer to add IP addresses in /etc/hosts then

### To configure static IP addresses for HAIP using the SF Oracle RAC script-based installer

1. From the SF Oracle RAC menu, enter **3** to select the option **Oracle Network Configuration**.

```
1) Create Oracle Users and Groups
2) Create Storage for OCR and Voting disk
3) Oracle Network Configuration
4) Exit SF Oracle RAC Configuration
b) Back to previous menu
Choose option: [1-4,b,q] (1) 3
```

2. Enter **1** to select the option **Configure private IP addresses (For HAIP Configuration) - For Oracle RAC 11.2.0.2 and later**.

```
1) Configure private IP addresses (For HAIP Configuration)
- For Oracle RAC 11.2.0.2 and later
2) Configure private IP addresses (PrivNIC Configuration)
- For Oracle RAC prior to 11.2.0.2
3) Configure private IP addresses (MultiPrivNIC Configuration)
- For Oracle RAC prior to 11.2.0.2
4) Exit SF Oracle RAC Configuration
b) Back to previous menu
Choose option: 1-3,b,q (1) 1
```

   The installer discovers available LLT links.

3. Enter **y** to add the IP addresses to the `/etc/hosts` file.

```
Do you want the Installer to add IP addresses in /etc/hosts
file? [y,n,q] (y)
```

   Enter **n** if you choose to add the IP addresses to the file manually and provide the following information:

```
Enter the IP for eth1 on sys1: [b] 192.168.12.1
Enter the IP for eth2 on sys1: [b] 192.168.13.1
Enter the IP for eth1 on sys2: [b] 192.168.12.2
Enter the IP for eth2 on sys2: [b] 192.168.13.2
```

   Go to step 5.

4 Perform this step only if you enabled the installer to add the IP address to the `/etc/hosts` file in the previous step. Provide the private IP address and the private node name for the IP address that must be added to the file.

---

**Note:** All IP addresses must be in the same subnet, failing which Oracle Clusterware/Grid Infrastructure will not be able to communicate properly across the nodes. If the private IP address entries are already present in the `/etc/hosts` file on one of nodes in the cluster, the installer does not update the file with the specified IP addresses on any of the nodes in the cluster. Only those IP addresses that are not already present in the file on any node are added.

---

```
Enter the IP for eth1 on sys1: [b] 192.168.12.1
Enter Hostname alias for the above IP address: [b] sys1-haip1
Enter the private IP for eth2 on sys1: [b] 192.168.13.1
Enter Hostname alias for the above IP address: [b] sys1-haip2
Enter the IP for eth1 on sys2: [b] 192.168.12.2
Enter Hostname alias for the above IP address: [b] sys2-haip1
Enter the private IP for eth2 on sys2: [b] 192.168.13.2
Enter Hostname alias for the above IP address: [b] sys2-haip2
```

5 Enter the netmask information for the private network:

```
Enter the Netmask for private network: [b] (255.255.255.0)
```

The SF Oracle RAC installer now displays the configuration information.

6 Enter **y** to review and confirm the configuration information. The installer updates the`/etc/hosts` file (if you chose an installer-based update).

7 If you chose to add the IP address information to the `/etc/hosts` file manually, proceed to update the file as described in the following procedure.

## Configuring static IP addresses for HAIP manually

You can configure static IP addresses manually using native operating system commands. See the operating system documentation for instructions.

For example, to configure a static IP address:

Edit the file `/etc/sysconfig/network-scripts/ifcfg-`*`interface_name`*

```
DEVICE=interface_name
ONBOOT=yes
BOOTPROTO=static
```

```
IPADDR=ip-address
NETMASK=netmask
```

## Verifying that multicast is functional on all private network interfaces

Multicast network communication on the private interconnect network must be enabled and functioning on all nodes otherwise the installation or upgrade of Oracle Grid Infrastructure may fail.

For more information, see the Oracle Metalink document: 1212703.1

## Creating Oracle Clusterware/Grid Infrastructure and Oracle database home directories manually

You can create the Oracle Clusterware/Grid Infrastructure and Oracle database home directories on the local file system or on a local Veritas file system, or on a Veritas cluster file system. When the installer prompts for the home directories at the time of installing Oracle Clusterware/Grid Infrastructure and Oracle database, it creates the directories locally on each node, if they do not exist.

**Note:** Veritas recommends that Oracle Clusterware and Oracle database binaries be installed local to each node in the cluster. For Oracle Grid Infrastructure binaries, Oracle requires that they be installed only on a local file system. Refer to the Oracle documentation for size requirements.

Table 13-2 lists the Oracle RAC directories you need to create:

**Table 13-2** List of directories

| Directory | Description |
|---|---|
| Oracle Grid Infrastructure Home Directory (GRID_HOME) | The path to the home directory that stores the Oracle Grid Infrastructure binaries. The Oracle Universal Installer (OUI) installs Oracle Grid Infrastructure and Oracle ASM into this directory, also referred to as GRID_HOME.<br>The directory must be owned by the installation owner of Oracle Grid Infrastructure (oracle or grid), with the permission set to 755.<br>The path to the grid home directory must be the same on all nodes.<br>Follow Oracle Optimal Flexible Architecture (OFA) guidelines while choosing the path. |

**Table 13-2** List of directories *(continued)*

| Directory | Description |
|---|---|
| Oracle base directory (ORACLE_BASE) | The base directory that contains all the Oracle installations. Create separate Oracle base directories for the grid user and the Oracle user.<br><br>It is recommended that installations of multiple databases maintain an Optimal Flexible Architecture (OFA) configuration.<br><br>The path to the Oracle base directory must be the same on all nodes. The permission on the Oracle base directory must be at least 755. |
| Oracle home directory (ORACLE_HOME) | The directory in which the Oracle database software is installed. The path to the Oracle home directory must be the same on all nodes. The permission on the Oracle home directory must be at least 755. |

Use one of the following options to create the directories:

Local file system — See "To create the directories on the local file system" on page 205.

Cluster File System — See "To create the file system and directories on cluster file system for Oracle database" on page 207.

### To create the directories on the local file system

1. Log in as the root user on each node.
2. Create a local file system and mount it using one of the following methods:
   - Using native operating system commands
     For instructions, see the operating system documentation.
   - Using Veritas File System (VxFS) commands

As the root user, create a VxVM local diskgroup on each node.

```
# vxdg init vxvm_dg \
dg_name
```

Create separate volumes for Oracle Clusterware/Oracle Grid Infrastructure binaries and Oracle binaries.

```
# vxassist -g vxvm_dg make clus_volname size
# vxassist -g vxvm_dg make ora_volname size
```

Create the file systems with the volumes.

```
# mkfs -t vxfs /dev/vx/rdsk/vxvm_dg/clus_volname
# mkfs -t vxfs /dev/vx/rdsk/vxvm_dg/ora_volname
```

Mount the file system.

```
# mount -t vxfs /dev/vx/dsk/vxvm_dg/clus_volname \
clus_home
# mount -t vxfs /dev/vx/dsk/vxvm_dg/ora_volname \
oracle_home
```

3 Create the directories for Oracle RAC.

```
# mkdir -p grid_base
# mkdir -p clus_home
# mkdir -p oracle_base
# mkdir -p oracle_home
```

4 Set appropriate ownership and permissions for the directories.

```
# chown -R grid:oinstall grid_base
# chmod -R 775 grid_base
# chown -R grid:oinstall clus_home
# chmod -R 775 clus_home

# chown -R oracle:oinstall oracle_base
# chmod -R 775 oracle_base
# chown -R oracle:oinstall oracle_home
# chmod -R 775 oracle_home
```

5 Add the resources to the VCS configuration.

See "To add the storage resources created on VxFS to the VCS configuration" on page 206.

6 Repeat all the steps on each node of the cluster.

**To add the storage resources created on VxFS to the VCS configuration**

1 Change the permissions on the VCS configuration file:

```
# haconf -makerw
```

2 Configure the VxVM volumes under VCS:

```
# hares -add dg_resname DiskGroup cvm
# hares -modify dg_resname DiskGroup vxvm_dg -sys node_name
# hares -modify dg_resname Enabled 1
```

3 Set up the file system under VCS:

```
# hares -add clusbin_mnt_resname Mount cvm

# hares -modify clusbin_mnt_resname MountPoint \
"clus_home"

# hares -modify clusbin_mnt_resname BlockDevice \
"/dev/vx/dsk/vxvm_dg/clus_volname" -sys node_name
# hares -modify clusbin_mnt_resname FSType vxfs
# hares -modify clusbin_mnt_resname FsckOpt "-n"
# hares -modify clusbin_mnt_resname Enabled 1
# hares -add orabin_mnt_resname Mount cvm

# hares -modify orabin_mnt_resname MountPoint \
"oracle_home"

# hares -modify orabin_mnt_resname BlockDevice \
"/dev/vx/dsk/vxvm_dg/ora_volname" -sys node_name
# hares -modify orabin_mnt_resname FSType vxfs
# hares -modify orabin_mnt_resname FsckOpt "-n"
# hares -modify orabin_mnt_resname Enabled 1
```

4 Link the parent and child resources:

```
# hares -link clusbin_mnt_resname vxvm_dg
# hares -link orabin_mnt_resname vxvm_dg
```

5 Repeat all the steps on each node of the cluster.

**To create the file system and directories on cluster file system for Oracle database**

Perform the following steps on the CVM master node in the cluster.

1 As the root user, create a VxVM shared disk group:

```
# vxdg -s init cvm_dg dg_name
```

2 Create the volume for Oracle database:

```
# vxassist -g cvm_dg make ora_volname
  size
```

3 Create the Oracle base directory and the Oracle home directory.

```
# mkdir -p oracle_base
# mkdir -p oracle_home
```

4 Create file system with the volume:

```
# mkfs -t vxfs /dev/vx/rdsk/cvm_dg/ora_volname
```

5 Mount the file system. Perform this step on each node.

```
# mount -t vxfs -o cluster /dev/vx/dsk/cvm_dg/ora_volname \
oracle_home
```

6 Change the ownership and permissions on all nodes of the cluster.

```
# chown -R oracle:oinstall oracle_base
# chmod -R 775 oracle_base
# chown -R oracle:oinstall oracle_home
# chmod -R 775 oracle_home
```

7 Add the CVMVolDg and CFSMount resources to the VCS configuration.

See "To add the CFSMount and CVMVolDg resources to the VCS configuration using CLI" on page 208.

**To add the CFSMount and CVMVolDg resources to the VCS configuration using CLI**

1 Change the permissions on the VCS configuration file:

```
# haconf -makerw
```

2 Configure the CVM volumes under VCS:

```
# hares -add dg_resname CVMVolDg cvm

# hares -modify dg_resname Critical 0

# hares -modify dg_resname CVMDiskGroup cvm_dg

# hares -modify dg_resname CVMVolume -add ora_volname

# hares -modify dg_resname CVMActivation sw
```

**3** Set up the file system under VCS:

```
# hares -add orabin_mnt_resname CFSMount cvm
# hares -modify orabin_mnt_resname Critical 0
# hares -modify orabin_mnt_resname MountPoint \
"oracle_home"
# hares -modify orabin_mnt_resname BlockDevice  \
"/dev/vx/dsk/cvm_dg/ora_volname"
```

**4** Link the parent and child resources:

```
# hares -link dg_resname cvm_clus

# hares -link orabin_mnt_resname dg_resname

# hares -link orabin_mnt_resname vxfsckd
```

**5** Enable the resources:

```
# hares -modify dg_resname Enabled 1

# hares -modify orabin_mnt_resname Enabled 1

# haconf -dump -makero
```

**6** Verify the resource configuration in the main.cf file.

**7** Verify that the resources are online on all systems in the cluster.

```
# hares -state dg_resname

# hares -state orabin_mnt_resname
```

---

**Note:** At this point, the crsorabin_voldg resource is reported offline, and the underlying volumes are online. Therefore, you need to manually bring the resource online on each node.

---

To bring the resource online manually:

```
# hares -online dg_resname -sys node_name
```

# Setting up user equivalence

You must establish grid user and Oracle user equivalence on all nodes to allow the Oracle Universal Installer to securely copy files and run programs on the nodes in the cluster without requiring password prompts.

Set up passwordless SSH communication between the cluster nodes for the Oracle user and the grid user.

For more information, see the Oracle documentation.

Chapter 14

# Installing Oracle RAC

This chapter includes the following topics:


- About installing Oracle RAC
- Installing the Oracle Clusterware/Grid Infrastructure software
- Configuring LLT links in the GPnP profile
- Installing the Oracle RAC database software
- Verifying the Oracle Clusterware/Grid Infrastructure and database installation


## About installing Oracle RAC

You can install Oracle RAC on shared storage or locally on each node.

Use one of the following ways to install Oracle RAC:

| | |
|---|---|
| SF Oracle RAC installer | The installer starts the installation process and prompts for information that is required by the Oracle Universal Installer. The Oracle Universal Installer launches with these installation values pre-filled and installs Oracle RAC.<br><br>You need to invoke the Veritas script-based installer to start the installation. |
| Oracle Universal Installer | The Oracle Universal Installer installs Oracle RAC. The installation values must be manually entered at the time of installation.<br><br>You need to invoke the Oracle Universal Installer to install Oracle RAC. |

| | |
|---|---|
| Response files | The SF Oracle RAC script-based installer supports silent installation of Oracle RAC using its own response files. You need to modify the SF Oracle RAC response files to include the path information of the Oracle RAC software binaries and the Oracle RAC response files.<br><br>For more information and instructions:<br><br>For instructions, see the chapter "Installing Oracle RAC using a response file" in this document. |

**Note:** The instructions in this chapter use variables and sample values whereever required. Replace these variables and sample values with values that conform to your installation requirements.

Before you start the installation:

- Keep the Oracle worksheets handy as you perform the installation tasks.
  See "Required installation information for Oracle Clusterware/Grid Infrastructure" on page 382.
  See "Required installation information for Oracle database" on page 385.
- Review your Oracle installation manuals and the appropriate Oracle support Web sites for additional information required during the installation.

# Installing the Oracle Clusterware/Grid Infrastructure software

This section provides instructions for installing Oracle Clusterware/Grid Infrastructure using the SF Oracle RAC installer. The SF Oracle RAC installer prompts for information required to invoke the Oracle Universal Installer and launches it. The responses provided to the SF Oracle RAC installer are pre-filled in the Oracle Universal Installer wizard. When you step through the installation, review or change these installation values in the Oracle Universal Installer.

**Note:** Before you begin the installation, verify that the nodes in the cluster are connected with network links using similar network devices. For example, if you use eth0 as a public link on one node in the cluster, all other nodes in the cluster must also use eth0 as the public link. Similarly, if you use eth1 as a private link on one node in the cluster, all other nodes in the cluster must also use eth1 as the private link.

Oracle Grid Infrastructure software is installed on each node in the GRID_HOME location.

---

**Note:** If you want to install Oracle Clusterware/Grid Infrastructure on VxFS or CFS, make sure that you created the appropriate storage before proceeding with the installation.

See "Creating Oracle Clusterware/Grid Infrastructure and Oracle database home directories manually" on page 204.

---

See "Installing Oracle Clusterware/Grid Infrastructure using the Oracle Universal Installer" on page 213.

## Installing Oracle Clusterware/Grid Infrastructure using the Oracle Universal Installer

This section provides instructions for installing the Oracle Clusterware/Grid Infrastructure software using the Oracle Universal Installer. The software is installed on each node in the Oracle Clusterware/Grid Infrastructure home directory.

**To install Oracle Clusterware/Grid Infrastructure using the Oracle Universal Installer**

1. Log in as the Oracle grid user. On the first node, set the DISPLAY variable.
   - For Bourne Shell (bash), type:

     ```
     $ DISPLAY=10.20.12.150:0.0;export DISPLAY
     ```

     where 10.20.12.150 is the IP address of X server where you want to export the display for the installer.
   - For C Shell (csh or tcsh), type:

     ```
     $ setenv DISPLAY 10.20.12.150:0.0
     ```

     where 10.20.12.150 is the IP address of X server where you want to export the display for the installer.
2. Start the Oracle Universal Installer on the first node.
3. Enter the required information when prompted by the Oracle Universal Installer.

   See "Required installation information for Oracle Clusterware/Grid Infrastructure" on page 382.

4 Review the configuration summary presented by the Oracle Universal Installer. The Oracle Universal Installer begins the Oracle Clusterware/Grid Infrastructure installation.

---

**Note:** If you want to save the Oracle Grid Infrastructure installation configuration into a response file for future installations, click the **Save Response File** option on the Summary page of the Oracle Universal Installer.

---

5 At the end of the Oracle Clusterware/Grid Infrastructure installation, run the following configuration scripts as the root user from each node in the cluster, in the listed order:

- orainstRoot.sh (located in the oraInventory directory)
  Ensure that the script exists on each node before proceeding.
- root.sh (located in the CRS_HOME or GRID_HOME directory, depending on your Oracle RAC version)

---

**Note:** You may skip this step if you selected the option **Automatically run configuration scripts**.

---

6 Return to the Oracle Universal Installer window and click **OK** to continue. The Oracle Universal Installer informs you that the Oracle Clusterware/Grid Infrastructure installation was successful. The Oracle Clusterware/Grid Infrastructure daemons are started on the node.

# Configuring LLT links in the GPnP profile

Update the GPnP profile to include the remaining LLT links that were not added to the profile during the Oracle Grid Infrastructure installation.

**To configure the LLT links in the GPnP profile**

1. View the currently configured interfaces:

```
# $GRID_HOME/bin/oifcfg getif
eth0 10.2.156.0         global      public
eth1 192.168.12.0       global      cluster_interconnect
```

   The interfaces that are currently stored in the GPnP profile, their subnets, and their role (public or cluster_interconnect) are displayed.

2. Add the remaining LLT links to the GPnP profile:

```
# $GRID_HOME/bin/oifcfg setif -global \
eth2/192.168.12.0:cluster_interconnect
```

   If you are using multiple IP addresses on different subnet for cluster interconnect (for load balancing), add the remaining interface subnets to the GPnP profile.

```
# $GRID_HOME/bin/oifcfg setif -global \
eth2/192.168.2.0:cluster_interconnect
# $GRID_HOME/bin/oifcfg setif -global \
eth1/192.168.2.0:cluster_interconnect
```

3. Verify that the correct interface subnet is in use:

```
# $GRID_HOME/bin/oifcfg getif
eth0 10.2.156.0         global      public
eth1 192.168.12.0       global      cluster_interconnect
eth2 192.168.12.0       global      cluster_interconnect
eth1 192.168.2.0        global      cluster_interconnect
eth2 192.168.2.0        global      cluster_interconnect
```

   Make sure all the LLT links are configured and listed in the GPnP profile.

# Installing the Oracle RAC database software

Before you start the installation of Oracle database, make sure that Oracle Clusterware/Grid Infrastructure is up and running. Veritas recommends you to install the Oracle database locally on each node.

---

**Note:** If you want to install Oracle database on VxFS or CFS, make sure that you created the appropriate storage before proceeding with the installation.

See "Creating Oracle Clusterware/Grid Infrastructure and Oracle database home directories manually" on page 204.

---

Install the software using one of the following methods:

| | |
|---|---|
| SF Oracle RAC script-based installer | Using script-based installer:<br>See "Installing the Oracle RAC database using the SF Oracle RAC script-based installer" on page 216.<br>Using response files:<br>See the chapter *Installation of SF Oracle RAC and Oracle RAC using a response file* in this document. |
| Oracle Universal Installer | See "Installing the Oracle RAC database using the Oracle Universal Installer" on page 219. |

# Installing the Oracle RAC database using the SF Oracle RAC script-based installer

The SF Oracle RAC installer performs the following tasks:

- Verifies the status of Oracle Clusterware/Grid Infrastructure on all nodes
- Invokes the Oracle Universal Installer to install the Oracle database
- Verifies the Oracle RAC database installation

**To install the Oracle RAC database using the SF Oracle RAC script-based installer**

1. Return to the SF Oracle RAC installer and type **4** to select the option **Install Oracle Clusterware/Grid Infrastructure and Database**.

```
1)  Configure SF Oracle RAC sub-components
2)  Prepare to Install Oracle
3)  Install Oracle Clusterware/Grid Infrastructure and Database
4)  Post Oracle Installation Tasks
5)  Exit SF Oracle RAC Configuration
Choose option: [1-5,q] (1) 3
```

2. Select the option **Install Oracle Database**.

```
1)  Install Oracle Clusterware/Grid Infrastructure
2)  Install Oracle Database
3)  Exit SF Oracle RAC Configuration
b)  Back to previous menu
Choose option: [1-3,b,q] (1) 2
```

3. Set the DISPLAY environment variable that is required for the Oracle Universal Installer.

```
Enter the DISPLAY environment variable: [b] 10.20.12.150:0.0
```

where **10.20.12.150** is the IP address of X server where you want to export the display for the installer.

4. Enter Oracle UNIX user name. The Oracle UNIX user name was previously set up during the pre-installation process.

```
Enter Oracle UNIX user name:  [b] (oracle)
```

5. Enter Oracle UNIX group name. The Oracle UNIX group name was previously set up during the pre-installation process.

```
Enter Oracle UNIX group name:  [b] (oinstall)
```

6. Enter the full path of the Oracle base directory.

```
Enter absolute path of Oracle Base directory: [b]
```

7 Enter the full path of the Oracle Clusterware/Grid Infrastructure home directory.

```
Enter absolute path of Oracle Clusterware/Grid Infrastructure
Home directory: [b]
```

8 Enter the full path of the Oracle database home directory.

```
Enter absolute path of Oracle Database Home directory: [b]
```

If the Oracle RAC database home directory you specified does not exist, the installer creates the directory locally on each node and sets appropriate permissions for the Oracle user.

9 Enter the full path of the database installation image.

```
Enter absolute path of Oracle Database install image: [b]
```

The installer determines the version of the Oracle software from the binaries.

10 Enter **y** to proceed with the installation.

11 Review and confirm the configuration information.

The installer verifies that Oracle Clusterware/Grid Infrastructure is running and invokes the Oracle Universal Installer:

12 Enter the required information when prompted by the Oracle Universal Installer.

See "Required installation information for Oracle database" on page 385.

13 Review the configuration summary presented by the Oracle Universal Installer. The Oracle Universal Installer begins the Oracle database installation.

14 Run the root.sh script as the root user on the cluster nodes:

Return to the Oracle Universal Installer window and click **OK** to continue.

15 Return to the SF Oracle RAC installer and press **Return** to continue. The installer verifies the Oracle database installation.

---

**Note:** After the installation completes successfully, the installer prompts you to relink the database binaries. Veritas recommends you to relink the SF Oracle RAC libraries only after completing all the required patch additions, if any. See the Oracle documentation for patch updates that may be required.

---

## Installing the Oracle RAC database using the Oracle Universal Installer

The following procedure describes how to install the Oracle RAC database using the Oracle Universal Installer. Veritas recommends that you install the Oracle RAC database locally on each node.

**To install Oracle RAC database using the Oracle Universal Installer**

1. To launch and use the Oracle Universal Installer, follow the instructions in the Oracle documentation.
2. Enter the required information when prompted by the Oracle Universal Installer.

   See "Required installation information for Oracle database" on page 385.
3. Review the configuration summary presented by the Oracle Universal Installer.

   **Note:** If you want to save the Oracle RAC database installation configuration into a response file for future installations, click the **Save Response File** option on the Summary page of the Oracle Universal Installer.

   The Oracle Universal Installer begins the Oracle database installation.
4. Completed the tasks as prompted by the Oracle Universal Installer.

# Verifying the Oracle Clusterware/Grid Infrastructure and database installation

- The following procedure verifies the Oracle Clusterware/Grid Infrastructure and Oracle RAC database installation by verifying that the Oracle processes are running on all nodes.
  To verify the installation, run the following command from any node in the cluster. Verify in the command output that the Oracle Clusterware/Grid Infrastructure processes are online on the nodes.

```
# $GRID_HOME/bin/crsctl stat res -t


--------------------------------------------------------------
NAME           TARGET  STATE     SERVER    STATE_DETAILS
--------------------------------------------------------------
Local Resources
--------------------------------------------------------------
ora.LISTENER.lsnr
```

```
               ONLINE  ONLINE    sys1
               ONLINE  ONLINE    sys2
ora.asm
               OFFLINE OFFLINE   sys1
               OFFLINE OFFLINE   sys2

.
.
.
--------------------------------------------------------------
Cluster Resources
--------------------------------------------------------------
ora.LISTENER_SCAN1.lsnr
      1        ONLINE  ONLINE       sys2
ora.LISTENER_SCAN2.lsnr
      1        ONLINE  ONLINE       sys1
ora.LISTENER_SCAN3.lsnr
      1        ONLINE  ONLINE       sys1
.
.
.
```

- To verify the Oracle RAC database installation, check the oraInventory logs.

Chapter 15

# Performing an automated Oracle RAC installation

This chapter includes the following topics:


- About installing Oracle RAC using response files
- Before you install
- Installing Oracle RAC
- Response file variable definitions for Oracle RAC
- Sample response file for installing Oracle RAC


## About installing Oracle RAC using response files

You can perform a silent pre-configuration of your systems by running the SF Oracle RAC installer with the `-responsefile` option. This capability in tandem with the Oracle response files for Oracle Clusterware and Oracle database installation enables you to standardize and automate Oracle RAC deployments in your cluster.

## Before you install

Make sure that you complete the tasks in the following procedure before starting the silent installation of Oracle RAC.

**To prepare the systems for installing Oracle RAC using response files**

1. Make sure that the systems meet the installation requirements.

   For information on requirements, see the chapter *System requirements* in this document.

2. Complete the following preparatory tasks on the nodes manually or by using the SF Oracle RAC installer:

   - Identify the public virtual IP addresses for use by Oracle
   - Plan for the storage and the private network configuration for Oracle RAC
   - Set the kernel parameters
   - Verify the user "nobody" exists
   - Create Oracle user and groups
   - Set up Oracle user equivalence
   - Edit the Oracle user profile

   For instructions, see the chapter *Before installing Oracle RAC* in this document.

3. Create response files for Oracle Clusterware/Grid Infrastructure and Oracle database installation using the response file template provided by Oracle RAC.

   For instructions on using the response file template, see the Oracle RAC documentation.

---

**Note:** Keep at hand the full path of the directory where these response files will be saved. The full path of the Oracle Clusterware/Grid Infrastructure response file must be set in the SF Oracle RAC response file variable `$CFG{crs_responsefile};`. The full path of the Oracle database response file must be set in the SF Oracle RAC response file variable `$CFG{db_responsefile};`.

---

4. Create an SF Oracle RAC response file.

   For information on various options available for creating a response file:

   Make sure that you provide the full path information of the Oracle Clusterware/Grid Infrastructure and Oracle database response files.

   For guidelines on creating a response file:

   For information on the list of required and optional variables:

   See "Response file variable definitions for Oracle RAC" on page 224.

   For a sample response file:

   See "Sample response file for installing Oracle RAC" on page 243.

5. Make sure that the Oracle user has appropriate read and write permissions on the response files.

6. Make sure that passwordless communication between Oracle users is set up on the nodes of the cluster.

# Installing Oracle RAC

To perform a silent installation of Oracle RAC, run the SF Oracle RAC installer.

The installer supports completion of the following tasks using the response file:

- Creates storage for OCR and voting disk
- Configures the private network for Oracle RAC
- Installs Oracle Clusterware/Grid Infrastructure by leveraging the corresponding Oracle response file
- Installs Oracle database by leveraging the Oracle database response file
- Relinks the Oracle RAC libraries with SF Oracle RAC libraries
- Configures the CSSD agent

The sample procedure assumes that Oracle user equivalence is established on all the nodes in the cluster.

**To perform Oracle pre-installation and installation on the nodes**

1. Navigate to the directory containing the SF Oracle RAC installer:

   ```
   # cd /opt/VRTS/install
   ```

2. Start the installation:

   ```
   # ./installer -responsefile /tmp/response_file
   ```

   Where `/tmp/response_file` is the full path name of the SF Oracle RAC response file.

3. Run the root.sh script as the root user on the cluster nodes.

   **Note:** Do not run the script simultaneously on your cluster nodes.

4. Complete the following Oracle post-installation tasks:

   - Add any patches or patchsets required by Oracle RAC .
   - Create the Oracle RAC database.
   - Add the Oracle UDP IPC private IP addresses to the Oracle `init.ora` file if the database cache fusion traffic is configured to use Oracle UDP IPC private IP addresses and if these addresses are configured as a PrivNIC or MultiPrivNIC resource for high availability.
   - Configure the Oracle RAC database for manual startup if you want the Oracle RAC database to be managed by VCS using the Oracle agent.
   - Configure the VCS service groups for Oracle RAC.
   - Verify the cluster.
   - Remove the temporary communication permissions.

   For instructions:

   See the chapter *Performing Oracle RAC post-installation tasks*.

# Response file variable definitions for Oracle RAC

The variable definitions for Oracle RAC are grouped in tabular format for the following Oracle tasks:

| | |
|---|---|
| Creating Oracle user and group | See Table 15-1 on page 225. |

| | |
|---|---|
| Creating storage for OCR and voting disk<br>See "Creating storage for OCR and voting disk " on page 184. | See Table 15-2 on page 227. |
| Configuring the private IP address for HAIP | See Table 15-3 on page 231. |
| Configuring the private IP address and MultiPrivNIC resource under VCS | See Table 15-5 on page 236. |
| Installing Oracle Clusterware | See Table 15-6 on page 239. |
| Installing Oracle database | See Table 15-7 on page 240. |
| Configuring CSSD resource | See Table 15-8 on page 241. |
| Relinking Oracle RAC libraries | See Table 15-9 on page 242. |

**Note:** Some of the variable definitions may occur in multiple sections, for example `$CFG{oracle_user}`. These variables need not be repeated if all the tasks are performed as a single installation activity. However, if you perform these tasks independently, make sure that all the required variables, as indicated in the table for each task, are supplied in the response file.

Table 15-1 lists the variables that are used to create the Oracle user and group.

**Table 15-1** Variables for creating Oracle user and group

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{create_oracle_user_group} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that Oracle user and group will be created.<br>The value 0 indicates that Oracle user and group will not be created. |
| $CFG{grid_user} | Scalar | Required<br>Defines the name of the grid user. |
| $CFG{oracle_user} | Scalar | Required<br>Defines the name of the Oracle user. |

**Table 15-1** Variables for creating Oracle user and group *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{oracle_uid} | Scalar | Required<br>Defines the user ID of the Oracle user. |
| $CFG{oracle_group} | Scalar | Required<br>Defines the primary group of the Oracle user. |
| $CFG{oracle_gid} | Scalar | Required<br>Defines the group ID of the Oracle user. |
| $CFG{oracle_user_home} | Scalar | Required<br>Defines the full path of the Oracle user's home directory. |
| $CFG{oracle_secondary_group} | List | Optional<br>Defines the list of secondary groups for the Oracle user. |
| $CFG{oracle_secondary_gid} | List | Optional<br>Defines the list of secondary group IDs for the Oracle user. The elements of this variable must be in the same order as that of the elements in the variable `$CFG{oracle_secondary_group}`. |

Table 15-2 lists the variables that are used to create the storage for OCR and voting disk.

**Table 15-2** Variables for creating storage for OCR and voting disk

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{create_ocr_vote_storage} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that the storage for OCR and voting disk will be created.<br>The value 0 indicates that the storage for OCR and voting disk will not be created. |
| $CFG{enable_mirroring} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that the storage for OCR and voting disk is mirrored. Provide two disks as input for the variable `$CFG{ocrvotedisks}`.<br>The value 0 indicates that the storage for OCR and voting disk is not mirrored. |

**Table 15-2** Variables for creating storage for OCR and voting disk *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{ocrvotedgoption} | Scalar | Required<br>Defines one of the following values:<br>■ 0: Indicates that a new non-FSS disk group will be created for OCR and voting disk storage.<br>■ 1: Indicates that an existing non-FSS disk group will be used to create the storage for OCR and voting disk.<br>■ 2: Indicates that a new FSS disk group will be created for OCR and voting disk storage.<br>■ 3: Indicates that an existing FSS disk group will be used to create the storage for OCR and voting disk.<br>**Note:** If you choose to use an existing disk group, use the `$CFG{ocrvotedgname}` variable to specify the name of an existing disk group that has a minimum of two disks (for mirroring).<br>**Note:** If you choose to create a disk group, you must set the following variables: `$CFG{ocrvotedisks}`, `$CFG{ocrvotedgname}` |
| $CFG{ocrvotescheme} | Scalar | Required<br>Defines the storage scheme to be used for OCR and voting disk.<br>The value 1 indicates Clustered File System.<br>The value 0 indicates CVM raw volumes. |

**Table 15-2** Variables for creating storage for OCR and voting disk *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{enable_sep_filesys} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that OCR and voting disk are located on separate file systems. Provide values for the following variables: `$CFG{ocrvolname}`, `$CFG{ocrvolsize}`, `$CFG{votevolname}`, `$CFG{votevolsize}`, `$CFG{ocrmount}`, `$CFG{votemount}`<br>The value 0 indicates that OCR and voting disk are located on the same file system. Provide a single volume name using the variable `$CFG{ocrvotevolname}`, a single mount point using the variable `$CFG{ocrvotemount}`, and the size using the variable `$CFG{ocrvotevolsize}`. |
| $CFG{ocrvotedisks} | List | Required<br>Defines the list of shared disks to be used for OCR and voting disk.<br>**Note:** If choose to create an FSS disk group, the attribute includes the node information, for example, `$CFG{ocrvotedisks}=[qw(node_1:disk_1 node_2:disk_2...)]`, where `node_x` indicates the node to which the disk is physically connected. |
| $CFG{ocrvotedgname} | Scalar | Required<br>Defines the name of the disk group to be used for OCR and voting disk. |
| $CFG{ocrvotevolname} | Scalar | Required<br>Defines the volume name for OCR and voting disk. This variable must be used only if you have set the storage scheme to 1 (Clustered File System). |

**Table 15-2** Variables for creating storage for OCR and voting disk *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{ocrvotevolsize} | Scalar | Required<br>Defines the size of the OCR and voting disk volume. This variable must be used only if you have set the storage scheme to 1 (Clustered File System). |
| $CFG{ocrvotemount} | Scalar | Required if you have chosen to locate OCR and voting disk on the same file system.<br>Defines the full path to the CFS mount point. This variable must be used only if you have set the storage scheme to 1 (Clustered File System). |
| $CFG{ocrmount} | Scalar | Required if you have chosen to locate OCR and voting disk on separate file systems.<br>Defines the full path to the CFS mount point for OCR. This variable must be used only if you have set the storage scheme to 1 (Clustered File System). |
| $CFG{votemount} | Scalar | Required if you have chosen to locate OCR and voting disk on separate file systems.<br>Defines the full path to the CFS mount point for voting disk. This variable must be used only if you have set the storage scheme to 1 (Clustered File System). |
| $CFG{ocrvolname} | Scalar | Required<br>Defines the volume name for OCR. This variable must be used only if you have set the storage scheme to 0 (CVM Raw Volumes). |
| $CFG{ocrvolsize} | Scalar | Required<br>Defines the size of the OCR volume. This variable must be used only if you have set the storage scheme to 0 (CVM Raw Volumes). |

**Table 15-2** Variables for creating storage for OCR and voting disk *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{votevolname} | Scalar | Required<br>Defines the volume name for voting disk. This variable must be used only if you have set the storage scheme to 0 (CVM Raw Volumes). |
| $CFG{votevolsize} | Scalar | Required<br>Defines the size of the voting disk volume. This variable must be used only if you have set the storage scheme to 0 (CVM Raw Volumes). |
| $CFG{oracle_user} | Scalar | Required<br>Defines the name of the Oracle user. |
| $CFG{oracle_group} | Scalar | Required<br>Defines the primary group of the Oracle user. |

Table 15-3 lists the variables that are used to configure the private IP address for HAIP.

**Table 15-3** Variables for configuring the private IP address for HAIP

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{*host*}{haip} | Scalar | Required<br>Defines the IP address to be configured for HAIP on the node.<br>Repeat this variable for each node in the cluster. For example, if you have two nodes in the cluster, you must provide this variable for each node.<br>For example:<br>`$CFG{sys1}{haip} = "192.168.12.1"`<br>`$CFG{sys2}{haip} = "192.168.12.2"` |

**Table 15-3** Variables for configuring the private IP address for HAIP *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{*host*}{hostname_for_ip} | Scalar | Required<br>Defines the private node name of the IP address to be configured for HAIP on the node.<br>Repeat this variable for each node in the cluster. For example, if you have two nodes in the cluster, you must provide this variable for each node.<br>For example:<br>`$CFG{sys1}{hostname_for_haip} = "sys1-haip";`<br>`$CFG{sys2}{hostname_for_haip} = "sys2-haip";` |
| $CFG{nic_netmask} | Scalar | Required<br>Defines the netmask for the private network. |
| $CFG{nic_add_ip_to_files} | Scalar | Required<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the IP addresses are added to the /etc/hosts file.<br>The value 0 indicates that the IP addresses may already be present in the file. |

Table 15-4 lists the variables that are used to configure the private IP address and PrivNIC resource under VCS.

**Table 15-4** Variables for configuring the private IP address and PrivNIC resource under VCS

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{config_privnic} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that the PrivNIC and private IP address information will be configured for Oracle Clusterware.<br>The value 0 indicates that the PrivNIC and private IP address information will not be configured for Oracle Clusterware. |
| $CFG{privnic_resname} | Scalar | Required<br>Defines the PrivNIC resource name in the main.cf file. |
| $CFG{privnic_interface_priority} | String | Required<br>Defines the priority that determines which NIC will be used in the event of a failover. Set the priority in decreasing order.<br>For example, the following priority setting indicates that eth2 will be given priority in the event of a failover:<br>**$CFG{privnic_interface_priority}="eth2 eth3";** |
| $CFG{*host1*}{privnicip} | Scalar | Required<br>Defines the IP address to be configured for the PrivNIC resource on the node.<br>Repeat this variable for each node in the cluster. For example, if you have two nodes in the cluster, you must provide this variable for each node.<br>For example:<br>`$CFG{sys1}{privnicip} ="192.168.12.1"`<br>`$CFG{sys2}{privnicip} ="192.168.12.2"` |

**Table 15-4** Variables for configuring the private IP address and PrivNIC resource under VCS *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{nic_reuseip} | Scalar | Required<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the existing IP addresses in the /etc/hosts file will be used.<br>The value 0 indicates that the IP addresses will not be reused. |
| $CFG{*host*}{hostname_for_ip} | Scalar | Required<br>Defines the private node name of the IP address (*hostname_for_ip*) for the PrivNIC resource and the node (*system*) for which the resource is configured.<br>Repeat this variable for each node in the cluster. For example, if you have two nodes in the cluster, you must provide this variable for each node.<br>For example:<br>`$CFG{sys1}{hostname_for_ip} ="sys1-priv"`<br>`$CFG{sys2}{hostname_for_ip} ="sys2-priv"` |
| $CFG{nic_netmask} | Scalar | Required<br>Defines the netmask for the private network. |

**Table 15-4** Variables for configuring the private IP address and PrivNIC resource under VCS *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{nic_add_ip_to_files} | Scalar | Required<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the IP addresses are added to the /etc/hosts file.<br>**Note:** Make sure that the IP addresses for the NIC resource are not already present in the files or set the *$CFG{nic_reuseip}* and *$CFG{nic_reusealias}* variables, otherwise the network configuration step fails.<br>The value 0 indicates that the IP addresses may already be present in the file. |
| $CFG{nic_reconfigure_existing_resource} | Scalar | Optional<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the existing PrivNIC resource in the main.cf file will be deleted and reconfigured.<br>The value 0 indicates that the existing PrivNIC resource in the main.cf file will be reused. |
| $CFG{nic_reusealias} | Scalar | Required<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the installer will not check the `/etc/hosts` file to determine whether the host name alias for the private IP addresses exist or not. The installer assumes that the host names alias information is present in the file. Make sure that the alias information is present in the file.<br>The value 0 indicates that the installer checks whether the host name alias information is present in the `/etc/hosts` file. Make sure that the alias information is present in the file otherwise the installation fails. |

Table 15-5 lists the variables that are used to configure the private IP addresses and the MultiPrivNIC resource under VCS.

**Table 15-5** Variables for configuring the private IP addresses and the MultiPrivNIC resource under VCS

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{config_multiprivnic} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that the MultiPrivNIC and private IP address information will be configured for Oracle Clusterware.<br>The value 0 indicates that the MultiPrivNIC and private IP address information will not be configured for Oracle Clusterware. |
| $CFG{multiprivnic_resname} | Scalar | Required<br>Defines the MultiPrivNIC resource name in the main.cf file. |
| $CFG{nic_add_ip_to_files} | Scalar | Required<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the IP addresses are added to the /etc/hosts file.<br>**Note:** Make sure that the IP addresses for the NIC resource are not already present in the files or set the *$CFG{nic_reuseip}* and *$CFG{nic_reusealias}* variables, otherwise the network configuration step fails.<br>The value 0 indicates that the IP addresses may already be present in the file. |

**Table 15-5** Variables for configuring the private IP addresses and the MultiPrivNIC resource under VCS *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{*host1*}{*NIC1*}{multiprivnicip} | List | Required<br>Defines the list of IP addresses for the MultiPrivNIC resource.<br>**Note:** The private IP addresses must be configured for each node and each interface in the cluster.<br>For example, if you have two nodes sys1 and sys2 in the cluster:<br>`$CFG{sys1}{eth1}`<br>`{multiprivnicip}="192.168.12.1";`<br>`$CFG{sys1}{eth2}`<br>`{multiprivnicip}=="192.168.2.1";`<br>`$CFG{sys2}{eth1}`<br>`{multiprivnicip}="192.168.12.2";`<br>`$CFG{sys2}{eth2}`<br>`{multiprivnicip}="192.168.2.2";` |
| $CFG{*host1*}{*NIC1*}{hostname_for_ip} | List | Required<br>Defines the list of private node names (*hostname_for_ip*) for the IP addresses configured in the MultiPrivNIC resource for the interface (*inf*) on the node (*system*) .<br>**Note:** The private IP address must be configured for each node and each interface in the cluster.<br>For example, if you have two nodes sys1 and sys2 in the cluster:<br>`$CFG{sys1}{eth1}`<br>`{hostname_for_ip}="sys1-priv";`<br>`$CFG{sys1}{eth2}`<br>`{hostname_for_ip}="sys1-priv1";`<br>`$CFG{sys2}{eth1}`<br>`{hostname_for_ip}="sys2-priv";`<br>`$CFG{sys2}{eth2}`<br>`{hostname_for_ip}="sys2-priv1";` |

**Table 15-5** Variables for configuring the private IP addresses and the MultiPrivNIC resource under VCS *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{nic_netmask} | Scalar | Required<br>Defines the netmask for the private network. |
| $CFG{nic_reconfigure_existing_resource} | Scalar | Optional<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the existing MultiPrivNIC resource in the main.cf file will be deleted and reconfigured.<br>The value 0 indicates that the existing MultiPrivNIC resource in the main.cf file will be reused. |
| $CFG{nic_reuseip} | Scalar | Required<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the existing IP addresses in the /etc/hosts file will be used.<br>The value 1 indicates that the existing IP addresses in the /etc/hosts or /etc/inet/ipnodes files will be used.<br>The value 0 indicates that the IP addresses will not be reused. |
| $CFG{nic_reusealias} | Scalar | Required<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the installer will not check the `/etc/hosts` file to determine whether the host name alias for the private IP addresses exist or not. The installer assumes that the host names alias information is present in the file. Make sure that the alias information is present in the file.<br>The value 0 indicates that the installer checks whether the host name alias information is present in the `/etc/hosts` file. Make sure that the alias information is present in the file otherwise the installation fails. |

Table 15-6 lists the variables that are used to install Oracle Clusterware.

