

Veritas Storage Foundation[™] for Oracle RAC Installation and Configuration Guide

HP-UX 11i v3

5.0



Veritas SF Oracle RAC Installation and Configuration Guide

Copyright © 2007 Symantec Corporation. All rights reserved.

SF 5.0 for Oracle RAC

Symantec, the Symantec logo, Storage Foundation are trademarks or registered trademarks of Symantec Corporation or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.

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 Symantec Corporation and its licensors, if any.

THIS 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, SYMANTEC CORPORATION 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” and “commercial computer software documentation” as defined in FAR Sections 12.212 and DFARS Section 227.7202.

Symantec Corporation
20330 Stevens Creek Blvd.
Cupertino, CA 95014
www.symantec.com

Third-party legal notices

Third-party software may be recommended, distributed, embedded, or bundled with this Symantec product. Such third-party software is licensed separately by its copyright holder. All third-party copyrights associated with this product are listed in the accompanying release notes.

HP-UX is a registered trademark of Hewlett-Packard Development Company, L.P.
Oracle is a registered trademark of Oracle Corporation.

Licensing and registration

Veritas Storage Foundation for Oracle RAC is a licensed product. See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide* for license installation instructions.

Technical support

For technical assistance, visit http://www.symantec.com/business/support/assistance_care.jsp and select phone or email support. Use the Knowledge Base search feature to access resources such as TechNotes, product alerts, software downloads, hardware compatibility lists, and our customer email notification service.

Contents

Section I SF Oracle RAC concepts and requirements

Chapter 1 Introducing SF Oracle RAC

About SF Oracle RAC	19
How SF Oracle RAC works (High-level perspective)	20
Component products and processes of SF Oracle RAC	22
Communication infrastructure	23
Cluster interconnect communication channel	25
Low-level communication: Port relationship between GAB and processes	27
Cluster Volume Manager	27
Cluster File System	28
Oracle Disk Manager	29
Veritas Cluster Server	30
RAC extensions	31
About I/O fencing	33
Preventing data corruption with I/O fencing	33
SCSI-3 persistent reservations	33
I/O fencing components	34
I/O fencing operations	35
I/O fencing communication	35
Additional features of SF Oracle RAC	36

Chapter 2 Planning SF Oracle RAC installation and configuration

Important pre-installation information	37
About SF Oracle RAC component features	38
Symantec Product Authentication Service	38
Veritas Cluster Management Console	41
Notification for VCS events	42
Global Clusters	42
Veritas Volume Replicator	43
Installation requirements	43
Hardware requirements	43
Supported software	45
Supported operating systems	45

- Supported HP-UX operating environments 45
- Typical SF Oracle RAC cluster setup 47
 - Preparing SF Oracle RAC cluster setup for optional features 49
- About CVM and CFS in an SF Oracle RAC environment 53
 - About CVM 54
 - About CFS 54
 - Coordinating CVM and CFS configurations 56
 - About shared disk groups 56
 - About raw volumes versus CFS for data files 58
- Overview of SF Oracle RAC installation and configuration tasks 58
 - Preparing to install and configure SF Oracle RAC 59
 - Installing SF Oracle RAC and configuring its components 60
 - Installing Oracle RAC and creating Oracle RAC database 61
 - Setting up VCS to manage RAC resources 61
 - Setting up disaster recovery in SF Oracle RAC environment (optional) 62
 - Setting up backup and recovery feature for SF Oracle RAC (optional) 62

Section II SF Oracle RAC installation and upgrade

Chapter 3 Preparing to install and configure SF Oracle RAC

- About preparing to install and configure SF 5.0 for Oracle RAC 65
- Preparing to install SF 5.0 for Oracle RAC 65
 - Obtaining SF Oracle RAC license keys 66
 - Synchronizing time settings on cluster nodes 66
 - Setting up inter-system communication 66
 - Mounting the product disc 67
 - Setting the PATH variable 67
 - Setting the MANPATH variable 68
 - Setting up shared storage 68
 - Removing pre-existing license keys 68
 - Verifying the systems before installation 69
- Gathering information to install and configure SF Oracle RAC 70

Chapter 4 Installing and configuring SF Oracle RAC components

- About installing and configuring SF Oracle RAC 78
 - Overview of installation and configuration tasks 79
- Invoking SF Oracle RAC installation program 80
 - Mounting the product disc 80
 - Starting software installation 80
 - Specifying systems for installation 81