**Table 15-6** Variables for installing Oracle Clusterware

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{install_oracle_clusterware} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that Oracle Clusterware will be configured.<br>The value 0 indicates that Oracle Clusterware will not be configured. |
| $CFG{oracle_user} | Scalar | Required<br>Defines the name of the Oracle user. |
| $CFG{oracle_group} | Scalar | Required<br>Defines the primary group of the Oracle user. |
| $CFG{oracle_base} | Scalar | Required<br>Defines the base directory for the Oracle RAC installation. |
| $CFG{crs_home} | Scalar | Required<br>Defines the Oracle Clusterware home directory.<br>The value in this variable must be the same as that of the 'ORACLE_HOME' variable in the Oracle Clusterware response file. |
| $CFG{crs_installpath} | Scalar | Required<br>Defines the full path of the Oracle Clusterware installation binaries. |
| $CFG{oracle_version} | Scalar | Required<br>Defines the version of the Oracle RAC binaries (for example, 11.2.0.3.0). This definition is overridden if a different Oracle RAC version is detected during the installation. |

**Table 15-6** Variables for installing Oracle Clusterware *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{crs_responsefile} | Scalar | Required<br>Defines the full path of the Oracle Clusterware response file. |

Table 15-7 lists the variables that are used to install Oracle database.

**Table 15-7** Variables for installing Oracle database

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{install_oracle_database} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that the Oracle RAC database will be configured.<br>The value 0 indicates that the Oracle RAC database will not be configured. |
| $CFG{oracle_user} | Scalar | Required<br>Defines the name of the Oracle user. |
| $CFG{oracle_group} | Scalar | Required<br>Defines the primary group of the Oracle user. |
| $CFG{oracle_base} | Scalar | Required<br>Defines the base directory for the Oracle RAC installation. |
| $CFG{crs_home} | Scalar | Required<br>Defines the Oracle Clusterware home directory.<br>The value in this variable must be the same as that of the 'ORACLE_HOME' variable in the Oracle Clusterware response file. |

**Table 15-7** Variables for installing Oracle database *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{db_home} | Scalar | Required<br>Defines the Oracle RAC database home directory.<br>The value in this variable must be the same as that of the 'ORACLE_HOME' variable in the Oracle RAC database response file. |
| $CFG{db_installpath} | Scalar | Required<br>Defines the full path of the Oracle RAC database installation binaries. |
| $CFG{oracle_version} | Scalar | Required<br>Defines the version of the Oracle RAC binaries (for example, 11.2.0.3.0). This definition is overridden if a different Oracle RAC version is detected during the installation. |
| $CFG{db_responsefile} | Scalar | Required<br>Defines the full path of the Oracle database response file. |

Table 15-8 lists the variables that are used to configure CSSD resource.

**Table 15-8** Variables for configuring CSSD resource

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{config_cssd_agent} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that the CSSD agent will be configured after Oracle RAC installation.<br>The value 0 indicates that the CSSD agent will not be configured after Oracle RAC installation. |

**Table 15-8** Variables for configuring CSSD resource *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{reconfigure_cssd_resource} | Scalar | Required<br>Defines a boolean value 0 or 1.<br>The value 1 indicates that the SF Oracle RAC installer deletes the existing CSSD resource from the main.cf file and reconfigures it.<br>The value 0 indicates that the SF Oracle RAC installer does not delete and reconfigure the resource. The installer exits the task with the message that the resource exists. |

Table 15-9 lists the variables that are used to relink Oracle RAC libraries.

**Table 15-9** Variables for relinking Oracle RAC libraries

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{relink_oracle_database} | Scalar | Required<br>Defines a Boolean value 0 or 1.<br>The value 1 indicates that the SF Oracle RAC libraries will be relinked with the Oracle RAC database after Oracle RAC installation.<br>The value 0 indicates that the SF Oracle RAC libraries will not be relinked with the Oracle RAC database after Oracle RAC installation. |
| $CFG{oracle_user} | Scalar | Required<br>Defines the name of the Oracle user. |
| $CFG{oracle_group} | Scalar | Required<br>Defines the primary group of the Oracle user. |

**Table 15-9** Variables for relinking Oracle RAC libraries *(continued)*

| Variable | List or Scalar | Description |
|---|---|---|
| $CFG{crs_home} | Scalar | Required<br>Defines the Oracle Clusterware home directory.<br>The value in this variable must be the same as that of the 'ORACLE_HOME' variable in the Oracle Clusterware response file. |
| $CFG{db_home} | Scalar | Required<br>Defines the Oracle RAC database home directory.<br>The value in this variable must be the same as that of the 'ORACLE_HOME' variable in the Oracle RAC database response file. |
| $CFG{oracle_version} | Scalar | Required<br>Defines the version of the Oracle RAC binaries (for example, 11.2.0.3.0). This definition is overridden if a different Oracle RAC version is detected during the installation. |

# Sample response file for installing Oracle RAC

The following sample response file installs Oracle RAC 11g Release 2 and performs the following Oracle RAC pre-installation and installation tasks on two nodes, sys1 and sys2, in the cluster:

- Creates a disk group for OCR and voting disk storage
- Creates OCR and voting disk storage on raw volumes
- Configures the PrivNIC agent for high availability on both nodes
- Installs Oracle Clusterware
- Installs Oracle database
- Relinks the Oracle RAC libraries with SF Oracle RAC libraries
- Configures the CSSD agent

```
#
# Configuration Values:
#
our %CFG;

$CFG{prod}="ENTERPRISE80";
$CFG{systems}=[ qw(sys1 sys2) ];
$CFG{opt}{configure}=1;

$CFG{config_privnic}=1;
$CFG{privnic_resname}="ora_priv";
$CFG{privnic_interface_priority}="eth2 eth3";
$CFG{sys1}{privnicip}=" 192.168.12.1";
$CFG{sys1}{hostname_for_ip}="sys1-priv";
$CFG{sys2}{privnicip}=" 192.168.12.2";
$CFG{sys2}{hostname_for_ip}="sys2-priv";
$CFG{nic_netmask}="255.255.255.0";
$CFG{nic_add_ip_to_files}=1;
$CFG{nic_reuseip}=1;
$CFG{nic_reusealias}=1;

$CFG{create_ocr_vote_storage}=1;
$CFG{ocrvotedgoption}=0;
$CFG{oracle_group}="dba";
$CFG{grid_user}="grid";
$CFG{oracle_user}="oracle";
$CFG{ocrvolname}="ocrvotevol";
$CFG{ocrvolsize}=640;
$CFG{enable_mirroring} = 1;
$CFG{ocrvotedgname}="ocrvotedg";
$CFG{ocrvotedisks}=[qw(Disk_1 Disk_2)];
$CFG{ocrvotescheme}=0;
$CFG{votevolname}="votevol";
$CFG{votevolsize}=500;

$CFG{install_oracle_clusterware}=1;
$CFG{install_oracle_database}=1;
$CFG{oracle_base} = "/u01/app/oracle";
$CFG{crs_home}="/u01/app/oracle/product/11.2.0/crshome";
$CFG{db_home}="/u01/app/oracle/product/11.2.0/dbhome_1";
$CFG{crs_installpath}="/cdrom/oracle/clusterware";
$CFG{db_installpath}="/cdrom/oracle/database";
$CFG{crs_responsefile}="/oracle/crs.rsp";
```

```
$CFG{db_responsefile} = "/oracle/db.rsp";

$CFG{config_cssd_agent}=1;
$CFG{relink_oracle_database}=1;
$CFG{uploadlogs}=0;

1;
```

Chapter 16

# Performing Oracle RAC post-installation tasks

This chapter includes the following topics:


- Adding Oracle RAC patches or patchsets
- Configuring the CSSD resource
- Preventing automatic startup of Oracle Clusterware/Grid Infrastructure
- Relinking the SF Oracle RAC libraries with Oracle RAC
- Updating the CSS misscount setting
- Creating the Oracle RAC database
- Configuring VCS service groups for Oracle RAC
- Preventing automatic database startup
- Removing existing PrivNIC or MultiPrivNIC resources
- Removing permissions for communication


## Adding Oracle RAC patches or patchsets

If you installed Oracle 11.2.0.1, apply the Oracle patch 8649805 to integrate Oracle Clusterware/Grid Infrastructure with VCS.

To install the required patches or patchsets, review the notes that accompany the patch or patchset.

Before installing any Oracle RAC patch or patchset software:

- Review the latest information on supported Oracle RAC patches and patchsets:

https://www.veritas.com/support/en_US/article.DOC4039

- You must have installed the base version of the Oracle RAC software.

# Configuring the CSSD resource

You must configure the CSSD resource to ensure that the CSSD dependencies on the resources that manage OCR and voting disk and the private IP address are satisfied before Oracle Clusterware/Grid Infrastructure starts.

---

**Note:** It is mandatory to use CSSD agent in SF Oracle RAC installations. Using the CSSD agent ensures adequate handling of inter-dependencies, thus preventing the premature startup of Oracle Clusterware/Grid Infrastructure, which causes cluster failures.

---

Before you configure the CSSD resource, make sure that the following requirements are satisfied:

1. Oracle Clusterware/Grid Infrastructure is up and running.
2. OCR and voting disk is configured on CVM raw volumes or CFS and managed by VCS.

---

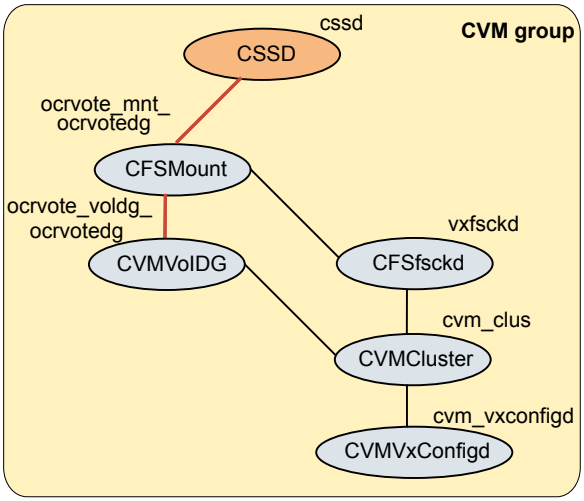
**Note:** The SF Oracle RAC installer configures the OCR, voting disk, and PrivNIC/MultiPrivNIC resources in the cvm group. If you configured one of these resources manually, make sure that these resources are placed in the cvm group. If the resources are not in the same group, configure the CSSD resource manually.

See "Configuring the CSSD resource manually" on page 250.

---

Figure 16-1 illustrates the configuration performed by the SF Oracle RAC installer when you configure external redundancy for OCR and voting disk for Oracle RAC. In the figure, the CSSD resource is configured under the CVM group.

**Figure 16-1** CSSD configuration by SF Oracle RAC installer for external redundancy




Use one of the following ways to configure the CSSD resource:

| | |
|---|---|
| SF Oracle RAC script-based installer | See "Configuring the CSSD resource using the SF Oracle RAC script-based installer" on page 248. |
| Manual | See "Configuring the CSSD resource manually" on page 250. |

## Configuring the CSSD resource using the SF Oracle RAC script-based installer

Configure the CSSD resource using the SF Oracle RAC installer if the OCR and voting disk storage is configured on CFS.

---

**Note:** If the OCR and voting disk storage is configured on ASM disk groups, configure the CSSD resource manually.

---

The installer performs the following configuration tasks:

- Adds the CSSD resource to the VCS configuration in the cvm group.

  ---

  **Note:** If the CSSD resource already exists, the installer enables reconfiguration of the resource by deleting the existing resource.

  ---

- Sets the dependency of the CSSD resource on the CFSMount or CVMVolDg resources that manage OCR and voting disk.
- Disables automatic startup of Oracle Clusterware/Grid Infrastructure.
- Enables the CSSD resource and saves the new configuration.

**To configure the CSSD resource using the SF Oracle RAC script-based installer**

1. Start the SF Oracle RAC installer, if it is not already running. Select the option **Post Oracle Installation Tasks**.

```
1)  Configure SF Oracle RAC sub-components
2)  Prepare to Install Oracle
3)  Install Oracle Clusterware/Grid Infrastructure and Database
4)  Post Oracle Installation Tasks
5)  Exit SF Oracle RAC Configuration
Choose option: [1-5,q] (1) 4
```

2. Select the option **Configure CSSD agent**.
3. Enter the name of the CSSD resource.

```
Please enter the name of CSSD resource: [b] (cssd)
```

4. Enter the full path to the Oracle Clusterware/Grid Infrastructure home directory.

```
Enter absolute path of Oracle Clusterware/Grid Infrastructure
Home directory: [b]
```

   The agent uses this information to locate the clusterware process binaries. If the information is not provided, the agent reads the information from the Oracle configuration file.

   The installer verifies that Oracle Clusterware/Grid Infrastructure is running on all the nodes. It also checks that the PrivNIC/MultiPrivNIC (for supported versions), OCR, and voting disk resources are configured in the same service group.

5. Press **Return** to continue with the configuration. The installer reads the resource and group mappings for the CSSD agent from the VCS configuration file and displays the information.

6 Enter **y** to continue with the configuration. Review the messages as the installer configures the CSSD agent and sets the appropriate dependencies.

7 Press **Return** to return to the installer menu.

At the end of the configuration, the installer disables automatic startup of Oracle Clusterware/Grid Infrastructure.

## Configuring the CSSD resource manually

Add the `cssd` resource to the VCS configuration and set CSSD dependencies on the resources that manage the OCR and voting disk for Oracle Clusterware/Grid Infrastructure.

**To configure the CSSD resource**

1 Change the permission on the VCS configuration file to read-write mode:

```
# haconf -makerw
```

2 Add the CSSD resource to the cvm group:

```
# hares -add cssd_resname CSSD cvm_grpname
```

3 Modify the CSSD resource attributes:

```
# hares -modify cssd_resname \
CRSHOME "/u01/app/11.2.0/grid"
# hares -modify cssd_resname Critical 0
# hares -override cssd_resname OnlineWaitLimit
# hares -modify cssd_resname OnlineWaitLimit 5
# hares -override cssd_resname OfflineWaitLimit
# hares -modify cssd_resname OfflineWaitLimit 3
```

4 Enable the CSSD resource:

```
# hares -modify cssd_resname Enabled 1
```

5 Set the dependency of the CSSD resource on the CFSMount or CVMVolDg resources that manage OCR and voting disk.

If you configured external redundancy for OCR and voting disk:

- If you configured OCR and voting disk on CVM raw volumes:

```
# hares -link cssd ocrvotevol_resname
```

- If you configured OCR and voting disk on CFS:

  ```
  # hares -link cssd ocrvotemnt_resname
  ```

If you configured normal redundancy for OCR and voting disk, additionally configure atleast resource dependency for OCR and voting disk:

- ```
  # hares -link cssd_resname ocrmnt_1_resname,\
  ocrmnt_2_resname,ocrmnt_3_resname -min x
  # hares -link cssd_resname votemnt_1_resname,\
  votemnt_2_resname,votemnt_3_resname -min y
  ```

---

**Note:**

See the Oracle documentation to determine the minimum OCR (x) and voting disks (y) required by Oracle Clusterware/Grid.

---

6 If the Oracle Clusterware/Grid Infrastructure and the Oracle database binaries are on CFS, set the dependencies between the CSSD resource and the CFSMount resources for the binaries manually:

```
# hares -link cssd clusbin_mnt_resname
# hares -link cssd orabin_mnt_resname
```

7 Change the permission on the VCS configuration file to read-only mode:

```
# haconf -dump -makero
```

# Preventing automatic startup of Oracle Clusterware/Grid Infrastructure

Perform this step only if you have manually configured the CSSD resource. If you have configured the CSSD resource using the installer, the installer disables automatic startup of Oracle Clusterware/Grid Infrastructure at the end of the CSSD configuration.

The use of the CSSD agent is mandatory to ensure adequate handling of service group inter-dependencies and thereby prevent the premature startup of Oracle Clusterware/Grid Infrastructure. Therefore, disable automatic startup of Oracle Clusterware/Grid Infrastructure when the system starts.
```

**To prevent automatic startup of Oracle Clusterware/Grid Infrastructure**

1 Log in as the root user on each node in the cluster.

2 Disable automatic startup of Oracle Clusterware/Grid Infrastructure.

```
# $GRID_HOME/bin/crsctl disable crs
```

# Relinking the SF Oracle RAC libraries with Oracle RAC

If you added or upgraded the Oracle patches, you must relink the SF Oracle RAC libraries—VCSMM and ODM—with Oracle. Relinking the libraries enables coordinated exchange of cluster membership information and protection of data.

Use one of the following ways to relink the libraries:

| | |
|---|---|
| SF Oracle RAC script-based installer | See "Relinking the SF Oracle RAC libraries with Oracle RAC using the SF Oracle RAC script-based installer" on page 252. |
| Manual | See "Relinking SF Oracle RAC libraries with Oracle RAC manually" on page 253. |

## Relinking the SF Oracle RAC libraries with Oracle RAC using the SF Oracle RAC script-based installer

Perform the steps in the following procedure to relink the libraries using the SF Oracle RAC script-based installer.

**To relink the SF Oracle RAC libraries with Oracle RAC**

1. Return to the SF Oracle RAC installer menu, and select the option **Post Oracle Installation**.

```
1)  Configure SF Oracle RAC sub-components
2)  Prepare to Install Oracle
3)  Install Oracle Clusterware/Grid Infrastructure and Database
4)  Post Oracle Installation Tasks
5)  Exit SF Oracle RAC Configuration
Choose option: [1-5,q] (1) 4
```

2. Select the option **Relink Oracle Database Binary**.

```
1)  Configure CSSD agent
2)  Relink Oracle Database Binary
3)  Exit SF Oracle RAC Configuration
b)  Back to previous menu
Choose option: [1-3,b,q] (1) 2
```

3. Provide the Oracle environment information—user name, group name, CRS_HOME (or GRID_HOME), ORACLE_HOME. Based on this information, the installer detects the version of Oracle installed.

---

**Note:** You need to provide the Oracle environment information only if you quit the installer after installing Oracle Clusterware/Grid Infrastructure and the Oracle database.

---

```
Enter Oracle UNIX user name: [b] oracle
Enter Oracle UNIX group name: [b] (oinstall)
Enter absolute path of Oracle Clusterware/Grid Infrastructure
Home directory: [b]
Enter absolute path of Oracle Database Home directory: [b]
.
.
Do you want to continue? [y,n,q] (y)
```

4. Review and confirm the Oracle database information.

   The installer starts relinking the libraries.

# Relinking SF Oracle RAC libraries with Oracle RAC manually

The relinking process involves the following tasks:

- Linking Veritas Membership library
- Linking the ODM library

The Oracle RAC and Veritas library locations for linking the Oracle libraries with SF Oracle RAC are listed in the following tables. You may use these tables as reference when you relink the libraries as described in this section.

The following table lists the Oracle RAC and Veritas library locations for the supported Oracle RAC versions.

**Table 16-1** Oracle RAC and Veritas library locations for VCSMM and ODM

| Libraries | Oracle RAC library | Veritas library |
|---|---|---|
| VCSMM | `$GRID_HOME/lib/libskgxn2.so` | `/etc/ORCLcluster/lib/libskgxn2.so` |
| ODM | For Oracle RAC 18c:<br>`$ORACLE_HOME/rdbms/lib/odm/libodm<version>.so` | `/opt/VRTSodm/lib64/libodm.so` |

## Linking Veritas Membership library

You need to link the SF Oracle RAC libraries with Oracle RAC to enable coordinated exchange of cluster membership information and protection of data integrity. Oracle provides an API called `skgxn` (system kernel generic interface node membership) to obtain information on membership. SF Oracle RAC implements the API as a library linked to Oracle Grid Infrastructure after the Oracle Grid Infrastructure installation. Oracle uses the linked skgxn library (libskgxn) to make ioctl calls to VCSMM, which in turn obtains membership information for clusters and instances.

**To link Veritas Membership library**

1. Log in as the Oracle grid user.
2. Change to the `$GRID_HOME/lib` directory:

```
$ cd $GRID_HOME/lib
```

3. If the `$GRID_HOME/lib/libskgxn2.so` library is not linked to `/etc/ORCLcluster/lib/libskgxn2.so`, perform the following steps to link the library:
   - Back up Oracle's libskgxn2 library located in the $GRID_HOME directory:

```
$ mv libskgxn2.so libskgxn2.so.oracle.`date `+%m_%d_%Y-%H_%M_%S'`
```

- As the grid user, link the Veritas Membership library:

```
$ ln -s /etc/ORCLcluster/lib/libskgxn2.so libskgxn2.so
```

## Linking the ODM library

Perform the steps in the procedure on each node if the Oracle libraries are on local storage. If the Oracle libraries are installed on shared storage, copy the libraries on one node only. Use the mount command to check that the file system containing the Oracle libraries are mounted.

**To link the Veritas ODM library**

1. Log in as the Oracle user.
2. Change to the `$ORACLE_HOME/lib` directory:

```
$ cd $ORACLE_HOME/lib
```

3. Back up Oracle's ODM library:

```
$ mv libodm<version>.so libodm<version>.so.oracle-`date +%m_%d_%Y-%H_%M_%
```

4. Link the Veritas ODM library with Oracle's `libodm` library:

   To link the Veritas Extension for ODM with the Oracle database manually, see the following Technote:

   https://www.veritas.com/support/en_US/article.TECH225500

   To disable direct NFS:

```
$ cd $ORACLE_HOME/rdbms/lib
$ make -f ins_rdbms.mk dnfs_off OR (make -f ins_rdbms.mk dnfs_off)

$ cd $ORACLE_HOME/rdbms/lib/odm
$ ln -s /opt/VRTSodm/lib64/libodm.so libodm18.so
```

# Updating the CSS misscount setting

Oracle RAC requires the CSS misscount value to be set to 600 seconds with vendor clusterware.

Check the current CSS misscount setting:

```
# $GRID_HOME/bin/crsctl get css misscount
```

If the current CSS misscount setting is not set to 600 seconds, update the setting as described in the following procedure.

**To update the CSS misscount setting**

1. Shut down Oracle Clusterware/Grid Infrastructure on all but one node.
2. Log in as the root user on the active node and set the CSS misscount interval to 600 seconds:

```
# $GRID_HOME/bin/crsctl set css misscount 600
```

3. Restart Oracle Clusterware/Grid Infrastructure.
4. Verify that the setting is updated:

```
# $GRID_HOME/bin/crsctl get css misscount
600
```

# Creating the Oracle RAC database

Create the Oracle RAC database on CVM raw volumes or CFS. Use your own tools or scripts, or review the guidelines on using the Oracle DBCA (Database Creation Assistant) tool to create the database.

For instructions on creating an Oracle RAC database:

See "About creating a test database" on page 445.

For more information, see the Oracle RAC documentation.

---

**Note:** If you plan to configure global clusters, then set up the Oracle RAC database only on the primary site. On the secondary site, the database will be replicated.

---

# Configuring VCS service groups for Oracle RAC

Before you configure VCS service groups for Oracle databases, review the following information:

- Supported types of database management
  See "Supported types of database management" on page 257.
- Sample service group configurations
  See "Sample service group configurations" on page 257.

To configure the VCS service groups for Oracle RAC:

| | |
|---|---|
| Traditional Oracle RAC databases | See "Configuring VCS service groups manually for traditional Oracle databases" on page 262. |
| Container Oracle RAC databases | See "Configuring VCS service groups manually for container Oracle databases" on page 266. |

# Supported types of database management

Table 16-2 lists the database management options in Oracle RAC that are supported by SF Oracle RAC.

**Table 16-2** Supported types of database management

| Management type | Description |
|---|---|
| Policy-managed databases | In policy-managed databases, administrators specify the server pool on which the database instances run. Oracle Grid Infrastructure determines the server on which the database instances run.<br>For more information, see the Oracle documentation. |
| Administrator-managed databases | In administrator-managed databases environments, the administrator specifies the servers on which the databases instances run.<br>For more information, see the Oracle documentation. |

# Sample service group configurations

You can set up the Oracle database to be managed by one of the following clusterwares:

- Cluster Server

  **Note:** Veritas recommends that the Oracle database be configured under VCS.

  When the database is configured under VCS:
  You can choose to configure the service group in a way that insulates all the databases from failure in any of the databases in the group.
  VCS manages the start and stop sequence of the applications and the database.
  See "Sample service group configurations with the VCS Oracle agent" on page 258.

- Oracle Clusterware/Grid Infrastructure
  See "Sample service group configurations without the VCS Oracle agent" on page 260.

## Sample service group configurations with the VCS Oracle agent

This section illustrates sample service group configurations with the VCS Oracle agent for multiple databases.

Figure 16-2 illustrates a service group configuration with the VCS Oracle agent.

**Figure 16-2** Service group configuration with the VCS Oracle agent

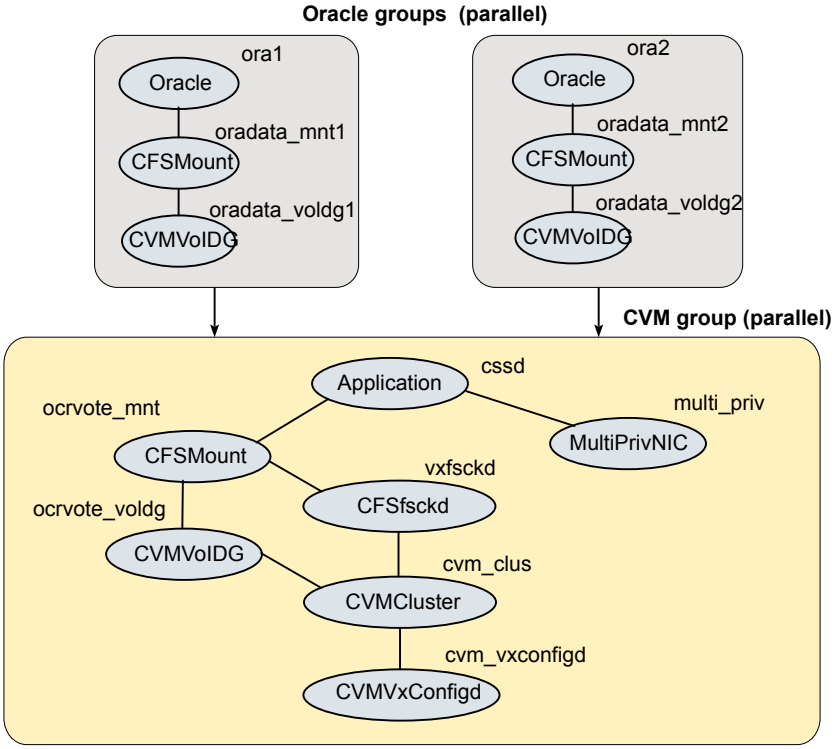



Figure 16-2 illustrates a service group configuration for container and pluggable databases with the VCS Oracle agent.

**Figure 16-3** Service group configuration for container and pluggable databases with the VCS Oracle agent

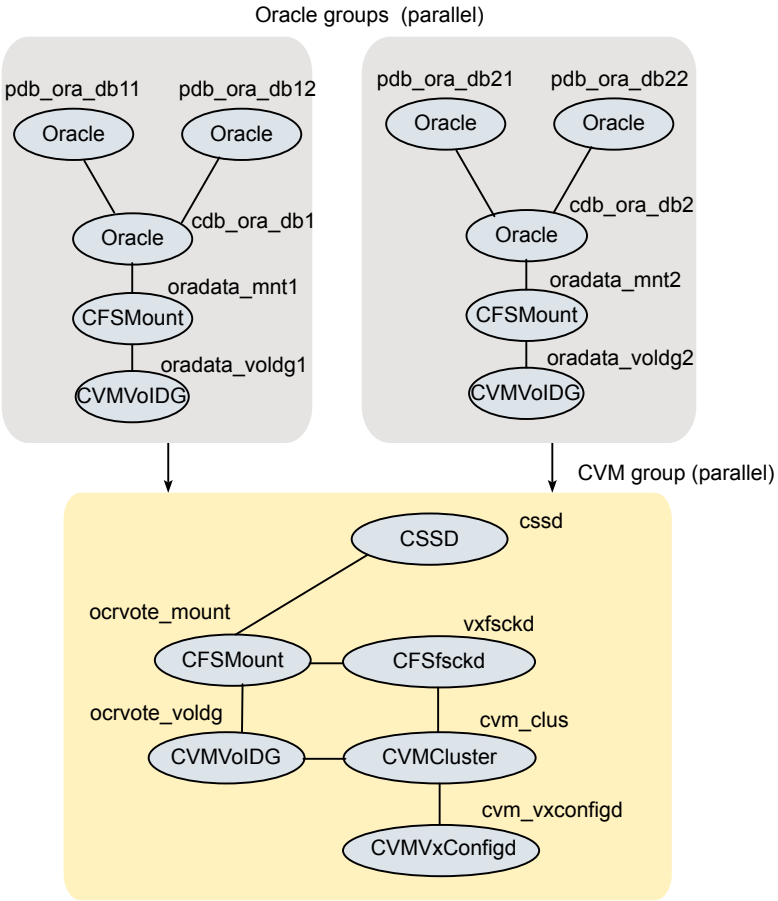
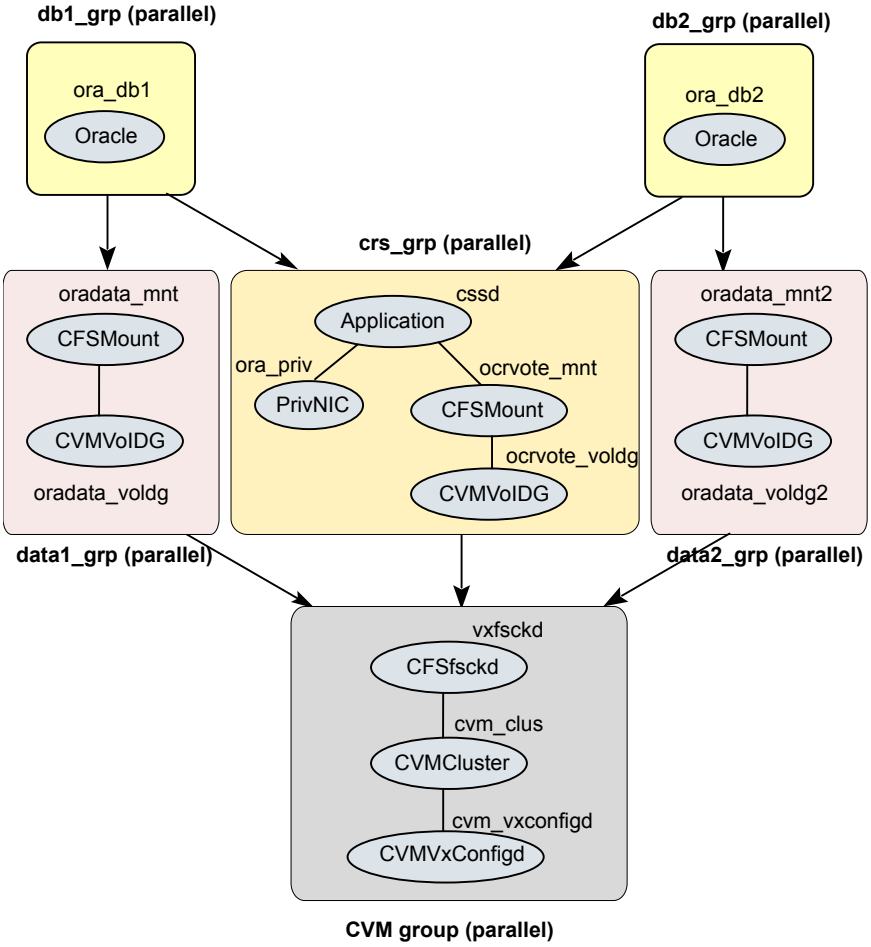



Figure 16-4 illustrates an alternate service group configuration with the VCS Oracle agent.

**Figure 16-4** Service group configuration with the VCS Oracle agent (alternate configuration)




## Sample service group configurations without the VCS Oracle agent

Figure 16-5 illustrates a sample configuration in the absence of the VCS Oracle agent for single database configurations.

**Figure 16-5** Service group configuration without VCS Oracle agent (single database)

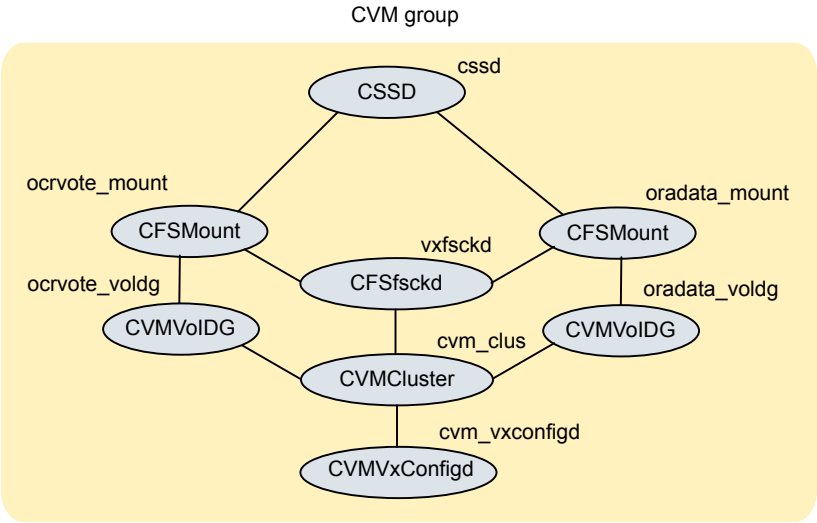
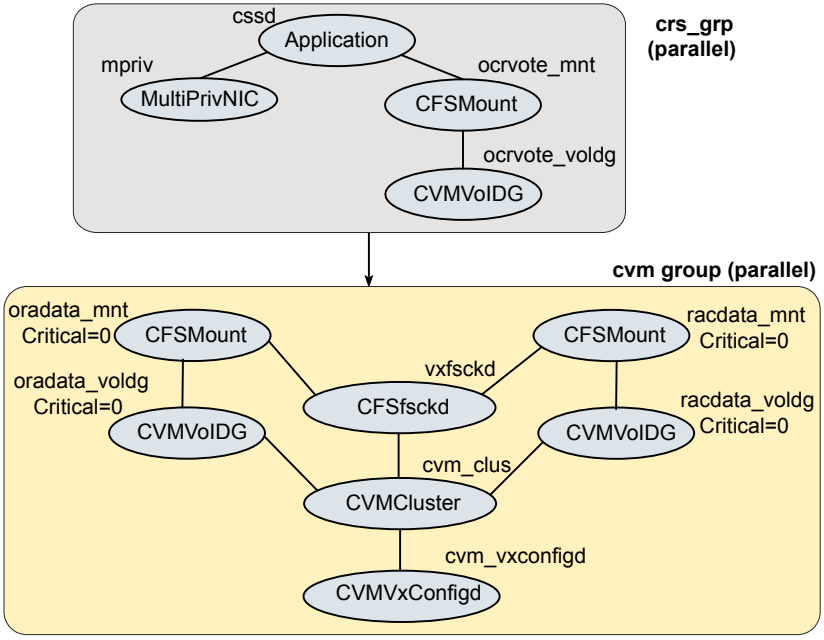



Figure 16-6 illustrates a sample service group configuration in the absence of the VCS Oracle agent for multiple database configurations.

**Figure 16-6** Service group configuration without the VCS Oracle agent (multiple databases)




In a service group configuration without the VCS Oracle agent, Oracle Clusterware/Grid Infrastructure controls the database. An online local firm dependency exists between the Oracle Clusterware/Grid Infrastructure group and the CVM group. When the system starts, the CVM group brings up the volume and mount points for the databases. The Oracle Clusterware/Grid Infrastructure group brings up the OCR and voting disk, configures the private IP address for Oracle Clusterware/Grid Infrastructure, and starts Oracle Clusterware/Grid Infrastructure. Oracle Clusterware/Grid Infrastructure starts the database and the application is brought online.

---

**Note:** In a service group configuration without the VCS Oracle agent, when the system starts, all volumes and mount points MUST be online for the dependent service groups to be online.

---

## Configuring VCS service groups manually for traditional Oracle databases

This section describes the steps to configure the VCS service group manually for traditional Oracle databases.

See Figure 16-2 on page 258.

The following procedure assumes that you have created the database.

**To configure the Oracle service group manually for traditional Oracle databases**

1. Change the cluster configuration to read-write mode:

```
# haconf -makerw
```

2. Add the service group to the VCS configuration:

```
# hagrp -add oradb_grpname
```

3. Modify the attributes of the service group:

```
# hagrp -modify oradb_grpname Parallel 1

# hagrp -modify oradb_grpname SystemList node_name1 0  node_name2 1

# hagrp -modify oradb_grpname AutoStartList node_name1 node_name2
```

4. Add the CVMVolDg resource for the service group:

```
# hares -add oradbdg_resname CVMVolDg oradb_grpname
```

5. Modify the attributes of the CVMVolDg resource for the service group:

```
# hares -modify oradbdg_resname CVMDiskGroup oradb_dgname
# hares -modify oradbdg_resname CVMActivation sw
# hares -modify oradbdg_resname CVMVolume oradb_volname
```

6. Add the CFSMount resource for the service group:

```
# hares -add oradbmnt_resname CFSMount oradb_grpname
```

7. Modify the attributes of the CFSMount resource for the service group:

```
# hares -modify oradbmnt_resname MountPoint "oradb_mnt"
# hares -modify oradbmnt_resname BlockDevice \
"/dev/vx/dsk/oradb_dgname/oradb_volname"
```

8. Add the Oracle RAC database instance to the service group:

```
# hares -add db_resname Oracle oradb_grpname
```

**9** Modify the attributes of the Oracle resource for the service group:

```
# hares -modify db_resname Owner oracle
# hares -modify db_resname Home "db_home"
# hares -modify db_resname StartUpOpt SRVCTLSTART
# hares -modify db_resname ShutDownOpt SRVCTLSTOP
```

**10** For administrator-managed databases, perform the following steps:

- Localize the Sid attribute for the Oracle resource:

```
# hares -local db_resname Sid
```

- Set the Sid attributes for the Oracle resource on each system:

```
# hares -modify db_resname Sid oradb_sid_node1 -sys node_name1
# hares -modify db_resname Sid oradb_sid_node2 -sys node_name2
```

**11** For policy-managed databases, perform the following steps:

- Modify the attributes of the Oracle resource for the service group:

```
# hares -modify db_resname DBName db_name
# hares -modify db_resname ManagedBy POLICY
```

- Set the Sid attribute to the Sid prefix for the Oracle resource on all systems:

```
# hares -modify db_resname Sid oradb_sid_prefix
```

---

**Note:** The Sid prefix is displayed on the confirmation page during database creation. The prefix can also be determined by running the following command :

```
# grid_home/bin/crsctl status resource ora.db_name.db -f | grep
GEN_USR_ORA_INST_NAME@ | tail -1 | sed 's/.*=//' | sed
's/_[0-9]$//'
```

---

- Set the IntentionalOffline attribute for the resource to 1 and make sure that the health check monitoring is disabled:

```
# hares -override db_resname IntentionalOffline
# hares -modify db_resname IntentionalOffline 1
# hares -modify db_resname MonitorOption 0
```

**12** For pluggable databases, perform the following steps:

- Modify the attributes of the Oracle resource for the service group:

```
# hares -modify db_resname DBName db_name
# hares -modify pdb_resname PDBName pdb_name
# hares -modify pdb_resname ManagedBy POLICY
```

- Set the Sid attribute to the Sid prefix for the Oracle resource on all systems:

```
# hares -modify db_resname Sid oradb_sid_prefix
# hares -modify pdb_resname Sid orapdb_sid_prefix
```

---

**Note:** The Sid prefix is displayed on the confirmation page during database creation. The prefix can also be determined by running the following command :

```
# grid_home/bin/crsctl status resource ora.pdb_name.db -f |
grep GEN_USR_ORA_INST_NAME@ | tail -1 | sed 's/.*=//' | sed
's/_[0-9]$//'
```

---

- Set the IntentionalOffline attribute for the resource to 1 and make sure that the health check monitoring is disabled:

```
# hares -override db_resname IntentionalOffline
# hares -modify db_resname IntentionalOffline 1
# hares -modify db_resname MonitorOption 0

# hares -override pdb_resname IntentionalOffline
# hares -modify pdb_resname IntentionalOffline 1
# hares -modify pdb_resname MonitorOption 0
```

**13** Set the dependencies between the CFSMount resource and the CVMVolDg resource for the Oracle service group:

```
# hares -link oradbmnt_resname oradbdg_resname
```

**14** Set the dependencies between the Oracle resource and the CFSMount resource for the Oracle service group:

```
# hares -link db_resname oradbmnt_resname
```

**15** Create an online local firm dependency between the oradb1_grp service group and the cvm service group:

```
# hagrp -link oradb_grpname cvm_grpname online local firm
```

**16** Enable the Oracle service group:

```
# hagrp -enableresources oradb_grpname
```

**17** Change the cluster configuration to the read-only mode:

```
# haconf -dump -makero
```

**18** Bring the Oracle service group online on all the nodes:

```
# hagrp -online oradb_grpname -any
```

---

**Note:** For policy-managed databases: When VCS starts or when the administrator attempts to bring the Oracle resource online, if the server is not part of the server pool associated with the database, the resource will remain offline. If Oracle Grid Infrastructure decides to move the server from the server pool, the database will be brought offline by the Oracle Grid Infrastructure and the oracle resource moves to offline state.

---

For more information and instructions on configuring the service groups using the CLI:

See the *Cluster Server Administrator's Guide*.

## Configuring VCS service groups manually for container Oracle databases

This section describes the steps to configure the VCS service group manually for container Oracle databases.

See Figure 16-2 on page 258.

The following procedure assumes that you have created the database.

**To configure the Oracle service group manually for container Oracle databases**

**1** Change the cluster configuration to read-write mode:

```
# haconf -makerw
```

**2** Add the service group to the VCS configuration:

```
# hagrp -add oradb_grpname
```

3 Modify the attributes of the service group:

```
# hagrp -modify oradb_grpname Parallel 1

# hagrp -modify oradb_grpname SystemList node_name1 0  node_name2 1

# hagrp -modify oradb_grpname AutoStartList node_name1 node_name2
```

4 Add the CVMVolDg resource for the service group:

```
# hares -add oradbdg_resname CVMVolDg oradb_grpname
```

5 Modify the attributes of the CVMVolDg resource for the service group:

```
# hares -modify oradbdg_resname CVMDiskGroup oradb_dgname
# hares -modify oradbdg_resname CVMActivation sw
# hares -modify oradbdg_resname CVMVolume oradb_volname
```

6 Add the CFSMount resource for the service group:

```
# hares -add oradbmnt_resname CFSMount oradb_grpname
```

7 Modify the attributes of the CFSMount resource for the service group:

```
# hares -modify oradbmnt_resname MountPoint "oradb_mnt"
# hares -modify oradbmnt_resname BlockDevice \
"/dev/vx/dsk/oradb_dgname/oradb_volname"
```

8 Add the container and pluggable Oracle RAC resources to the service group:

```
# hares -add cdb_resname Oracle oradb_grpname

# hares -add pdb_resname Oracle oradb_grpname
```

**9** Modify the attributes of the container and pluggable Oracle resources for the service group:

For container Oracle resource:

```
# hares -modify cdb_resname Owner oracle
# hares -modify cdb_resname Home "db_home"
# hares -modify cdb_resname StartUpOpt SRVCTLSTART
# hares -modify cdb_resname ShutDownOpt SRVCTLSTOP
```

For pluggable Oracle resource:

```
# hares -modify pdb_resname Owner oracle
# hares -modify pdb_resname Home "db_home"
# hares -modify pdb_resname StartUpOpt STARTUP
# hares -modify pdb_resname ShutDownOpt IMMEDIATE
```

**10** For container databases that are administrator-managed, perform the following steps:

- Localize the Sid attribute for the container Oracle resource:

```
# hares -local cdb_resname Sid
```

- Set the Sid attributes for the container Oracle resource on each system:

```
# hares -modify cdb_resname Sid oradb_sid_node1 -sys node_name1
# hares -modify cdb_resname Sid oradb_sid_node2 -sys node_name2
```

For pluggable databases that reside in administrator-managed container databases, perform the following steps:

- Localize the Sid attribute for the pluggable Oracle resource:

```
# hares -local pdb_resname Sid
```

- Set the Sid attributes for the pluggable Oracle resource on each system:

```
# hares -modify pdb_resname Sid oradb_sid_node1 -sys node_name1
# hares -modify pdb_resname Sid oradb_sid_node2 -sys node_name2
```

- Set the PDBName attribute for the pluggable Oracle database:

```
# hares -modify pdb_resname PDBName pdbname
```

**11** For container databases that are policy-managed, perform the following steps:

- Modify the attributes of the container Oracle resource for the service group:

```
# hares -modify cdb_resname DBName db_name
# hares -modify cdb_resname ManagedBy POLICY
```

- Set the Sid attribute to the Sid prefix for the container Oracle resource on all systems:

```
# hares -modify cdb_resname Sid oradb_sid_prefix
```

---

**Note:** The Sid prefix is displayed on the confirmation page during database creation. The prefix can also be determined by running the following command :

```
# grid_home/bin/crsctl status resource ora.db_name.db -f | grep
GEN_USR_ORA_INST_NAME@ | tail -1 | sed 's/.*=//' | sed
's/_[0-9]$//'
```

---

- Set the IntentionalOffline attribute for the resource to 1 and make sure that the health check monitoring is disabled:

```
# hares -override cdb_resname IntentionalOffline
# hares -modify cdb_resname IntentionalOffline 1
# hares -modify cdb_resname MonitorOption 0
```

For pluggable databases that reside in policy-managed container databases, perform the following steps:

- Modify the attributes of the pluggable Oracle resource for the service group:

```
# hares -modify pdb_resname DBName db_name
# hares -modify pdb_resname ManagedBy POLICY
```

- Set the Sid attribute to the Sid prefix for the pluggable Oracle resource on all systems:

```
# hares -modify pdb_resname Sid oradb_sid_prefix
```

---

**Note:** The Sid prefix is displayed on the confirmation page during database creation. The prefix can also be determined by running the following command :

```
# grid_home/bin/crsctl status resource ora.db_name.db -f | grep
GEN_USR_ORA_INST_NAME@ | tail -1 | sed 's/.*=//' | sed
's/_[0-9]$//'
```

---

- Set the IntentionalOffline attribute for the resource to 1 and make sure that the health check monitoring is disabled:

```
# hares -override pdb_resname IntentionalOffline
# hares -modify pdb_resname IntentionalOffline 1
# hares -modify pdb_resname MonitorOption 0
```

- Set the PDBName attribute for the pluggable Oracle database:

```
# hares -modify pdb_resname PDBName pdbname
```

**12** Set the dependency between the pluggable database resources and the corresponding container database resource:

```
# hares -link pdb_resname cdb_resname
```

Repeat this step for each pluggable database resource in a container database.

**13** Repeat steps 8 to 12 for each container database.

**14** Set the dependencies between the CFSMount resource and the CVMVolDg resource for the Oracle service group:

```
# hares -link oradbmnt_resname
      oradbdg_resname
```

**15** Set the dependencies between the Oracle resource and the CFSMount resource for the Oracle service group:

```
# hares -link db_resname
      oradbmnt_resname
```

**16** Create an online local firm dependency between the oradb1_grp service group and the cvm service group:

```
# hagrp -link oradb_grpname
      cvm_grpname online local firm
```

**17** Enable the Oracle service group:

```
# hagrp -enableresources oradb_grpname
```

**18** Change the cluster configuration to the read-only mode:

```
# haconf -dump -makero
```

**19** Bring the Oracle service group online on all the nodes:

```
# hagrp -online oradb_grpname -any
```

---

**Note:** For policy-managed databases: When VCS starts or when the administrator attempts to bring the Oracle resource online, if the server is not part of the server pool associated with the database, the resource will remain offline. If Oracle Grid Infrastructure decides to move the server from the server pool, the database will be brought offline by the Oracle Grid Infrastructure and the oracle resource moves to offline state.

---

For more information and instructions on configuring the service groups using the CLI:

See the *Cluster Server Administrator's Guide*.

# Managing database restart after failure

When a database instance faults, it is observed that both Oracle Clusterware and VCS (through the Oracle agent) respond to the resource fault. When Oracle Clusterware detects an instance failure, it attempts to restart the instance based on the value set for the `RESTART_ATTEMPTS` attribute in the database resource profile of Oracle Clusterware. At the same time, VCS detects the instance failure and faults the resource and takes action depending on the VCS configuration.

To avoid both clusterwares from independently responding to the fault, you need to modify the appropriate resource parameters as described in the following sections.

**To manage database restart after failure**

**1** Log into one of the nodes in the cluster as the root user.

**2** Perform one of the following steps:

- Disable Oracle Clusterware/Grid Infrastructure from restarting the instance by modifying the `RESTART_ATTEMPTS` attribute as follows:

```
# crsctl modify resource db_resname -attr "RESTART_ATTEMPTS=0"
```

For Oracle RAC 12.1.0.2:

```
# crsctl modify resource db_resname \
-attr "RESTART_ATTEMPTS=0" -unsupported
```

- Prevent VCS from faulting the resource until Oracle Clusterware/Grid Infrastructure exhausts its restart attempts.

  ```
  # haconf -makerw
  # hares -modify db_resname ToleranceLimit = tl_value
  # haconf -dump -makero
  ```

  where *tl_value* is a value greater than or equal to the value of the `RESTART_ATTEMPTS` attribute.

## Location of VCS log files

You may want to review the log files at `/var/VRTSvcs/log/engine_A.log` for errors or status messages. When large amounts of data are written, multiple log files may be written, such as engine_B.log, engine_C.log, and so on. The engine_A.log contains the most recent data.

# Preventing automatic database startup

Configure the Oracle RAC database for manual startup if you want the Oracle RAC database to be managed by VCS using the Oracle agent. If you configure the VCS service groups for Oracle, you need to prevent the Oracle database from starting automatically. The Oracle Clusterware/Grid Infrastructure and Oracle agent may attempt to start the database instance at the same time if the database mount is available. To prevent the Oracle database from starting automatically, you must change the management policy for the database from automatic to manual using the Oracle `SRVCTL` command. The command changes the AUTO_START attribute of the Oracle database and instance resources.

**To prevent automatic database startup**

1. Register the database, if not already registered.

   ```
   $ srvctl add database -d db_name -o db_home \
   -p location_parameterfile -y manual
   ```

2. Once the database is registered, change the management policy for the database to manual.

   ```
   $ srvctl modify database -db db_name -policy manual
   ```

# Removing existing PrivNIC or MultiPrivNIC resources

If you upgrade to any of the supported Oracle RAC versions, you need to remove existing PrivNIC or MultiPrivNIC resources that you may have configured.

**To remove existing PrivNIC or MultiPrivNIC resources**

1. Freeze the cvm service group and delete PrivNIC/MultiPrivNIC. If Oracle database is under VCS control, freeze the database service group.

```
# haconf -makerw
# hagrp -freeze cvm
# hagrp -freeze oracle_group
# hares -delete PrivNIC_MultiPrivNIC_res
# haconf -dump -makero
```

2. Unfreeze the cvm service group. If Oracle database is under VCS control, unfreeze the database service group.

```
# haconf -makerw
# hagrp -unfreeze cvm
# hagrp -unfreeze oracle_group
# haconf -dump -makero
```

# Removing permissions for communication

Make sure you completed the installation of Veritas InfoScale Enterprise and the verification of disk support for I/O fencing. If you used `rsh`, remove the temporary `rsh` access permissions that you set for the nodes and restore the connections to the public network.

If the nodes use `ssh` for secure communications, and you temporarily removed the connections to the public network, restore the connections.

Chapter 17

# Upgrading Oracle RAC

This chapter includes the following topics:


- Preparing to upgrade Oracle RAC
- Upgrading Oracle RAC binaries
- Migrating the Oracle RAC database


## Preparing to upgrade Oracle RAC

Perform the following tasks before upgrading Oracle RAC.

For supported Oracle RAC upgrade paths, refer to the Oracle documentation.

The upgrade procedure assumes that the beginning configuration includes the following components, and that these components are running on the cluster nodes:

- The SF RAC component of a supported InfoScale Enterprise version
- A supported version of the operating system

---

**Note:** When you upgrade to a different version of Oracle RAC, make sure that the full path of the Oracle Clusterware/Grid Infrastructure home directory and the Oracle database home directory is different from the path where the existing version of Oracle RAC resides.

---

**To prepare for upgrade Oracle RAC**

1. Take a hot or cold backup of the existing database.
2. Back up the existing Oracle home and central inventory.
3. Verify that the cluster configuration is compatible for upgrading Oracle RAC.

4 If the Oracle RAC database is under VCS control, freeze the Oracle service groups to prevent VCS from reporting the resource as faulted when Oracle RAC stops and starts the database during the upgrade:

```
# haconf -makerw

# hagrp -freeze oradb_grpname -persistent
```

5 Freeze the cvm service group to prevent VCS from reporting the resource as faulted when Oracle Clusterware is stopped and started during the upgrade:

```
# hagrp -freeze cvm_grpname -persistent

# haconf -dump -makero
```

# Upgrading Oracle RAC binaries

Review your Oracle installation manuals and the appropriate Oracle support Web sites before upgrading Oracle RAC.

**To upgrade Oracle RAC binaries**

1 Upgrade the Oracle grid infrastructure.

For instructions, see the Oracle RAC documentation.

2 Make sure that the Oracle grid infrastructure is running.

3 Install the Oracle RAC database binaries.

For instructions, see the Oracle RAC documentation.

4 Complete the following post-installation tasks:

- Add Oracle RAC patches or patchsets.
  See "Adding Oracle RAC patches or patchsets" on page 246.
- Relink the SF Oracle RAC libraries with Oracle RAC.

---

**Note:** If you have multiple ORACLE_HOME directories, ensure that you relink the Oracle libraries in all ORACLE_HOME directories.

---

See "Relinking the SF Oracle RAC libraries with Oracle RAC" on page 252.

# Migrating the Oracle RAC database

For instructions on migrating the existing Oracle RAC database, see the Oracle metalink documentation.

After migrating the database, complete the post-upgrade tasks:

See "Performing post-upgrade tasks" on page 276.

## Performing post-upgrade tasks

Perform the steps in the following procedure to complete the upgrade.

**To perform post-upgrade tasks**

1. Change the cluster configuration to read-write mode:

   ```
   # haconf -makerw
   ```

2. Add the Oracle UDP IPC private IP addresses to the Oracle `init.ora` file if the database cache fusion traffic is configured to use Oracle UDP IPC private IP addresses and if these addresses are configured as a PrivNIC or MultiPrivNIC resource for high availability.

3. Modify the Oracle RAC configuration to prevent automatic startup of Oracle Clusterware.

   See "Preventing automatic startup of Oracle Clusterware/Grid Infrastructure" on page 251.

4. Modify the Oracle RAC database configuration to prevent automatic database startup if you want the Oracle RAC database to be managed by VCS using the Oracle agent.

   See "Preventing automatic database startup" on page 272.

5. Unfreeze the VCS service groups that were frozen earlier.

   As root user, enter:

   ```
   # hagrp -unfreeze oradb_grpname -persistent

   # hagrp -unfreeze cvm_grpname -persistent

   # haconf -dump -makero
   ```

Section 5

# Adding and removing nodes


- Chapter 18. Adding a node to SF Oracle RAC clusters
- Chapter 19. Removing a node from SF Oracle RAC clusters

Chapter 18

# Adding a node to SF Oracle RAC 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 to the cluster using the SF Oracle RAC response file
- Configuring private IP addresses for Oracle RAC on the new node
- Adding the new node to Oracle RAC
- 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 SF Oracle RAC and create a cluster, you can add and remove nodes from the cluster.You can create clusters of up to 16 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 SF Oracle RAC cluster.

**Table 18-1** 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 279. |
| Add a new node to the cluster. | See "Adding a node to a cluster using the Veritas InfoScale installer" on page 282.<br>See "Adding the node to a cluster manually" on page 293. |
| 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 302. |
| Prepare the new node for installing Oracle. | See "Preparing the new nodes for installing Oracle RAC using the installer" on page 286. |
| Add the node to Oracle. | See "Adding the new node to Oracle RAC" on page 316. |
| 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 317.<br>See "Updating the Storage Foundation for Databases (SFDB) repository after adding a node" on page 318. |

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 SF Oracle RAC 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 SF Oracle RAC.

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 SF Oracle RAC 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:

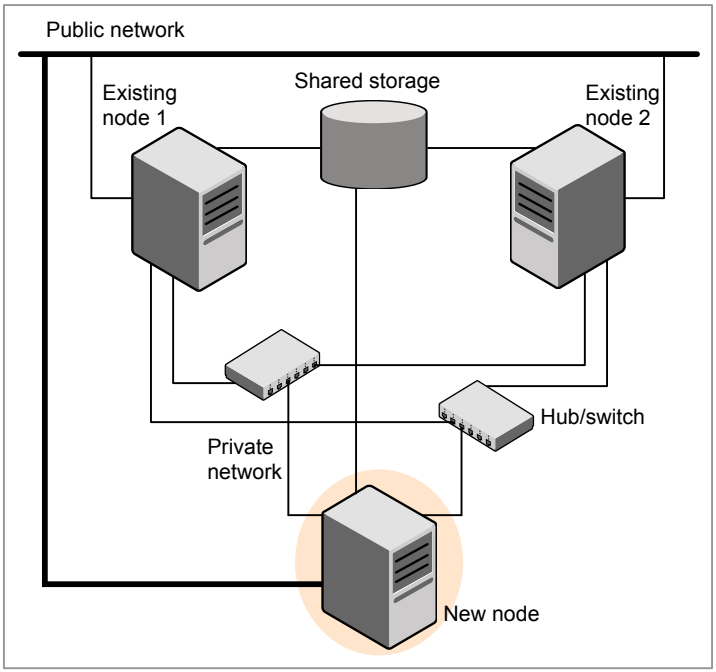
   ```
   # vxdctl protocolversion
   Cluster running at protocol 260
   ```

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 18-1.

**Figure 18-1** Adding a node to a two-node cluster using two switches



**To set up the hardware**

1. Connect the SF Oracle RAC 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.
   - If you already use independent hubs, connect the two Ethernet controllers on the new node to the independent hubs.

   Figure 18-1 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 SF Oracle RAC 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 SF Oracle RAC when prompted.

   ```
   Would you like to configure InfoScale Enterprise after installation?
    [y,n,q] (n) n
   ```

3 Restart the node if the installer asks for a reboot.

# 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/vcsmmtab`

  `/etc/vx/.uuids/clusuuid`

  `/etc/sysconfig/llt`

  `/etc/sysconfig/gab`

  `/etc/sysconfig/vxfen`

  `/etc/sysconfig/vcsmm`

  `/etc/sysconfig/amf`
- 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 SF Oracle RAC processes and configures CVM and CFS on the new node.

At the end of the process, the new node joins the SF Oracle RAC cluster.

## Prerequisite

Enable the required LLT ports on the firewall.

See "Enabling LLT ports in firewall" on page 534.

---

**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 329.

---

**Caution:** If you plan to use the Veritas InfoScale installer for completing the Oracle pre-installation tasks on the new node, do not quit the installer after adding the node to the cluster. If you quit the installer, you must perform the Oracle pre-installation tasks manually.

---

**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. If the `cssd` type resource is configured as a critical resource, the cvm group will be brought offline on the new node. Modify the configuration to make the `cssd` type resource non-critical.

   On one of the nodes in the existing cluster, configure the `cssd` type resource as a non-critical resource:

   ```
   # haconf -makerw
   # hares -modify cssd Critical 0
   # haconf -dump -makero
   ```

3 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.

4 Enter the name of a node in the existing SF Oracle RAC 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
```

5 Review and confirm the cluster information.

6 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.

7 Enter the name of the network interface that you want to configure as the first private heartbeat link.

---

**Note:** 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. The LLT configuration for the new node must be the same as that of the existing cluster.

---

```
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.

---

8 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.

9 Review and confirm the information.

10 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
```

11 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.

12 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.

---

**Note:** Do not quit the installer if you want to perform the Oracle pre-installation tasks using the installer.

---

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
```

13 Complete the preparatory steps for installing Oracle RAC on the new nodes.

See "Preparing the new nodes for installing Oracle RAC using the installer" on page 286.

**14** Add the node to Oracle RAC.

See "Adding the new node to Oracle RAC" on page 316.

**15** Confirm that the new node has joined the SF Oracle RAC cluster using `lltstat -n` and `gabconfig -a` commands.

## Preparing the new nodes for installing Oracle RAC using the installer

The installer performs the following tasks:

- Creates the Oracle user and groups on the new node
- Configures the private IP addresses or the PrivNIC or MultiPrivNIC resources (if they are configured in the existing cluster).
- If the CFSMount and CVMVolDg resources for OCR and voting disk are configured under the `cvm` service group, the installer brings them online after adding the node to the cluster.
  If the resources are configured in any other service group, make sure that you modify the service group to include the new node and bring the service group online.

  **Note:** If OCR and voting disk are not configured under VCS, manually mount the OCR and voting disk after you finish the steps in the following procedure.

- Starts the CVM group on the new node.

**To prepare to install Oracle RAC on the new node using the installer**

1. After you configure SF Oracle RAC on the new node, the installer displays the following options for configuring Oracle RAC:

```
1)  Create Oracle Users and Groups
2)  Configure private IP addresses (HAIP Configuration)
- For Oracle RAC 11.2.0.2 and later
3)  Finish
```

---

**Note:** Options 3 and 4 are displayed depending on whether or not you configured the private IP addresses as PrivNIC or MultiPrivNIC resources in the existing cluster.

---

Enter **1** to select the option **Create Oracle User and Group** from the SF Oracle RAC installer menu.

See "Creating Oracle user and groups on the new node" on page 289.

2. Configure private IP addresses for HAIP, if configured in the existing cluster.

Enter **2** to select the option **Configure private IP addresses (HAIP Configuration)**.

See "Configuring private IP addresses for Oracle RAC on the new node" on page 315.

3. Configure the PrivNIC resource for Oracle Clusterware (only if the IP addresses on the existing cluster are configured as PrivNIC resources).

You must manually update the PrivNIC resource configuration in the following cases:

- If the PrivNIC resource on the existing cluster is configured under a group other than `cvm`.
- If the `Device` attribute in the PrivNIC resource configuration on the existing cluster is not configured for each node as follows:

```
Device@sys1 = {eth1=0, eth2=1}
Device@sys2 = {eth1=0, eth2=1}
```

Enter **3** to select the option **Configure private IP addresses (PrivNIC configuration)**.

See "Configuring the PrivNIC resource for Oracle Clusterware" on page 289.

4. Configure the MultiPrivNIC resource for Oracle Clusterware and Oracle UDP IPC (only if the IP addresses on the existing cluster are configured as MultiPrivNIC resources).

You must manually update the MultiPrivNIC resource configuration in the following cases:

- If the MultiPrivNIC resource on the existing cluster is configured under a group other than `cvm`.
- If the `Device` attribute in the MultiPrivNIC resource configuration on the existing cluster is not configured for each node as follows:

  ```
  Device@sys1 = {eth1=0, eth2=1, eth3=2}
  Device@sys2 = {eth1=0, eth2=1, eth3=2}
  ```

Enter **4** to select the option **Configure private IP addresses (MultiPrivNIC configuration)**.

See "Configuring the MultiPrivNIC resource for Oracle Clusterware and Oracle UDP IPC" on page 291.

**5** Select **Finish** to start the `cvm` group on the new node.

---

**Note:** The `cssd` type resource appears FAULTED until the new node is added to Oracle Clusterware.

---

**6** Verify that all the GAB ports are up:

```
# gabconfig -a

GAB Port Memberships
===============================================================
Port a gen   1b0802 membership 0123
Port b gen   1b0820 membership 0123
Port d gen   1b0809 membership 0123
Port f gen   1b082f membership 0123
Port h gen   1b0831 membership 0123
Port m gen   1b0827 membership 0123
Port o gen   1b080b membership 0123
Port u gen   1b082d membership 0123
Port v gen   1b0829 membership 0123
Port w gen   1b082b membership 0123
Port y gen   1b0828 membership 0123
```

**7** Complete the following additional preparatory tasks using the instructions in the chapter "Preparing to install Oracle RAC":

- Identify public virtual IP addresses for use by Oracle RAC.
- Set the kernel parameters.

- Verify that the user "nobody" exists.
- Set up Oracle user equivalence for remote shell and remote copy environments.
- Edit the Oracle user profile.
- If the OCR and voting disk resources are not configured under VCS, mount the OCR and voting disk manually.

**8** Create Oracle Clusterware and Oracle RAC database home directories manually.

See "Creating Oracle Clusterware/Grid Infrastructure and Oracle database home directories on the new node " on page 308.

## Creating Oracle user and groups on the new node

Perform the steps in the following procedure to create Oracle user and groups on the new node.

---

**Note:** Set the password of the Oracle user manually before you configure secure shell or remote shell connection on the node.

---

**To create Oracle user and groups on the new node**

**1** Enter the Oracle user name that is used for Oracle RAC operations on the existing cluster.

---

**Note:** If the Oracle user and groups already exist on the new node, make sure that the UID and GID of the Oracle user and groups are the same as that on the current cluster.

---

```
Enter Oracle UNIX user name:  [b] oracle
```

The installer obtains the existing group and identifier information based on the Oracle user name.

**2** Review and confirm the information.

The installer adds the user and groups to the new node.

**3** Press **Return** to continue with the other configuration tasks.

## Configuring the PrivNIC resource for Oracle Clusterware

Perform this step only if the private IP addresses are configured as PrivNIC resources in the existing cluster.

---

**Note:** Make sure that the network interface names of the private interconnects on the new node are the same as those of the existing cluster. For maximum failover options, all available LLT links are used for PrivNIC configuration.

---

Review the pre-configuration information displayed by the installer and ensure that you meet the requirements.

**To configure the PrivNIC resource for Oracle Clusterware**

1. Enter the name for the PrivNIC resource. The installer displays the names of the existing PrivNIC resources. Specify the name of an existing resource.

```
Enter the PrivNIC resource name: [b] (ora_priv)
```

2. Enter **y** if you want the installer to add the IP address to the `/etc/hosts` file.

   You can also add the IP address to the file manually after the configuration.

   The installer displays the existing PrivNIC resource configuration on the cluster.

```
Resource name: ora_priv
System: sys1
        Private Interfaces: eth1 eth2
        Private IP address: 192.168.12.1
        Alias for above IP: sys1-priv
System: sys2
        Private Interfaces: eth1 eth2
        Private IP address: 192.168.12.2
        Alias for above IP: sys2-priv
Is this information correct? [y,n,q] (y)
```

3. Review and confirm the information.

4. Enter the private IP address and its private node name for the new node.

```
Enter the private IP for sys5: [b] 192.168.12.5
Enter Hostname alias for the above IP address: [b] sys5-priv
```

   The installer displays the resource configuration for the new node.

```
Resource name: ora_priv
System: sys5
        Private Interfaces: eth1 eth2
        Private IP address: 192.168.12.5
        Alias for above IP: sys5-priv
Is this information correct? [y,n,q] (y)
```

5 Review and confirm the information.

The installer updates the existing PrivNIC resource with the resource configuration for the new node and updates the `/etc/hosts` file on the new node as well as on the existing nodes (if you chose to update the file through the installer).

6 Press **Return** to continue with the other configuration tasks.

## Configuring the MultiPrivNIC resource for Oracle Clusterware and Oracle UDP IPC

Perform this step only if the private IP addresses are configured as MultiPrivNIC resources in the existing cluster.

---

**Note:** Make sure that you configure the same interfaces (as those on the existing cluster) for private interconnects on the new node. For maximum failover options, all available LLT links are used for MultiPrivNIC configuration.

---

Review the pre-configuration information displayed by the installer and ensure that you meet the requirements.

### To configure the MultiPrivNIC resource for Oracle Clusterware and Oracle UDP IPC

1. Enter the name for the MultiPrivNIC resource. The installer displays the names of the existing MultiPrivNIC resources. Specify the name of an existing resource.

```
Enter the MultiPrivNIC resource name: [b] (multi_priv)
```

2. Enter **y** if you want the installer to add the IP address to the `/etc/hosts` file.

   You can also add the IP address to the file manually after the configuration.

   The installer displays the existing MultiPrivNIC resource configuration on the cluster.

```
Resource name: multi_priv
System: sys1
        Private Interfaces: eth1 eth2
        Private IPs on eth1: 192.168.12.1
        Aliases for above IPs: sys1-priv
        Private IPs on eth2: 192.168.2.1
        Aliases for above IPs: sys1-priv1
System: sys2
        Private Interfaces: eth1 eth2
        Private IPs on eth1: 192.168.12.2
        Aliases for above IPs: sys2-priv
        Private IPs on eth2: 192.168.2.2
        Aliases for above IPs: sys2-priv1
Is this information correct? [y,n,q] (y)
```

3. Review and confirm the information.

4. Enter the private IP address and the corresponding private node name for the eth1 interface on the new node.

```
Enter IP addresses for sys5 for eth1
separated by space: [b,q,?]  192.168.12.5
Enter Hostname aliases for the above IP addresses
separated by space: [b,q,?] sys5-priv
```

5 Enter the private IP address and the corresponding private node name for the eth2 interface on the new node.

```
Enter IP addresses for sys5 for eth2
separated by space: [b,q,?]  192.168.2.6
Enter Hostname aliases for the above IP addresses
separated by space: [b,q,?] sys5-priv1
```

The installer displays the resource configuration for the new node.

```
Resource name: multi_priv
System: sys5
        Private Interfaces: eth1 eth2
        Private IPs on eth1: 192.168.12.5
        Aliases for above IPs: sys5-priv
        Private IPs on eth2: 192.168.2.6
        Aliases for above IPs: sys5-priv1
Is this information correct? [y,n,q] (y)
```

6 Review and confirm the information.

The installer updates the existing MultiPrivNIC resource with the resource configuration for the new node and updates the `/etc/hosts` file on the new node as well as on the existing nodes (if you chose to update the file through the installer).

7 Press **Return** to continue with the other configuration tasks.

# 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 18-2** 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 294. |
| Configure the cluster processes on the new node. | See "Configuring cluster processes on the new node" on page 295. |

**Table 18-2** 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.<br><br>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 300. |
| Start VCS. | See "To start VCS on the new node" on page 301. |
| Configure CVM and CFS. | See "Configuring Cluster Volume Manager (CVM) and Cluster File System (CFS) on the new node" on page 302. |
| 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 304. |
| Complete the preparatory steps for installing Oracle RAC on the new nodes. | See "Preparing the new node manually for installing Oracle RAC" on page 305. |
| Add the node to Oracle RAC. | See "Adding the new node to Oracle RAC" on page 316. |

## 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.

**To configure cluster processes on the new node 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 Oracle; refer to the Oracle 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 Copy the following files from one of the nodes in the existing cluster to the new node:

```
/etc/sysconfig/llt
/etc/sysconfig/gab
/etc/sysconfig/vcs
/etc/sysconfig/vcsmm
```

8 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
```

9 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
```

10 Start the LLT, GAB, VCSMM, and ODM drivers on the new node:

For supported Linux distributions:

```
# systemctl start llt
# systemctl start gab
# systemctl start vcsmm
# systemctl start vxgms
# systemctl start vxglm
# systemctl restart vxodm
```

11 On the new node, verify that the GAB port memberships are a, d, and o:

```
# gabconfig -a


GAB Port Memberships
===============================================================
Port a gen    df204 membership 012
Port d gen    df20a membership 012
Port o gen    df207 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 18-3 uses the following information for the following command examples.

**Table 18-3** 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 301.

2 Start fencing on the new node:

3 On the new node, verify that the GAB port memberships are a, b, d, and o:

```
# gabconfig -a

GAB Port Memberships
===================================================================
Port a gen    df204 membership 012
Port b gen    df20d membership 012
Port d gen    df20a membership 012
Port o gen    df207 membership 012
```

## After adding the new node

If you added the new node to the cluster manually, before you starting VCS you must create the file `cssd-pretend-offline` on the new node and make the `cssd`

resource non-critical. Failing this, the `cssd` resource lapses into an UNKNOWN state until Oracle Clusterware is installed on the new node, thus preventing the `cvm` group from coming online.

If you used the product installer to add the new node to the cluster, you will not need to perform this step.

---

**Note:** The `cssd` resource will remain in FAULTED/OFFLINE state till Oracle Clusterware is installed on the new node.

---

Start VCS on the new node.

**To start VCS on the new node**

1. If you installed the new node manually:
   - On one of the nodes in the existing cluster, configure the `cssd` resource as a non-critical resource:

     ```
     # haconf -makerw
     # hares -modify cssd Critical 0
     # haconf -dump -makero
     ```

   - Create the file `cssd-pretend-offline` on the new node:

     ```
     # touch /var/VRTSvcs/lock/cssd-pretend-offline
     ```

2. Start VCS on the new node:

   ```
   # hastart
   ```

   VCS brings the CVM group online.

3. Verify that the CVM group is online:

   ---

   **Note:** The CVM group reports `Partial` state till Oracle Clusterware/Grid Infrastructure is installed on the new node.

   ---

   ```
   # hagrp -state
   ```

## 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.

## 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 SF Oracle RAC 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 SF Oracle RAC cluster:

```
# haconf -dump -makero
```

# 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
```

## Preparing the new node manually for installing Oracle RAC

Complete the following preparatory tasks manually before you install Oracle RAC on the new node.

**To prepare to install Oracle RAC on the new node**

1 Create Oracle user and groups.

2 Configure IP addresses for Oracle HAIP.

See "Configuring private IP addresses for Oracle RAC on the new node" on page 315.

3 Start VCS on the new node.

See "Starting VCS on the new node" on page 307.

4 Create the Oracle Clusterware/Grid Infrastructure and Oracle RAC database home directories for installation.

See "Creating Oracle Clusterware/Grid Infrastructure and Oracle database home directories on the new node " on page 308.

5 Complete the following additional preparatory tasks using the instructions in the chapter "Preparing to install Oracle RAC":

- Identify public virtual IP addresses for use by Oracle RAC.
- Set the kernel parameters.
- Verify that the user "nobody" exists.
- Set up Oracle user equivalence for remote shell and remote copy environments.
- Edit the Oracle user profile.

## Configuring the PrivNIC resource for Oracle Clusterware

This section provides instructions for configuring private IP address and PrivNIC resource for Oracle Clusterware.

Identify a private IP address that you want to use for the new node. Make sure that the IP address is in the same subnet as the existing cluster.

The procedure uses the following IP address for the new node:

On sys5: 192.168.12.5

**To configure the PrivNIC resource for Oracle Clusterware**

1. Make a backup copy of the main.cf file. For example:

```
# cd /etc/VRTSvcs/conf/config

# cp main.cf main.cf.2node
```

2. Add the private IP address to the the `ora_priv` resource on the active node:

```
# haconf -makerw

# hares -modify priv_resname Device -add nic1_node1 0 \
-sys nodenew_name

# hares -modify priv_resname Device -add nic2_node1 1 \
-sys nodenew_name

# hares -modify priv_resname Address "privnic_ip_newnode" \
-sys nodenew_name

# haconf -dump -makero
```

## Configuring the MultiPrivNIC resource for Oracle Clusterware and UDP IPC

This section provides instructions for configuring private IP addresses and MultiPrivNIC resource for Oracle Clusterware and UDP IPC.

Identify the private IP addresses that you want to use for the new node. Make sure that the IP addresses are in the same subnet as the existing cluster.

The procedure uses the following IP addresses for the new node:

| | |
|---|---|
| IP addresses for the new node | 192.168.12.5 |
| | 192.168.2.6 |

**To configure the MultiPrivNIC resource for Oracle Clusterware**

- Use the following commands to add private IP into multi_priv resource on the active node:

```
# haconf -makerw

# hares -modify multipriv_resname Device -add eth1 0 \
-sys nodenew_name

# hares -modify multipriv_resname Device -add eth2 1 \
-sys nodenew_name

# hares -modify multipriv_resname Address -add \
multipriv_ip_newnode 0 -sys nodenew_name

# hares -modify multipriv_resname Address -add \
multipriv_udpip1_newnode 1 -sys nodenew_name

# haconf -dump -makero
```

## Starting VCS on the new node

Before you start VCS, create the file `cssd-pretend-offline` on the new node and make the `cssd` resource non-critical. Failing this, the `cssd` resource lapses into an UNKNOWN state until Oracle Clusterware is installed on the new node, thus preventing the `cvm` group from coming online.

---

**Note:** The `cssd` resource will remain in FAULTED/OFFLINE state till Oracle Clusterware is installed on the new node.

---

**To start VCS on the new node**

1 On one of the nodes in the existing cluster, configure the `cssd` resource as a non-critical resource:

```
# haconf -makerw
# hares -modify cssd Critical 0
# haconf -dump -makero
```

2 Create the file `cssd-pretend-offline` on the new node:

```
# touch /var/VRTSvcs/lock/cssd-pretend-offline
```

3 Start VCS on the new node:

```
# hastart
```

## Creating Oracle Clusterware/Grid Infrastructure and Oracle database home directories on the new node

The Oracle Clusterware/Grid Infrastructure and Oracle database home directories must be located on the same storage as that on the existing nodes.

Depending on the storage in the existing cluster, use one of the following options to create the directories:

| | |
|---|---|
| Local file system | See "To create the directories on the local file system" on page 308. |
| Cluster File System | See "To create the file system and directories on cluster file system for Oracle Clusterware and Oracle database" on page 310. |

**To create the directories on the local file system**

1 Log in as the root user on the node.

2 Create a local file system and mount it using one of the following methods:

- Using native operating system commands
  For instructions, see the operating system documentation.
- Using Veritas File System (VxFS) commands

As the root user, create a VxVM local diskgroup on each node.

```
# vxdg init vxvm_dg \
dg_name
```

| | |
|---|---|
| Create separate volumes for Oracle Clusterware/Oracle Grid Infrastructure binaries and Oracle binaries. | `# vxassist -g vxvm_dg make clus_volname size`<br>`# vxassist -g vxvm_dg make ora_volname size` |
| Create the file systems with the volumes. | `# mkfs -t vxfs /dev/vx/rdsk/vxvm_dg/clus_volname`<br>`# mkfs -t vxfs /dev/vx/rdsk/vxvm_dg/ora_volname` |
| Mount the file system. | `# mount -t vxfs /dev/vx/dsk/vxvm_dg/clus_volname \ clus_home`<br>`# mount -t vxfs /dev/vx/dsk/vxvm_dg/ora_volname \ oracle_home` |

**3** Create the directories for Oracle RAC.

```
# mkdir -p grid_base
# mkdir -p clus_home
# mkdir -p oracle_base
# mkdir -p oracle_home
```

**4** Set appropriate ownership and permissions for the directories.

```
# chown -R grid:oinstall grid_base
# chmod -R 775 grid_base
# chown -R grid:oinstall clus_home
# chmod -R 775 clus_home

# chown -R oracle:oinstall oracle_base
# chmod -R 775 oracle_base
# chown -R oracle:oinstall oracle_home
# chmod -R 775 oracle_home
```

**5** Add the resources to the VCS configuration.

See "To add the storage resources created on VxFS to the VCS configuration" on page 310.

**6** Repeat all the steps on each new node of the cluster.

**To add the storage resources created on VxFS to the VCS configuration**

1. Change the permissions on the VCS configuration file:

```
# haconf -makerw
```

2. Configure the VxVM volumes under VCS:

```
# hares -add dg_resname DiskGroup cvm
# hares -modify dg_resname DiskGroup vxvm_dg -sys nodenew_name
# hares -modify dg_resname Enabled 1
```

3. Set up the file system under VCS:

```
# hares -add clusbin_mnt_resname Mount cvm

# hares -modify clusbin_mnt_resname MountPoint \
"clus_home"

# hares -modify clusbin_mnt_resname BlockDevice \
"/dev/vx/dsk/vxvm_dg/clus_volname" -sys nodenew_name
# hares -modify clusbin_mnt_resname FSType vxfs
# hares -modify clusbin_mnt_resname FsckOpt "-n"
# hares -modify clusbin_mnt_resname Enabled 1
# hares -add orabin_mnt_resname Mount cvm

# hares -modify orabin_mnt_resname MountPoint \
"oracle_home"

# hares -modify orabin_mnt_resname BlockDevice \
"/dev/vx/dsk/vxvm_dg/ora_volname" -sys nodenew_name
# hares -modify orabin_mnt_resname FSType vxfs
# hares -modify orabin_mnt_resname FsckOpt "-n"
# hares -modify orabin_mnt_resname Enabled 1
```

4. Link the parent and child resources:

```
# hares -link clusbin_mnt_resname vxvm_dg
# hares -link orabin_mnt_resname vxvm_dg
```

5. Repeat all the steps on each new node of the cluster.

**To create the file system and directories on cluster file system for Oracle Clusterware and Oracle database**

Perform the following steps on the CVM master node in the cluster.

1. Create the Oracle base directory, clusterware home directory, and the Oracle home directory.

```
# mkdir -p oracle_base
# mkdir -p oracle_home
# mkdir -p clus_home
# mkdir -p grid_base
```

2. Mount the file systems. Perform this step on each new node.

```
# mount -t vxfs -o cluster /dev/vx/dsk/cvm_dg/ora_volname \
oracle_home
```

# Adding a node to the cluster using the SF Oracle RAC response file

Perform the steps in the following procedure to add new nodes to an SF Oracle RAC cluster using the SF Oracle RAC response file.

### To add a node to the cluster using a response file

1. If the cssd resource is configured as a critical resource, the cvm group will be brought offline on the new node. Modify the configuration to make the cssd resource non-critical.

   On one of the nodes in the existing cluster, configure the cssd resource as a non-critical resource:

   ```
   # haconf -makerw
   # hares -modify cssd Critical 0
   # haconf -dump -makero
   ```

2. Create a response file using one of the available options.

   **Note:** Make sure that you replace the host names in the response file appropriately.

   For information on various options available for creating a response file:

   For response file variable definitions:

   See "Response file variables to add a node to an SF Oracle RAC cluster" on page 313.

   For a sample response file:

   See "Sample response file for adding a node to an SF Oracle RAC cluster" on page 314.

3. Navigate to the product directory on the installation media that contains the SF Oracle RAC installation program.

4. Start the installer:

   ```
   # ./installer -responsefile /tmp/response_file
   ```

   Where `/tmp/response_file` is the full path name of the response file.

   Depending on the fencing configuration in the existing cluster, the installer configures fencing on the new node.

   The installer then starts all the required SF Oracle RAC 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.

5 Verify that all the GAB ports are up:

```
# gabconfig -a
GAB Port Memberships
===============================================================
Port a gen   1b0802 membership 0123
Port b gen   1b0820 membership 0123
Port d gen   1b0809 membership 0123
Port f gen   1b082f membership 0123
Port h gen   1b0831 membership 0123
Port m gen   1b0827 membership 0123
Port o gen   1b080b membership 0123
Port u gen   1b082d membership 0123
Port v gen   1b0829 membership 0123
Port w gen   1b082b membership 0123
Port y gen   1b0828 membership 0123
```

6 Complete the following additional preparatory tasks using the instructions in the chapter "Preparing to install Oracle RAC":

- Identify public virtual IP addresses for use by Oracle RAC.
- Set the kernel parameters.
- Verify that the user "nobody" exists.
- Set up Oracle user equivalence for remote shell and remote copy environments.
- If the OCR and voting disk resources are not configured under VCS, mount the OCR and voting disk manually.

7 Create Oracle Clusterware and Oracle RAC database home directories manually.

See "Creating Oracle Clusterware/Grid Infrastructure and Oracle database home directories on the new node " on page 308.

# Response file variables to add a node to an SF Oracle RAC cluster

Table 18-4 lists the response file variables that you can define to add a node to an SF Oracle RAC cluster.

**Table 18-4** Response file variables for adding a node to an SF Oracle RAC cluster

| Variable | Description |
|---|---|
| $CFG{addnode_create_oracle_user_group} | Defines whether or not to create oracle user and group after adding a node<br>List or scalar: scalar<br>Optional or required: optional |
| $CFG{oracle_user} | Defines the oracle user name to create after adding a node.<br>List or scalar: scalar<br>Optional or required: optional |
| $CFG{addnode_config_haip} | Defines whether or not to configure HAIP for configuring the private network for Oracle RAC after adding a node to the cluster.<br>List or scalar: scalar<br>Optional or required: optional |
| $CFG{opt}{addnode} | Adds a node to an existing cluster.<br>List or scalar: scalar<br>Optional or required: required |
| $CFG{newnodes} | Specifies the new nodes to be added to the cluster.<br>List or scalar: list<br>Optional or required: required |

# Sample response file for adding a node to an SF Oracle RAC cluster

The following example shows a response file for upgrading SF Oracle RAC.

```
our %CFG;

$CFG{clustersystems}=[ qw(sys1) ];
$CFG{newnodes}=[ qw(sys5) ];
$CFG{nic_add_ip_to_files}=1;
$CFG{opt}{addnode}=1;
$CFG{opt}{configure}=1;
```

```
$CFG{opt}{vr}=1;
$CFG{prod}="ENTERPRISE80";
$CFG{sys5}{eth1}{haipip}="192.168.12.3";
$CFG{sys5}{eth1}{hostname_for_haip}="sys5-haip1";
$CFG{sys5}{eth2}{haipip}="192.168.13.3";
$CFG{sys5}{eth2}{hostname_for_haip}="sys5-haip2";
$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 private IP addresses for Oracle RAC on the new node

**Note:** IPv6 addresses are not supported in this release.

The following tables lists the available options for configuring the private network for Oracle RAC on the new node. Use one of the following options to configure the private network.

**Table 18-5** Options for configuring the private network for Oracle RAC

| Option | Description |
|---|---|
| Link aggregation or NIC bonding | See "Using link aggregation or NIC bonding" on page 315. |
| Oracle High Availability IP (HAIP) | See "Using Oracle High Availability IP (HAIP)" on page 316. |

## Using link aggregation or NIC bonding

To configure link aggregation or NIC bonding for the new node, see the operating system documentation.