Installing SF 5.0 for Oracle RAC	81
Licensing SF Oracle RAC	81
Installing SF Oracle RAC depots	82
Rebooting the nodes after installation	83
Invoking SF Oracle RAC configuration program	83
Mounting the product disc	83
Starting software configuration	84
Specifying systems for configuration	84
Choosing the configuration task	85
Performing basic system checks	85
Preparing private interconnects for LLT	86
Configuring SF Oracle RAC components	87
Configuring the Veritas Cluster Server and optional features	87
Setting permissions for database administration	95
Configuring the cluster volume manager	96
Configuring VVR on each node	97
Starting the VAILAgent	98
Creating SF Oracle RAC configuration files	99
Configuring standalone hosts for Storage Foundation	99
Starting SF Oracle RAC	101
Checking shared disks for I/O fencing	101
Testing the shared disks for SCSI-3	102
Setting up I/O fencing for SF Oracle RAC	104
Verifying the SF Oracle RAC configuration	105
Initializing disks	106
Setting up coordinator disk groups	107
Stopping SF Oracle RAC on all nodes	108
Configuring /etc/vxfendg disk group for I/O fencing	108
Updating /etc/vxfenmode file	109
Starting I/O fencing	109
Modifying VCS configuration to enable I/O fencing	110
Starting SF Oracle RAC on all nodes	111
Verifying I/O fencing configuration	112
Removing permissions for communication	112
Unmounting the product disc	112
Verifying SF Oracle RAC installation using VCS configuration file	113
About VCS configuration file after SF Oracle RAC installation	113
Sample main.cf file after SF Oracle RAC installation	114

Chapter 5 Upgrading SF Oracle RAC

About upgrading to SF 5.0 for Oracle RAC on HP-UX 11i v3	117
Upgrading SF Oracle RAC	119
Upgrading SF Oracle RAC to version 5.0 for HP-UX 11i v3	119

Upgrading CVM protocol and disk group version	126
---	-----

Section III Setting up SF Oracle RAC with Oracle 10g R2

Chapter 6 Preparing to install Oracle 10g RAC

About preparing to install Oracle 10g RAC	129
About Oracle 10g RAC in a Veritas SF Oracle RAC environment	131
Cluster Ready Services	131
Oracle Cluster Registry	131
Application Resources	132
Identifying storage for Oracle 10g RAC components	133
Getting the information to create Oracle user and group id	133
Identifying storage for CRS component	134
Identifying storage for OCR and Vote-disk	134
Identifying storage for Oracle 10g binaries	135
Identifying storage for CRS component	135
Getting the information to configure	
Oracle user's environment	136
Getting the path information of the installer	136
Performing the Oracle 10g preinstallation tasks	136
Creating Oracle user and group id	137
Creating disk groups, volumes, and mount points for Oracle	138
Configuring private IP addresses for CRS	141
Editing the VCS configuration for additional CRS resources	143
Creating public virtual IP addresses for Oracle	145
Starting VCS, CVM, and CFS on all nodes	145
Verifying GAB port membership	145

Chapter 7 Installing Oracle 10g RAC

About installing Oracle 10g RAC	147
Setting Oracle environment variables and invoking	
the Oracle installer	148
Installing CRS	150
Installing Oracle 10g Binaries	152
Completing Oracle 10g post-installation tasks	153
Adding Oracle 10g patches	153
Relinking the SF Oracle RAC libraries to Oracle 10g	154
Creating Oracle 10g database	156
Configuring Oracle 10g service group in a VCS configuration	156

Chapter 8 Upgrading and migrating to Oracle 10g R2 software

	Before you upgrade and migrate to Oracle 10g R2	157
	Migrating the database	158
	Applying Oracle patchsets	158
Chapter 9	Configuring VCS service groups for Oracle 10g	
	About VCS service group for Oracle 10g dependencies	161
	Configuring CVM and Oracle service groups	164
	Configuring CVM service group for Oracle 10g manually	164
	Configuring the service groups using the wizard	165
	Location of VCS log files	174
Chapter 10	Adding and removing SF Oracle RAC nodes for Oracle 10g	
	Adding a node to an Oracle 10g cluster	175
	Checking system requirements for new node	176
	Physically adding a new system to the cluster	176
	Installing SF Oracle RAC on the new system	176
	Running vxinstall to start VxVM	178
	Configuring LLT, GAB, VCSMM, ODM, and VXFEN drivers	179
	Configuring the new node	180
	Using the Oracle add node procedure	182
	Removing a node from an Oracle 10g cluster	183
	Using the Oracle remove node procedure	183
	Removing SF 5.0 for Oracle RAC	183
Chapter 11	Uninstalling SF Oracle RAC with Oracle 10g	
	About uninstalling SF 5.0 for Oracle RAC from Oracle 10g cluster	187
	Preparing to uninstall SF 5.0 for Oracle RAC from Oracle 10g cluster	189
	Taking the Oracle resources offline	189
	Removing the Oracle database (Optional)	190
	Uninstalling Oracle 10g (optional)	190
	Unlinking the Oracle Binary From Veritas Libraries	190
	Stopping applications using CFS (outside of VCS control)	192
	Unmounting CFS file systems (outside of VCS control)	192
	Stopping VCS	192
	Removing SF 5.0 for Oracle RAC from Oracle 10g cluster	192
	Removing the SF Oracle RAC depots	193
	Removing other configuration files (optional)	194
Section IV	Disaster Recovery with SF Oracle RAC	
Chapter 12	Preparing for global clustering	