**Note:** Ensure that a link configured under an aggregated link or NIC bond is not configured as a separate LLT link.

When LLT is configured over a bonded interface, perform one of the following steps to prevent GAB from reporting jeopardy membership:

- Configure an additional network interface under LLT in addition to the bonded NIC.
- Add the following line in the `/etc/llttab` file:

  ```
  set-dbg-minlinks 2
  ```

  Restart LLT.

  ```
  # systemctl start llt
  ```

## Using Oracle High Availability IP (HAIP)

If you configured Oracle HAIP for IP address failover in the existing cluster, ensure that you configure the private network interface with a static IP address for the new node.

To configure a static IP address for the new node:

Edit the file `/etc/sysconfig/network-scripts/ifcfg-interface_name`

```
DEVICE=interface_name
ONBOOT=yes
BOOTPROTO=none
IPADDR=ip-address
NETMASK=netmask
```

# Adding the new node to Oracle RAC

Install Oracle Clusterware and Oracle RAC database on the node using the Oracle RAC add node procedure.

For instructions, see the Oracle RAC documentation.

After installing Oracle Clusterware and Oracle RAC database, perform the following post-installation tasks:

1. Delete the file `/var/VRTSvcs/lock/cssd-pretend-offline` on the new node.

   Clear the fault and probe the cssd resource on the new node to bring the cssd resource online.

   ```
   # hares -clear cssd
   # hares -probe cssd -sys sys5
   ```

2. If the cssd resource was configured as a non-critical resource, reconfigure it as a critical resource.

```
# haconf -makerw
# hares -modify cssd Critical 1
# haconf -dump -makero
```

3. If cssd is not configured under the cvm group, add the new node information to the service group containing the cssd resource.
4. Add new Oracle RAC database instances for the new node.

   For instructions, see the Oracle RAC documentation.
5. Update the Oracle RAC database service groups to include the new database instances in the VCS configuration file.
6. For other service groups that are configured under VCS, manually update the service group configuration for the new node.

# 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*

# 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 19

# Removing a node from SF Oracle RAC 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 SF Oracle RAC cluster. The following table provides a summary of the tasks required to remove a node to an existing SF Oracle RAC cluster.

**Table 19-1** Tasks for removing a node from a cluster

| Step | Description |
|---|---|
| Prepare to remove the node:<br>■ Back up the configuration file.<br>■ Check the status of the nodes and the service groups.<br>■ Take the service groups offline and removing the database instances. | See "Removing a node from a cluster" on page 325. |
| Remove the node from the cluster. | See "Removing a node from a cluster" on page 325. |
| Modify the cluster configuration on remaining nodes.<br>■ Edit the /etc/llthosts file.<br>■ Edit the /etc/gabtab file.<br>■ Modify the VCS configuration to remove the node.<br>■ Modify the CVM configuration to remove the node. | See "Modifying the VCS configuration files on existing nodes" on page 327.<br>See "Modifying the Cluster Volume Manager (CVM) configuration on the existing nodes to remove references to the deleted node" on page 329. |
| 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 329. |
| 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 330. |
| 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 331. |

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 the Oracle RAC service groups offline (if under VCS control) on the node you want to remove.

   ```
   # hagrp -offline oracle_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.

3. Remove the Oracle RAC database software from the node.

   For instructions, see the Oracle RAC documentation.

4. Remove Oracle Clusterware from the node.

   For instructions, see the Oracle RAC document.

**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 SF Oracle RAC on the node using the Veritas InfoScale Enterprise installer.

   ```
   # cd /opt/VRTS/install
   ```

   ```
   # ./installer -stop sys5
   ```

   The installer stops all SF Oracle RAC 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 327.

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 329.

# 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

**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 crsgrp 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 SF Oracle RAC 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 for Oracle RAC 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 `sys5` is as follows:

- The existing cluster `clus1` comprises three nodes `sys1`, `sys2`, and `sys5` and hosts a single database.
- The Oracle database is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- Only one private IP address is configured for Oracle Clusterware. The private IP address is managed by the PrivNIC agent for high availability.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.

---

**Note:** The following sample file shows in **bold** the configuration information that is removed when the node "sys5" is removed from the cluster.

---

```
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"
include "OracleTypes.cf"
include "PrivNIC.cf"


cluster clus1 (
        UserNames = { admin = bopHo }
        Administrators = { admin }
        UseFence = SCSI3
        )
```

```
system sys1 (
        )
system sys2 (
        )
system sys5 (
       )
```

**Note:** In the following group oradb1_grp, the sys5 node must be removed.

```
group oradb1_grp (
        SystemList = { sys1 = 0, sys2 = 1, sys5 = 2 }
        AutoFailOver = 0
        Parallel = 1
        AutoStartList = { sys1, sys2, sys5 }
        )
```

**Note:** In the following Oracle resource, the sys5 node information must be removed.

```
Oracle ora1 (
          Critical = 0
          Sid @sys1 = vrts1
          Sid @sys2 = vrts2
          Sid @sys5 = vrts3
          Owner = oracle
          Home = "/app/oracle/orahome"
          StartUpOpt = "SRVCTLSTART"
          ShutDownOpt = "SRVCTLSTOP"
          )

CFSMount oradata_mnt (
         Critical = 0
         MountPoint = "/oradata"
         BlockDevice = "/dev/vx/dsk/oradatadg/oradatavol"
         )

CVMVolDg oradata_voldg (
         Critical = 0
         CVMDiskGroup = oradatadg
         CVMVolume = { oradatavol }
         CVMActivation = sw
         )
```

```
requires group cvm online local firm
ora1 requires oradata_mnt
oradata_mnt requires oradata_voldg
```

**Note:** In the following CVM and CVMCluster resources, the sys5 node information must be removed.

```
group cvm (
        SystemList = { sys1 = 0, sys2 = 1, sys5 =2}
        AutoFailOver = 0
        Parallel = 1
        AutoStartList = { sys1, sys2, sys5 }
        )

        CSSD cssd (
                Critical = 0
                CRSHOME = "/u01/app/grid/product/11.2.0/gridhome"
                OnlineWaitLimit = 5
                OfflineWaitLimit = 3
                )

       CFSMount ocrvote_mnt (
                Critical = 0
                MountPoint = "/ocrvote"
                BlockDevice = "/dev/vx/dsk/ocrvotedg/ocrvotevol"
                MountOpt= "mincache=direct"
                )

      CVMVolDg ocrvote_voldg (
               Critical = 0
               CVMDiskGroup = ocrvotedg
               CVMVolume = { ocrvotevol }
               CVMActivation = sw
               )


      CFSfsckd vxfsckd (
               )
```

```
CVMCluster cvm_clus (
        CVMClustName = clus1
        CVMNodeId = { sys1 = 0, sys2 = 1, sys5 =2 }
        CVMTransport = gab
        CVMTimeout = 200
        )


CVMVxconfigd cvm_vxconfigd (
         Critical = 0
         CVMVxconfigdArgs = { syslog }
         )
```

**Note:** In the following PrivNIC resource, the sys5 node information must be removed.

```
PrivNIC ora_priv (
        Critical = 0
        Device@sys1 = { eth1 = 0, eth2 = 1}
        Device@sys2 = { eth1 = 0, eth2 = 1}
        Device@sys5 = { eth1 = 0, eth2 = 1}
        Address@sys1 = "192.168.12.1"
        Address@sys2 = "192.168.12.2"
        Address@sys5 = "192.168.12.5"
        NetMask = "255.255.255.0"
        )


cssd requires ocrvote_mnt
cssd requires ora_priv
ocrvote_mnt requires ocrvote_voldg
ocrvote_mnt requires vxfsckd
ocrvote_voldg requires cvm_clus
vxfsckd requires cvm_clus
cvm_clus requires cvm_vxconfigd
```

Section 6

# Configuration of disaster recovery environments


- Chapter 20. Configuring disaster recovery environments

Chapter 20

# Configuring disaster recovery environments

This chapter includes the following topics:


- Disaster recovery options for SF Oracle RAC
- Hardware requirements for campus cluster
- Supported replication technologies for global clusters
- About setting up a campus cluster for disaster recovery
- About setting up a global cluster environment for SF Oracle RAC
- About configuring a parallel global cluster using Volume Replicator (VVR) for replication


## Disaster recovery options for SF Oracle RAC

SF Oracle RAC 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:

See "Supported replication technologies for global clusters" on page 337.

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 338.

See "About setting up a global cluster environment for SF Oracle RAC" on page 340.

See "About configuring a parallel global cluster using Volume Replicator (VVR) for replication" on page 341.

For complete details for configuring your disaster recovery environment once clusters are installed and configured:

See the *Veritas InfoScale™ Disaster Recovery Implementation Guide*.

# Hardware requirements for campus cluster

Table 20-1 lists the hardware requirements for campus clusters.

**Table 20-1** Hardware requirements for campus clusters

| Item | Description |
|---|---|
| Storage | ■ The storage switch (to which each host on a site connects) must have access to storage arrays at all the sites.<br>■ Volumes must be mirrored with storage allocated from at least two sites.<br>■ DWDM links are recommended between sites for storage links. DWDM works at the physical layer and requires multiplexer and de-multiplexer devices.<br>■ The storage and networks must have redundant-loop access between each node and each storage array to prevent the links from becoming a single point of failure. |
| Network | ■ Oracle requires that all nodes use the IP addresses from the same subnet.<br>■ Veritas recommends a common cross-site physical infrastructure for storage and LLT private networks. |
| I/O fencing | I/O fencing requires placement of a third coordinator point at a third site. The DWDM can be extended to the third site or the iSCSI LUN at the third site can be used as the third coordination point. Alternatively Coordination Point Server can be deployed at the third remote site as an arbitration point. |

# Supported replication technologies for global clusters

SF Oracle RAC supports the following hardware-based replication and software-based replication technologies for global cluster configurations:

| | |
|---|---|
| Hardware-based replication | ■ EMC SRDF<br>■ Hitachi TrueCopy<br>■ IBM Metro Mirror<br>■ IBM SAN Volume Controller (SVC)<br>■ EMC MirrorView |
| Software-based replication | ■ Volume Replicator<br>■ Oracle Data Guard |

# 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 for Oracle RAC (SF Oracle RAC)

The following high-level tasks illustrate the setup steps for a parallel campus cluster in an SF for Oracle RAC environment.

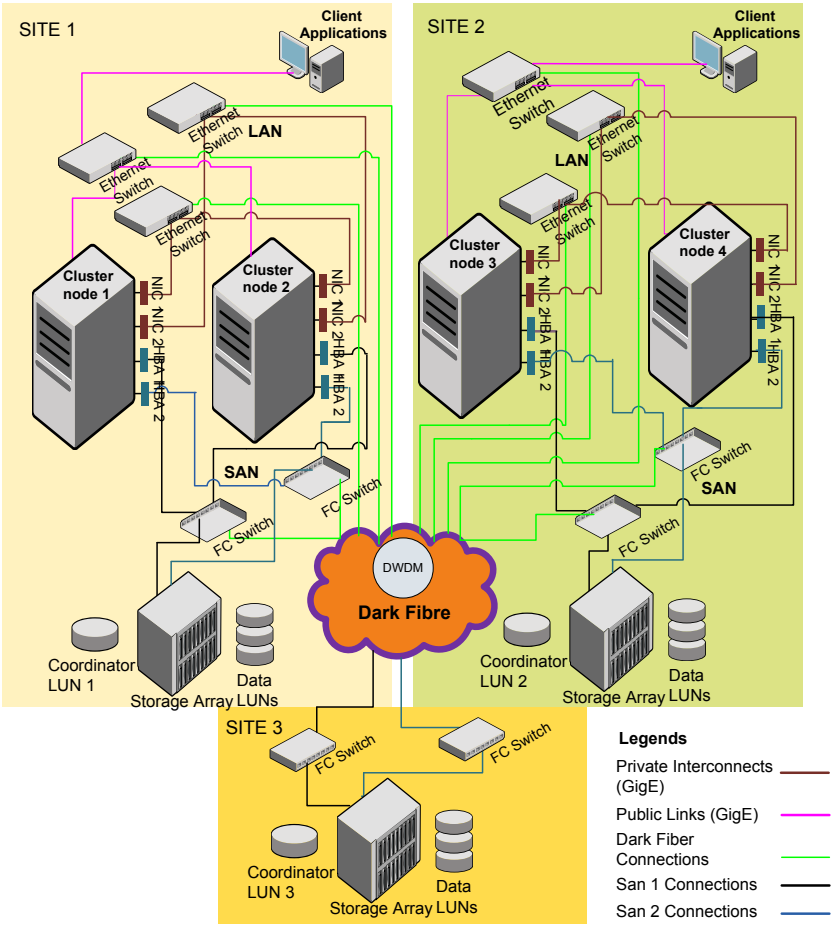
**Table 20-2** Tasks for setting up a parallel campus cluster for disaster recovery

| 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 Oracle RAC Clusterware and database binaries | See "About preparing to install Oracle RAC" on page 176. |
| Configure VxVM disk groups for campus cluster | See the *Veritas InfoScale™ Disaster Recovery Implementation Guide*. |

**Table 20-2** 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*.<br>For SFCFSHA, see your database documentation. |
| Install Oracle RAC Clusterware and database binaries | See "About installing Oracle RAC" on page 211. |
| Configure VCS service groups | See the *Veritas InfoScale™ Disaster Recovery Implementation Guide*. |

**Figure 20-1** 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 SF Oracle RAC

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 SF Oracle RAC on each cluster.

- Configure a SF Oracle RAC cluster at the primary site
- Configure an SF Oracle RAC 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 SF Oracle RAC 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 SF Oracle RAC, whether you are using keyless licensing or installing keys manually. You must have the GCO option for a global cluster and VVR enabled.
  Review SF Oracle RAC requirements and licensing information.
- Both clusters have SF Oracle RAC 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 SF Oracle RAC 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 7

# Installation reference


- Appendix A. Installation scripts
- Appendix B. Tunable files for installation
- Appendix C. Sample installation and configuration values
- Appendix D. Configuration files
- Appendix E. Configuring the secure shell or the remote shell for communications
- Appendix F. Automatic Storage Management
- Appendix G. Creating a test database
- Appendix H. High availability agent information
- Appendix I. SF Oracle RAC deployment scenarios
- Appendix J. Configuring LLT over UDP
- Appendix K. 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.<br>See "About configuring secure shell or remote shell communication modes before installing products" on page 421. |
| -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/llttab 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.

See “Performing a postcheck on a node” on page 98.

Appendix B

# Tunable files for installation

This appendix includes the following topics:


- About setting tunable parameters using the installer or a response file
- Setting tunables for an installation, configuration, or upgrade
- Setting tunables with no other installer-related operations
- Setting tunables with an un-integrated response file
- Preparing the tunables file
- Setting parameters for the tunables file
- Tunables value parameter definitions


## About setting tunable parameters using the installer or a response file

You can set non-default product and system tunable parameters using a tunables file. With the file, you can set tunables such as the I/O policy or toggle native multi-pathing. The tunables file passes arguments to the installer script to set tunables. With the file, you can set the tunables for the following operations:

- When you install, configure, or upgrade systems.

  ```
  # ./installer -tunablesfile tunables_file_name
  ```

  See "Setting tunables for an installation, configuration, or upgrade" on page 354.

- When you apply the tunables file with no other installer-related operations.

```
# ./installer -tunablesfile tunables_file_name -settunables [
sys1 sys2 ...]
```

See "Setting tunables with no other installer-related operations" on page 355.

- When you apply the tunables file with an un-integrated response file.

```
# ./installer -responsefile response_file_name -tunablesfile
tunables_file_name
```

See "Setting tunables with an un-integrated response file" on page 356.

For more information on response files, see the *chapter: About response files*.

You must select the tunables that you want to use from this guide.

See "Tunables value parameter definitions" on page 358.

# Setting tunables for an installation, configuration, or upgrade

You can use a tunables file for installation procedures to set non-default tunables. You invoke the installation script with the `tunablesfile` option. The tunables file passes arguments to the script to set the selected tunables. You must select the tunables that you want to use from this guide.

See "Tunables value parameter definitions" on page 358.

---

**Note:** Certain tunables only take effect after a system reboot.

---

**To set the non-default tunables for an installation, configuration, or upgrade**

1. Prepare the tunables file.

   See "Preparing the tunables file" on page 357.

2. Make sure the systems where you want to install SF Oracle RAC meet the installation requirements.
3. Complete any preinstallation tasks.
4. Copy the tunables file to one of the systems where you want to install, configure, or upgrade the product.
5. Mount the product disc and navigate to the directory that contains the installation program.

6 Start the installer for the installation, configuration, or upgrade. For example:

```
# ./installer -tunablesfile /tmp/tunables_file
 -settunables [sys1 sys2 ...]
```

Where /tmp/*tunables_file* is the full path name for the tunables file.

7 Proceed with the operation. When prompted, accept the tunable parameters.

Certain tunables are only activated after a reboot. Review the output carefully to determine if the system requires a reboot to set the tunable value.

8 The installer validates the tunables. If an error occurs, exit the installer and check the tunables file.

# Setting tunables with no other installer-related operations

You can use the installer to set tunable parameters without any other installer-related operations. You must use the parameters described in this guide. Note that many of the parameters are product-specific. You must select the tunables that you want to use from this guide.

See "Tunables value parameter definitions" on page 358.

---

**Note:** Certain tunables only take effect after a system reboot.

---

**To set tunables with no other installer-related operations**

1 Prepare the tunables file.

See "Preparing the tunables file" on page 357.

2 Make sure the systems where you want to install SF Oracle RAC meet the installation requirements.

3 Complete any preinstallation tasks.

4 Copy the tunables file to one of the systems that you want to tune.

5 Mount the product disc and navigate to the directory that contains the installation program.

6 Start the installer with the `-settunables` option.

```
# ./installer -tunablesfile tunables_file_name -settunables [
sys123 sys234 ...]
```

Where /tmp/*tunables_file* is the full path name for the tunables file.

7 Proceed with the operation. When prompted, accept the tunable parameters.

Certain tunables are only activated after a reboot. Review the output carefully to determine if the system requires a reboot to set the tunable value.

8 The installer validates the tunables. If an error occurs, exit the installer and check the tunables file.

# Setting tunables with an un-integrated response file

You can use the installer to set tunable parameters with an un-integrated response file. You must use the parameters described in this guide. Note that many of the parameters are product-specific. You must select the tunables that you want to use from this guide.

See "Tunables value parameter definitions" on page 358.

---

**Note:** Certain tunables only take effect after a system reboot.

---

**To set tunables with an un-integrated response file**

1 Make sure the systems where you want to install SF Oracle RAC meet the installation requirements.

2 Complete any preinstallation tasks.

3 Prepare the tunables file.

See "Preparing the tunables file" on page 357.

4 Copy the tunables file to one of the systems that you want to tune.

5 Mount the product disc and navigate to the directory that contains the installation program.

6 Start the installer with the `-responsefile` and `-tunablesfile` options.

```
# ./installer -responsefile response_file_name -tunablesfile
tunables_file_name
```

Where *response_file_name* is the full path name for the response file and *tunables_file_name* is the full path name for the tunables file.

7 Certain tunables are only activated after a reboot. Review the output carefully to determine if the system requires a reboot to set the tunable value.

8 The installer validates the tunables. If an error occurs, exit the installer and check the tunables file.

For more information on response files, see the *chapter: About response files*.

# Preparing the tunables file

A tunables file is a Perl module and consists of an opening and closing statement, with the tunables defined between. Use the hash symbol at the beginning of the line to comment out the line. The tunables file opens with the line "our %TUN;" and ends with the return true "1;" line. The final return true line only needs to appear once at the end of the file. Define each tunable parameter on its own line.

You can use the installer to create a tunables file template, or manually format tunables files you create.

**To create a tunables file template**

- Start the installer with the `-tunables` option. Enter the following:

  ```
  # ./installer -tunables
  ```

  You see a list of all supported tunables, and the location of the tunables file template.

**To manually format tunables files**

- Format the tunable parameter as follows:

  ```
  $TUN{"tunable_name"}{"system_name"|"*"}=value_of_tunable;
  ```

For the *system_name*, use the name of the system, its IP address, or a wildcard symbol. The *value_of_tunable* depends on the type of tunable you are setting. End the line with a semicolon.

The following is an example of a tunables file.

```
#
# Tunable Parameter Values:
#
our %TUN;

$TUN{"tunable1"}{"*"}=1024;
$TUN{"tunable3"}{"sys123"}="SHA256";

1;
```

# Setting parameters for the tunables file

Each tunables file defines different tunable parameters. The values that you can use are listed in the description of each parameter. Select the tunables that you want to add to the tunables file and then configure each parameter.

See "Tunables value parameter definitions" on page 358.

Each line for the parameter value starts with $TUN. The name of the tunable is in curly brackets and double-quotes. The system name is enclosed in curly brackets and double-quotes. Finally define the value and end the line with a semicolon, for example:

```
$TUN{"dmp_daemon_count"}{"node123"}=16;
```

In this example, you are changing the dmp_daemon_count value from its default of 10 to 16. You can use the wildcard symbol "*" for all systems. For example:

```
$TUN{"dmp_daemon_count"}{"*"}=16;
```

# Tunables value parameter definitions

When you create a tunables file for the installer you can only use the parameters in the following list.

Prior to making any updates to the tunables, refer to the *Storage Foundation Cluster File System High Availability Administrator's Guide* for detailed information on product tunable ranges and recommendations.

Table B-1 describes the supported tunable parameters that can be specified in a tunables file.

**Table B-1** Supported tunable parameters

| Tunable | Description |
|---|---|
| autoreminor | (Veritas Volume Manager) Enable reminoring in case of conflicts during disk group import. |
| autostartvolumes | (Veritas Volume Manager) Enable the automatic recovery of volumes. |
| dmp_cache_open | (Dynamic Multi-Pathing) Whether the first open on a device performed by an array support library (ASL) is cached. |
| dmp_daemon_count | (Dynamic Multi-Pathing) The number of kernel threads for DMP administrative tasks. |

**Table B-1** Supported tunable parameters *(continued)*

| Tunable | Description |
|---|---|
| dmp_delayq_interval | (Dynamic Multi-Pathing) The time interval for which DMP delays the error processing if the device is busy. |
| dmp_fast_recovery | (Dynamic Multi-Pathing) Whether DMP should attempt to obtain SCSI error information directly from the HBA interface. This tunable must be set after Dynamic Multi-Pathing is started. |
| dmp_health_time | (Dynamic Multi-Pathing) The time in seconds for which a path must stay healthy. |
| dmp_log_level | (Dynamic Multi-Pathing) The level of detail to which DMP console messages are displayed. |
| dmp_low_impact_probe | (Dynamic Multi-Pathing) Whether the low impact path probing feature is enabled. |
| dmp_lun_retry_timeout | (Dynamic Multi-Pathing) The retry period for handling transient errors. |
| dmp_monitor_fabric | (Dynamic Multi-Pathing) Whether the Event Source daemon (`vxesd`) uses the Storage Networking Industry Association (SNIA) HBA API. This tunable must be set after Dynamic Multi-Pathing is started. |
| dmp_monitor_ownership | (Dynamic Multi-Pathing) Whether the dynamic change in LUN ownership is monitored. |
| dmp_native_support | (Dynamic Multi-Pathing) Whether DMP does multi-pathing for native devices. |
| dmp_path_age | (Dynamic Multi-Pathing) The time for which an intermittently failing path needs to be monitored before DMP marks it as healthy. |
| dmp_pathswitch_blks_shift | (Dynamic Multi-Pathing) The default number of contiguous I/O blocks sent along a DMP path to an array before switching to the next available path. |
| dmp_probe_idle_lun | (Dynamic Multi-Pathing) Whether the path restoration kernel thread probes idle LUNs. |

**Table B-1** Supported tunable parameters *(continued)*

| Tunable | Description |
|---|---|
| dmp_probe_threshold | (Dynamic Multi-Pathing) The number of paths will be probed by the restore daemon. |
| dmp_restore_cycles | (Dynamic Multi-Pathing) The number of cycles between running the check_all policy when the restore policy is check_periodic. |
| dmp_restore_interval | (Dynamic Multi-Pathing) The time interval in seconds the restore daemon analyzes the condition of paths. |
| dmp_restore_policy | (Dynamic Multi-Pathing) The policy used by DMP path restoration thread. |
| dmp_restore_state | (Dynamic Multi-Pathing) Whether kernel thread for DMP path restoration is started. |
| dmp_retry_count | (Dynamic Multi-Pathing) The number of times a path reports a path busy error consecutively before DMP marks the path as failed. |
| dmp_scsi_timeout | (Dynamic Multi-Pathing) The timeout value for any SCSI command sent via DMP. |
| dmp_sfg_threshold | (Dynamic Multi-Pathing) The status of the subpaths failover group (SFG) feature. |
| dmp_stat_interval | (Dynamic Multi-Pathing) The time interval between gathering DMP statistics. |
| fssmartmovethreshold | (Veritas Volume Manager) The file system usage threshold for SmartMove (percent). This tunable must be set after Veritas Volume Manager is started. |
| max_diskq | (Veritas File System) Specifies the maximum disk queue generated by a single file. The installer can only set the system default value of max_diskq. Refer to the tunefstab(4) manual page for setting this tunable for a specified block device. |

**Table B-1** Supported tunable parameters *(continued)*

| Tunable | Description |
|---|---|
| read_ahead | (Veritas File System) The 0 value disables read ahead functionality, the 1 value (default) retains traditional sequential read ahead behavior, and the 2 value enables enhanced read ahead for all reads. The installer can only set the system default value of read_ahead. Refer to the tunefstab(4) manual page for setting this tunable for a specified block device. |
| read_nstream | (Veritas File System) The number of parallel read requests of size read_pref_io that can be outstanding at one time. The installer can only set the system default value of read_nstream. Refer to the tunefstab(4) manual page for setting this tunable for a specified block device. |
| read_pref_io | (Veritas File System) The preferred read request size. The installer can only set the system default value of read_pref_io. Refer to the tunefstab(4) manual page for setting this tunable for a specified block device. |
| reclaim_on_delete_start_time | (Veritas Volume Manager) Time of day to start reclamation for deleted volumes. This tunable must be set after Veritas Volume Manager is started. |
| reclaim_on_delete_wait_period | (Veritas Volume Manager) Days to wait before starting reclamation for deleted volumes. This tunable must be set after Veritas Volume Manager is started. |
| same_key_for_alldgs | (Veritas Volume Manager) Use the same fencing key for all disk groups. This tunable must be set after Veritas Volume Manager is started. |
| sharedminorstart | (Veritas Volume Manager) Start of range to use for minor numbers for shared disk groups. This tunable must be set after Veritas Volume Manager is started. |

**Table B-1** Supported tunable parameters *(continued)*

| Tunable | Description |
|---|---|
| storage_connectivity | (Veritas Volume Manager) The CVM storage connectivity type. This tunable must be set after Veritas Volume Manager is started. |
| usefssmartmove | (Veritas Volume Manager) Configure SmartMove feature (all, thinonly, none). This tunable must be set after Veritas Volume Manager is started. |
| vol_cmpres_enabled | (Veritas Volume Manager) Allow enabling compression for Volume Replicator. |
| vol_cmpres_threads | (Veritas Volume Manager) Maximum number of compression threads for Volume Replicator. |
| vol_default_iodelay | (Veritas Volume Manager) Time to pause between I/O requests from VxVM utilities (10ms units). This tunable requires a system reboot to take effect. |
| vol_fmr_logsz | (Veritas Volume Manager) Maximum size of bitmap Fast Mirror Resync uses to track changed blocks (KBytes). This tunable requires a system reboot to take effect. |
| vol_max_adminio_poolsz | (Veritas Volume Manager) Maximum amount of memory used by VxVM admin I/O's (bytes). This tunable requires a system reboot to take effect. |
| vol_max_nmpool_sz | (Veritas Volume Manager) Maximum name pool size (bytes). |
| vol_max_rdback_sz | (Veritas Volume Manager) Storage Record readback pool maximum (bytes). |
| vol_max_wrspool_sz | (Veritas Volume Manager) Maximum memory used in clustered version of Volume Replicator . |
| vol_maxio | (Veritas Volume Manager) Maximum size of logical VxVM I/O operations (kBytes). This tunable requires a system reboot to take effect. |

**Table B-1** Supported tunable parameters *(continued)*

| Tunable | Description |
|---|---|
| vol_maxioctl | (Veritas Volume Manager) Maximum size of data passed into the VxVM ioctl calls (bytes). This tunable requires a system reboot to take effect. |
| vol_maxparallelio | (Veritas Volume Manager) Number of I/O operations vxconfigd can request at one time. This tunable requires a system reboot to take effect. |
| vol_maxspecialio | (Veritas Volume Manager) Maximum size of a VxVM I/O operation issued by an ioctl call (kBytes). This tunable requires a system reboot to take effect. |
| vol_min_lowmem_sz | (Veritas Volume Manager) Low water mark for memory (bytes). |
| vol_nm_hb_timeout | (Veritas Volume Manager) Volume Replicator timeout value (ticks). |
| vol_rvio_maxpool_sz | (Veritas Volume Manager) Maximum memory requested by Volume Replicator (bytes). |
| vol_stats_enable | (Veritas Volume Manager) Enable VxVM I/O stat collection. |
| vol_subdisk_num | (Veritas Volume Manager) Maximum number of subdisks attached to a single VxVM plex. This tunable requires a system reboot to take effect. |
| voldrl_max_drtregs | (Veritas Volume Manager) Maximum number of dirty VxVM regions. This tunable requires a system reboot to take effect. |
| voldrl_max_seq_dirty | (Veritas Volume Manager) Maximum number of dirty regions in sequential mode. This tunable requires a system reboot to take effect. |
| voldrl_min_regionsz | (Veritas Volume Manager) Minimum size of a VxVM Dirty Region Logging (DRL) region (kBytes). This tunable requires a system reboot to take effect. |

**Table B-1** Supported tunable parameters *(continued)*

| Tunable | Description |
|---|---|
| voldrl_volumemax_drtregs | (Veritas Volume Manager) Max per volume dirty regions in log-plex DRL. |
| voldrl_volumemax_drtregs_20 | (Veritas Volume Manager) Max per volume dirty regions in DCO version 20. |
| voldrl_dirty_regions | (Veritas Volume Manager) Number of regions cached for DCO version 30. |
| voliomem_chunk_size | (Veritas Volume Manager) Size of VxVM memory allocation requests (bytes). This tunable requires a system reboot to take effect. |
| voliomem_maxpool_sz | (Veritas Volume Manager) Maximum amount of memory used by VxVM (bytes). This tunable requires a system reboot to take effect. |
| voliot_errbuf_dflt | (Veritas Volume Manager) Size of a VxVM error trace buffer (bytes). This tunable requires a system reboot to take effect. |
| voliot_iobuf_limit | (Veritas Volume Manager) Maximum total size of all VxVM I/O trace buffers (bytes). This tunable requires a system reboot to take effect. |
| voliot_iobuf_max | (Veritas Volume Manager) Maximum size of a VxVM I/O trace buffer (bytes). This tunable requires a system reboot to take effect. |
| voliot_max_open | (Veritas Volume Manager) Maximum number of VxVM trace channels available for vxtrace commands. This tunable requires a system reboot to take effect. |
| volpagemod_max_memsz | (Veritas Volume Manager) Maximum paging module memory used by Instant Snapshots (Kbytes). |
| volraid_rsrtransmax | (Veritas Volume Manager) Maximum number of VxVM RAID-5 transient reconstruct operations in parallel. This tunable requires a system reboot to take effect. |
| vxfs_mbuf | (Veritas File System) Maximum memory used for VxFS buffer cache. This tunable requires a system reboot to take effect. |

**Table B-1** Supported tunable parameters *(continued)*

| Tunable | Description |
|---|---|
| vxfs_ninode | (Veritas File System) Number of entries in the VxFS inode table. This tunable requires a system reboot to take effect. |
| write_nstream | (Veritas File System) The number of parallel write requests of size write_pref_io that can be outstanding at one time. The installer can only set the system default value of write_nstream. Refer to the tunefstab(4) manual page for setting this tunable for a specified block device. |
| write_pref_io | (Veritas File System) The preferred write request size. The installer can only set the system default value of write_pref_io. Refer to the tunefstab(4) manual page for setting this tunable for a specified block device. |

Appendix C

# Sample installation and configuration values

This appendix includes the following topics:


- About the installation and configuration worksheets
- SF Oracle RAC worksheet
- Oracle RAC worksheet
- Replicated cluster using VVR worksheet
- Replicated cluster using SRDF worksheet
- Required installation information for Oracle Clusterware/Grid Infrastructure
- Required installation information for Oracle database


## About the installation and configuration worksheets

The installation programs prompt you for information during the installation and configuration of SF Oracle RAC and Oracle RAC. The installation and configuration worksheets provide sample values that you can use as examples of the information required when you run the Veritas installation programs or the Oracle Universal Installer.

Veritas recommends using the worksheets to record values for your systems before you begin the installation and configuration process.

# SF Oracle RAC worksheet

This section provides worksheets for installing and configuring SF Oracle RAC, its component products, and features.

Table C-1 contains the sample values that may be used when you install and configure SF Oracle RAC. Enter the SF Oracle RAC values for your systems in the following table:

**Table C-1** SF Oracle RAC worksheet

| Installation information | Sample value | Assigned value |
|---|---|---|
| Number of nodes in the cluster | 2 | |
| Host names for Primary cluster | sys1 and sys2<br>**Note:** Do not use the underscore character in host names. Host names that use the underscore character are not compliant with RFC standards and cause issues. | |
| Host names for added or removed node | sys5<br>**Note:** Do not use the underscore character in host names. Host names that use the underscore character are not compliant with RFC standards and cause issues. | |
| Primary cluster name | clus1 | |
| Primary cluster ID number | 101 | |
| Private network links | eth1, eth2 | |
| Cluster Manager NIC (Primary NIC) | eth0 | |
| Cluster Manager IP | 10.10.12.1, 10.10.12.2 | |
| Netmask for the virtual IP address | 255.255.240.0 | |
| VCS user name (not required if you configure your cluster in secure mode) | VCS usernames must not exceed 1024 characters.<br>Example: smith | |

**Table C-1** SF Oracle RAC worksheet *(continued)*

| Installation information | Sample value | Assigned value |
|---|---|---|
| VCS user password | VCS passwords must not exceed 512 characters. | |
| VCS user privileges | Users have three levels of privileges: A=Administrator, O=Operator, or G=Guest.<br>Example: A | |
| Domain-based address of SMTP server | smtp.example.com | |
| Email address of SMTP notification recipients | admin@example.com | |
| Minimum severity of events for SMTP email notification | Events have four levels of severity:<br>■ I=Information<br>■ W=Warning<br>■ E=Error<br>■ S=SevereError<br>Example: I<br>The severity levels are defined as follows:<br>■ Information - Important events that exhibit normal behavior<br>■ Warning - Deviation from normal behavior<br>■ Error - A fault<br>■ Severe Error -Critical error that can lead to data loss or corruption | |
| SNMP trap daemon port number the console | 162 | |
| System name for the SNMP console | system2 | |

**Table C-1** SF Oracle RAC worksheet *(continued)*

| Installation information | Sample value | Assigned value |
|---|---|---|
| Minimum severity of events for SNMP trap notification | Events have four levels of severity:<br>■ I=Information<br>■ W=Warning<br>■ E=Error<br>■ S=SevereError<br>Example: I<br>The severity levels are defined as follows:<br>■ Information - Important events that exhibit normal behavior<br>■ Warning - Deviation from normal behavior<br>■ Error - A fault<br>■ Severe Error -Critical error that can lead to data loss or corruption | |
| Vxfen disks<br>These values are required if SCSI-3 disks are used as coordination points for your configuration. | /dev/sdc, /dev/sdd, /dev/sde | |
| Vxfen disk group | vxfencoordg | |

# Cluster Server component information

Table C-2 displays the information that is required to configure the Cluster Server component.

**Table C-2** Cluster Server component information

| Information | Example | Assigned values |
|---|---|---|
| Name of the cluster | The name must begin with a letter of the alphabet (a-z, A-Z) and contain only the characters a through z, A through Z, and 1 through 0, hyphen (-), and underscore (_).<br>Example: clus1 | |
| Unique ID number for the cluster | Number in the range of 0-65535. Within the site that contains the cluster, each cluster must have a unique ID.<br>Example: 101 | |
| Device names of the NICs used by the private networks among systems | You can choose a network interface card that is not part of any aggregated interface, or you can choose an aggregated interface.<br>The interface names that are associated with each NIC for each network link must be the same on all nodes.<br>For example:<br>■ eth1<br>■ eth2<br>Do not use the network interface card that is used for the public network, which is typically eth0. | |

# I/O fencing information

Table C-3 displays the information that is required to configure I/O fencing.

**Table C-3** I/O fencing information

| Information | Sample values | Assigned values |
|---|---|---|
| The name of three disks that form the coordinator disk group. | The following are examples of disk names:<br>■ /dev/sdc<br>■ /dev/sdd<br>■ /dev/sde | |
| The names for each disk in the coordinator disk group (if using DMP). | The following are examples:<br>■ /dev/vx/rdmp/sdc3<br>■ /dev/vx/rdmp/sdd3<br>■ /dev/vx/rdmp/sde3 | |

# SF Oracle RAC add user information

Table C-4 displays the information that is required to add VCS users. If you configure SF Oracle RAC cluster in secure mode, you need to add VCS users.

**Note:** Adding VCS users is optional.

**Table C-4** SF Oracle RAC add user information

| Information | Examples | Assigned values |
|---|---|---|
| User name | smith | |
| User password | ***** | |

**Table C-4** SF Oracle RAC add user information *(continued)*

| Information | Examples | Assigned values |
|---|---|---|
| User privilege | Users have three levels of privileges:<br>■ A=Administrator<br>■ O=Operator<br>■ G=Guest<br>Example: `A`<br>VCS privilege levels include:<br>■ Administrators— Can perform all operations, including configuration options on the cluster, service groups, systems, resources, and users.<br>■ Operators—Can perform specific operations on a cluster or a service group.<br>■ Guests—Can view specified objects. | |

# Global cluster information

Table C-5 displays the information that is required to configure global clusters.

**Note:** Global clusters are an optional feature that requires a license.

**Table C-5** Global cluster information

| Information | Example | Assigned values |
|---|---|---|
| Name of the public NIC | You must specify appropriate values for the NIC when you are prompted.<br>Example: `eth0` | |
| Virtual IP address of the NIC | You must specify appropriate values for the virtual IP address when you are prompted.<br>Example: `10.10.12.1` | |

**Table C-5** Global cluster information *(continued)*

| Information | Example | Assigned values |
|---|---|---|
| Netmask for the virtual IP address | You must specify appropriate values for the netmask when you are prompted.<br>Example: `255.255.255.0` | |

# Oracle RAC worksheet

This section provides a worksheet with sample values for installing and configuring Oracle RAC.

**Note:** *Italicized* text in parenthesis indicates the corresponding variable that is used in procedures. When you perform the steps in the procedures, ensure that you replace the variables with the values that are assigned to them in this worksheet.

Table C-6 displays the sample worksheets that may be used as reference when you perform the corresponding tasks.

**Table C-6** Required information for Oracle RAC

| Sample value sheet | Go to |
|---|---|
| Oracle user and group | See Table C-7 on page 374. |
| Public IP addresses and host names | See Table C-8 on page 375. |
| PrivNIC and MultiPrivNIC | |
| Oracle High Availability IP (HAIP) | See Table C-9 on page 375. |
| Oracle RAC home directories | See Table C-10 on page 376. |
| OCR and voting disk | See Table C-11 on page 378. |
| CSSD and Oracle database configuration | See Table C-12 on page 379. |

Table C-7 displays sample values that may be used when you create Oracle users and groups.

**Table C-7** Sample value sheet - Oracle user and group

| Information | Sample value | Assigned value |
|---|---|---|
| Oracle user name<br>(*user_name*) | ■ For Oracle Clusterware<br>**oracle**<br>■ For Oracle Grid Infrastructure<br>**grid** | |
| Oracle user ID<br>(*user_id*) | **1000** | |
| Oracle group name - Primary group<br>(*grp_name*) | **oinstall** (for inventory group as primary group) | |
| Oracle group name - Secondary group<br>(*grp_name_sec*) | **dba** (for dba group as secondary group) | |
| Oracle group ID - Primary group ID<br>(*grp_id*) | **1000** (for inventory group as primary group) | |
| Oracle group ID - Secondary group ID<br>(*grp_id_sec*) | **1001** (for dba group as secondary group) | |
| Oracle user home directory<br>(*usr_home_ora*) | **/home/oracle** | |
| Grid user home directory<br>(*usr_home_grid*) | **/home/grid** | |

Table C-8 displays sample values for public IP addresses and host names.

**Table C-8** Sample value sheet - Public IP addresses and host names

| Information | Sample value | Assigned value |
|---|---|---|
| Name of a node in the cluster<br>(*node_name1*)<br>(*node_name2*) | For a two-node cluster:<br>■ **sys1**<br>■ **sys2** | |
| Name of a new node added to the cluster<br>(*nodenew_name*) | **sys5** | |
| Virtual IP address | **10.10.10.10** | |
| Virtual IP alias | **sys1-vip** | |
| SCAN IP addresses | A minimum of three addresses is recommended.<br>**10.10.10.20**<br>**10.10.10.21**<br>**10.10.10.22** | |
| SCAN name | **clus1_scan** | |

The following table displays sample values that may be used when you use Oracle HAIP for Oracle RAC.

**Table C-9** Sample value sheet - Oracle HAIP

| Information | Sample value | Assigned value |
|---|---|---|
| Private IP address for first node in the cluster (sys1) | **192.168.12.2** | |
| Hostname alias for first node in the cluster (sys1) | **sys1-haip1** | |
| Private IP address for second node in the cluster (sys2) | **192.168.12.2** | |

**Table C-9** Sample value sheet - Oracle HAIP *(continued)*

| Information | Sample value | Assigned value |
|---|---|---|
| Hostname alias for second node in the cluster (sys2) | **sys2-haip1** | |
| Netmask | **255.255.255.0** | |

Table C-10 displays sample values that may be used when you create the Oracle RAC home directories.

**Table C-10** Sample value sheet - Oracle RAC home directories

| Information | Sample value | Assigned value |
|---|---|---|
| Disk for each node that contains the Oracle Clusterware and database binaries<br>*disk_name* | **Disk_1** | |
| VxVM local disk group name<br>*dg_name* | ▪ **bindg_sys1** (on sys1)<br>▪ **bindg_sys2** (on sys2) | |
| CVM disk group name<br>*cvm_dg* | **bindg** | |
| Volume name for Oracle Clusterware binaries<br>*clus_volname* | For Oracle Grid Infrastructure<br>**gridbinvol** | |
| Oracle Clusterware home directory<br>*clus_home* | For Oracle Grid Infrastructure, path to *grid_home*<br>**/u01/app/grid/product/11.2.0/gridhome** | |
| Volume name for Oracle Database binaries<br>*ora_volname* | **orabinvol** | |

**Table C-10** Sample value sheet - Oracle RAC home directories *(continued)*

| Information | Sample value | Assigned value |
|---|---|---|
| VCS resource name for disk groups containing the Oracle Clusterware/Grid Infrastructure and Oracle Database home directories<br>*dg_resname* | **bin_dg** | |
| Oracle Clusterware binary mount point resource name<br>*clusbin_mnt_resname* | For Oracle Grid Infrastructure<br>**gridbin_mnt** | |
| Oracle Database binary mount point resource name<br>*orabin_mnt_resname* | **orabin_mnt** | |
| Oracle base directory for Oracle Clusterware<br>*oracle_base* | **/u01/app/oracle** | |
| Oracle base directory for Oracle Grid Infrastructure<br>*grid_base* | **/u01/app/grid** | |
| Oracle base directory for Database<br>*oracle_base* | **/u01/app/oracle** | |
| Oracle Database home directory<br>*db_home* | **/u02/app/oracle/product/11.2.0/dbhome_1** | |
| Oracle Grid Infrastructure home directory<br>*grid_home* | **/u01/app/11.2.0/grid** | |

Table C-11 displays sample values that may be used when you create the storage for OCR and voting disk.

**Table C-11** Sample value sheet - OCR and voting disk

| Information | Sample value | Assigned value |
|---|---|---|
| Disks for creating a shared disk group for OCR and voting disk<br>*disk_name2*<br>*disk_name3* | **Disk_2**<br>**Disk_3** | |
| Shared disk group for OCR and voting disk<br>*ocrvote_dgname* | **ocrvotedg** | |
| OCR volume on CVM raw volumes<br>*ocr_volname* | **ocrvol** | |
| Voting disk volume on CVM raw volumes<br>*vote_volname* | **votevol** | |
| OCR and voting disk volume on CFS<br>*ocrvote_volname* | **ocrvotevol** | |
| Volume options for OCR and voting disk | nmirror=2 | |
| File system on shared volume (CFS) | /dev/vx/rdsk/ocrvotedg/ocrvotevol | |
| Mount point for shared file system<br>*ocrvote_mnt* | **/ocrvote** | |
| CVMVolDg resource name for OCR and voting disk<br>*ocrvotevol_resname* | **ocrvote_voldg_ocrvotedg** | |

**Table C-11** Sample value sheet - OCR and voting disk *(continued)*

| Information | Sample value | Assigned value |
|---|---|---|
| CFSMount resource name for OCR and voting disk<br>*ocrvotemnt_resname* | **ocrvote_mnt_ocrvotedg** | |
| Mount point for archive logs | **/oradata** | |
| CVM group name<br>*cvm_grpname* | **cvm** | |

Table C-12 displays sample values that may be used when you configure the CSSD agent and the Oracle database.

**Table C-12** Sample value sheet - CSSD and Oracle database configuration

| Information | Sample value | Assigned value |
|---|---|---|
| CVM group name<br>*cvm_grpname* | **cvm** | |
| CSSD group name<br>*cssd_grpname* | **cssd** | |
| VCS service group for Oracle Database<br>*oradb_grpname* | **oradb_grp** | |
| VCS resource name for Oracle<br>*db_resname* | **oradb** | |
| Oracle database name<br>*db_name* | **db** | |
| Global Database Name | db.example.com (database name.domain) | |
| Disk group name for Oracle Database<br>*oradb_dgname* | **oradatadg** (Shared) | |

**Table C-12** Sample value sheet - CSSD and Oracle database configuration *(continued)*

| Information | Sample value | Assigned value |
|---|---|---|
| Volume name for Oracle Database<br>*oradb_volname* | **oradatavol** | |
| Mount point for Oracle Database<br>*oradb_mnt* | **/oradata** (Shared) | |
| CVMVolDg resource for Oracle Database<br>*oradbdg_resname* | **oradata_voldg** | |
| CFSMount resource for Oracle Database<br>*oradbmnt_resname* | **oradata_mnt** | |
| SID prefix for policy-managed databases<br>*oradb_sid_prefix* | **db** | |
| SID on first node in the cluster<br>*oradb_sid_node1* | **db1** | |
| SID on second node in the cluster<br>*oradb_sid_node2* | **db2** | |

# Replicated cluster using VVR worksheet

Table C-13 contains the sample values that may be used when you install and configure CVM and VVR. If applicable, enter the CVM/VVR values for your systems in the following table:

**Table C-13** Replicated cluster using VVR worksheet

| Installation information | Sample value | Assigned value |
|---|---|---|
| Host names for Secondary Cluster | mercury, jupiter | |
| Secondary Cluster Name | rac_cluster102 | |
| Secondary cluster ID number | 102 | |
| Primary Cluster Address | 10.10.10.101 | |
| Primary Cluster Logowner IP | 10.10.9.101 | |
| Secondary Cluster Address | 10.11.10.102 | |
| Secondary Cluster Logowner IP | 10.11.9.102 | |
| RVG Name | rac1_rvg | |
| Global Database Name | vrts | |
| Database Resource Name | ora1 | |
| Database Group Name (depends on cvm, includes resources oracle agent etc.) | oradb1_grp | |
| Srl Volume Name | rac1_srl | |
| Resolvable Virtual Hostname of the cluster on Primary Site (for VVR) | rac_clus101_priv | |
| Resolvable Virtual Hostname of the cluster on Secondary Site (for VVR) | rac_clus102_priv | |
| Private IP addresses for Secondary Cluster | 192.168.12.3 - 192.168.12.4 | |

# Replicated cluster using SRDF worksheet

Table C-14 contains the sample values that may be used when you install and configure CVM and VVR. If applicable, enter the CVM/VVR values for your systems in the following table:

**Table C-14** Replicated cluster using SRDF worksheet

| Installation information | Sample value | Assigned value |
|---|---|---|
| Host names for Secondary Cluster | mercury, jupiter | |
| Secondary Cluster Name | rac_cluster102 | |
| Secondary cluster ID number | 102 | |
| Primary Cluster Address | 10.10.10.101 | |
| Secondary Cluster Address | 10.11.10.102 | |
| Global Database Name | vrts | |
| Database Resource Name | ora1 | |
| Database Group Name (depends on cvm, includes resources oracle agent etc.) | oradb1_grp | |
| Private IP addresses for Secondary Cluster | 192.168.12.3 - 192.168.12.4 | |

# Required installation information for Oracle Clusterware/Grid Infrastructure

This section describes the information required by the Oracle Universal Installer for installing the Oracle Clusterware/Grid Infrastructure software.

The following table lists the information required by the Oracle Universal Installer when you install Oracle Grid Infrastructure.

**Table C-15** Required installation information for Oracle Grid Infrastructure

| OUI menu | Description |
|---|---|
| **Select installation option** | Select the option **Install and Configure Grid Infrastructure for a Cluster**. |
| **Select cluster type** | Select the option **Configure a standard cluster**. |

**Table C-15** Required installation information for Oracle Grid Infrastructure *(continued)*

| OUI menu | Description |
|---|---|
| **Select installation type** | Select the option **Advanced Installation**. |
| **Grid Plug and Play information** | Provide the following information:<br>■ Name of the cluster<br>■ SCAN name<br>The SCAN address on the domain name server (DNS) must resolve to three addresses (recommended) or at least one address.<br>■ SCAN port |
| **Cluster Node Information** | Provide the public host name and the corresponding virtual host name of the nodes on which you want to install Oracle Grid Infrastructure. |
| **Specify network interface usage** | Identify the planned use for each interface: Public, Private, or Do Not use.<br>Review the following notes for configuring the network interface:<br>■ Make sure that the same private interfaces that you specified at the time of configuring PrivNIC and MultiPrivNIC are listed on the screen.<br>■ Mark the interfaces for the subnet containing the private IP addresses managed by the PrivNIC/MultiPrivNIC agents as ‘Private’.<br>The interfaces that are **Private** are stored in GPnP profile as a 'cluster_interconnect' for Oracle Grid Infrastructure communication and database cache fusion traffic.<br>■ For Flex ASM, mark the interfaces for the subnet containing the private IP addresses as **ASM & Private**. |
| **Grid Infrastructure Management Repository** | Select the option **No**. |
| **Storage option information** | For cluster file system with Oracle database: Select the option **Shared File System**.<br>For Oracle Flex ASM: Select the option **Use Oracle Flex ASM for storage**.<br>For Flex ASM, store the OCR and voting disk information on an ASM disk group. Select the appropriate disk paths, for example, `/dev/vx/rdsk/ocrasmdg/ocrasmvol` |

**Table C-15** Required installation information for Oracle Grid Infrastructure *(continued)*

| OUI menu | Description |
|---|---|
| **OCR storage option**<br>(Cluster file system only) | Enter the full path of the location where you want to store the OCR information.<br>For example, if you are storing the OCR information on CFS, enter: `/ocrvote/ocr`.<br>**Note:** Select the option **External Redundancy** when you install Oracle Clusterware/Grid Infrastructure. Mirroring is performed by CVM if you have chosen to mirror the OCR volume.<br>If you are storing the OCR information on an ASM disk group that uses CVM raw volumes, enter: `/dev/vx/rdsk/ocrvotedg/ocrvol`<br>**Note:** Select the option **External** when you create the ASM disk group and instances. Mirroring is performed by CVM if you have chosen to mirror the OCR volume. |
| **Voting Disk storage option**<br>(Cluster file system only) | Enter the full path of the location where you want to store the voting disk information.<br>For example, if you are storing the voting disk information on CFS, enter: `/ocrvote/vote`<br>**Note:** Select the option **External Redundancy** when you install Oracle Clusterware/Grid Infrastructure. Mirroring is performed by CVM if you have chosen to mirror the voting disk volume.<br>If you are storing the voting disk information on an ASM disk group that uses CVM raw volumes, enter: `/dev/vx/rdsk/ocrvotedg/votevol`<br>**Note:** Select the option **External** when you create the ASM disk group and instances. Mirroring is performed by CVM if you have chosen to mirror the voting disk volume. |
| **Failure Isolation** | Select the option **Do not use Intelligent Platform Management Interface (IPMI)** |
| **Privileged Operating System Groups** | Review or select the appropriate operating system groups. |
| **Specify installation location** | Enter the full path to the Oracle base directory and the Oracle Grid Infrastructure home directory. |

**Table C-15** Required installation information for Oracle Grid Infrastructure *(continued)*

| OUI menu | Description |
|---|---|
| **Create inventory** | Enter the full path to the Oracle inventory directory where you want to store the installation files. |
| **Root script execution configuration** | When the software is configured, some configuration scripts need to be run as the root user. Select the option **Automatically run configuration scripts** if you want the installer to run the scripts automatically. |

# Required installation information for Oracle database

This section describes the information required by the Oracle Universal Installer for installing the Oracle database software.

Table C-16 lists the information required by the Oracle Universal Installer when you install Oracle database software.

**Table C-16** Required installation information for Oracle database - Oracle RAC 11g Release 2 and later versions

| OUI menu | Description |
|---|---|
| **Select installation option** | Select the option **Install database software only**. |
| **Grid Installation Options** | Select the option **Oracle Real Application Clusters database installation**. |
| **Node selection** | Select **Real Application Clusters database installation**.<br>Select the nodes on which the Oracle RAC database software must be installed. |
| **Select database edition** | Select **Enterprise Edition**. |
| **Specify installation location** | Review or enter the ORACLE_BASE and ORACLE_HOME directory paths.<br>The Oracle Universal Installer runs product-specific prerequisite checks. Any items that are flagged must be manually checked and configured. |

**Table C-16** Required installation information for Oracle database - Oracle RAC 11g Release 2 and later versions *(continued)*

| OUI menu | Description |
|---|---|
| **Operating System Groups** | Review or select the appropriate operating system groups. |

Appendix D

# Configuration files

This appendix includes the following topics:


- About VCS configuration file
- About the LLT and GAB configuration files
- About I/O fencing configuration files
- Sample configuration files


## About VCS configuration file

This section provides a high-level overview of the contents of the VCS configuration file after the SF Oracle RAC installation. Review the configuration file after the SF Oracle RAC installation and before the Oracle installation.

The configuration file includes the following information:

- The "include" statements list the various VCS types files (`.cf`) for SF Oracle RAC.
  The files are located in the /etc/VRTSvcs/conf/config directory. These files define the agents that control the resources in the cluster.
- The cluster definition, with the cluster name provided during installation (for example, rac_cluster101), includes the names of users and administrators of the cluster. The UseFence = SCSI3 attribute is not automatically present; you must manually add it after the installation.
- The main.cf includes the cvm service group. The service group includes definitions for monitoring the CFS and the CVM resources. The CVMCluster agent resource definition indicates that the nodes use GAB for messaging operations.
  The cvm group has the Parallel attribute set to 1. This value enables the resources to run in parallel on each node in the system list.

# 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/llttab files. GAB requires /etc/gabtab file.

Table D-1 lists the LLT configuration files and the information that these files contain.

**Table D-1** LLT configuration files

| File | Description |
|---|---|
| /etc/sysconfig/llt | |

**Table D-1** LLT configuration files *(continued)*

| File | Description |
|---|---|
| | This file stores the start and stop environment variables for LLT:<br>■ LLT_START—Defines the startup behavior for the LLT module after a system reboot. Valid values include:<br>1—Indicates that LLT is enabled to start up.<br>0—Indicates that LLT is disabled to start up.<br>■ LLT_STOP—Defines the shutdown behavior for the LLT module during a system shutdown. Valid values include:<br>1—Indicates that LLT is enabled to shut down.<br>0—Indicates that LLT is disabled to shut down.<br>The installer sets the value of these variables to 1 at the end of SF Oracle RAC configuration.<br>If you manually configured VCS, make sure you set the values of these environment variables to 1.<br>Assign the buffer pool memory for RDMA operations:<br>■ 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.<br>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.<br>Define the number of RDMA queue pairs per LLT link<br>■ 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.<br>Enable or disable the adaptive window feature:<br>■ 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.<br>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"'.<br>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"'<br>Configure LLT over TCP<br>■ 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 |

**Table D-1** LLT configuration files *(continued)*

| File | Description |
|---|---|
| | value to maximum 64 connections. |
| /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.<br>For example, the file /etc/llthosts contains the entries that resemble:<br>`0 sys1`<br>`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.<br>For example, the file /etc/lltab contains the entries that resemble:<br>`set-node sys1`<br>`set-cluster 2`<br>`link eth1 eth1 - ether - -`<br>`link eth2 eth2 - ether - -`<br>If you use aggregated interfaces, then the file contains the aggregated interface name instead of the eth-*MAC_address*.<br>`set-node sys1`<br>`set-cluster 2`<br>`link eth1 eth-00:04:23:AC:12:C4 - ether - -`<br>`link eth2 eth-00:04:23:AC:12:C5 - ether - -`<br>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.<br>If you configured a low priority link under LLT, the file also includes a "link-lowpri" line.<br>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 D-2 lists the GAB configuration files and the information that these files contain.

**Table D-2** GAB configuration files

| File | Description |
|---|---|
| /etc/sysconfig/gab | This file stores the start and stop environment variables for GAB:<br>■ GAB_START—Defines the startup behavior for the GAB module after a system reboot. Valid values include:<br>1—Indicates that GAB is enabled to start up.<br>0—Indicates that GAB is disabled to start up.<br>■ GAB_STOP—Defines the shutdown behavior for the GAB module during a system shutdown. Valid values include:<br>1—Indicates that GAB is enabled to shut down.<br>0—Indicates that GAB is disabled to shut down.<br>The installer sets the value of these variables to 1 at the end of SF Oracle RAC configuration. |
| /etc/gabtab | After you install SF Oracle RAC, the file /etc/gabtab contains a `gabconfig(1)` command that configures the GAB driver for use.<br>The file /etc/gabtab contains a line that resembles:<br>`/sbin/gabconfig -c -nN`<br>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.<br>**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 I/O fencing configuration files

Table D-3 lists the I/O fencing configuration files.

**Table D-3** I/O fencing configuration files

| File | Description |
|---|---|
| /etc/sysconfig/vxfen | This file stores the start and stop environment variables for I/O fencing:<br>■ VXFEN_START—Defines the startup behavior for the I/O fencing module after a system reboot. Valid values include:<br>1—Indicates that I/O fencing is enabled to start up.<br>0—Indicates that I/O fencing is disabled to start up.<br>■ VXFEN_STOP—Defines the shutdown behavior for the I/O fencing module during a system shutdown. Valid values include:<br>1—Indicates that I/O fencing is enabled to shut down.<br>0—Indicates that I/O fencing is disabled to shut down.<br>The installer sets the value of these variables to 1 at the end of SF Oracle RAC configuration. |
| /etc/vxfendg | This file includes the coordinator disk group information.<br>This file is not applicable for server-based fencing. |

**Table D-3** I/O fencing configuration files *(continued)*

| File | Description |
|---|---|
| /etc/vxfenmode | This file contains the following parameters:<br>■ vxfen_mode<br>  ■ scsi3—For disk-based fencing.<br>  ■ customized—For server-based fencing.<br>  ■ disabled—To run the I/O fencing driver but not do any fencing operations.<br>■ vxfen_mechanism<br>This parameter is applicable only for server-based fencing. Set the value as cps.<br>■ scsi3_disk_policy<br>  ■ dmp—Configure the vxfen module to use DMP devices<br>  The disk policy is dmp by default. If you use iSCSI devices, you must set the disk policy as dmp.<br>**Note:** You must use the same SCSI-3 disk policy on all the nodes.<br>■ List of coordination points<br>This list is required only for server-based fencing configuration.<br>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.<br>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.<br>■ autoseed_gab_timeout<br>This parameter enables GAB automatic seeding of the cluster even when some cluster nodes are unavailable.<br>This feature is applicable for I/O fencing in SCSI3 and customized mode.<br>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.<br>-1—Turns the GAB auto-seed feature off. This setting is the default.<br>■ detect_false_pesb<br>0—Disables stale key detection.<br>1—Enables stale key detection to determine whether a preexisting split brain is a true condition or a false alarm.<br>Default: 0<br>**Note:** This parameter is considered only when `vxfen_mode=customized`. |

**Table D-3** 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.<br>**Note:** The /etc/vxfentab file is a generated file; do not modify this file.<br>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:<br>■ DMP disk:<br>`/dev/vx/rdmp/sdx3 HITACHI%5F1724-100%20%20FAStT%5FDISKS%5F6 00A0B8000215A5D000006804E795D0A3`<br>`/dev/vx/rdmp/sdy3 HITACHI%5F1724-100%20%20FAStT%5FDISKS%5F6 00A0B8000215A5D000006814E795D0B3`<br>`/dev/vx/rdmp/sdz3 HITACHI%5F1724-100%20%20FAStT%5FDISKS%5F6 00A0B8000215A5D000006824E795D0C3`<br>For server-based fencing, the /etc/vxfentab file also includes the security settings information. |

# Sample configuration files

SF Oracle RAC provides several sample configuration files illustrating various scenarios. You may use the sample files as a guideline for setting up your cluster environment. These sample files are located at `/etc/VRTSvcs/conf/sample_rac/`.

This section briefly describes each of the sample files and illustrates the service group configuration for each of them. The section does not include a copy of the main.cf files.

The following sample files are discussed in this section:

- sfrac02_main.cf file
- sfrac03_main.cf file
- sfrac04_main.cf file
- sfrac05_main.cf file
- sfrac06_main.cf file

- sfrac07_main.cf and sfrac08_main.cf files
- sfrac09_main.cf and sfrac10_main.cf files
- sfrac11_main.cf file
- sfrac12_main.cf and sfrac13_main.cf files
- sfrac14_main.cf file
- sfrac15_main.cf and sfrac16_main.cf files

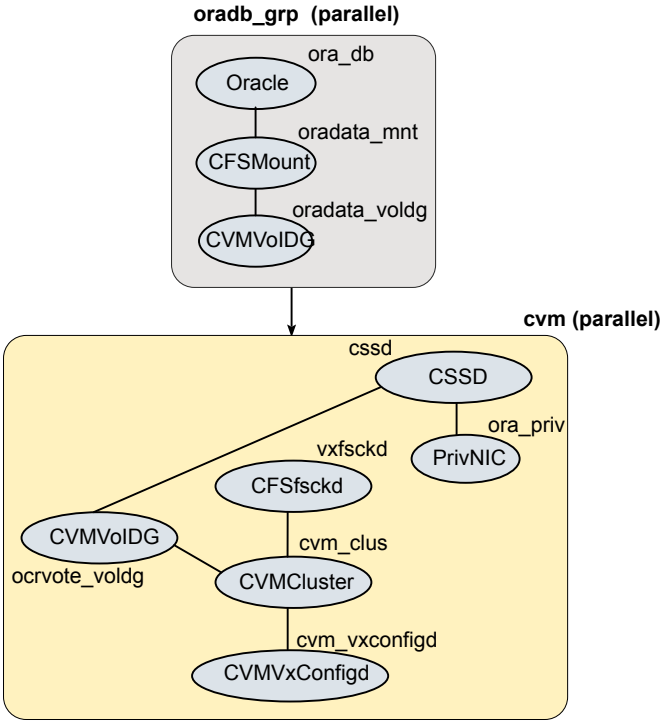
## sfrac02_main.cf file

This sample file describes the following configuration:

- A two node SF Oracle RAC cluster hosting single database.
- The Oracle database is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion.
  The private IP address is managed by the PrivNIC agent for high availability.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CVM raw volumes.

Figure D-1 illustrates the configuration.

**Figure D-1** Service group configuration for sfrac02_main.cf file



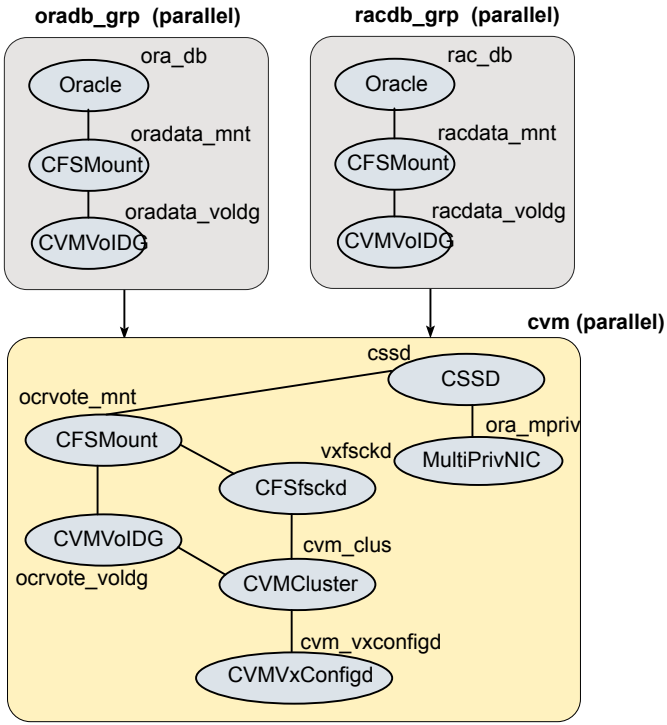

## sfrac03_main.cf file

This sample file describes the following configuration:

- A two node SF Oracle RAC cluster hosting two databases.
- The Oracle database is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- One IP address (on eth1) is shared by Oracle Clusterware and one of the databases for cache fusion.
  The second IP address (on eth2) is used by the second database for cache fusion.
  The private IP addresses are managed by the MultiPrivNIC agent for high availability.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.

Figure D-2 illustrates the configuration.

**Figure D-2** Service group configuration for sfrac03_main.cf file




## sfrac04_main.cf file

This sample file describes the following configuration:

- A two node SF Oracle RAC cluster hosting two databases.
- The Oracle database is stored on CFS.
- Database is not managed by VCS. Oracle Clusterware starts, stops, and monitors the databases.
  The CRSResource agent monitors the status of the database, the VIP resource, and the listener resource configured under Oracle Clusterware.

---

**Note:** The CFSMount and CVMVolDg resources for Oracle database can not be set as critical resources in the group.

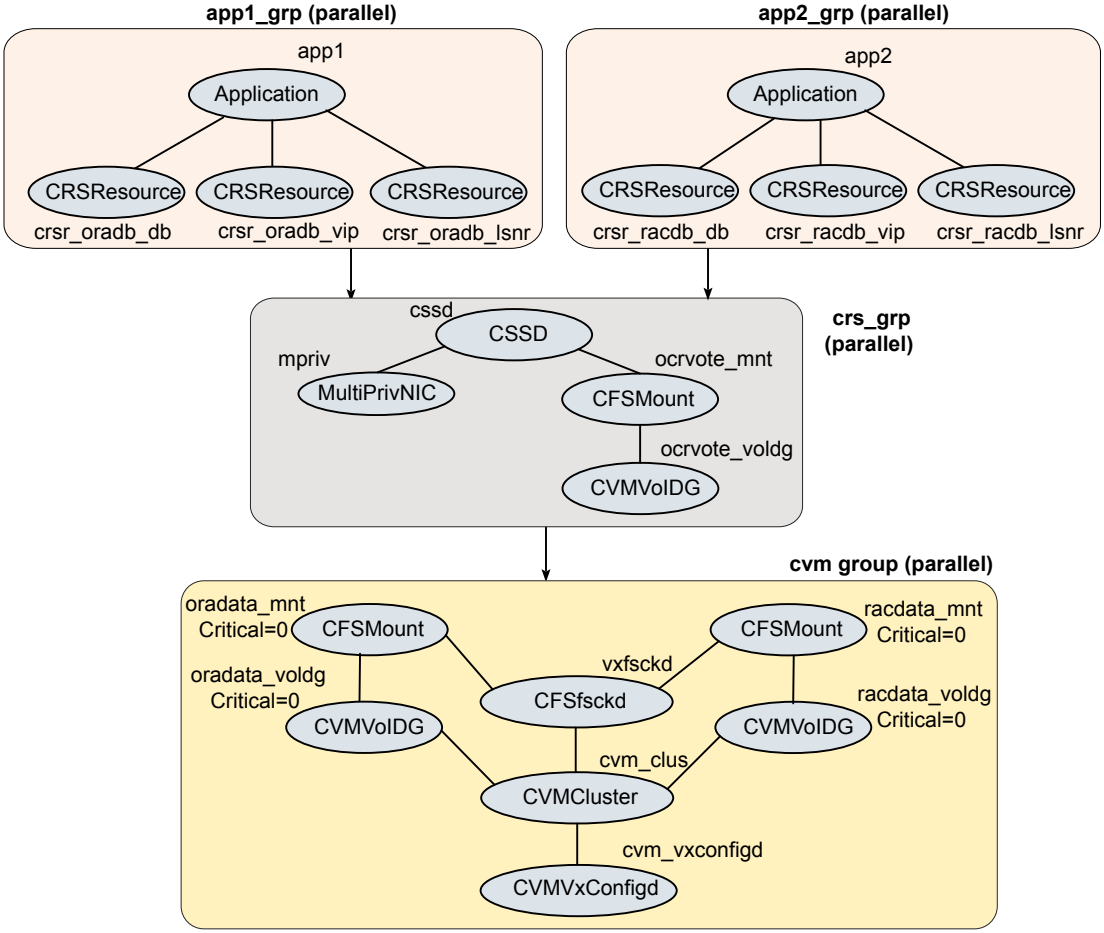
The CRSResource agent appears FAULTED until Oracle Clusterware brings up the database.

---

- The database uses the Oracle UDP IPC for database cache fusion.
- One IP address (on eth1) is shared by Oracle Clusterware and one of the databases for cache fusion.
  The second IP address (on eth2) is used by the second database for cache fusion.
  The private IP addresses are managed by the MultiPrivNIC agent for high availability.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.

Figure D-3 illustrates the configuration.

**Figure D-3** Service group configuration for sfrac04_main.cf file




## sfrac05_main.cf file

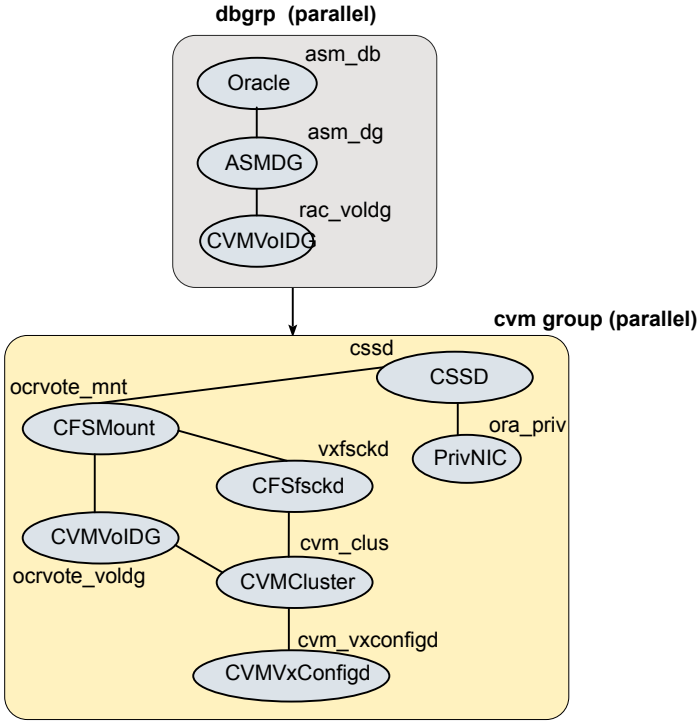
This sample file describes the following configuration:

- A two node SF Oracle RAC cluster hosting single database.
- The Oracle database is stored on ASM.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.

- A common IP address is used by Oracle Clusterware and database cache fusion. The private IP address is managed by the PrivNIC agent for high availability.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.

Figure D-4 illustrates the configuration.

**Figure D-4** Service group configuration for sfrac05_main.cf file




## sfrac06_main.cf file

This sample file describes the following configuration:

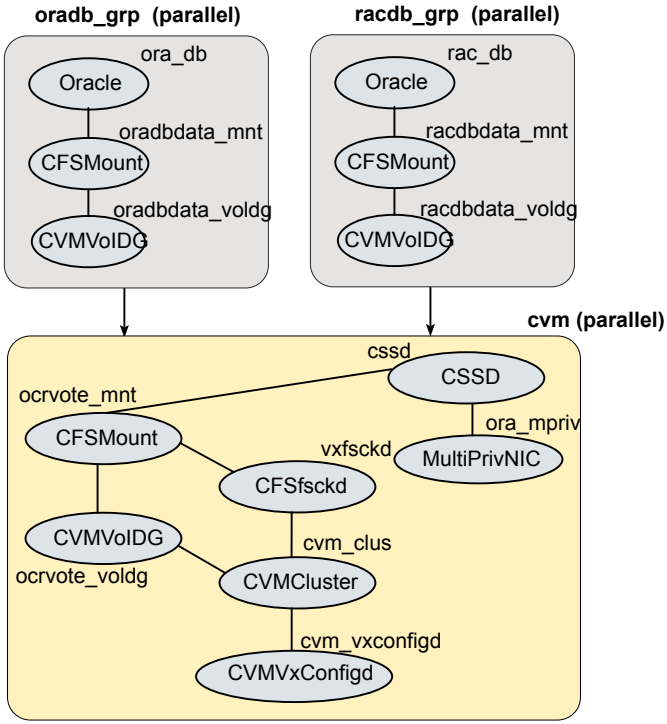
- A two node SF Oracle RAC cluster hosting two databases.
- The Oracle database is stored on CFS.
- The database is managed by the VCS agent for Oracle. The agent starts, stops, and monitors the database.
- The database uses Oracle UDP IPC for cache fusion. A dedicated link is used for each database .
- One private IP address (on eth1) is used by Oracle Clusterware.

The private IP address on eth2 is used by one of the databases for cache fusion. The private IP address on NIC3 is used by the other database for cache fusion. The private IP addresses are managed by the MultiPrivNIC agent for high availability.

- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.

Figure D-5 illustrates the configuration.

**Figure D-5** Service group configuration for sfrac06_main.cf file




## sfrac07_main.cf and sfrac08_main.cf files

The sample configuration, sfrac07_main.cf, describes a disaster recovery configuration for the primary site. The sample configuration, sfrac08_main.cf, describes a disaster recovery configuration for the secondary site. The configuration uses VVR for replicating data between the sites.

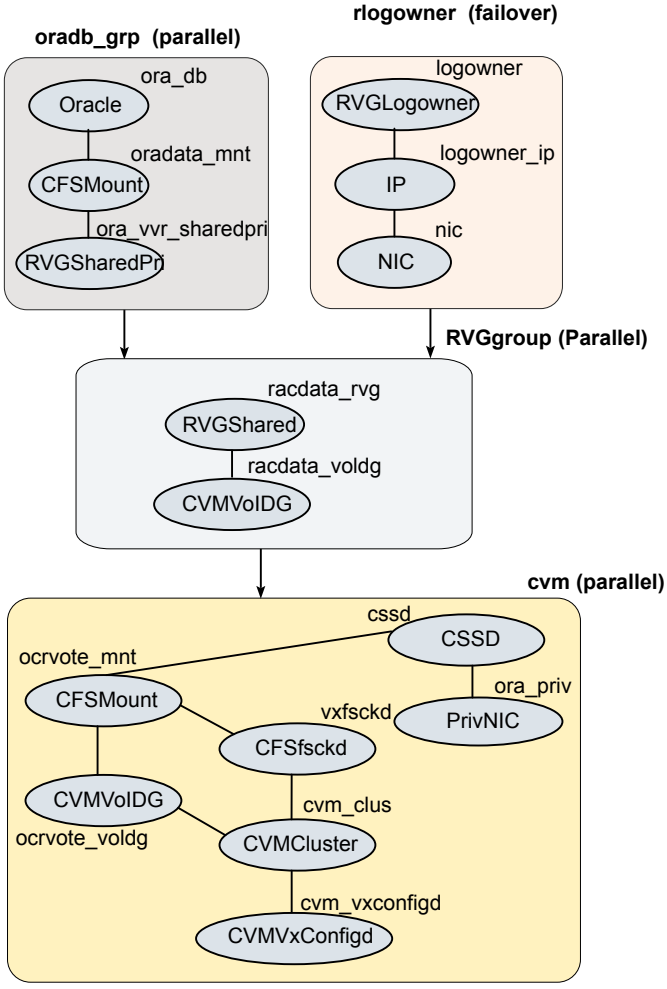
This sample file describes the following configuration:

- Two SF Oracle RAC clusters, comprising two nodes each, hosted at different geographical locations.

- A single Oracle database that is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion. The private IP address is managed by the PrivNIC agent for high availability.
- One virtual IP address must be configured under the `ClusterService` group on each site for inter-cluster communication.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.
- Volume Replicator (VVR) is used to replicate data between the sites.
- The shared volumes replicated across the sites are configured under the RVG group.
- The replication link used by VVR for communicating log information between sites are configured under the `rlogowner` group. This is a failover group that will be online on only one of the nodes in the cluster at each site.
- The database group will be online on the primary cluster. The RVGSharedPri resource determines where the database group will be brought online.
- The database group is configured as a global group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute.

Figure D-6 illustrates the configuration. The service group configuration is the same on the primary and secondary site. The availability of groups (online/offline) differ between the sites.

**Figure D-6** Service group configuration for sfrac07_main.cf and sfrac08_main.cf files




## sfrac09_main.cf and sfrac10_main.cf files

The sample configuration, sfrac09_main.cf, describes a disaster recovery configuration for the primary site. The sample configuration, sfrac10_main.cf, describes a disaster recovery configuration for the secondary site. The sample configuration uses EMC SRDF technology for replicating data between the sites.

---

**Note:** You can use other supported hardware-based replication technologies with this configuration.

---

This sample file describes the following configuration:

- Two SF Oracle RAC clusters, comprising two nodes each, hosted at different geographical locations.
- A single Oracle database that is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion. The private IP address is managed by the PrivNIC agent for high availability.
- One virtual IP address must be configured under the `ClusterService` group on each site for inter-cluster communication.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.
- EMC SRDF is used to replicate data between the sites.
- The SRDF disk groups that are replicated across the sites using SRDF technology and the replication mode are specified under the SRDF resource in the database group. The CVM disk group that comprises the SRDF disk group must be configured under the CVMVolDg resource in the database group.
- The database group will be online on the primary cluster. The SRDF resource determines where the database group will be brought online.
- The database group is configured as a global group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute.

Figure D-7 illustrates the configuration on the primary site.

**Figure D-7** Service group configuration for sfrac09_main.cf file

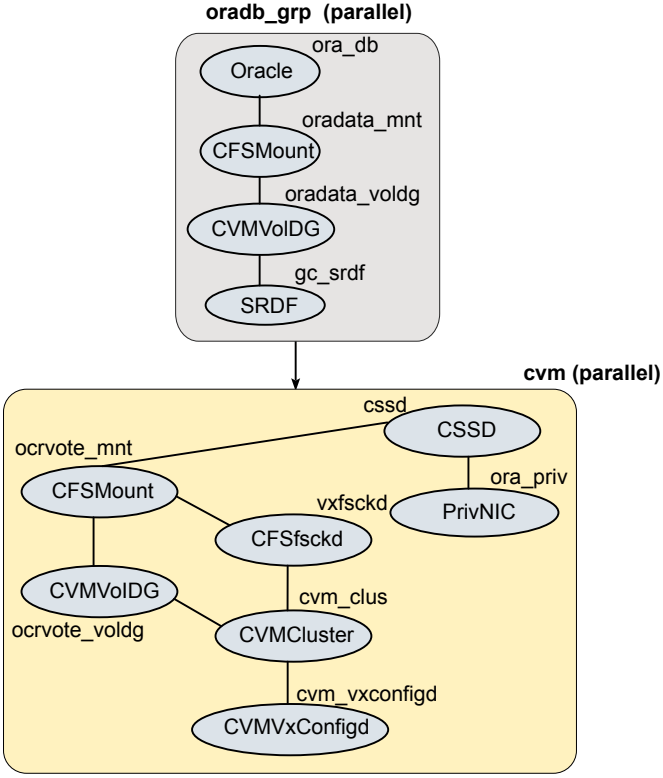
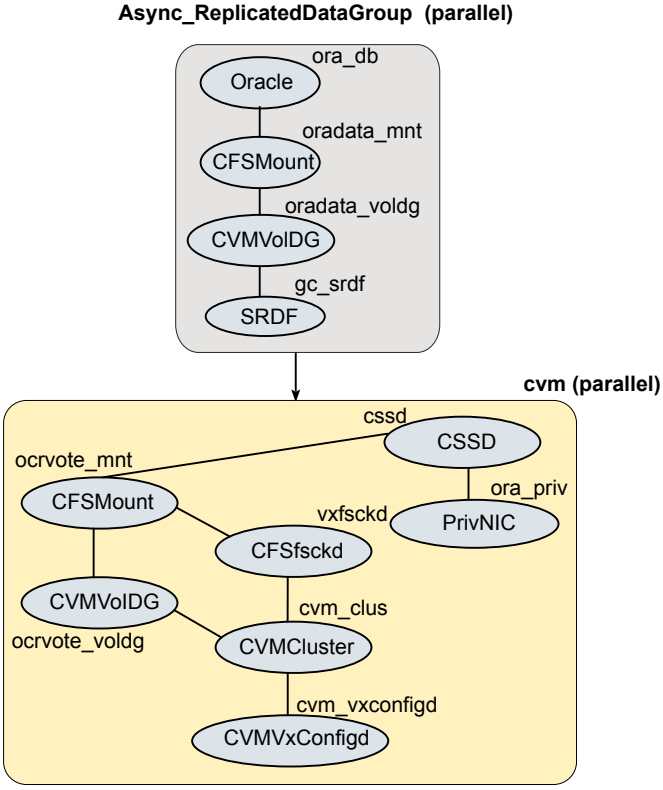



Figure D-8 illustrates the configuration on the secondary site.

**Figure D-8** Service group configuration for sfrac10_main.cf file




## sfrac11_main.cf file

This sample file describes the following configuration:

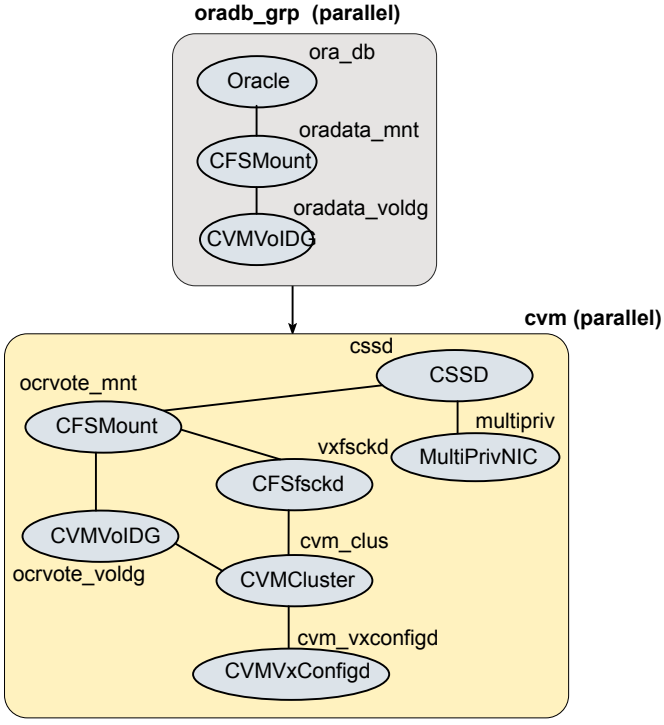
- An SF Oracle RAC campus cluster with four nodes hosted across two sites.
- Each site comprises two nodes of the cluster hosting a single database.
- The Oracle database is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- The IP address on eth1 is used by Oracle Clusterware.
  The second IP address on NIC2 is used for Oracle database cache fusion.

The private IP addresses are managed by the MultiPrivNIC agent for high availability.

- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.
- Group the hosts at each physical site into separate logical system zones using the SystemZones attribute.

Figure D-9 illustrates the configuration.

**Figure D-9** Service group configuration for sfrac11_main.cf file




## sfrac12_main.cf and sfrac13_main.cf files

The sample configuration, sfrac12_main.cf, describes a disaster recovery configuration for the primary site. The sample configuration, sfrac13_main.cf, describes a disaster recovery configuration for the secondary site with fire-drill capability. The sample configuration uses Hitachi True Copy technology for replicating data between the sites.

---

**Note:** You can use other supported hardware-based replication technologies with this configuration.

---

This sample file describes the following configuration:

- Two SF Oracle RAC clusters, comprising two nodes each, hosted at different geographical locations.
- A single Oracle database that is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion. The private IP address is managed by the PrivNIC agent for high availability.
- One virtual IP address must be configured under the `ClusterService` group on each site for inter-cluster communication.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.
- Hitachi True Copy is used to replicate data between the sites.
- The HTC disk groups that are replicated across the sites using HTC technology and the replication mode are specified under the HTC resource in the database group. The CVM disk group that comprises the HTC disk group must be configured under the CVMVolDg resource in the database group.
- The database group will be online on the primary cluster. The HTC resource determines where the database group will be brought online.
- The database group is configured as a global group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute.
- The database group oradb_grp_fd on the secondary is configured for fire drill.
- When the group oradb_grp_fd is brought online, the HTCSnap creates a snapshot of the disk group configured under the HTC resource in the database group oradg_grp.
  Further, the Oracle database and the associated volumes and mount points configured under the service group oradb_grp_fd are brought online using the snapshots created by HTCSnap.

Figure D-10 illustrates the configuration on the primary site.

**Figure D-10** Service group configuration for sfrac12_main.cf file

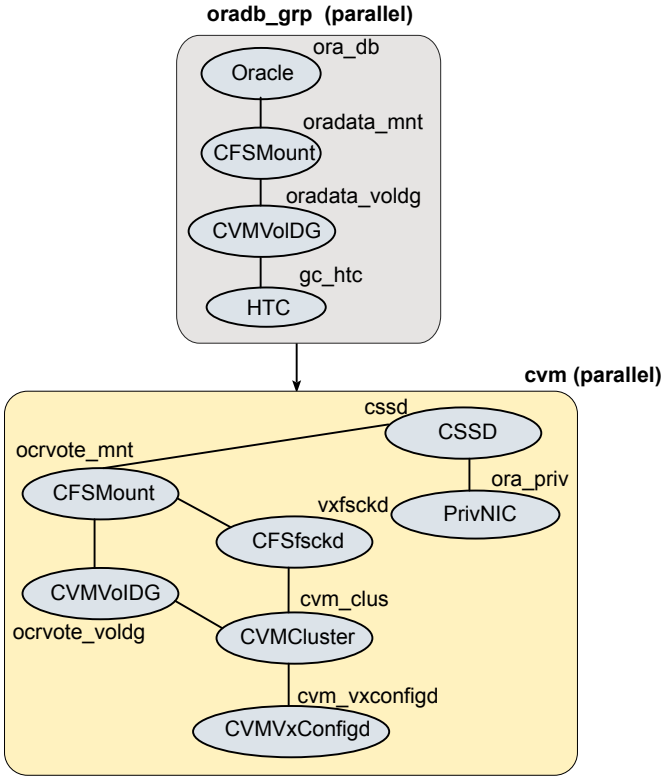
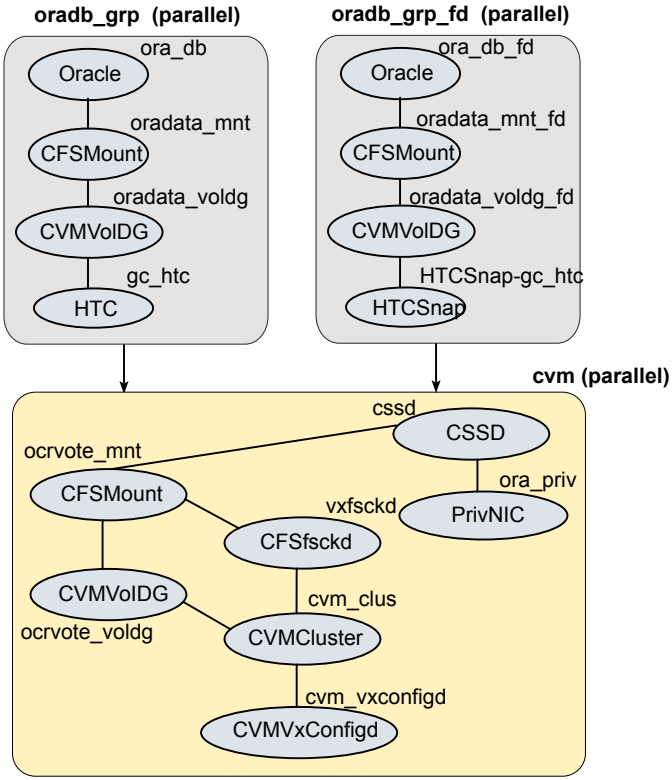



Figure D-11 illustrates the configuration on the secondary site.

**Figure D-11** Service group configuration for sfrac13_main.cf file



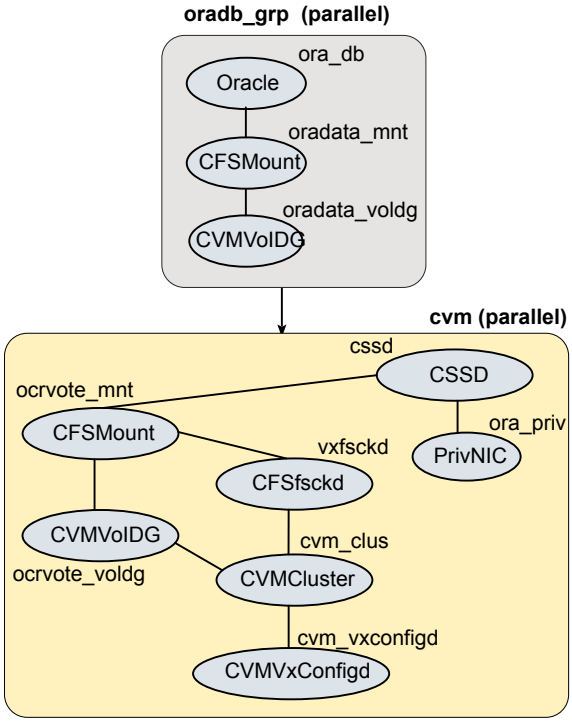

## sfrac14_main.cf file

This sample file describes the following configuration:

- A two node SF Oracle RAC cluster hosting single database.
- The Oracle database is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion.
  The private IP address is managed by the PrivNIC agent for high availability.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.

Figure D-12 illustrates the configuration.

**Figure D-12** Service group configuration for sfrac14_main.cf file



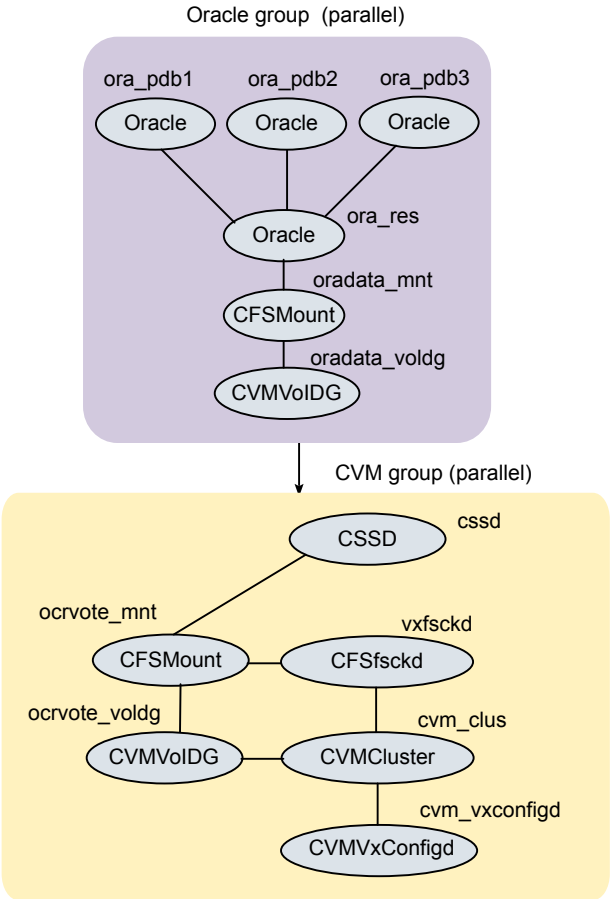

## sfrac15_main.cf and sfrac16_main.cf files

The sample files describe the following configuration:

- A two-node SF Oracle RAC cluster hosting a single container database, which contains 3 pluggable databases.
- Oracle Cluster Registry (OCR) and voting disk stored on CFS.
- Oracle container database and pluggable databases are stored on CFS.
- Container and pluggable databases are managed by VCS.
  The VCS Oracle agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion.
- The cluster is set up using the Secure Cluster option.

Figure D-13 illustrates the configuration.

**Figure D-13** Service group configuration for sfrac15_main.cf and sfrac16_main.cf files



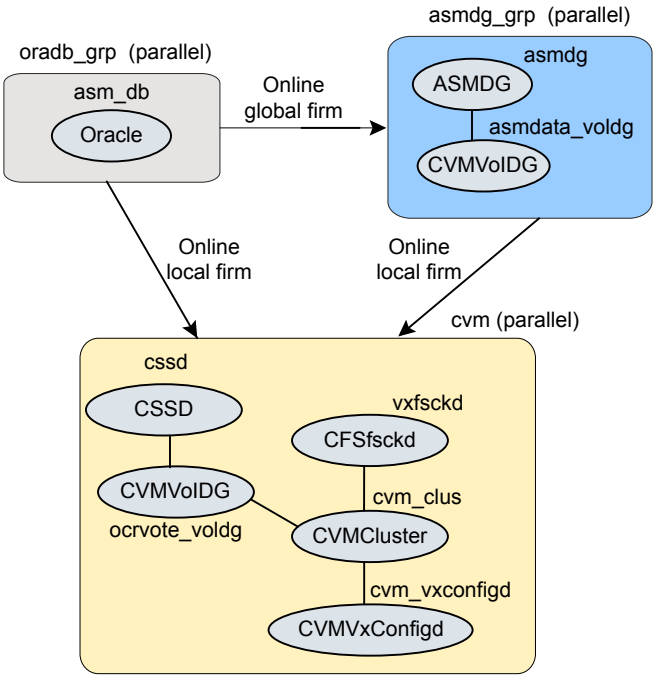

## sfrac17_main.cf file

The sample file describes the following configuration:

- A four-node SF Oracle RAC cluster hosting a single database.
- Oracle Cluster Registry (OCR) and voting disk stored on ASM in Flex ASM cluster.
- Oracle database is stored on ASM (on separate ASM disk groups).
- The database is managed by VCS.
  The VCS Oracle agent starts, stops, and monitors the database.

Figure D-14 illustrates the configuration.

**Figure D-14** Service group configuration for sfrac17_main.cf file




## 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 420.

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 415.
- 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 417.

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 E

# 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.

---

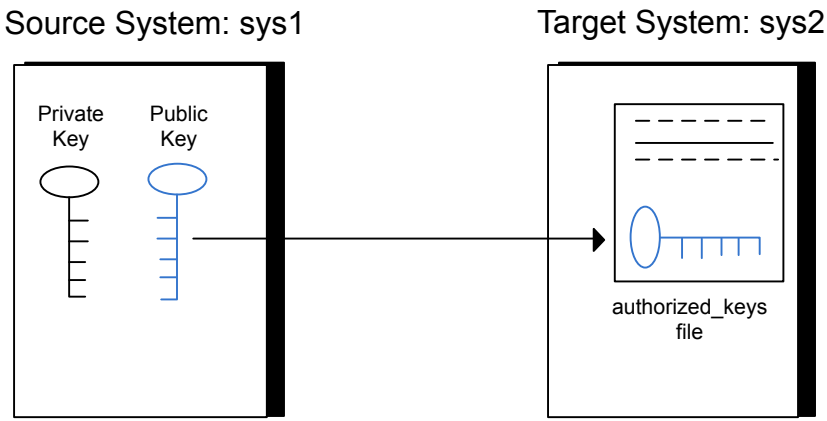
# Manually configuring passwordless ssh

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 E-1 illustrates this procedure.

**Figure E-1** Creating the DSA key pair and appending it to target systems



If you are installing Oracle, you must configure a DSA key and an RSA key for the Oracle user in addition to the DSA key required for the root user to install SF Oracle RAC.

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 E-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 E-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:<br><hostname><br><user>:<password>@<hostname><br><user>:<password>@<hostname>:<br><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 422.

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 F

# Automatic Storage Management

This appendix includes the following topics:


- About ASM in SF Oracle RAC environments
- ASM configuration with SF Oracle RAC
- Configuring ASM in SF Oracle RAC environments
- Configuring Flex ASM in SF Oracle RAC environments


## About ASM in SF Oracle RAC environments

ASM is an integrated storage management solution from Oracle RAC that combines file system and volume management capabilities.

ASM can be configured with Cluster Volume Manager (CVM) for better performance and availability. CVM mirrored volumes with dynamic multi-pathing improves data access performance and offers continuous data availability in large heterogeneous SAN environments. You can create CVM disk groups and volumes for use as ASM disks groups and configure the ASM disk groups to be managed by the Veritas ASMDG agent. The ASMDG agent mounts, unmounts, and monitors the ASM disk groups.

The following are supported on ASM for that version:

The following are not supported on ASM for that version:

ASM provides storage for Oracle Cluster Registry devices (OCR), voting disk, data files, control files, online redo logs, archive log files, and backup files.

**Note:** ASM does not support Oracle binaries, trace files, alert logs, export files, tar files, core files, application binaries and data.

# ASM configuration with SF Oracle RAC

Configure ASM disk groups over CVM volumes in SF Oracle RAC environments. The CVM volumes are mirrored for high availability of data and leverage Dynamic Multi-Pathing to access the shared storage.

Figure F-1 illustrates the configuration of ASM disk groups over CVM.

**Figure F-1** ASM disk groups over CVM

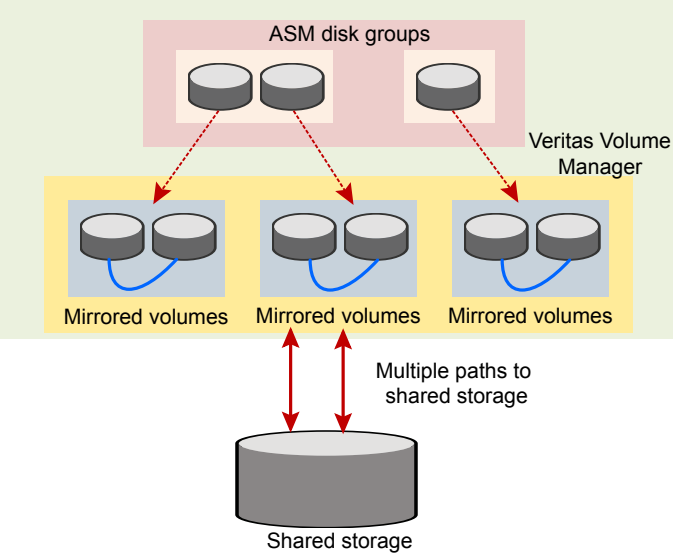
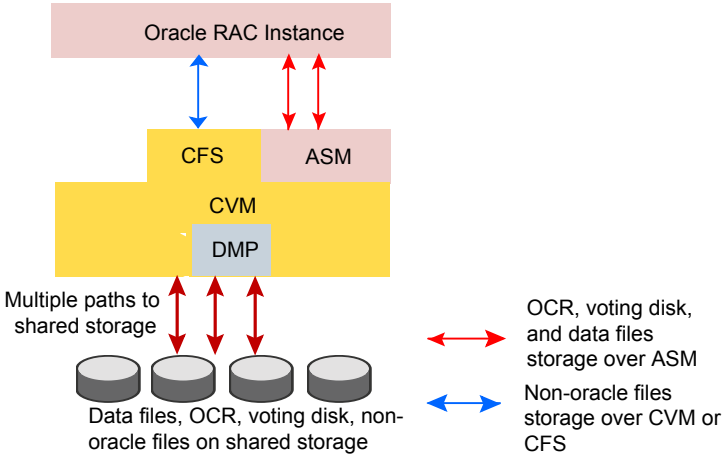



Figure F-2 illustrates the following supported configuration for Oracle RAC:

- ASM disk groups configured over CVM volumes
- Oracle Clusterware and database binaries stored locally
- Oracle database files stored on ASM configured over CVM
  The Oracle databases are managed by Oracle Clusterware.
- Oracle Cluster Registry and voting disk stored on ASM over CVM volumes

**Figure F-2** Supported Oracle RAC configuration




# Configuring ASM in SF Oracle RAC environments

Before you configure ASM, review the planning guidelines for ASM:

---

**Note:** Make sure that you have installed SF Oracle RAC and Oracle RAC before configuring ASM.

---

The ASM home directory is the same as the Oracle Grid Infrastructure home directory (GRID_HOME).

**To configure ASM in SF Oracle RAC installations**

1. Create the required database storage for ASM by creating CVM disk groups and volumes for use as ASM disks.

   See "Creating database storage for ASM" on page 435.

2. Create ASM disk groups and instances.

   See "Creating ASM disk groups and instances" on page 435.

3. Verify the ASM setup.

   See "Verifying the ASM setup" on page 436.

4 Create the Oracle database. For instructions, see the Oracle RAC documentation.

5 Configure VCS service groups for the Oracle database.

See "Configuring VCS service groups for database instances on ASM" on page 437.

## Creating database storage for ASM

This step creates the database storage on ASM using CVM volumes. To create the storage for Oracle databases on ASM, first create the required CVM disk groups and volumes. Then, use these CVM volumes to create ASM disk groups for storing the database files.

**To create database storage on ASM**

1 Log in as the root user to the CVM master:

To determine the CVM master:

```
# vxdctl -c mode
```

2 Initialize the disks as VxVM disks:

```
# vxdisksetup -i sdx
```

3 Create the CVM disk group and volume:

```
# vxdg -s init ora_asm_dg sdx

# vxassist -g ora_asm_dg make ora_asm_vol 2000M
```

4 Set the permissions for the Oracle user on the volume:

```
# vxedit -g ora_asm_dg  \
set group=dba user=oracle mode=660 ora_asm_vol
```

Configure ASM using either ASMCA or OEM.

## Creating ASM disk groups and instances

You can create ASM disk groups using ASM Configuration Assistant (ASMCA) or manually.

The following tasks are performed:

- The ASM disk group and instance is created.

- The ASM instance is started and the disk group is mounted on all nodes in the cluster. The default ASM instance name is +ASM*n* where n is the instance number depending on the number of ASM instances.

  **Note:** For ASM instances that use a pfile, if you restart a node, make sure that the underlying volumes are available before the ASM disk group is mounted. Then, update the ASM init.ora parameter with the full path that contains the VxVM volumes created for ASM.

  For ASM instances that use an spfile, the parameter is updated automatically by the Oracle Configuration Assistant.

  In either case, dependencies must be managed manually. Veritas recommends the use of the ASMDG agent for easier management of ASM disk groups.

SF Oracle RAC requires the following settings when you create ASM disk groups and instances:

| | |
|---|---|
| Type of parameter file for the ASM instance | server parameter file (spfile) |
| Disk Discovery Path | Enter the full path that contains the VxVM volumes created for ASM, for example `/dev/vx/rdsk/ora_asm_dg/ora_asm_vol`. |
| Redundancy | Select the option **External**. Mirroring is performed by CVM. |
| Select member disks | Select the VxVM disks you want to use; You may need to select the **Force** checkbox next to the ASM disks you want to use if the disk group creation fails. |

For detailed instructions, see the Oracle documentation.

# Verifying the ASM setup

Verify that the database services for ASM are up and running after the installation.

**To verify the ASM installation**

1 Change to the Oracle Grid Infrastructure home directory:

```
# cd $GRID_HOME/bin
```

2 Verify the status of ASM on all the nodes in the cluster:

```
# ./srvctl status asm
ASM instance +ASM1 is running on node sys1
```

The sample output shows that there is one ASM instance running on the local node.

## Configuring VCS service groups for database instances on ASM

This section describes how to configure the Oracle service group using the CLI for databases on ASM.

For sample service group illustration:

See "sfrac05_main.cf file" on page 400.

For a sample configuration file describing the configuration, see the file `sfrac05_main.cf` in the directory `/etc/VRTSvcs/conf/sample_rac/`.

The following procedure assumes that you have created the database.

**To configure the Oracle service group using the CLI**

1 Change the cluster configuration to read-write mode:

```
# haconf -makerw
```

2 Add the service group to the VCS configuration:

```
# hagrp -add dbgrp
```

3 Modify the attributes of the service group:

```
# hagrp -modify dbgrp Parallel 1

# hagrp -modify dbgrp SystemList sys1 0  sys2 1

# hagrp -modify dbgrp AutoStartList sys1 sys2
```

4 Add the CVMVolDg resource for the service group:

```
# hares -add oradata_voldg CVMVolDg dbgrp
```

**5** Modify the attributes of the CVMVolDg resource for the service group:

```
# hares -modify oradata_voldg CVMDiskGroup ora_asm_dg
# hares -modify oradata_voldg CVMActivation sw
# hares -modify oradata_voldg CVMVolume ora_asm_vol
```

**6** Add the ASMDG resource for the service group:

```
# hares -add asmdg ASMDG dbgrp
```

**7** Modify the attributes of the ASMDG resource for the service group.

**Note:** The $ASM_HOME variable refers to the ASM home directory.

```
# hares -modify asmdg DiskGroups ASM_RAC_DG
# hares -modify asmdg Home "$ASM_HOME"
# hares -local asmdg  Sid
# hares -modify asmdg Sid "+ASM1" -sys sys1
# hares -modify asmdg Sid "+ASM2" -sys sys2
# hares -modify asmdg Owner grid
```

**8** Add the Oracle RAC database instance to the service group:

```
# hares -add asmdb Oracle dbgrp
```

**9** Modify the attributes of the Oracle resource for the service group:

```
# hares -modify asmdb Owner oracle
# hares -local asmdb Sid
# hares -modify asmdb Sid oradb1 -sys sys1
# hares -modify asmdb Sid oradb2 -sys sys2
# hares -modify asmdb Home "$ORACLE_HOME"
# hares -modify asmdb StartUpOpt SRVCTLSTART
# hares -modify asmdb ShutDownOpt SRVCTLSTOP
```

**10** Set the dependencies between the ASMDG resource and the CVMVolDg resource for the Oracle service group:

```
# hares -link asmdg oradata_voldg
```

11 Set the dependencies between the Oracle resource and the ASMDG resource for the Oracle service group:

```
# hares -link asmdb asmdg
```

12 Create an online local firm dependency between the dbgrp service group and the cvm service group:

```
# hagrp -link dbgrp cvm online local firm
```

13 Enable the Oracle service group:

```
# hagrp -enableresources dbgrp
```

14 Change the cluster configuration to the read-only mode:

```
# haconf -dump -makero
```

15 Bring the Oracle service group online on all the nodes:

```
# hagrp -online dbgrp -any
```

# Configuring Flex ASM in SF Oracle RAC environments

Make sure that you have installed SF Oracle RAC before configuring ASM.

**To configure Flex ASM in SF Oracle RAC environments**

1 Create the required storage for Flex ASM by creating CVM disk groups and volumes for use as ASM disks.

See "Creating storage for Flex ASM" on page 440.

2 Verify the Flex ASM setup.

See "Verifying the Flex ASM setup" on page 440.

3 Create the Oracle database. For instructions, see the Oracle RAC documentation.

4 Configure VCS service groups for the Oracle database.

See "Configuring VCS service groups for database instances on ASM for Flex ASM cluster" on page 441.

# Creating storage for Flex ASM

This step creates the storage for Flex ASM using CVM volumes.

To create the storage for Flex ASM, first create the required CVM disk groups and volumes. Then, use these CVM volumes to create ASM storage for OCR and voting disks.

---

**Note:** Keep the OCR and voting disk information on ASM disk groups separate from the ASM disk groups that contain the Oracle database.

---

**To create storage for Flex ASM**

1. Log in as the root user to the CVM master node.

   To determine the CVM master:

   ```
   # vxdctl -c mode
   ```

2. Initialize the disks as VxVM disks.

   ```
   # vxdisksetup -i sdx
   ```

3. Create the CVM disk group and volume.

   ```
   # vxdg -s init ocrasmdg sdx
   # vxassist -g ora_asm_dg make ora_asm_vol 10000M
   ```

4. Set the permissions for the Oracle user on the volume

   ```
   # vxedit -g ocrasmdg \
   set group=oinstall user=grid mode=660 ocrasmvol
   ```

Use the CVM disk groups and volumes as storage for OCR and voting disk information.

For detailed instructions, see the Oracle documentation.

# Verifying the Flex ASM setup

Verify that the Flex ASM cluster is set up.

**To verify the Flex ASM setup**

1 Change to the Oracle Grid Infrastructure home directory.

```
# cd $GRID_HOME/bin
```

2 Verify the cluster mode on any node of the cluster.

```
# ./asmcmd showclustermode
ASM cluster : Flex mode enabled
```

The sample output indicates that Flex ASM cluster is set up.

## Configuring VCS service groups for database instances on ASM for Flex ASM cluster

This section describes how to configure the Oracle service group using the CLI for databases on ASM for a Flex ASM cluster.

For sample service group illustration:

See "sfrac17_main.cf file" on page 413.

For a sample configuration file describing the configuration, see the file `sfrac17_main.cf` in the directory `/etc/VRTSvcs/conf/sample_rac/`.

The following procedure assumes that you have created the database.

**To configure VCS service groups for database instances on ASM for Flex ASM cluster**

1 Change the cluster configuration to read-write mode.

```
# haconf -makerw
```

2 Add the service group to the VCS configuration.

```
# hagrp -add oraasm_grp
```

3 Modify the attributes of the service group.

```
# hagrp -modify oraasm_grp Parallel 1
# hagrp -modify oraasm_grp SystemList \
sys01 0 sys02 1 sys03 2 sys04 3
# hagrp -modify oraasm_grp AutoStartList \
sys01 sys02 sys03 sys04
```

**4** Add the CVMVolDg resource for the service group.

```
# hares -add asmdata_voldg CVMVolDg oraasm_grp
```

**5** Modify the attributes of the CVMVolDg resource for the service group.

```
# hares -modify asmdata_voldg CVMDiskGroup asmdg
# hares -modify asmdata_voldg CVMActivation sw
# hares -modify asmdata_voldg CVMVolume asmvol
```

**6** Add the ASMDG resource for the service group.

```
# hares -add asmdg ASMDG oraasm_grp
```

**7** Modify the attributes of the ASMDG resource for the service group.

---

**Note:** The $GRID_HOME variable refers to the Grid home directory.

---

```
# hares -modify asmdg DiskGroups ASMDG
# hares -modify asmdg Home "$GRID_HOME"
# hares -local asmdg Sid
# hares -modify asmdg Sid "+ASM1" -sys sys01
# hares -modify asmdg Sid "+ASM2" -sys sys02
# hares -modify asmdg Sid "+ASM3" -sys sys03
# hares -modify asmdg Sid "+ASM4" -sys sys04
# hares -modify asmdg Owner grid
```

**8** Modify ToleranceLimit to 5 for ASMDG resource.

```
# hares -override asmdg ToleranceLimit
# hares -modify asmdg ToleranceLimit 5
```

**9** Set the dependencies between the ASMDG resource and the CVMVolDg resource for the Oracle service group.

```
# hares -link asmdg asmdata_voldg
```

**10** Create an online local firm dependency between the oraasm_grp service group and the cvm service group.

```
# hagrp -link oraasm_grp cvm online local firm
```

**11** Enable the Oracle ASM service group.

```
# hagrp -enableresources oraasm_grp
```

**12** Change the cluster configuration to the read-only mode.

```
# haconf -dump -makero
```

**13** Bring the Oracle ASM service group online on all the nodes.

```
# hagrp -online oraasm_grp -any
```

**To configure Oracle Database Service group for Flex ASM setup using CLI**

**1** Change the cluster configuration to read-write mode.

```
# haconf -makerw
```

**2** Add the service group to the VCS configuration.

```
# hagrp -add oradb_grp
```

**3** Modify the attributes of the service group.

```
# hagrp -modify oradb_grp Parallel 1
# hagrp -modify oradb_grp SystemList \
sys01 0 sys02 1 sys03 2 sys04 3
# hagrp -modify oradb_grp AutoStartList \
sys01 sys02 sys03 sys04
```

**4** Add the Oracle resource for the service group.

```
# hares -add asmdb Oracle oradb_grp
```

5 Modify the attributes of the ORACLE resource for the service group.

---

**Note:** The $ORACLE_HOME variable refers to the Oracle home directory.

---

```
# hares -modify asmdb Home "$ORACLE_HOME"
# hares -local asmdb Sid
# hares -modify asmdb Sid "racdb1" -sys sys01
# hares -modify asmdb Sid "racdb2" -sys sys02
# hares -modify asmdb Sid "racdb3" -sys sys03
# hares -modify asmdb Sid "racdb4" -sys sys04
# hares -modify asmdb Owner oracle
# hares -modify asmdb StartUpOpt SRVCTLSTART
# hares -modify asmdb ShutDownOpt SRVCTLSTOP
```

6 Create an online local firm dependency between the oraasm_grp service group and the cvm service group.

```
# hagrp -link oradb_grp cvm online local firm
```

7 Create an online global form dependency between oradb_grp service group and the orasm_grp service group.

```
# hagrp -link oradb_grp oraasm_grp online global firm
```

8 Enable the Oracle Database service group.

```
# hagrp -enableresources oradb_grp
```

9 Change the cluster configuration to the read-only mode.

```
# haconf -dump -makero
```

10 Bring the Oracle Database service group online on all the nodes.

```
# hagrp -online oradb_grp -any
```

Appendix G

# Creating a test database

This appendix includes the following topics:


- About creating a test database
- Creating a database for Oracle


## About creating a test database

A test database can be created and used for both testing and troubleshooting purposes.

The following optional procedures describe the methods for creating a test database.

## Creating a database for Oracle

Before you begin to create the database, ensure that the following prerequisites are met:

- The Clusterware/Grid daemons must be running.
  To verify the status of Oracle Clusterware, type the following command:

  ```
  # $GRID_HOME/bin/crsctl stat res -t
  ```

- All private IP addresses on each node must be up.
  Use the `ping` command to verify that all private IP addresses on each node are up.

Refer to your Oracle documentation for instructions on how to install the Oracle database.

You can create the database on one of the following types of storage:

- Cluster File System (CFS)

See "Creating the database storage on CFS" on page 446.

- ASM
  See "Creating database storage for ASM" on page 435.

## Creating the database storage on CFS

You can configure the database storage on Veritas cluster file system. The following procedure provides the steps to create the file system.

---

**Note:** SmartIO provides application templates to optimize caching for databases running on VxFS file systems. For more information, see the *Veritas InfoScale SmartIO for Solid State Drives Solutions Guide*, topic "About SmartIO caching for Oracle databases on VxFS file systems."

---

**To create the database storage on CFS**

1. Create a disk group (for example, oradatadg):

   ```
   # vxdg -s init oradatadg Disk_1
   ```

2. Create a single shared volume (for example, oradatavol) that is large enough to contain a file system for all tablespaces.

   For example, assuming 6.8 GB is required for database storage, type:

   ```
   # vxassist -g oradatadg make oradatavol 6800M
   ```

3. Start the volume in the disk group:

   ```
   # vxvol -g oradatadg startall
   ```

4. Create a VxFS file system in this volume. From one node, type:

   ```
   # mkfs -t vxfs /dev/vx/rdsk/oradatadg/oradatavol
   ```

5. Create a mount point for the shared file system:

   ```
   # mkdir /oradata
   ```

6. From the same node, mount the file system:

   ```
   # mount -t vxfs -o cluster /dev/vx/dsk/oradatadg/oradatavol \
   /oradata
   ```

7 Set the "Oracle"user as the owner of the file system, and set "755" as the permissions:

```
# chown oracle:oinstall /oradata
# chmod 755 /oradata
```

8 On the other node(s), complete step 5 and step 6.

9 Create the database using the Oracle documentation.

Appendix H

# High availability agent information

This appendix includes the following topics:


- About agents
- CVMCluster agent
- CVMVxconfigd agent
- CVMVolDg agent
- CFSMount agent
- CFSfsckd agent
- CSSD agent
- VCS agents for Oracle
- CRSResource 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 SF Oracle RAC agent are described in this appendix.

## VCS agents included within SF Oracle RAC

SF Oracle RAC includes the following VCS agents:

- CVMCluster agent
- CVMVxconfigd agent
- CVMVolDg agent
- CFSMount agent

An SF Oracle RAC 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*.

## VCS agents for Oracle included within SF Oracle RAC

SF Oracle RAC includes the following VCS agents for Oracle:

- Oracle agent
  The Oracle agent monitors the database processes.
- Netlsnr agent
  The Netlsnr agent brings the listener services online, monitors their status, and takes them offline.

- PrivNIC agent
  The PrivNIC agent provides high availability to a single private IP address across LLT Ethernet interfaces for a system.
- MultiPrivNIC agent
  The MultiPrivNIC agent provides high availability to multiple private IP addresses across LLT Ethernet interfaces for a system.
- CSSD agent
  The CSSD (Cluster Synchronization Services daemon) agent provides the resources to monitor Oracle Clusterware. The agent ensures that the dependency of cssd on the OCR and the VOTE resources and the PrivNIC (optional) resource are satisfied.
- ASMDG agent
  The ASMDG agent mounts and unmounts the ASM disk groups onto an ASM instance.

See the *Cluster Server Agent for Oracle Installation and Configuration Guide*

# 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 H-1 describes the entry points used by the CVMCluster agent.

**Table H-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 H-2 describes the user-modifiable attributes of the CVMCluster resource type.

**Table H-2** CVMCluster agent attributes

| Attribute | Description |
|---|---|
| CVMClustName | Name of the cluster.<br>■ Type and dimension: string-scalar |
| CVMNodeAddr | List of host names and IP addresses.<br>■ 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.<br>■ Type and dimension: string-association |
| CVMTransport | Specifies the cluster messaging mechanism.<br>■ Type and dimension: string-scalar<br>■ Default = gab<br>**Note:** Do not change this value. |
| PortConfigd | The port number that is used by CVM for vxconfigd-level communication.<br>■ Type and dimension: integer-scalar |
| PortKmsgd | The port number that is used by CVM for kernel-level communication.<br>■ Type and dimension: integer-scalar |
| CVMTimeout | Timeout in seconds used for CVM cluster reconfiguration.<br>■ Type and dimension: integer-scalar<br>■ 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.<br>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.<br>■ Type and dimension: string<br>■ Default = 0 |

**Table H-2** CVMCluster agent attributes *(continued)*

| Attribute | Description |
|---|---|
| CVMDGSubClust | Enables or disables the application isolation feature.<br>The values for the attribute are as follows:<br>■ 1<br>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.<br>■ 0<br>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 SF Oracle RAC 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 SF Oracle RAC 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 H-3 describes the entry points for the CVMVxconfigd agent.

**Table H-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. |

**Table H-3** CVMVxconfigd entry points *(continued)*

| Entry Point | Description |
|---|---|
| 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 H-4 describes the modifiable attributes of the CVMVxconfigd resource type.

**Table H-4** CVMVxconfigd agent attribute

| Attribute | Description |
|---|---|
| CVMVxconfigdArgs | List of the arguments that are sent to the `online` entry point.<br>Veritas recommends always specifying the `syslog` option.<br>■ Type and dimension: keylist |

**Table H-4** CVMVxconfigd agent attribute *(continued)*

| Attribute | Description |
|---|---|
| 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.<br>This attribute includes the following keys:<br>■ Mode: Define this attribute to enable or disable intelligent resource monitoring.<br>Valid values are as follows:<br>■ 0—Does not perform intelligent resource monitoring<br>■ 2—Performs intelligent resource monitoring for online resources and performs poll-based monitoring for offline resources<br>Default: 0<br>■ MonitorFreq: This key value specifies the frequency at which the agent invokes the monitor agent function. The value of this key is an integer.<br>Default: 1<br>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.<br>After the resource registers with the AMF kernel driver, the agent calls the monitor agent function as follows:<br>■ After every (MonitorFreq x MonitorInterval) number of seconds for online resources<br>■ After every (MonitorFreq x OfflineMonitorInterval) number of seconds for offline resources<br>■ 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.<br>Default: 3.<br>■ Type and dimension: integer-association<br>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 H-5 describes the entry points used by the CVMVolDg agent.

**Table H-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.<br>Starts all volumes and volume sets in the shared disk group specified by the CVMVolume attribute.<br>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.<br>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.<br>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.<br>**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 H-6 describes the user-modifiable attributes of the CVMVolDg resource type.

**Table H-6** CVMVolDg agent attributes

| Attribute | Description |
|---|---|
| CVMDiskGroup (required) | Shared disk group name.<br>■ Type and dimension: string-scalar |

**Table H-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.<br>■ Type and dimension: string-keylist |
| CVMActivation (required) | Activation mode for the disk group.<br>■ Type and dimension: string-scalar<br>■ Default = `sw (shared-write)`<br>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.<br>■ 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.<br>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.<br>The value 0 indicates that the agent will not deport the shared disk group when the CVMVolDg resource is taken offline.<br>■ Type and dimension: integer-scalar<br>■ Default = 0<br>**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.<br>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.<br>The deport operation fails if the shared disk group contains open volumes. |

**Table H-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.<br>■ Type and dimension: integer-scalar<br>■ 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 ora_voldg (
        Critical = 0
```

```
CVMDiskGroup = oradatadg
CVMVolume = { oradata1, oradata2 }
CVMActivation = sw
 )
```

# 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 H-7 provides the entry points for the CFSMount agent.

**Table H-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. |
| 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. |

**Table H-7** CFSMount agent entry points *(continued)*

| Entry Point | Description |
|---|---|
| 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 H-8 lists user-modifiable attributes of the CFSMount Agent resource type.

**Table H-8** CFSMount Agent attributes

| Attribute | Description |
|---|---|
| MountPoint | Directory for the mount point.<br>■ Type and dimension: string-scalar |
| BlockDevice | Block device for the mount point.<br>■ 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.<br>■ Type and dimension: string-keylist |

**Table H-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.<br>This attribute includes the following keys:<br>■ Mode: Define this attribute to enable or disable intelligent resource monitoring.<br>Valid values are as follows:<br>■ 0—Does not perform intelligent resource monitoring<br>■ 1—Performs intelligent resource monitoring for offline resources and performs poll-based monitoring for online resources<br>■ 2—Performs intelligent resource monitoring for online resources and performs poll-based monitoring for offline resources<br>■ 3—Performs intelligent resource monitoring for both online and for offline resources<br>Default: 0<br>■ MonitorFreq: This key value specifies the frequency at which the agent invokes the monitor agent function. The value of this key is an integer.<br>Default: 1<br>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.<br>After the resource registers with the AMF kernel driver, the agent calls the monitor agent function as follows:<br>■ After every (MonitorFreq x MonitorInterval) number of seconds for online resources<br>■ After every (MonitorFreq x OfflineMonitorInterval) number of seconds for offline resources<br>■ 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.<br>Default: 3.<br>■ Type and dimension: integer-association |

**Table H-8** CFSMount Agent attributes *(continued)*

| Attribute | Description |
|---|---|
| MountOpt<br>(optional) | Options for the mount command. To create a valid MountOpt attribute string:<br>■ Use the VxFS type-specific options only.<br>■ Do not use the -o flag to specify the VxFS-specific options.<br>■ Do not use the `-t` vxfs file system type option.<br>■ Be aware the cluster option is not required.<br>■ Specify options in comma-separated list:<br>`ro`<br>`ro,cluster`<br>`blkclear,mincache=closesync`<br>■ 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.<br>■ 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;
        )
```

# 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 H-9 describes the CFSfsckd agent entry points.

**Table H-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 H-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 H-10 lists user-modifiable attributes of the CFSfsckd Agent resource type.

**Table H-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.<br>This attribute includes the following keys:<br>■ Mode: Define this attribute to enable or disable intelligent resource monitoring.<br>Valid values are as follows:<br>■ 0—Does not perform intelligent resource monitoring<br>■ 1—Performs intelligent resource monitoring for offline resources and performs poll-based monitoring for online resources<br>■ 2—Performs intelligent resource monitoring for online resources and performs poll-based monitoring for offline resources<br>■ 3—Performs intelligent resource monitoring for both online and for offline resources<br>Default: 0<br>■ MonitorFreq: This key value specifies the frequency at which the agent invokes the monitor agent function. The value of this key is an integer.<br>Default: 1<br>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.<br>After the resource registers with the AMF kernel driver, the agent calls the monitor agent function as follows:<br>■ After every (MonitorFreq x MonitorInterval) number of seconds for online resources<br>■ After every (MonitorFreq x OfflineMonitorInterval) number of seconds for offline resources<br>■ 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.<br>Default: 3.<br>■ Type and dimension: integer-association |

# 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 (
)
```

# CSSD agent

The CSSD agent starts, stops, and monitors Oracle Clusterware/Grid Infrastructure. It ensures that the OCR, the voting disk and the private IP address resources required by Oracle Clusterware/Grid Infrastructure are online before Oracle Clusterware/Grid Infrastructure starts. For this purpose, the cssd resource must be configured as a parent of the resources that manage the OCR, the voting disk, and the private IP address used by Oracle Clusterware/Grid Infrastructure. Using the CSSD agent in SF Oracle RAC installations ensures adequate handling of inter-dependencies and thus prevents the premature startup of Oracle Clusterware/Grid Infrastructure, which causes cluster failures.

The automatic startup of Oracle Clusterware/Grid Infrastructure must be disabled when the system starts. This prevents the premature startup of Oracle Clusterware/Grid Infrastructure during system startup. Thus, VCS ensures that Oracle Clusterware/Grid Infrastructure is started using the CSSD agent only when all the dependent resources of the cssd resource configured under VCS are online.

During system shutdown, the agent stops Oracle Clusterware/Grid Infrastructure before the OCR and voting disk resources are taken offline by VCS. This ensures that Oracle Clusterware/Grid Infrastructure does not panic the nodes in the cluster due to unavailability of the required resources.

The CSSD agent supports IMF monitoring. By default, IMF monitoring is enabled with a monitoring value of 3.

---

**Note:** It is mandatory to use CSSD agent in SF Oracle RAC installations. You must configure the CSSD agent after installing Oracle Clusterware/Grid Infrastructure.

---

## Functions of the CSSD agent

Table H-11 describes the functions of the CSSD agent.

**Table H-11** CSSD agent functions

| Function | Description |
|---|---|
| Online | Starts Oracle Clusterware/Grid Infrastructure. |
| Offline | Stops Oracle Clusterware/Grid Infrastructure.<br>**Note:** The CSSD agent detects an intentional offline of the CSSD resource only when the command `crsctl stop crs [-f]` is used to stop Oracle Clusterware/Grid Infrastructure. |
| Monitor | Checks the status of Oracle Clusterware/Grid Infrastructure.<br>The Oracle Clusterware/Grid Infrastructure processes monitored are as follows:<br>■ `cssd`<br>■ `crsd`<br>■ `evmd`<br>■ `ohasd`<br>**Note:** The `ohasd` daemon is not monitored by IMF.<br>The agent restarts `cssd`, `crsd`, and `evmd` processes if the `RestartDaemons` attribute is set to 1.<br>**Note:** If you want to temporarily disable monitoring of Oracle Clusterware/Grid Infrastructure processes, ensure that you first stop the CSSD agent:<br>See "Disabling monitoring of Oracle Grid Infrastructure processes temporarily " on page 471. |
| Clean | Forcibly stops Oracle Clusterware/Grid Infrastructure on nodes running Oracle RAC using the following command:<br>`# crsctl stop crs -f` |
| imf_init | Initializes the agent to interface with the AMF kernel driver, which is the IMF notification module for CSSD agent. This function runs when the agent starts up. |
| 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). |

# Attributes of the CSSD agent

Table H-12 lists the attributes of the CSSD agent:

**Table H-12** Attributes for CSSD resource

| Attribute Name | Dimension | Description |
|---|---|---|
| Critical | Boolean | Optional<br>A Boolean value 0 or 1.<br>1: Indicates that the resource is critical.<br>0: Indicates that the resource is non-critical. |
| RestartDaemons | Boolean | Optional<br>A Boolean value 0 or 1.<br>The default value is set to 1.<br>1: Indicates that the Oracle Clusterware/Grid Infrastructure processes `cssd`, `crsd`, and `evmd` are restarted if the status of these processes is unhealthy.<br>0: Indicates that the Oracle Clusterware/Grid Infrastructure processes `cssd`, `crsd`, and `evmd` are not restarted if the status of these processes is unhealthy. |

**Table H-12** Attributes for CSSD resource *(continued)*

| Attribute Name | Dimension | Description |
|---|---|---|
| IntentionalOffline | Boolean | Defines how VCS responds when Oracle Clusterware/Grid Infrastructure is intentionally stopped outside of VCS control.<br>If you stop Oracle Clusterware/Grid Infrastructure outside of VCS control, the agent behavior is as follows:<br>■ 0—The CSSD agent registers a fault and initiates the failover of the service group.<br>■ 1—The CSSD agent takes the CSSD resource offline. |
| CRSHOME | string-scalar | Optional<br>The agent uses this information to locate clusterware process binaries.<br>If the value is not provided, the agent reads the information from the Oracle configuration file. |
| IMF | integer-association | Optional<br>Resource-type level attribute that determines whether the CSSD agent must perform intelligent resource monitoring. You can also override the value of this attribute at resource-level.<br>By default, IMF monitoring is enabled with a monitoring value of 3. To disable monitoring, set the value to 0. |

## States of the CSSD agent

Table H-13 describes the states of the CSSD agent.