About preparing for global clustering	199
About global clustering	200
Global clustering for Oracle RAC	200
Replication in a shared disk environment	201
Setting up a secondary site	202
Obtaining and installing license keys for VVR	202
Installing SF Oracle RAC on the secondary site	202
Installing Oracle on the secondary site	204
Configuring VCS service groups for the secondary site	205

Chapter 13 Configuring global clustering

About configuring global clustering	207
Preparing clusters for replication	207
Adding the VVR resource types to the VCS configuration	208
Configuring global clustering	208
Defining the remote cluster and heartbeat Cluster Objects	211
Setting up replication	213
Creating the SRL volume on the primary site	213
Setting up replication objects on the primary site	214
Configuring replication for the secondary site	215
Starting replication of Oracle database volume	218
Configuring VCS to replicate the database volume	220
Modifying the VCS configuration for replication	220
Starting VCS on all nodes in both clusters	230
Migration and takeover of primary replication role	231
Migrating the role of primary site to the remote site	231
Taking over the primary role by the remote cluster	233

Section V Backup and recovery for SF Oracle RAC

Chapter 14 Configuring the repository database for Oracle

About configuring the repository database for Oracle	239
Creating and configuring the repository database for Oracle	239
Setting administrative permissions	241

Chapter 15 Using Storage Checkpoints and Storage Rollback

About Storage Checkpoints and Storage Rollback in SF Oracle RAC	243
Using Storage Checkpoints and Storage Rollback for backup and restore	244
About Storage Checkpoints and Storage Rollback	244

Determining Space Requirements for Storage Checkpoints	245
Performance of Storage Checkpoints	247
Backing up and recovering the database using Storage Checkpoints	248
Verifying a Storage Checkpoint using the command line	248
Backing up using a Storage Checkpoint	249
Recovering a database using a Storage Checkpoint	250
Guidelines for Oracle Recovery	251
Using the Storage Checkpoint Command Line Interface (CLI)	253
Overview of Commands	254
Examples of using the Command Line Interface	256
Prerequisites	256
Creating or updating the repository using dbed_update	256
Creating Storage Checkpoints using dbed_ckptcreate	257
Displaying Storage Checkpoints using dbed_ckptdisplay	258
Mounting Storage Checkpoints using dbed_ckptmount	263
Unmounting Storage Checkpoints using dbed_ckptumount	264
Performing Storage Rollback using dbed_ckptrollback	264
Removing Storage Checkpoints using dbed_ckptremove	266
Cloning the Oracle instance using dbed_clonedb	266

Chapter 16 Using FlashSnap for backup and recovery

About Veritas Database FlashSnap	271
Solving typical database problems with Database FlashSnap	272
About Database FlashSnap applications	273
Using Database FlashSnap	274
Using Database FlashSnap commands	275
Using Database FlashSnap options	275
Planning to use Database FlashSnap	276
Selecting the snapshot mode	276
Preparing hosts and storage for Database FlashSnap	277
Setting up hosts	277
Creating a snapshot mirror of a volume or volume set used by the database	279
Upgrading existing volumes to use Veritas Volume Manager 5.0	283
Summary of database snapshot steps	288
Creating a snapplan (dbed_vmchecksnap)	293
Creating multi-mirror snapshots	299
Validating a snapplan (dbed_vmchecksnap)	300
Displaying, copying, and removing a snapplan (dbed_vmchecksnap)	303
Displaying a snapplan	303
Copying a snapplan	304

Removing a snapplan	305
Creating a snapshot (dbed_vmsnap)	305
Backing up the database from snapshot	
volumes (dbed_vmclonedb)	308
Mounting the snapshot volumes and backing up	311
Cloning a database (dbed_vmclonedb)	313
Using Database FlashSnap to Clone a Database	313
Shutting down the clone database and	
unmounting file systems	317
Restarting a Clone Database	318
Recreating Oracle tempfiles	319
Resynchronizing the snapshot to your database	320
Removing a snapshot volume	321