**Table H-13** CSSD agent states

| State | Description |
|---|---|
| Online | Indicates that Oracle Clusterware/Grid Infrastructure is running. |
| Offline | Indicates that Oracle Clusterware/Grid Infrastructure is not running. |
| Unknown | Indicates the inability to determine the state of the resource due to incorrect attribute settings or other configuration issues. |

## Disabling monitoring of Oracle Grid Infrastructure processes temporarily

If you plan to temporarily disable the monitoring of Oracle Grid Infrastructure processes—`cssd`, `crsd`, `evmd`, `ohasd`—you need to first stop the CSSD agent and then manually stop the Oracle Grid Infrastructure processes.

For example, if you want to replace the `evmd` binary, you need to perform the following steps:

1. Stop the CSSD agent:

   ```
   # haagent -stop CSSD -force -sys node_name
   ```

   **Note:** If you fail to stop the CSSD agent, it attempts to restart Oracle Grid Infrastructure processes, which were taken offline manually outside of VCS.

2. Stop the Oracle Grid Infrastructure process, for example, `evmd`:

   ```
   # crsctl stop resource ora.evmd -init -f
   ```

3. Replace the `evmd` binary. For instructions, see the Oracle documentation.
4. Start the `evmd` resource:

   ```
   # crsctl start resource ora.evmd -init
   ```

5. Verify that the following command reports the status of all daemons as "healthy":

   ```
   # crsctl check crs
   ```

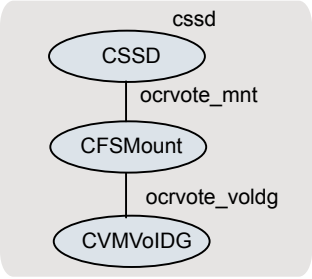
6. Start the CSSD agent:

   ```
   # haagent -start CSSD -sys node_name
   ```

# Sample service group configurations with the CSSD agent

Figure H-1 illustrates a basic service group configuration with the CSSD agent.

**Figure H-1** Basic service group configuration with the CSSD agent




In this basic configuration, the OCR and voting disk volumes/mount points and IP addresses are configured under the CSSD resource. This ensures that these resources are online before the CSSD agent starts Oracle Clusterware.

---

**Note:** Depending on whether the Oracle database is started by Oracle Clusterware or by the VCS agent for Oracle, you must configure the Oracle database mounts such that they are online before the database starts.

---

For more sample configurations:

See "Sample configuration files" on page 395.

# Type definition of the CSSD resource

The following extract shows the type definition of the CSSD resource in the `CSSD.cf` file.

```
type CSSD (
   static int IMF{} = { Mode=3, MonitorFreq=1, RegisterRetryLimit=3 }
   static str IMFRegList[] = { CRSHOME }
   static str ArgList[] = { State, IState, CRSHOME, RestartDaemons }
   static int ContainerOpts{} = { RunInContainer=1, PassCInfo=0 }
   static boolean IntentionalOffline = 1
   static int OfflineWaitLimit = 3
   static int OnlineWaitLimit = 5
   str CRSHOME
   boolean RestartDaemons = 1
)
```

## Sample configuration of the CSSD resource

The following extract from the `main.cf` file illustrates a sample CSSD agent configuration:

```
CSSD cssd (
    Critical = 0
    CRSHOME="/u01/app/grid/product/11.2.0/gridhome"
    )
```

# VCS agents for Oracle

The VCS agents for Oracle include the following agents that work together to make Oracle highly available:

- The Oracle agent monitors the database processes.
  See "Oracle agent functions" on page 473.
  See "Resource type definition for the Oracle agent" on page 481.
- The Netlsnr agent monitors the listener process.
  See "Netlsnr agent functions" on page 492.
- The ASMDG agent monitors the Oracle ASM disk groups.
  See "ASMDG agent functions" on page 498.
  See "Resource type definition for the ASMDG agent" on page 499.

Refer to the *Cluster Server Agent for Oracle Installation and Configuration Guide* for more details on the agent functions and the resource types.

## Oracle agent functions

The Oracle agent monitors the database processes.

Table H-14 lists the Oracle agent functions.

**Table H-14** Oracle agent functions

| Agent operation | Description |
|---|---|
| Online | Starts the Oracle database.<br><br>The agent uses the default startup option STARTUP_FORCE. For RAC clusters, you must manually change the value of the StartUpOpt attribute to SRVCTLSTART.<br><br>If you set the option to SRVCTLSTART, the agent uses the following `srvctl` command to start the Oracle database:<br><br>`srvctl start database -d` *`database_name`* |
| Offline | Stops the Oracle database.<br><br>The agent uses the default shutdown option IMMEDIATE. For RAC clusters, you must manually change the value of the ShutDownOpt attribute to SRVCTLSTOP.<br><br>If you set the option to SRVCTLSTOP, the agent uses the following `srvctl` command to stop the Oracle database:<br><br>`srvctl stop database -d` *`database_name`* |
| Monitor | Verifies the status of the Oracle processes. The Oracle agent provides two levels of monitoring: basic and detail.<br><br>See "Monitor options for the Oracle agent in traditional database and container database" on page 477. |
| oracle_imf_init | Initializes the agent to interface with the AMF kernel driver, which is the IMF notification module for Oracle agent. This function runs when the agent starts up. |
| oracle_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. |
| oracle_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). |

**Table H-14** Oracle agent functions *(continued)*

| Agent operation | Description |
|---|---|
| Clean | Forcibly stops the Oracle database.<br>If you set the shutdown option to SRVCTLSTOP, the agent uses the following `srvctl` command:<br>`srvctl stop database -d` *`database_name`*<br>If the process does not respond to the `srvctl` command, then the agent does the following:<br>■ Scans the process table for the processes that are associated with the configured instance<br>■ Kills the processes that are associated with the configured instance |
| Info | Provides the static and dynamic information about the state of the database.<br>**Note:** This attribute is not applicable PDB resources.<br>See "Info entry point for Cluster Server agent for Oracle" on page 479. |
| Action | Performs the predefined actions on a resource.<br>See "Action entry point for Cluster Server agent for Oracle" on page 480. |

## Startup and shutdown options for the pluggable database (PDB)

You can specify Startup and Shutdown options for the PDB resources that are configured.

Table H-15 lists the startup options that the agent supports in the PDB.

**Table H-15** Startup options for PDB

| Options | Description |
|---|---|
| STARTUP | Runs the command `STARTUP PLUGGABLE DATABASE <pdbname>;` if the pdbname is configured. |
| STARTUP_FORCE | Runs the command `STARTUP PLUGGABLE DATABASE <pdbname> FORCE;` if the pdbname is configured. |

**Table H-15** Startup options for PDB *(continued)*

| Options | Description |
|---|---|
| RESTRICTED | Runs the command `STARTUP PLUGGABLE DATABASE <pdbname> RESTRICT;` if the pdbname is configured. |
| OPEN_RO | Runs the command `STARTUP PLUGGABLE DATABASE <pdbname> OPEN READ ONLY;` if the pdbname is configured. |

During offline, shutdown of individual PDB will be attempted by running `shutdown immediate`.

See "Recommended startup modes for pluggable database (PDB) based on container database (CDB) startup modes" on page 476.

## Recommended startup modes for pluggable database (PDB) based on container database (CDB) startup modes

You can specify the startup options for the PDB based on the values configured in the CDB.

Table H-16 lists the shutdown options that the agent supports.

**Table H-16** CDB Startup options

| CDB startup mode | Corresponding PDB startup modes |
|---|---|
| SRVCTLSTART | When the CDB is set with the startup value as SRVCTLSTART then the PDB can be assigned with one of the following values:<br>■ STARTUP<br>■ STARTUP_FORCE<br>■ RESTRICTED<br>■ OPEN_RO |
| SRVCTLSTART_RO | When the CDB is assigned with the SRVCTLSTART_RO startup value then the PDB should be OPEN_RO. |

**Table H-16** CDB Startup options *(continued)*

| CDB startup mode | Corresponding PDB startup modes |
|---|---|
| STARTUP | When the CDB is set with the startup value as STARTUP then the PDB can be assigned with one of the following values:<br>■ STARTUP<br>■ STARTUP_FORCE<br>■ RESTRICTED<br>■ OPEN_RO |
| STARTUP_FORCE | When the CDB is set with the startup value as STARTUP_FORCE then the PDB can be assigned with one of the following values:<br>■ STARTUP<br>■ STARTUP_FORCE<br>■ RESTRICTED<br>■ OPEN_RO |
| RESTRICTED | When the CDB is assigned with the startup value as RESTRICTED then the PDB should be RESTRICTED. |
| RECOVERDB | When the CDB is set with the startup value as RECOVERDB then the PDB can be assigned with one of the following values:<br>■ STARTUP<br>■ STARTUP_FORCE<br>■ RESTRICTED<br>■ OPEN_RO |
| CUSTOM | The custom script should start the PDB as CUSTOM. |

## Monitor options for the Oracle agent in traditional database and container database

The Oracle agent provides two levels of monitoring: basic and detail. By default, the agent does a basic monitoring.

The basic monitoring mode has the following options:

- Process check
- Health check

The MonitorOption attribute of the Oracle resource determines whether the agent must perform basic monitoring in Process check or Health check mode.

Table H-17 describes the basic monitoring options.

**Table H-17** Basic monitoring options

| Option | Description |
|---|---|
| 0<br>(Default) | Process check<br>The agent scans the process table for the `ora_dbw0`, `ora_smon`, `ora_pmon`, `ora_lmon`, and `ora_lgwr` processes to verify that Oracle is running.<br>In this mode, the agent also supports intelligent resource monitoring. |
| 1 | Health check<br>The agent uses the Health Check APIs from Oracle to monitor the SGA and retrieve the information about the instance.<br>If you want to use the Oracle agent's intentional offline functionality, you must enable Health check monitoring.<br>The agent does not support intelligent resource monitoring in this mode. |

Review the following considerations if you want to configure basic monitoring:

- Basic monitoring of Oracle processes is user-specific. As a result, an Oracle instance started under the context of another user cannot be detected as online. For example, if an Oracle instance is started under the user "oraVRT" and the agent is configured for a user "oracle", the agent will not detect the instance started by "oraVRT" as online.
  This could lead to situations where issuing a command to online a resource on a node might online an already running instance on that node (or any other node).
  So, Veritas recommends that instances started outside SF Oracle RAC control be configured with the correct Owner attribute corresponding to the OS user for that instance.

In the detail monitoring mode, the agent performs a transaction on a test table in the database to ensure that Oracle database functions properly. The agent uses this test table for internal purposes. Veritas recommends that you do not perform any other transaction on the test table.

## Monitor for the pluggable database

During monitor, the Oracle agent connects to the CDB and fetches the status of all PDBs that are part of the CDB.

When IntentionalOffline is set for a CDB and PDB resource, and CDB is shut outside VCS then the PDB resource reports as Offline and not FAULTED. To enable this functionality, set the IntentionalOffline attribute to 1 for the PDB resource.

---

**Note:** Pluggable database will be failed over along with the Container database. Individual Pluggable database cannot be failed over using VCS.

---

## Info entry point for Cluster Server agent for Oracle

The supports the Info entry point, which provides static and dynamic information about the state of the database.

To invoke the Info entry point, type the following command:

```
# hares -value resource ResourceInfo [system]\
 [-clus cluster | -localclus]
```

For the above command to be executed successfully, you must perform the following command:

- Execute the resource refreshinfo using the following command:

```
# hares -refreshinfo <res-name> -sys <sys-name>
```

The entry point retrieves the following static information:

- Version
- InstanceNo
- InstanceName
- DatabaseName
- HostName
- StartupTime
- Parallel
- Thread
- InstanceRole

The entry point retrieves the following dynamic information:

- InstanceStatus
- Logins
- OpenMode
- LogMode
- ShutdownPending
- DatabaseStatus
- Shared Pool Percent free
- Buffer Hits Percent

You can add additional attributes by adding sql statements to the file /opt/VRTSagents/ha/bin/Oracle/resinfo.sql. For example:

```
select 'static:HostName:'||host_name from v$instance;
select 'dynamic:ShutdownPending:'||shutdown_pending from
v$instance;
```

The format of the selected record must be as follows:

```
attribute_type:userkey_name:userkey_value
```

The variable *attribute_type* can take the value static and/or dynamic.

## Action entry point for Cluster Server agent for Oracle

The supports the Action entry point, which enables you to perform predefined actions on a resource.

To perform an action on a resource, type the following command:

```
# hares -action res token [-actionargs arg1 ...] \
[-sys system] [-clus cluster]
```

You can also add custom actions for the agent.

For further information, refer to the *Cluster Server Agent Developer's Guide*.

See Table H-19 on page 481. describes the agent's predefined virtual fire drill actions.

Table H-18 describes the agent's predefined actions.

**Table H-18** Predefined agent actions

| Action | Description |
|---|---|
| VRTS_GetInstanceName | Retrieves the name of the configured instance. You can use this option for the Oracle and the Netlsnr resources. |
| VRTS_GetRunningServices | Retrieves the list of processes that the agent monitors. You can use this option for the Oracle and the Netlsnr resources. |
| DBRestrict | Changes the database session to enable the RESTRICTED mode. |
| DBUndoRestrict | Changes the database session to disable the RESTRICTED mode. |
| DBSuspend | Suspends a database. |
| DBResume | Resumes a suspended database. |
| DBTbspBackup | Backs up a tablespace; `actionargs` contains name of the tablespace to be backed up. |

**Note:** VRTS_GetRunningServices, DBRestrict, DBResume, DBSuspend, DBTbspBackup, and DBUndoRestrict action entry points are not supported for Oracle PDB resources.

Table H-19 lists the virtual fire drill actions of the lets you run infrastructure checks and fix specific errors.

**Table H-19** Predefined virtual fire drill actions

| Virtual fire drill action | Description |
|---|---|
| getid<br>(Oracle agent) | Verifies that the Oracle Owner exists on the node. |
| home.vfd<br>(Oracle agent) | Verifies the following:<br>■ ORACLE_HOME is mounted on the node and corresponding entry is in the fstab.<br>If the ORACLE_HOME is not mounted, the action entry point checks if any other resource has already mounted ORACLE_HOME.<br>■ Pfile is provided and it exists on the node.<br>■ Password file from $ORACLE_HOME/dbs/orapw[SID] is present. |
| owner.vfd<br>(Oracle agent) | Verifies the uid and gid of the Oracle Owner attribute. Checks if uid and gid of Owner attribute is the same on the node where the Oracle resource is currently ONLINE. |
| pfile.vfd<br>(Oracle agent) | Checks for the presence of pfile or spfile on the local disk. If both pfile and spfile are not present, the agent function exits. If the Oracle resource is online in the cluster, the agent function logs a message that the spfile must be on the shared storage because the Oracle resource is online. |
| tnsadmin.vfd<br>(Netlsnr agent) | Checks if listener.ora file is present. If the listener.ora file is not present, it checks if ORACLE_HOME is mounted and displays appropriate messages. |

# Resource type definition for the Oracle agent

The Oracle agent of the is represented by the Oracle resource type in SF Oracle RAC.

The following extract shows the type definition of the Oracle resource in the OracleTypes.cf file.

```
type Oracle (
static str AgentDirectory = "/opt/VRTSagents/ha/bin/Oracle"
static keylist SupportedActions = { VRTS_GetInstanceName, VRTS_GetRunningServices,
DBRestrict, DBUndoRestrict, DBResume, DBSuspend, DBTbspBackup, "home.vfd",
"owner.vfd", "getid", "pfile.vfd" }
static str ArgList[] = { Sid, Owner, Home, Pfile, StartUpOpt, ShutDownOpt, DBAUser,
DBAPword,  EnvFile, AutoEndBkup, User, Pword, Table, MonScript,
Encoding, MonitorOption, DBName, ManagedBy }
static int IMF{} = { Mode=3, MonitorFreq=5, RegisterRetryLimit=3 }
static str IMFRegList[] = { Home, Owner, Sid, MonitorOption }
str Sid
str Owner
str Home
str Pfile
str StartUpOpt = STARTUP_FORCE
str ShutDownOpt = IMMEDIATE
str DBName
str ManagedBy = "ADMIN"
str DBAUser
str DBAPword
str EnvFile
boolean AutoEndBkup = 1
str MonScript = "./bin/Oracle/SqlTest.pl"
str User
str Pword
str Table
str Encoding
int MonitorOption = 0
static boolean IntentionalOffline = 0
)
```

## Attribute definitions for the Oracle agent

The Oracle agent attributes are classified as required, optional, and internal.

You must assign values to the following required attributes.

**Table H-20** Required attributes for Oracle agent

| Required attributes | Type and dimension | Definition |
|---|---|---|
| Sid | string-scalar | The variable $ORACLE_SID that represents the Oracle instance. The Sid is considered case-sensitive by the Oracle agent and by the Oracle database server.<br>For a policy managed database, the Sid attribute should be set to Sid prefix. See "About the Sid attribute in a policy managed database" on page 492. |
| Owner | string-scalar | The Oracle user who has privileges to start or stop the database instance.<br>The agent also supports LDAP users as Oracle user. |
| Home | string-scalar | The $ORACLE_HOME path to Oracle binaries and configuration files. For example, you could specify the path as /opt/ora_home.<br>**Note:** Do not append a slash (/) at the end of the path. |

You can configure the following optional attributes if they are necessary for your configuration.

**Table H-21** Optional attributes for Oracle agent

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| DBAUser | string-scalar | The database user who has sysdba privileges to start or stop the database. |
| DBAPword | string-scalar | Encrypted password for DBAUser.<br>Encrypt passwords only when entering them using the command-line. Passwords must be encrypted using the VCS Encrypt Utility (`/opt/VRTSvcs/bin/vcsencrypt`). |

**Table H-21** Optional attributes for Oracle agent *(continued)*

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| StartUpOpt | string-scalar | Startup options for the Oracle instance. This attribute can take the following values for traditional database and container database:<br>■ STARTUP<br>■ STARTUP_FORCE<br>■ RESTRICTED<br>■ RECOVERDB<br>■ SRVCTLSTART<br>■ CUSTOM<br>■ SRVCTLSTART_FORCE<br>Default is STARTUP_FORCE.<br>Startup options for the Oracle instance in pluggable database. This attribute can take the following values:<br>■ STARTUP<br>■ STARTUP_FORCE<br>■ RESTRICTED<br>■ OPEN_RO<br>See "Startup and shutdown options for the pluggable database (PDB)" on page 475. |
| ShutDownOpt | string-scalar | Shut down options for the Oracle instance. This attribute can take the following values for traditional database and container database:<br>■ IMMEDIATE<br>■ TRANSACTIONAL<br>■ SRVCTLSTOP<br>■ CUSTOM<br>■ SRVCTLSTOP_TRANSACT<br>■ SRVCTLSTOP_ABORT<br>■ SRVCTLSTOP_IMMEDIATE<br>Default is IMMEDIATE.<br>Shut down options for the Oracle instance in pluggable database is IMMEDIATE.<br>See "Startup and shutdown options for the pluggable database (PDB)" on page 475. |

**Table H-21** Optional attributes for Oracle agent *(continued)*

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| EnvFile | string-scalar | The full path name of the file that is sourced by the entry point scripts. This file contains the environment variables set by the user for the Oracle database server environment such as LD_LIBRARY_PATH, NLS_DATE_FORMAT, and so on.<br>The syntax for the contents of the file depends on the login shell of Owner. File must be readable by Owner. The file must not contain any prompts for user input. |
| Pfile | string-scalar | The name of the initialization parameter file with the complete path of the startup profile.<br>You can also use the server parameter file. Create a one-line text initialization parameter file that contains only the SPFILE parameter. See the Oracle documentation for more information.<br>**Note:** This attribute is applicable only for traditional and container databases. |
| AutoEndBkup | boolean-scalar | Setting the AutoEndBkup attribute to a non-zero value takes the datafiles in the database out of the backup mode, during Online.<br>Default = 1<br>**Note:** If a node fails during a hot backup of container database or pluggable database, you must set AutoEndBkup attribute of the corresponding CDB resource to 1. When the AutoEndBkup is set to 1 for the CDB, it also ends the backup of both CDB and PDB during Online. |

**Table H-21** Optional attributes for Oracle agent *(continued)*

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| MonitorOption | integer-scalar | Monitor options for the Oracle instance. This attribute can take values 0 or 1.<br>For traditional and container databases:<br>■ 0—Process check monitoring (recommended)<br>■ 1—Health check monitoring<br>For pluggable databases:<br>■ 0—Basic monitoring<br>Default = 0<br>The agent supports intelligent resource monitoring only for traditional and CDBs when the attribute value is set to 0.<br>See "Monitor options for the Oracle agent in traditional database and container database" on page 477.<br>See "Monitor for the pluggable database" on page 479. |

**Table H-21** Optional attributes for Oracle agent *(continued)*

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| IMF | integer-association | |

**Table H-21** Optional attributes for Oracle agent *(continued)*

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| | | This resource-type level attribute determines whether the Oracle agent must perform intelligent resource monitoring. You can also override the value of this attribute at resource-level.<br>This attribute includes the following keys:<br>■ Mode: Define this attribute to enable or disable intelligent resource monitoring.<br>Valid values are as follows:<br>■ 0—Does not perform intelligent resource monitoring<br>■ 1—Performs intelligent resource monitoring for offline resources and performs poll-based monitoring for online resources<br>■ 2—Performs intelligent resource monitoring for online resources and performs poll-based monitoring for offline resources<br>■ 3—Performs intelligent resource monitoring for both online and for offline resources<br>Default: 3<br>■ MonitorFreq: This key value specifies the frequency at which the agent invokes the monitor agent function. The value of this key is an integer.<br>Default: 5<br>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.<br>After the resource registers with the AMF kernel driver, the agent calls the monitor agent function as follows:<br>■ After every (MonitorFreq x MonitorInterval) number of seconds for online resources<br>■ After every (MonitorFreq x OfflineMonitorInterval) number of seconds for offline resources<br>■ RegisterRetryLimit: If you enable intelligent resource monitoring, the agent invokes the oracle_imf_register agent function to register the resource with the AMF kernel driver. The value of the RegisterRetryLimit 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. |

**Table H-21** Optional attributes for Oracle agent *(continued)*

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| | | Default: 3<br>**Note:** IMF is supported only in traditional and container databases. |
| MonScript | string-scalar | Pathname to the script provided for detail monitoring. The default (basic monitoring) is to monitor the database PIDs only.<br>**Note:** Second-level monitoring is disabled if the value of the attribute MonScript is invalid or is set to an empty string.<br>The pathname to the supplied detail monitor script is /opt/VRTSagents/ha/bin/Oracle/SqlTest.pl.<br>MonScript also accepts a pathname relative to /opt/VRTSagents/ha. A relative pathname should start with "./", as in the path ./bin/Oracle/SqlTest.pl. |
| User | string-scalar | Internal database user. Connects to the database for detail monitoring. |
| LevelTwoMonitorFreq | integer-scalar | Specifies the frequency at which the agent for this resource type must perform second-level or detailed monitoring. You can also override the value of this attribute at resource-level.<br>The value indicates the number of monitor cycles after which the agent will monitor Oracle in detail. For example, the value 5 indicates that the agent will monitor Oracle in detail every five online monitor intervals.<br>If you manually upgraded to the SF Oracle RAC 8.0 agent, and if you had enabled detail monitoring in the previous version, set the value of the LevelTwoMonitorFreq attribute appropriately.<br>**Note:** If you set the AutoEndBkup attribute value to 0, then make sure that the LevelTwoMonitorFreq attribute value is 1 for detail monitoring.<br>Default = 0 |
| Pword | string-scalar | Encrypted password for internal database-user authentication.<br>Encrypt passwords only when entering them using the command-line. Passwords must be encrypted using the VCS Encrypt Utility (`/opt/VRTSvcs/bin/vcsencrypt`).<br>**Note:** This attribute is not applicable for PDB resources. |

**Table H-21** Optional attributes for Oracle agent *(continued)*

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| Table | string-scalar | Table for update by `User/Pword`. |
| Encoding | string-scalar | Specifies operating system encoding that corresponds to Oracle encoding for the displayed Oracle output.<br>Default is "".<br>**Note:** This attribute is not applicable for PDB resources. |
| IntentionalOffline | | This resource-type level attribute defines how VCS reacts when Oracle is intentionally stopped outside of VCS control.<br>If you stop Oracle traditional and container database out of VCS control, the agent behavior is as follows:<br>■ 0—The Oracle agent registers a fault and initiates the failover of the service group.<br>■ 1—The Oracle agent takes the Oracle resource offline when Health check monitoring is enabled.<br>If Health check monitoring is not enabled, the agent registers a fault and initiates the failover of the service group.<br>**Note:** If you want to use the intentional offline functionality of the agent, you must set the value of the MonitorOption attribute as 1 to enable Health check monitoring in the CDB.<br>**Note:** If a CDB resource is configured with IntentionalOffline, and the configured CDB is brought down outside the VCS control, then the PDB resources will also report as offline. To enable this functionality, set the IntentionalOffline attribute to 1 for the PDB resource.<br>See the *Cluster Server Administrator's Guide*. |
| DBName | string-scalar | Set this attribute only when the database is a policy managed RAC database. The value of this attribute must be set to the database unique name. |
| ManagedBy | string-scalar | Default value for this attribute is ADMIN. In a policy managed RAC database this attribute must be set to POLICY. |
| PDBName | string-scalar | This attribute must be configured for a PDB. And the value should be set for a PDB database name. For traditional and CDB database, do not set this attribute. |

**Table H-21** Optional attributes for Oracle agent *(continued)*

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| UseSystemD | Boolean | systemD is a system and service manager for Linux operating systems. It helps manage applications across Linux distributions that support systemD feature.<br><br>When Online Entry point is invoked with `UseSystemD` attribute enabled, the Oracle resource comes online as a service in system.slice. Without the `UseSystemD` attribute enabled, a typical online entry point starts the resource in user.slice.<br><br>Consideration: Upon a system reboot, if application resource has been started under user.slice, there is a possibility that the processes are not stopped gracefully. This may cause the database to corrupt. This behavior is different for each Oracle application version. Hence, enable UseSystemD attribute for Oracle on you system.<br><br>The application process can be viewed as `vcs-<resource_name>.service`, in /etc/systemd/system.<br><br>Set the `UseSystemD` value to 0 or 1. |
| SystemDAttrList | string-scalar | Ensure that `UseSystemD` is enabled before you set the key-value pairs in the `SystemDAttrList` attribute. For more information, refer to the Red Hat website to know more about the different limits you can apply to the application.<br><br>For example,<br><br>`# hares -modify oracle SystemDAttrList LimitLOCKS 44433`<br><br>Where, LimitLOCKS is the key and 44433 is the value for the key. |

Table H-22 lists the internal attribute for Oracle agent. This attribute is for internal use only. recommends not to modify the value of this attribute.

**Table H-22** Internal attributes for Oracle agent

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| AgentDirectory | static-string | Specifies the location of binaries, scripts, and other files related to the Oracle agent.<br><br>Default is /opt/VRTSagents/ha/bin/Oracle. |

### About the Sid attribute in a policy managed database

The SID attribute is a required attribute. This section provides information to define the SID attribute in a policy managed database.

The SID prefix comprises of the first 8 alphanumeric characters of the database unique name. It can be a combination of letters a-z; uppercase and lowercase and numbers 0-9.

The SID prefix cannot have operating system special characters. Therefore, avoid the use of special characters in the first 8 characters of the database unique name. Special characters are omitted if used in the first 8 characters. There is a single SID prefix for every database. The SID prefix for a database must be unique within the cluster.

For an Oracle RAC database, each instance has a unique identifier, ORACLE_SID, which consists of the SID prefix and an instance number. The ORACLE_SID for Oracle RAC database instances is generated differently, depending on how you choose to manage the database. If you select a policy-managed database, then Oracle generates the SID in the format name_#, where name is the first eight alphanumeric characters of DB_UNIQUE_NAME, and # is the instance number. If you select an admin-managed database, then DBCA generates the SID for the instance names in advance, and the SID is in the format name#.

To find the Sid prefix name, run the following command:

```
# ${GRID_HOME}/bin/crsctl status resource ora.${DBName}.db -f | grep
GEN_USR_ORA_INST_NAME@ | tail -1 | sed 's/.*=//' | sed 's/_[0-9]$//',
```

where *GRID_HOME* is grid home path and *DBName* is the database unique name.

---

**Note:** When a policy managed database is created, the Sid prefix is displayed on the confirmation page of the installation procedure.

---

See "Attribute definitions for the Oracle agent" on page 482.

## Netlsnr agent functions

The listener is a server process that listens to incoming client connection requests and manages traffic to the database. The Netlsnr agent brings the listener services online, monitors their status, and takes them offline.

The Netlsnr agent is IMF-aware.

Table H-23 lists the Netlsnr agent functions.

**Table H-23** Netlsnr agent functions

| Agent operation | Description |
|---|---|
| Online | Starts the listener process by using the following command:<br>`lsnrctl start $LISTENER` |
| Offline | Stops the listener process by using the following command:<br>`lsnrctl stop $LISTENER`<br>If the listener is configured with a password, the agent uses the password to stop the listener. |
| Monitor | Verifies the status of the listener process.<br>The Netlsnr agent provides two levels of monitoring, basic and detail:<br>■ In the basic monitoring mode, the agent scans the process table for the `tnslsnr` process to verify that the listener process is running.<br>■ In the detail monitoring mode, the agent uses the `lsnrctl status $LISTENER` command to verify the status of the Listener process. (Default) |
| netlsnr_imf_init | Initializes the agent to interface with the AMF kernel driver, which is the IMF notification module for Netlsnr agent. This function runs when the agent starts up. |
| netlsnr_imf_ getnotification | Gets notification about resource state change. This function runs after the agent registers with the AMF kernel module. This function continuously waits for notification and takes action on the resource upon notification. |
| netlsnr_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). |
| Clean | Scans the process table for `tnslsnr $LISTENER` and kills it. |
| Action | Performs the predefined actions on a resource.<br>See "Action entry point for Cluster Server agent for Oracle" on page 480. |

# Resource type definition for the Netlsnr agent

The Netlsnr agent of the is represented by the Netlsnr resource type in SF Oracle RAC.

```
type Netlsnr (
 static str AgentDirectory = "/opt/VRTSagents/ha/bin/Netlsnr"
 static keylist SupportedActions = { VRTS_GetInstanceName,
        VRTS_GetRunningServices, "tnsadmin.vfd" }
 static str ArgList[] = { Owner, Home, TnsAdmin, Listener,
        EnvFile, MonScript, LsnrPwd, Encoding }
 static int IMF{} = { Mode=3, MonitorFreq=5, RegisterRetryLimit=3 }
 static str IMFRegList[] = { Home, Owner, Listener }
 str Owner
 str Home
 str TnsAdmin
 str Listener = "LISTENER"
 str EnvFile
 str MonScript = "./bin/Netlsnr/LsnrTest.pl"
 str LsnrPwd
 str Encoding
 static boolean IntentionalOffline = 0
)
```

## Attribute definition for the Netlsnr agent

Review the description of the Netlsnr agent attributes. The agent attributes are classified as required, optional, and internal.

Table H-24 lists the required attributes for Netlsnr agent. You must assign values to the required attributes.

**Table H-24** Required attributes for Netlsnr agent

| Required attributes | Type and dimension | Definition |
|---|---|---|
| Owner | string-scalar | The Oracle user who has privileges to start or stop the listener process.<br>The agent also supports LDAP users as Oracle user. |
| Home | string-scalar | The $ORACLE_HOME path to Oracle binaries and configuration files. For example, you could specify the path as /opt/ora_home.<br>Do not append a slash (/) at the end of the path. |

Table H-25 lists the optional attributes for Netlsnr agent. You can configure the optional attributes if necessary.

**Table H-25** Optional attributes for Netlsnr agent

| Optional attributes | Type and dimension | Definition |
|---|---|---|
| TnsAdmin | string-scalar | The $TNS_ADMIN path to directory in which the Listener configuration file resides (listener.ora).<br>Default is /var/opt/oracle. |
| Listener | string-scalar | Name of Listener. The name for Listener is considered case-insensitive by the Netlsnr agent and the Oracle database server.<br>Default is LISTENER.<br>**Note:** If a grid infrastructure is set up, make sure that you do not use the default name for the database listener. Otherwise, you may encounter various issues, because the grid infrastructure starts a listener with its default name, which is LISTENER. |
| LsnrPwd | string-scalar | The SF Oracle RAC encrypted password used to stop and monitor the listener. This password is set in the Listener configuration file.<br>Encrypt passwords only when entering them using the command-line. Passwords must be encrypted using the SF Oracle RAC Encrypt utility. |
| EnvFile | string-scalar | Specifies the full path name of the file that is sourced by the entry point scripts. This file contains the environment variables set by the user for the Oracle listener environment such as LD_LIBRARY_PATH and so on.<br>The syntax for the contents of the file depends on the login shell of Owner. This file must readable by Owner. The file must not contain any prompts for user input. |

**Table H-25** Optional attributes for Netlsnr agent *(continued)*

| Optional attributes | Type and dimension | Definition |
|---|---|---|
| IMF | integer-association | This resource-type level attribute determines whether the Netlsnr agent must perform intelligent resource monitoring. You can also override the value of this attribute at resource-level.<br>This attribute includes the following keys:<br>■ Mode: Define this attribute to enable or disable intelligent resource monitoring.<br>Valid values are as follows:<br>■ 0—Does not perform intelligent resource monitoring<br>■ 1—Performs intelligent resource monitoring for offline resources and performs poll-based monitoring for online resources<br>■ 2—Performs intelligent resource monitoring for online resources and performs poll-based monitoring for offline resources<br>■ 3—Performs intelligent resource monitoring for both online and for offline resources<br>Default: 3<br>■ MonitorFreq: This key value specifies the frequency at which the agent invokes the monitor agent function. The value of this key is an integer.<br>Default: 5<br>You can set this attribute to a non-zero value in some cases where the agent requires to perform poll-based resource monitoring in addition to the intelligent resource monitoring.<br>After the resource registers with the AMF kernel driver, the agent calls the monitor agent function as follows:<br>■ After every (MonitorFreq x MonitorInterval) number of seconds for online resources<br>■ After every (MonitorFreq x OfflineMonitorInterval) number of seconds for offline resources<br>■ RegisterRetryLimit: If you enable intelligent resource monitoring, the agent invokes the netlsnr_imf_register agent function to register the resource with the AMF kernel driver. The value of the RegisterRetryLimit 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.<br>Default: 3 |

**Table H-25** Optional attributes for NetIsnr agent *(continued)*

| Optional attributes | Type and dimension | Definition |
|---|---|---|
| MonScript | string-scalar | Pathname to the script provided for detail monitoring. By default, the detail monitoring is enabled to monitor the listener process.<br>**Note:** If the value of the attribute MonScript is set to an empty string, the agent disables detail monitoring.<br>The pathname to the supplied detail monitoring script is /opt/VRTSagents/ha/bin/NetIsnr/LsnrTest.pl.<br>MonScript also accepts a pathname relative to /opt/VRTSagents/ha. A relative pathname should start with "./", as in the path ./bin/NetIsnr/LsnrTest.pl. |
| LevelTwo MonitorFreq | integer-scalar | Specifies the frequency at which the agent for this resource type must perform second-level or detailed monitoring.<br>If you enabled detail monitoring, then set the value of the LevelTwoMonitorFreq attribute.<br>Default = 0 |
| Encoding | string-scalar | Specifies operating system encoding that corresponds to Oracle encoding for the displayed Oracle output.<br>Default is "". |
| IntentionalOffline | | For future use.<br>Do not change the value of this attribute.<br>Default = 0 |

Table H-26 lists the internal attribute for NetIsnr agent. This attribute is for internal use only. recommends not to modify the value of this attribute.

**Table H-26** Internal attributes for NetIsnr agent

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| AgentDirectory | static-string | Specifies the location of binaries, scripts, and other files related to the NetIsnr agent.<br>Default is /opt/VRTSagents/ha/bin/NetIsnr. |

# ASMDG agent functions

The ASMDG agent mounts the ASM disk groups that the Oracle databases use, monitors the status, unmounts the ASM disk groups.

You must have specified the disk group names in the DiskGroups attribute of the ASMDG agent.

Table H-27 lists the ASMDG agent operations.

**Table H-27** ASMDG agent operations

| Agent operation | Description |
|---|---|
| Online | Mounts the specified Oracle ASM disk groups to an ASM instance by using the following SQL command:<br>`alter diskgroup dg_name1, dg_name2 mount` |
| Offline | Unmounts the specified Oracle ASM disk groups from an ASM instance by using the following SQL command:<br>`alter diskgroup dg_name1, dg_name2 dismount`<br>**Note:** The following Oracle message appears in the VCS log when an ASM instance with no ASM disk groups mounted is shut down:<br>ORA-15100: invalid or missing diskgroup name |
| Monitor | Verifies the status of the specified ASM disk groups.<br>The disk groups can be in one of the following states:<br>■ mounted<br>■ dismounted<br>■ unknown<br>■ broken<br>■ connected<br>If multiple ASM disk groups are configured for a resource, then the ASMDG agent returns the resource state considering the status of all the specified ASM disk groups.<br>**Note:** In case of Flex ASM, the ASMDG agent honors the cardinality configured for the ASM Instances. If the ASM instance is down on a particular node it is reported as offline if the number of ASM instances equal to the cardinality value and it is reported as Faulted if the number of ASM instances falls below the cardinality value. |

**Table H-27** ASMDG agent operations *(continued)*

| Agent operation | Description |
|---|---|
| Clean | Forcibly unmounts the Oracle ASM disk groups by using the following SQL command:<br>`alter diskgroup dg_name1, dg_name2 dismount force` |

# Resource type definition for the ASMDG agent

The ASMDG agent is represented by the ASMDG resource type in SF Oracle RAC. The following extract shows the type definition of the ASMDG resource in the OracleASMTypes.cf file.

```
type ASMDG (
 static str AgentDirectory = "/opt/VRTSagents/ha/bin/ASMDG"
 static str ArgList[] = { Sid, Owner, Home, DBAUser,
   DBAPword, DiskGroups, EnvFile, Encoding }
 static boolean IntentionalOffline = 1
 str Sid
 str Owner
 str Home
 str DBAUser
 str DBAPword
 keylist DiskGroups
 str EnvFile
 str Encoding
)
```

## Attribute definition for the ASMDG agent

The ASMDG agent attributes are classified as required, optional, and internal.

You must assign values to the following required attributes.

**Table H-28** Required attributes for ASMDG agent

| Required attributes | Type and dimension | Definition |
|---|---|---|
| DiskGroups | keylist | The ASM disk groups, where you store the Oracle database files. |
| Sid | string-scalar | The variable $ORACLE_SID that represents the ASM instance. |

**Table H-28** Required attributes for ASMDG agent *(continued)*

| Required attributes | Type and dimension | Definition |
|---|---|---|
| Owner | string-scalar | The Oracle user who has privileges to mount or unmount the ASM disk group.<br>The agent also supports LDAP users as Oracle user. |
| Home | string-scalar | The $ORACLE_HOME path to Oracle ASM binaries and configuration files. For example, you could specify the path as /opt/ora_home.<br>**Note:** Do not append a slash (/) at the end of the path. |

You can configure the following optional attributes if they are necessary for your configuration.

**Table H-29** Optional attributes for ASMDG agent

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| DBAUser | string-scalar | The ASM user who has sysasm privileges to start or stop the ASM instance. |
| DBAPword | string-scalar | Encrypted password for DBAUser.<br>Encrypt passwords only when entering them using the command-line. Passwords must be encrypted using the `vcsencrypt` utility. |

**Table H-29** Optional attributes for ASMDG agent *(continued)*

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| EnvFile | string-scalar | The full path name of the file that is sourced by the entry point scripts. This file contains the environment variables set by the user for the Oracle database server environment such as LD_LIBRARY_PATH, NLS_DATE_FORMAT, and so on.<br>The syntax for the contents of the file depends on the login shell of Owner. File must be readable by Owner. The file must not contain any prompts for user input. |
| Encoding | string-scalar | Specifies operating system encoding that corresponds to Oracle encoding for the displayed Oracle output. Default is "". |

The following attribute is for internal use only. recommends that you do not modify the value of this attribute.

**Table H-30** Internal attributes for ASMDG agent

| Optional Attributes | Type and Dimension | Definition |
|---|---|---|
| AgentDirectory | static-string | Specifies the location of binaries, scripts, and other files related to the ASMDG agent.<br>Default is /opt/VRTSagents/ha/bin/ASMDG. |

# CRSResource agent

The `CRSResource` agent provides an alternative mechanism for monitoring the Oracle database in the absence of the VCS Oracle agent. It checks the status of the Oracle Clusterware resources, which include the Oracle database instance, the listener, and the virtual IP address (VIP). The agent supports multiple database configurations and ensures that the Oracle database is online and available to an application when it starts.

**Note:** The `CRSResource` agent is supported only in admin-managed database environments.

## Functions of the CRSResource agent

Table H-31 describes the functions of the CRSResource agent.

**Table H-31** CRSResource agent entry points

| Entry point | Description |
|---|---|
| Monitor | Checks the status of the following Oracle Clusterware resources:<br>■ Oracle database<br>■ Virtual IP address<br>■ Listener<br>The monitor script checks the status of the resources every 60 seconds. |

## States of the CRSResource agent

Table H-32 lists the states of the CRSResource agent.

**Table H-32** States of the CRSResource agent

| State | Description |
|---|---|
| Online | Indicates that the Oracle database, VIP, and listener are running. |
| Faulted | Indicates that the Oracle database, VIP, and listener are not running or are faulted. |
| Unknown | Indicates the inability to determine the state of the resource due to incorrect attributes settings or other configuration issues. |

## Attributes of the CRSResource agent

Table H-33 lists the attributes of the CRSResource agent.