Section VI Troubleshooting SF Oracle RAC and optimizing I/O performance

Chapter 17 Investigating I/O performance for SF Oracle RAC: Storage Mapping

About Storage Mapping in SF Oracle RAC	325
Understanding Storage Mapping	326
Verifying Veritas Storage Mapping setup	327
Using vxstorage_stats	327
Displaying Storage Mapping information	328
Displaying I/O statistics information	329
Using dbed_analyzer	330
Obtaining Storage Mapping information for	
a list of tablespaces	331
Oracle File Mapping (ORAMAP)	332
Mapping components	332
Storage Mapping views	333
Verifying Oracle file mapping setup	334
Enabling Oracle file mapping	334
Accessing dynamic performance views	335
Using Oracle Enterprise Manager	336
About arrays for Storage Mapping and statistics	337

Chapter 18 Troubleshooting SF Oracle RAC

About troubleshooting SF Oracle RAC	339
Running scripts for engineering support analysis	339
Troubleshooting tips	340

Troubleshooting Oracle	341
Oracle log files	341
Oracle notes	341
Oracle troubleshooting topics	342
Troubleshooting fencing	343
Node is unable to join cluster while another node is being ejected	343
vxfsentsthdw fails when SCSI TEST UNIT READY command fails	344
Removing existing keys from disks	344
System panics to prevent potential data corruption	345
Clearing keys after split brain using vxfsenclearpre command	347
Adding or removing coordinator disks	348
Troubleshooting ODM	349
File system configured incorrectly for ODM shuts down Oracle	349
Troubleshooting VCSIPC	350
VCSIPC wait warning messages in Oracle trace/log files	350
VCSIPC errors in Oracle trace/log files	350
Troubleshooting CVM	351
Shared disk group cannot be imported	351
CVMVolDg does not go online even though CVMCluster is online	351
Troubleshooting interconnects	351
Restoring communication between host and disks after cable disconnection	352
Troubleshooting SF Oracle RAC checkpoint	352
SF Oracle RAC checkpoint feature requires DBA group to have MLOCK privileges	352

Section VII SF Oracle RAC Reference

Appendix A Sample VCS configuration files for SF Oracle RAC

About sample main.cf files	355
Sample main.cf for Oracle 10g without Oracle agent	356
Sample main.cf for Oracle 10g with Oracle agent	357
Sample main.cf for Oracle 10g for CVM/VVR primary site	360
Sample main.cf for Oracle 10g for CVM/VVR secondary site	363

Appendix B SF Oracle RAC agents

About configuring the VCS agents in SF Oracle RAC	369
About CVMCluster agent	370
Entry points for CVMCluster agent	370

Attribute definition for CVMCluster agent	370
Resource type definition for CVMCluster agent	370
Sample configuration for CVMCluster agent	371
About CVMVxconfigd agent	371
Entry points for CVMVxconfigd agent	372
Attribute definition for CVMVxconfigd agent	372
Resource type definition for CVMVxconfigd agent	372
Sample configuration for CVMVxconfigd agent	372
About the CVMVolDg agent	373
Entry points for CVMVolDg agent	373
Attribute definition for CVMVolDg agent	373
Resource type definition for CVMVolDg agent	374
Sample configuration for CVMVolDg agent	374
About the CFSSMount agent	374
Entry points for CFSSMount agent	374
Attribute definition for CFSSMount agent	375
Resource type Definition for CFSSMount agent	376
Sample configuration for CFSSMount agent	376
About PrivNIC agent	376
Entry point for PrivNIC agent	377
Attribute definition for PrivNIC agent	377
Resource type definition for PrivNIC agent	379
Sample configuration for PrivNIC agent	379
About CSSD agent	380

Appendix C Creating a starter database

About creating a starter database	381
Creating a starter database for Oracle 10g	381
Creating Oracle 10g database on raw volumes	382
Creating Oracle 10g database on CFS	383

Appendix D I/O fencing testing and scenarios

I/O fencing of shared storage	387
Verifying data storage arrays using the vxfentsthdw utility	388
General guidelines for using vxfentsthdw	388
vxfentsthdw options	389
How I/O fencing works in different event scenarios	394
About vxfenadm utility	398
Registration key formatting	399

Appendix E Configuring the Symantec License Inventory Agent

About the Symantec License Inventory Manager	402
--	-----

	When the Symantec License Inventory Agent is installed	403
	When the server and access points are installed	403
	What you can do with the agent after it is installed	403
	How to remove the agent	404
	How to order the Symantec License Inventory Manager license and media kit	405
Appendix F	Tunable kernel driver parameters	
	About tunable parameters	407
	About LMX tunable parameters	408
	About VXFEN tunable parameters	408
	Configuring the module parameters	409
Appendix G	Error messages	
	About error messages	411
	LMX error messages	411
	LMX critical error messages	412
	LMX non-critical error messages	413
	VxVM error messages	414
	VXFEN driver error messages	414
	VXFEN driver informational message	415
	Node ejection informational messages	415
Glossary		417
Index		421

SF Oracle RAC concepts and requirements

- [Chapter 1, “Introducing SF Oracle RAC” on page 19](#)
- [Chapter 2, “Planning SF Oracle RAC installation and configuration” on page 37](#)

Introducing SF Oracle RAC

This chapter contains the following topics:

- [About SF Oracle RAC](#)
- [How SF Oracle RAC works \(High-level perspective\)](#)
- [Component products and processes of SF Oracle RAC](#)
- [About I/O fencing](#)
- [Additional features of SF Oracle RAC](#)

About SF Oracle RAC

Veritas Storage Foundation™ for Oracle® RAC by Symantec provides a robust infrastructure for Oracle Real Application Clusters (RAC) that simplifies management of RAC databases. Storage Foundation for Oracle RAC (SF Oracle RAC) integrates existing Veritas storage management and clustering technologies into a flexible solution for administrators.

SF Oracle RAC is a storage management and clustering solution that enables you to:

- Create a standard approach toward application and database management in data centers. While other clusterware can only work with an Oracle database, SF Oracle RAC incorporates existing Veritas storage management and clustering technologies that provide flexible support for many types of applications and databases. Administrators can apply existing expertise of Veritas technologies toward this product.
- Set up an infrastructure for Oracle RAC that simplifies database management while fully integrating with Oracle Clusterware, formerly known as Oracle Cluster Ready Services (CRS).
- Enhance scalability and availability with access to eight RAC instances per database in a cluster.

- Transition from a local high availability site to a wide-area disaster recovery environment with primary and secondary sites.
- Back up and recover databases using volume-level and file system-level snapshot technologies. SF Oracle RAC enables full volume-level snapshots for off-host processing that reduce load on production systems, and file system-level snapshots that involve efficient backup and rollback processes.
- Prevent data corruption at the storage layer with robust split-brain protection.
- Increase scalability with high throughput and low latency technology for use by Oracle Cache Fusion.
- Share all types of files, in addition to Oracle database files, across nodes.
- Increase availability and performance with dynamic multipathing (DMP), which provides wide storage array support for protection from failures and performance bottlenecks in the HBAs and SAN switches.
- Model and test cluster configurations without affecting production systems using the simulator and firedrill clustering technologies.
- Optimize I/O performance through storage mapping technologies and tunable attributes.

How SF Oracle RAC works (High-level perspective)

Real Application Clusters (RAC) is a parallel database environment that takes advantage of the processing power of multiple computers. The Oracle database is the physical data stored in tablespaces on disk, while the Oracle instance is a set of processes and shared memory that provide access to the physical database. Specifically, the instance involves server processes acting on behalf of clients to read data into shared memory and make modifications to it, and background processes to write changed data to disk.

In traditional environments, only one instance accesses a database at a specific time. SF Oracle RAC enables all nodes to concurrently run Oracle instances and execute transactions against the same database. This software coordinates access to the shared data for each node to provide consistency and integrity. Each node adds its processing power to the cluster as a whole and can increase overall throughput or performance.

At a conceptual level, SF Oracle RAC is a cluster that manages applications (instances), networking, and storage components using resources contained in service groups. SF Oracle RAC clusters have many of the same properties as Veritas Cluster Server (VCS) clusters:

- Each node runs its own operating system.