**Table H-33** Attributes of the CRSResource agent

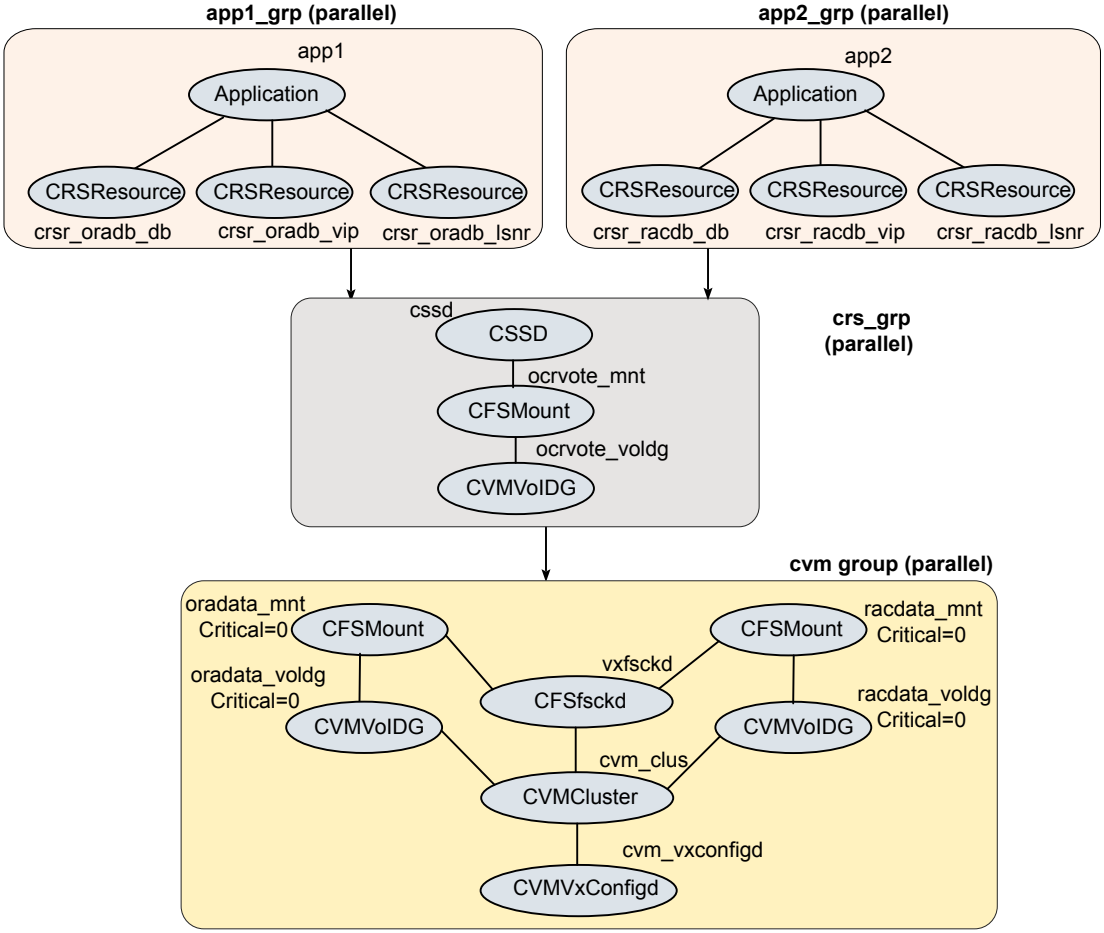
| Attribute | Dimension | Description |
|---|---|---|
| ResType | string-scalar | Indicates the type of the Oracle Clusterware resource.<br>The types of resources are as follows:<br>■ DB<br>■ VIP<br>■ Listener |
| DBHome | string-scalar | The path to ORACLE_HOME containing the Oracle binaries and configuration files. For example, you could specify the path as /app/oracle/orahome.<br>**Note:** Do not append a slash (/) at the end of the path. |
| CRSHome | string-scalar | The path to CRS_HOME containing the Oracle Clusterware binaries.<br>**Note:** Do not append a slash (/) at the end of the path. |
| DBName | string-scalar | The name of the database on which the services are configured.<br>This attribute is optional if you are configuring CRSResource to monitor VIP or listener. |
| SID | string-scalar | The name of the Oracle instance in the variable $ORACLE_SID. The SID is case-sensitive.<br>This attribute is optional if you are configuring CRSResource to monitor VIP or listener. |
| Owner | string-scalar | The name of the user for the Oracle database.<br>This attribute is optional if you are configuring CRSResource to monitor VIP or listener. |

# VCS service group dependencies with the CRSResource agent

In a service group configuration with the CRSResource agent, Oracle Clusterware controls the database. An online local firm dependency exists between the groups—Application group, Oracle Clusterware group, and the CVM group.

Figure H-2 shows a schematic illustration of the service group dependencies.

**Figure H-2** Service group dependencies with CRSResource agent




In the configuration:

- When the system starts, the CVM group brings up the volume and mount points for the databases. The Oracle Clusterware group brings up the OCR and voting disk, configures the private IP address for Oracle Clusterware, and starts Oracle Clusterware. Oracle Clusterware starts the database and the application is brought online. CRSResource comes online when the Oracle Clusterware resources (database/VIP/listener) are started by Oracle Clusterware.

---

**Note:** When the system starts, all volumes and mount points MUST be online for the dependent service groups to be online.

---

- The oradata_mnt and oradata_voldg resources are configured as non-critical resources (critical=0) for managing failure scenarios.
  See "How CRSResource agent handles failures" on page 505.
- When CRSResource faults for any of the Oracle Clusterware resources, the application is brought offline.

The limitations of this configuration are as follows:

- The CFSMount and CVMVolDg resources can not be set as critical resources in the group.
  If the mount points and volume disk groups for all the databases are configured as critical in a single service group, then failure of any of them results in the whole group being FAULTED or brought offline. To ensure that a resource failure does not affect other resources in the group, the attribute Critical is set to zero for the CFSMount and CVMVolDg resources.
  However, if any of the database mounts fail to come online or a volume does not start, the whole service group fails to come online.
- CRSResource reports as FAULTED until Oracle Clusterware brings up the database instance, VIP, and listener. Even after Oracle Clusterware starts the database instance, VIP and listener, CRSResource remains in the FAULTED state for the OfflineMonitorInterval period. The status of CRSResource cannot be changed.

## How CRSResource agent handles failures

The CRSResource agent ensures that faults in the resources of an application do not adversely impact other applications running on the system. To isolate failures, the oradata_mnt and oradata_voldg resources are configured as non-critical resources (critical=0). This ensures that storage issues in either of these resources do not affect the other databases and the dependent application continues to be online.

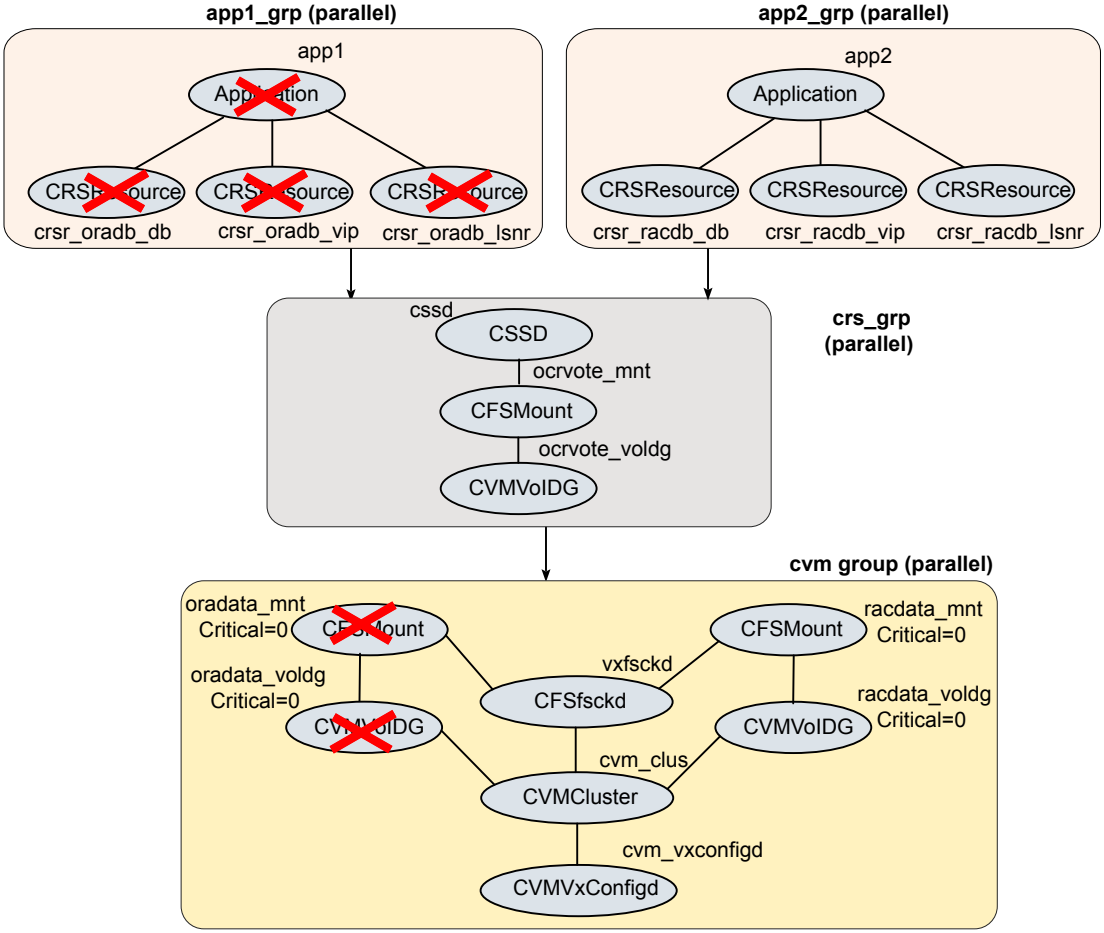
---

**Note:** The resources are considered non-critical only for the purpose of managing failure scenarios.

---

Figure H-3 illustrates a failure scenario.

**Figure H-3** How CRSResource agent handles failures




## Fault configurations with CRSResource agent

This section discusses scenarios that cause CRSResource to report FAULTED.

### Scenario 1: CRSResource fault at system startup

1. The CVM group brings online the volume and mount points for the databases when the system starts.
2. The Oracle Clusterware group brings up the OCR and voting disk, configures the private IP address for Oracle Clusterware and starts Oracle Clusterware.

Oracle Clusterware is unable to bring the database or other resources online within the start duration configured in the agent.

3. The agent starts the monitoring operation after the set duration.

   If the Oracle Clusterware resources are not brought online yet, the CRSResource appears faulted.

The following graphic illustrates the scenario.

**Figure H-4** CRSResource fault at system startup

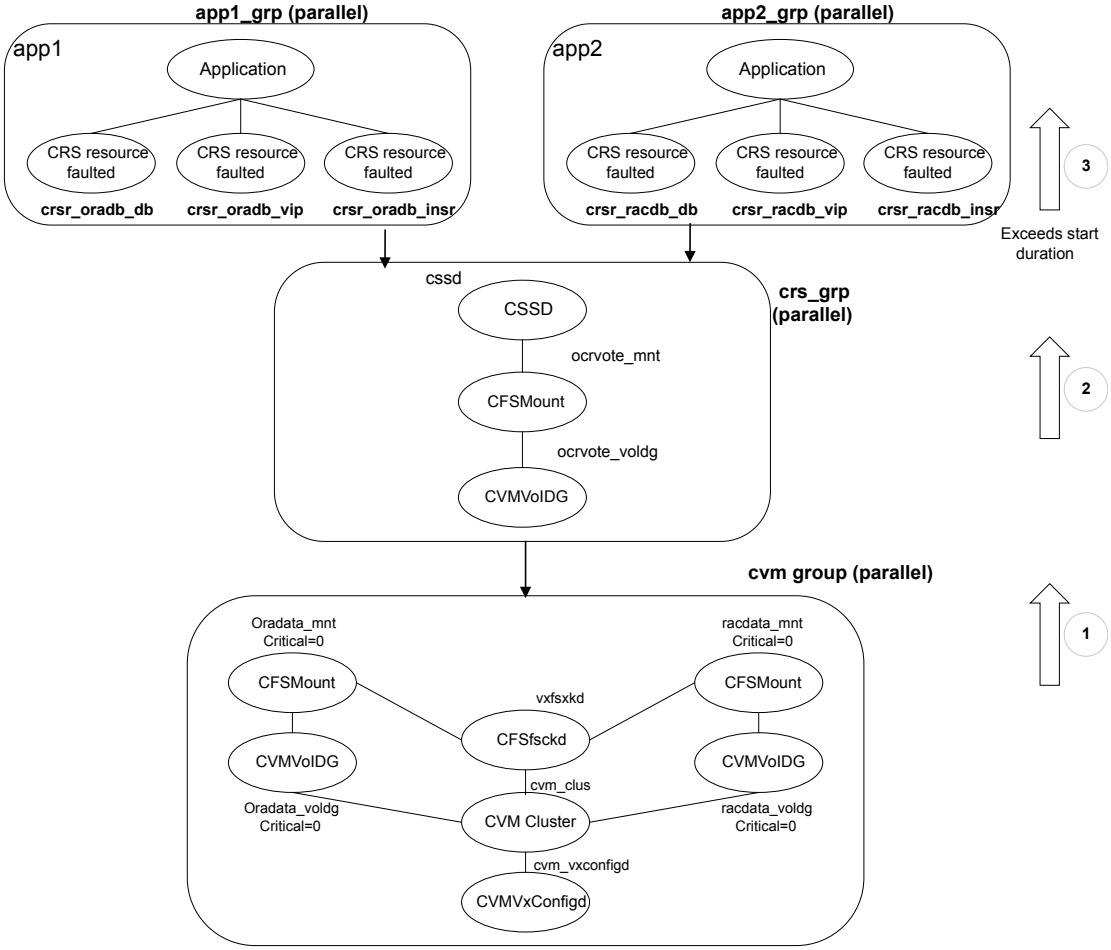
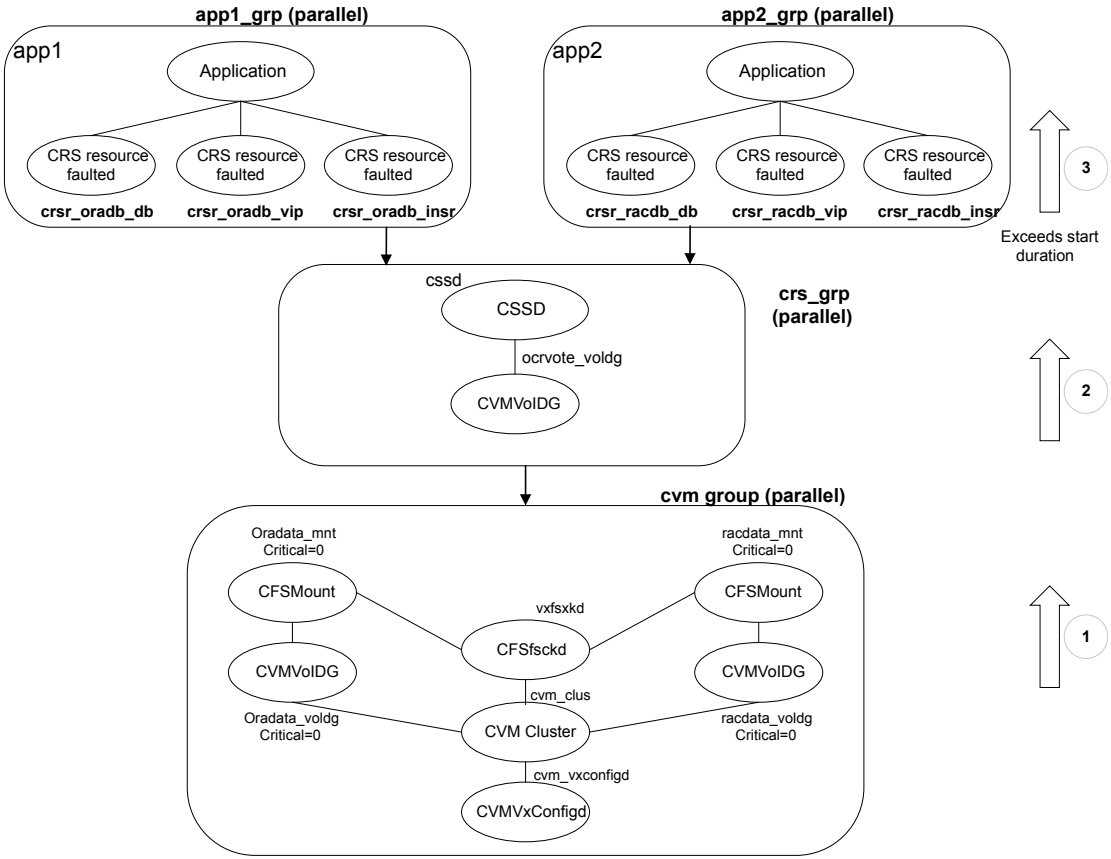


**Figure H-5** For Oracle RAC 12cR2 CRSResource fault at system startup




## Scenario 2: CRSResource fault when resource is brought offline

Any or all of the Oracle Clusterware resources (database, listener, or VIP) are brought offline due to a fault or for administrative purposes.

CRSResource reports FAULTED until Oracle Clusterware brings the resources online.

# Resource type definition for the CRSResource agent

The following extract from the `CRSResource.cf` file describes the type definition for the CRSResource agent.

```
type CRSResource (
     static int MonitorTimeout = 300
     static int OfflineMonitorInterval = 60
     static str ArgList[] = { ResType, DBHome, CRSHome,
                              DBName, SID, Owner }
     static str Operations = None
     str ResType
     str DBHome
     str CRSHome
     str DBName
     str SID
     str Owner
)
```

# Sample configuration for the CRSResource agent

The following sample configuration describes the DB, VIP, and Listener resource configurations for the CRSResource agent.

---

**Note:** The DBHome and CRSHome paths must be specified without a terminating slash (/).

---

```
CRSResource crsr_oradb_db (
    ResType = DB
    DBHome = "/app/oracle/orahome"
    CRSHome = "/app/crshome"
    DBName = oradb
    SID @sys1 = oradb1
    SID @sys2 = oradb2
    Owner = oracle
    )
CRSResource crsr_oradb_vip (
    ResType = VIP
    DBHome = "/app/oracle/orahome"
    CRSHome = "/app/crshome"
    )
CRSResource crsr_oradb_lsnr (
    ResType = Listener
```

```
DBHome = "/app/oracle/orahome"
CRSHome = "/app/crshome"
)
```

Appendix I

# SF Oracle RAC deployment scenarios

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:

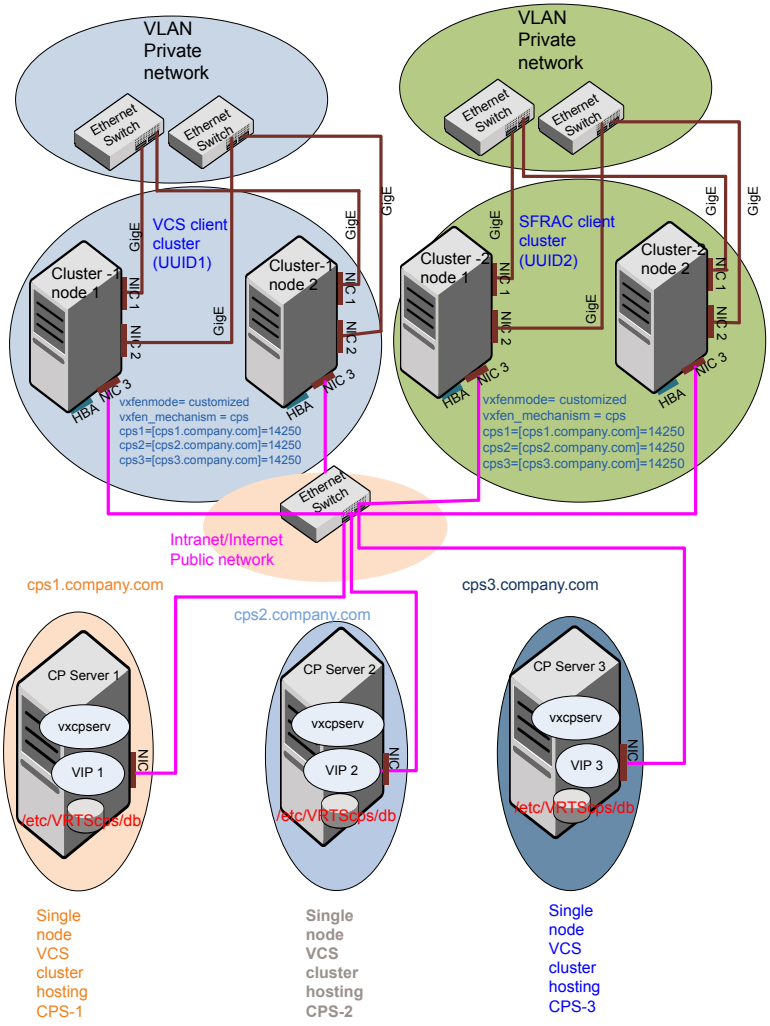
- Two unique client clusters that are served by 3 CP servers:
  See Figure I-1 on page 513.
- Client cluster that is served by highly available CP server and 2 SCSI-3 disks:
- Two node campus cluster that is served be remote CP server and 2 SCSI-3 disks:
- Multiple client clusters that are served by highly available CP server and 2 SCSI-3 disks:

### Two unique client clusters served by 3 CP servers

Figure I-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 I-1** Two unique client clusters served by 3 CP servers



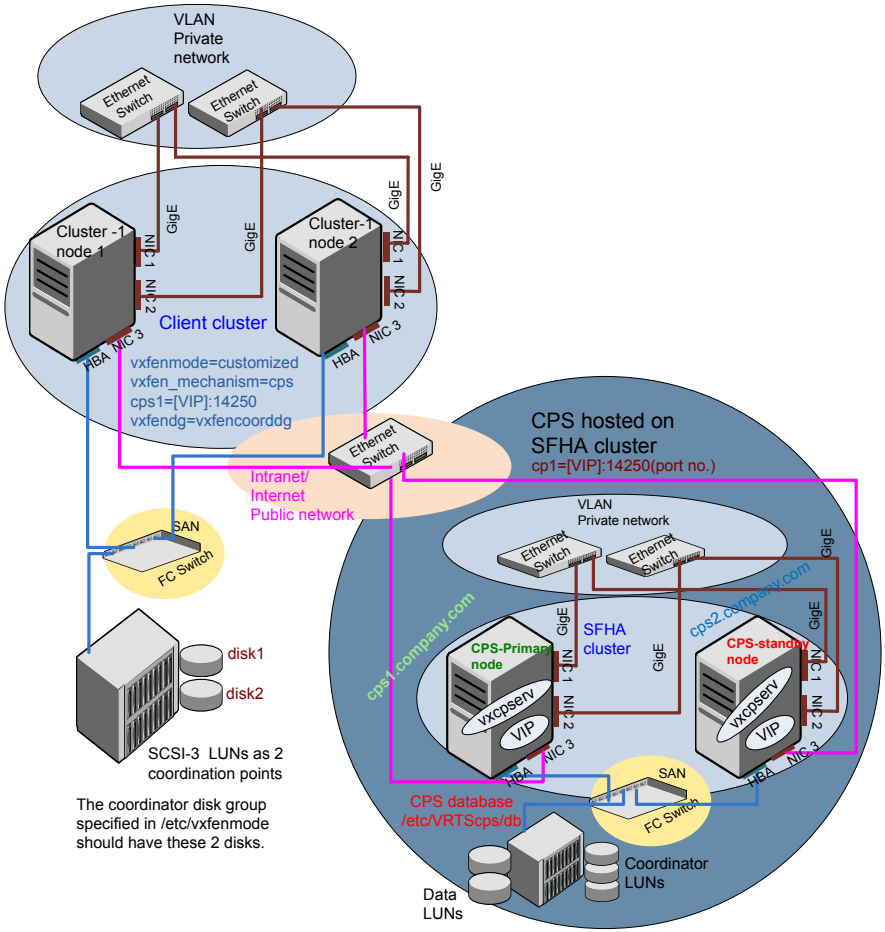

# Client cluster served by highly available CPS and 2 SCSI-3 disks

Figure I-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 I-2** Client cluster served by highly available CP server and 2 SCSI-3 disks



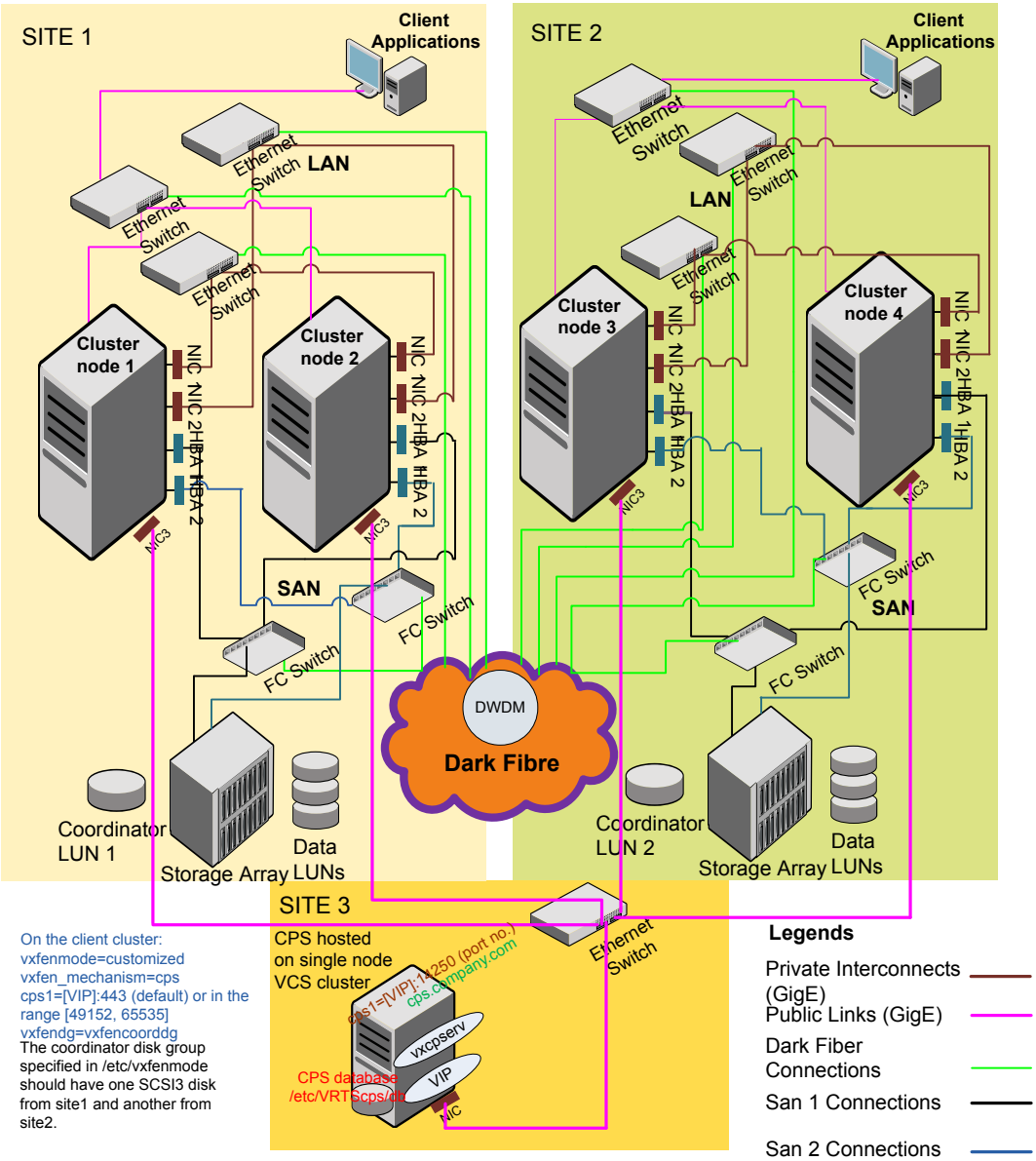
# Two node campus cluster served by remote CP server and 2 SCSI-3 disks

Figure I-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 I-3** Two node campus cluster served by remote CP server and 2 SCSI-3



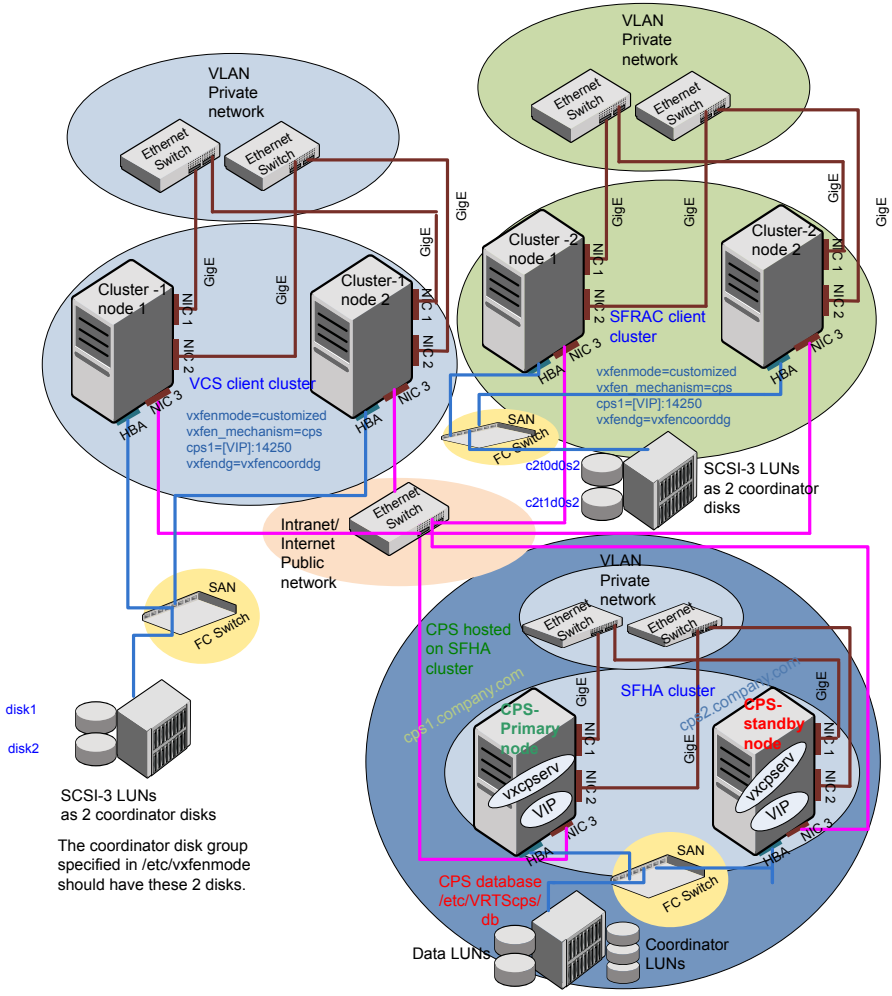
# Multiple client clusters served by highly available CP server and 2 SCSI-3 disks

Figure I-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 I-4** Multiple client clusters served by highly available CP server and 2 SCSI-3 disks



Appendix J

# 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

SF Oracle RAC 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/llttab explicitly depending on the subnet for each link.
  See "Broadcast address in the /etc/llttab file" on page 520.
- Make sure that each NIC has an IP address that is configured before configuring LLT.
- Make sure the IP addresses in the /etc/llttab 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 522.
- Set the broadcast address correctly for direct-attached (non-routed) links.
  See "Sample configuration: direct-attached links" on page 523.
- For the links that cross an IP router, disable broadcast features and specify the IP address of each link manually in the /etc/llttab file.
  See "Sample configuration: links crossing IP routers" on page 525.

## Broadcast address in the /etc/llttab file

The broadcast address is set explicitly for each link in the following example.

- Display the content of the /etc/llttab 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/llttab 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/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 523.
- See "Sample configuration: links crossing IP routers" on page 525.

Table J-1 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 J-1** 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.<br>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.<br>See "Selecting UDP ports" on page 522. |
| *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.<br>■ "-" is the default for clusters spanning routers. |

# The set-addr command in the /etc/llttab file

The `set-addr` command in the /etc/llttab 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 525.

Table J-2 describes the fields of the set-addr command.

**Table J-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.

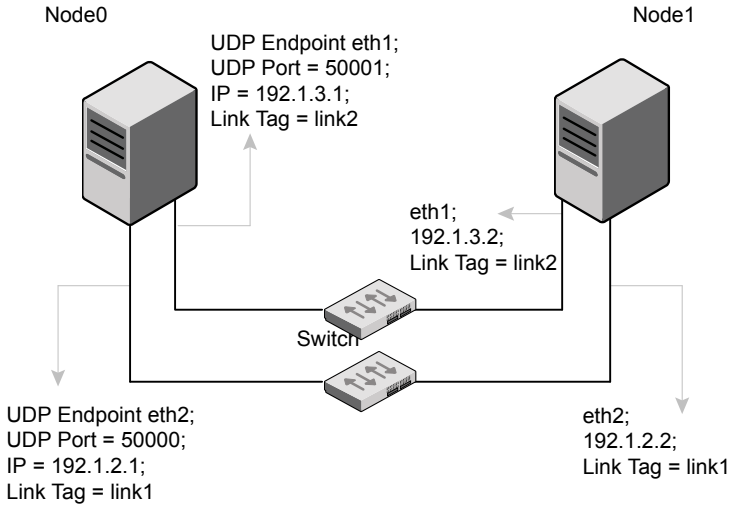
```
# cat /etc/llttab
set-node nodexyz
set-cluster 100

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 J-1 depicts a typical configuration of direct-attached links employing LLT over UDP.

**Figure J-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:

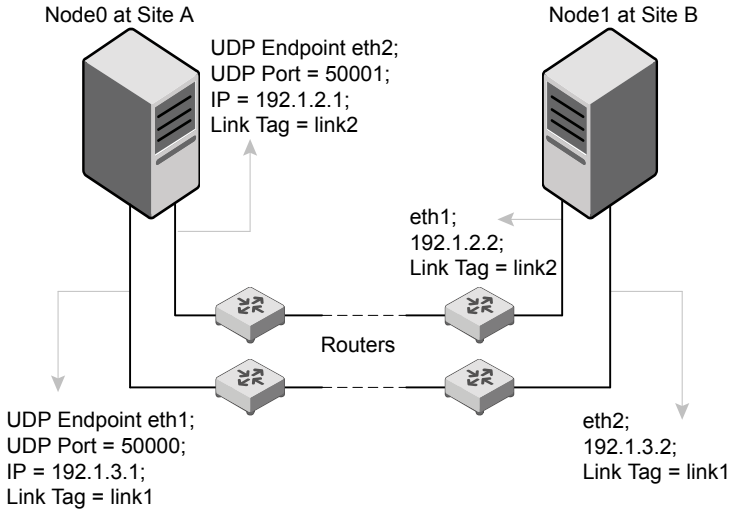
```
set-node Node1
set-cluster 1
#configure Links
#link tag-name device node-range link-type udp port MTU \
IP-address bcast-address
```

```
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 J-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 J-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

SF Oracle RAC 8.0 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 528.
- 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 530.

## 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 529.
- See "Sample configuration: links crossing IP routers" on page 530.

Note that some of the fields in Table J-3 differ from the command for standard LLT links.

Table J-3 describes the fields of the link command that are shown in the /etc/llttab file examples.

**Table J-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.<br>See "Selecting UDP ports" on page 528. |

**Table J-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 530.

Table J-4 describes the fields of the set-addr command.

**Table J-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:

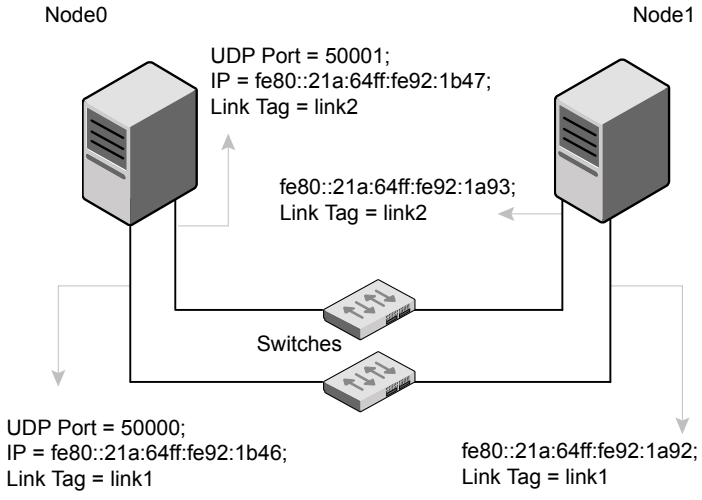
```
# netstat -au | more
Active Internet connections (servers and established)
Proto Recv-Q Send-Q Local Address Foreign Address         State
udp        0      0 *:32768                 *:*
udp        0      0 *:956                   *:*
udp        0      0 *:tftp                  *:*
udp        0      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 J-3 depicts a typical configuration of direct-attached links employing LLT over UDP.

**Figure J-3** 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 uses IPv6 multicast requests for peer node address discovery. So the addresses of peer nodes do not need to be specified in the /etc/llttab 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 -
```

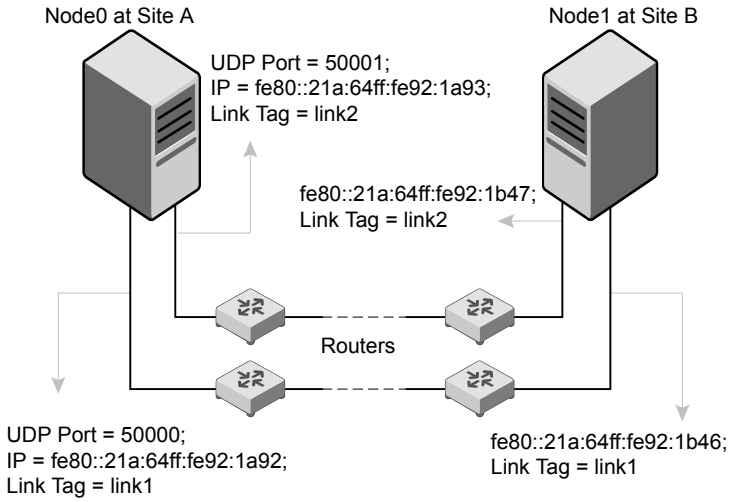
The file for Node 1 resembles:

```
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 J-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 J-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 534.

### 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.<br>Usage: set-udpports <initial_port_number><br>Example: set-udpports 60000<br>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.<br>Usage: set-udpthreads <number of threads per socket><br>Example: set-udpthreads 3 |

| Tunable | Description |
| --- | --- |
| set-udpsockets | Specifies how many sockets per link needs to be created.<br>Usage: set-udpsockets <number of sockets per link><br>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 K

# 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 537.

# 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 538.

# 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 539.

See "Configuring LLT over RDMA" on page 538.

# 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 K-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 539. |
| 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 540. |
| Configure RDMA over an Ethernet network | Load RDMA and Ethernet drivers. | See "Configuring RDMA over an Ethernet network" on page 541. |
| Configuring RDMA over an InfiniBand network | Load RDMA and InfiniBand drivers. | See "Configuring RDMA over an InfiniBand network" on page 543. |
| Tune system performance | Tune CPU frequency and boot parameters for systems. | See "Tuning system performance" on page 547. |
| Configure LLT manually | Configure LLT to use RDMA capability.<br>Alternatively, you can use the installer to automatically configure LLT to use RDMA. | See "Manually configuring LLT over RDMA" on page 548. |
| Verify LLT configuration | Run LLT commands to test the LLT over RDMA configuration. | See "Verifying LLT configuration" on page 553. |

# Choosing supported hardware for LLT over RDMA

To configure LLT over RDMA you need to use the hardware that is RDMA enabled.

**Table K-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 K-2** *(continued)*

| Hardware | Supported types | Reference |
|---|---|---|
| Network switch | Mellanox, InfiniBand switches<br>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 K-3** lists the drivers and utilities required for RDMA, InfiniBand or Ethernet network.

| Packages | RHEL | SUSE |
|---|---|---|
| Userland device drivers for RDMA operations | ▪ `libmthca`<br>▪ `libmlx4`<br>▪ `rdma`<br>▪ `librdmacm-utils` | ▪ `libmthca-rdmav2`<br>▪ `libmlx4-rdmav2`<br>▪ `ofed`<br>▪ `librdmacm` |

**Table K-3** lists the drivers and utilities required for RDMA, InfiniBand or Ethernet network. *(continued)*

| Packages | RHEL | SUSE |
|---|---|---|
| OpenSM related package (InfiniBand only) | ▪ `opensm`<br>▪ `opensm-libs`<br>▪ `libibumad` | ▪ `opensm`<br>▪ `libibumad3` |
| InfiniBand troubleshooting and performance tests | ▪ `Ibutils`<br>▪ `infiniband-diags`<br>▪ `Perftest` | ▪ `Ibutils`<br>▪ `infiniband-diags` |
| libibverbs packages for userland InfiniBand operations | ▪ `libibverbs-devel`<br>▪ `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 541.

See "Configuring RDMA and Ethernet drivers" on page 542.

See "Configuring IP addresses over Ethernet Interfaces" on page 542.

## 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. 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 544.

See “ Configuring the OpenSM service” on page 545.

See “Configuring IP addresses over InfiniBand Interfaces” on page 546.

## 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. 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.

## 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/llttab` explicitly depending on the subnet for each link.
  See "Broadcast address in the `/etc/llttab` file" on page 549.

- 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/llttab` 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 550.
- See the sample comfiguration for direct-attached (non-routed) links. See "Sample configuration: direct-attached links" on page 551.

## Broadcast address in the `/etc/llttab` file

The broadcast address is set explicitly for each link in the following example.

- Display the content of the `/etc/llttab` file on the first node sys1:

```
sys1 # cat /etc/llttab

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 K-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 K-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.<br>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.<br>See "Selecting UDP ports" on page 522. |
| *MTU* | "-" is the default, which has a value of 8192. Do not change this default value for the RDMA links. |
| *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
  ```

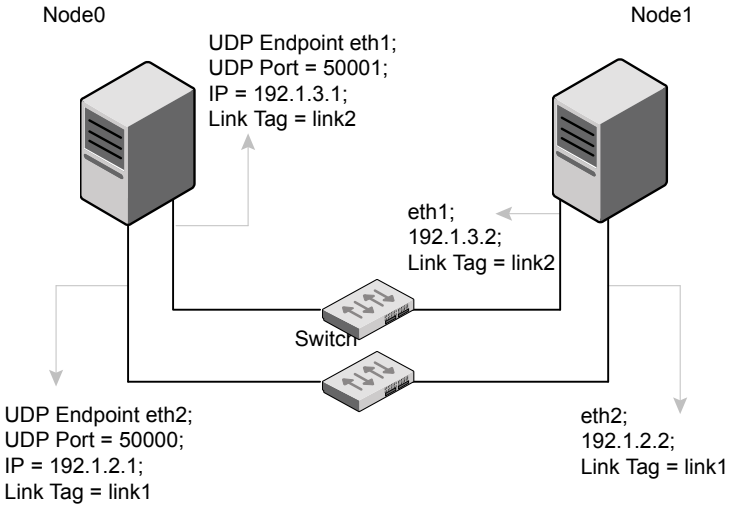
  For the second network interface on the node sys2:

  ```
  IP address=192.168.10.2, Broadcast address=192.168.10.255,
  Netmask=255.255.255.0
  ```

## Sample configuration: direct-attached links

Figure K-1 depicts a typical configuration of direct-attached links employing LLT over UDP.

**Figure K-1** A typical configuration of direct-attached links that uses LLT over RDMA




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 broadcasts 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-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/llttab

The following is a sample of LLT over RDMA in the etc/llttab 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 546.

## 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 545.

## 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.