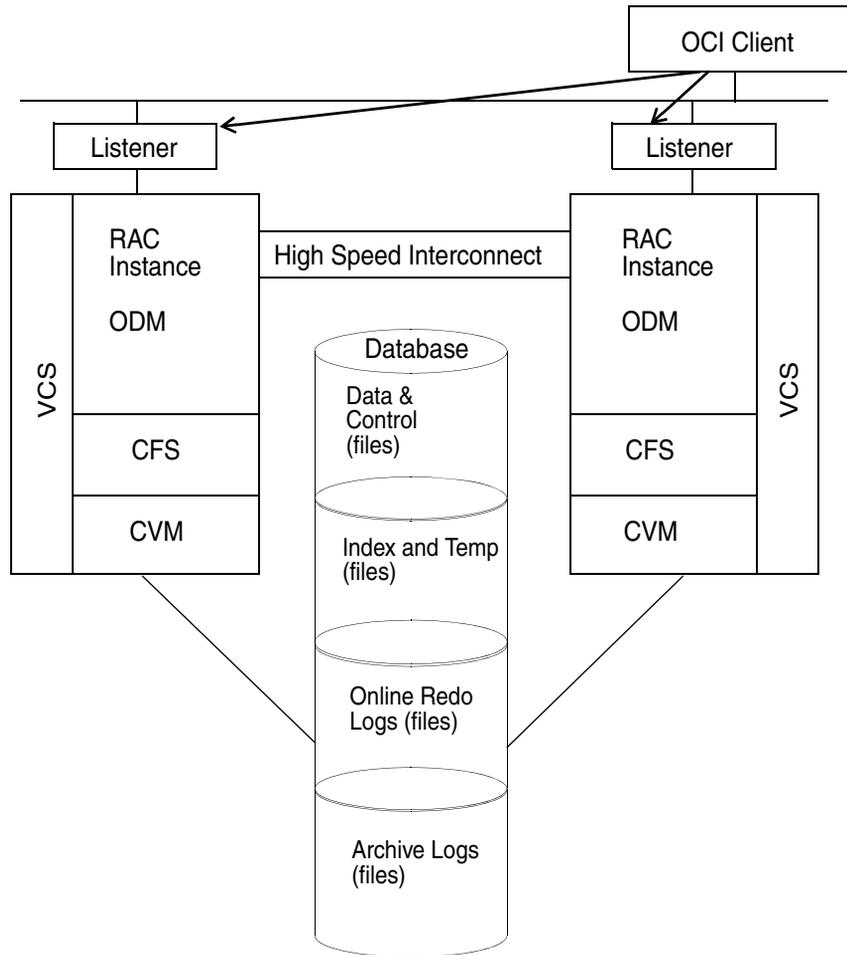
- A cluster interconnect enables cluster communications.
- A public network connects each node to a LAN for client access.
- Shared storage is accessible by each node that needs to run the application.

SF Oracle RAC adds these technologies to a failover cluster environment, which are engineered specifically to improve performance, availability, and manageability of Oracle RAC environments:

- Cluster File System (CFS) and Cluster Volume Manager (CVM) technologies to manage multi-instance database access to shared storage.
- An Oracle Disk Manager (ODM) library to maximize Oracle disk I/O performance.
- Interfaces to Oracle Clusterware (referred to as CRS—formerly Cluster Ready Services) and RAC for managing cluster membership and communication.

SF Oracle RAC provides an environment that can tolerate failures with minimal downtime and interruption to users. If a node fails as clients access the same database on multiple nodes, clients attached to the failed node can reconnect to a surviving node and resume access. Recovery after failure in the SF Oracle RAC environment is far quicker than recovery for a failover database because another Oracle instance is already up and running. The recovery process involves applying outstanding redo log entries from the failed node.

Figure 1-1 SF Oracle RAC architecture



Component products and processes of SF Oracle RAC

To understand how SF Oracle RAC manages database instances running in parallel on multiple nodes, review the architecture and communication

mechanisms that provide the infrastructure for Oracle RAC. [Table 1-1](#) highlights the SF Oracle RAC component products.

Table 1-1 SF Oracle RAC component products

Component product	Description
Cluster Volume Manager (CVM)	Enables simultaneous access to shared volumes based on technology from Veritas Volume Manager (VxVM). See “ Cluster Volume Manager ” on page 27.
Cluster File System (CFS)	Enables simultaneous access to shared file systems based on technology from Veritas File System (VxFS). See “ Cluster File System ” on page 28.
Database Accelerator	Provides the interface with the Oracle Disk Manager (ODM) API. See “ Oracle Disk Manager ” on page 29.
Cluster Server (VCS)	Uses technology from Veritas Cluster Server to manage Oracle RAC databases and infrastructure components. See “ Veritas Cluster Server ” on page 30.
RAC Extensions	Manages cluster membership and communications between cluster nodes. See “ RAC extensions ” on page 31.

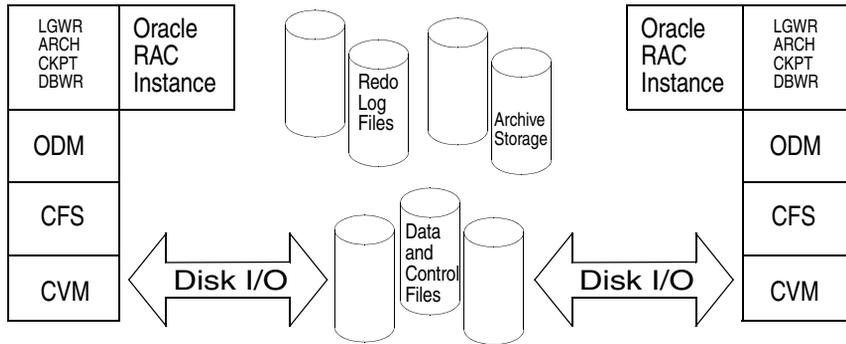
Communication infrastructure

To understand the communication infrastructure, review the data flow and communication requirements.

Data flow

The CVM, CFS, ODM, and Oracle Real Application Clusters (RAC) elements reflect the overall data flow, or data stack, from an instance running on a server to the shared storage. The various Oracle processes composing an instance -- such as DB Writers, Log Writer, Checkpoint, Archiver, and Server -- read and write data to the storage through the I/O stack in the diagram. Oracle communicates through the ODM interface to CFS, which in turn accesses the storage through the CVM.

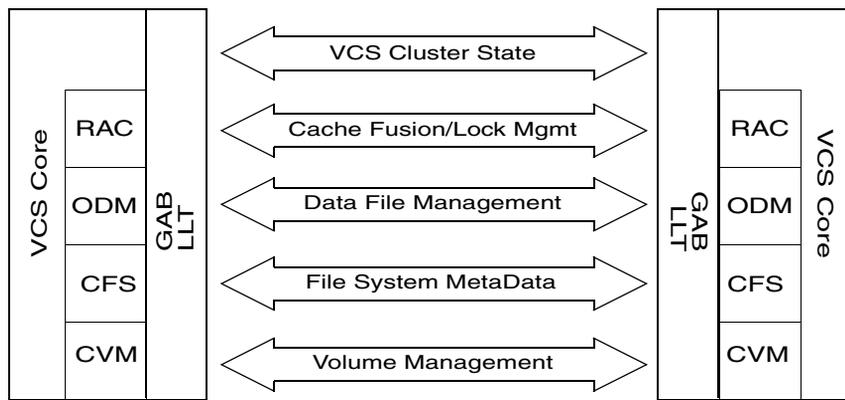
Figure 1-2 Data stack



Communication requirements

End-users on a client system are unaware that they are accessing a database hosted by multiple instances. The key to performing I/O to a database accessed by multiple instances is communication between the processes. Each layer or component in the data stack must reliably communicate with its peer on other nodes to function properly. RAC instances must communicate to coordinate protection of data blocks in the database. ODM processes must communicate to coordinate data file protection and access across the cluster. CFS coordinates metadata updates for file systems, while CVM coordinates the status of logical volumes and maps.

Figure 1-3 Communication stack



Cluster interconnect communication channel

The cluster interconnect provides the communication channel for all system-to-system communication, in addition to one-node communication between modules. Low Latency Transport (LLT) and Group Membership Services/Atomic Broadcast (GAB) make up the VCS communications package central to the operation of SF Oracle RAC.

In a standard operational state, significant traffic through LLT and GAB results from Lock Management and Cache Fusion, while traffic for other data is relatively sparse.

Low latency transport

LLT provides fast, kernel-to-kernel communications and monitors network connections. LLT functions as a high performance replacement for the IP stack and runs directly on top of the Data Link Protocol Interface (DLPI) layer. The use of LLT rather than IP removes latency and overhead associated with the IP stack. The major functions of LLT are traffic distribution, heartbeats, and support for RAC Inter-Process Communications (VCSIPC).

- Traffic distribution

LLT distributes (load-balances) internode communication across all available cluster interconnect links. All cluster communications are evenly distributed across as many as eight network links for performance and fault resilience. If a link fails, LLT redirects traffic to the remaining links.

- Heartbeats

LLT is responsible for sending and receiving heartbeat traffic over network links. The Group Membership Services function of GAB uses heartbeats to determine cluster membership.

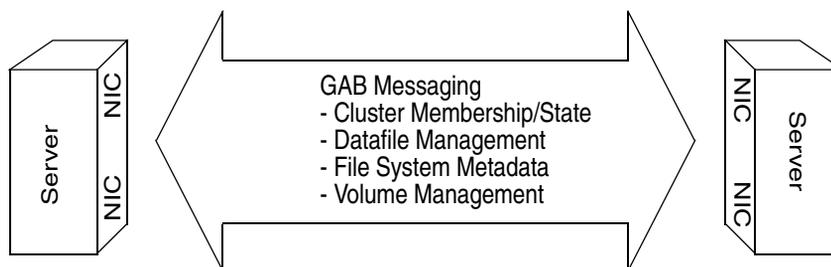
- VCSIPC

RAC Inter-Process Communications (VCSIPC) uses the VCSIPC shared library for these communications. VCSIPC leverages all features of LLT and uses LMX, an LLT multiplexer, to provide fast data transfer between Oracle processes on different nodes.

Group membership services/atomic broadcast

The GAB protocol is responsible for cluster membership and cluster communications.

Figure 1-4 Cluster communication



Cluster membership

At a high level, all nodes configured by the installer can operate as a cluster; these nodes form a cluster membership. In SF Oracle RAC, a cluster membership specifically refers to all systems configured with the same cluster ID communicating by way of a redundant cluster interconnect.

All nodes in a distributed system, such as SF Oracle RAC, must remain constantly alert to the nodes currently participating in the cluster. Nodes can leave or join the cluster at any time because of shutting down, starting up, rebooting, powering off, or faulting processes. SF Oracle RAC uses its cluster membership capability to dynamically track the overall cluster topology.

SF Oracle RAC uses LLT heartbeats to determine cluster membership:

- When systems no longer receive heartbeat messages from a peer for a predetermined interval, a protocol excludes the peer from the current membership.
- GAB informs processes on the remaining nodes that the cluster membership has changed; this action initiates recovery actions specific to each module. For example, CVM must initiate volume recovery and CFS must perform a fast parallel file system check.
- When systems start receiving heartbeats from a peer outside of the current membership, a protocol enables the peer to join the membership.

Cluster communications

GAB provides reliable cluster communication between SF Oracle RAC modules. GAB provides guaranteed delivery of point-to-point messages and broadcast messages to all nodes. Point-to-point messaging involves sending and acknowledging the message. Atomic-broadcast messaging ensures all systems within the cluster receive all messages. If a failure occurs while transmitting a broadcast message, GAB ensures all systems have the same information after recovery.

Low-level communication: Port relationship between GAB and processes

All components in SF Oracle RAC use GAB for communication. Each process wanting to communicate with a peer process on other nodes registers with GAB on a specific port. This registration enables communication and notification of membership changes. For example, the VCS engine (HAD) registers on port h. HAD receives messages from peer had processes on port h. HAD also receives notification when a node fails or when a peer process on port h becomes unregistered.

Some processes use multiple ports for specific communications requirements. For example, CVM uses multiple ports to allow communications by kernel and user-level functions in CVM independently.

Cluster Volume Manager

CVM is an extension of Veritas Volume Manager, the industry-standard storage virtualization platform. CVM extends the concepts of VxVM across multiple nodes. Each node recognizes the same logical volume layout, and more importantly, the same state of all volume resources.

CVM supports performance-enhancing capabilities, such as striping, mirroring, and mirror break-off (snapshot) for off-host backup. You can use standard VxVM commands from one node in the cluster to manage all storage. All other nodes immediately recognize any changes in disk group and volume configuration with no interaction.

CVM architecture

CVM is designed with a “master and slave” architecture. One node in the cluster acts as the configuration master for logical volume management, and all other nodes are slaves. Any node can take over as master if the existing master fails. The CVM master exists on a per-cluster basis and uses GAB and LLT to transport its configuration data.

Just as with VxVM, the Volume Manager configuration daemon, `vxconfigd`, maintains the configuration of logical volumes. This daemon handles changes to the volumes by updating the operating system at the kernel level. For example, if a mirror of a volume fails, the mirror detaches from the volume and `vxconfigd` determines the proper course of action, updates the new volume layout, and informs the kernel of a new volume layout. CVM extends this behavior across multiple nodes and propagates volume changes to the master `vxconfigd`. (You must perform operator-initiated changes on the master node.) The `vxconfigd` process on the master pushes these changes out to slave `vxconfigd` processes, each of which updates the local kernel.

CVM does not impose any write locking between nodes. Each node is free to update any area of the storage. All data integrity is the responsibility of the upper application. From an application perspective, standalone systems access logical volumes in the same way as CVM systems.

CVM imposes a “Uniform Shared Storage” model. All nodes must connect to the same disk sets for a given disk group. Any node unable to detect the entire set of physical disks for a given disk group cannot import the group. If a node loses contact with a specific disk, CVM excludes the node from participating in the use of that disk.

CVM communication

CVM communication involves various GAB ports for different types of communication.

Port w

Most CVM communication uses port w for `vxconfigd` communications. During any change in volume configuration, such as volume creation, plex attachment or detachment, and volume resizing, `vxconfigd` on the master node uses port w to share this information with slave nodes.

When all slaves use port w to acknowledge the new configuration as the next active configuration, the master updates this record to the disk headers in the VxVM private region for the disk group as the next configuration.

Port v

CVM uses port v for kernel-to-kernel communication. During specific configuration events, certain actions require coordination across all nodes. An example of synchronizing events is a resize operation. CVM must ensure all nodes see the new or old size, but never a mix of size among members.

CVM also uses this port to obtain cluster membership from GAB and determine the status of other CVM members in the cluster.

Cluster File System

CFS enables you to simultaneously mount the same file system on multiple nodes and is an extension of the industry-standard Veritas File System. Unlike other file systems which send data through another node to the storage, CFS is a true SAN file system. All data traffic takes place over the storage area network (SAN), and only the metadata traverses the cluster interconnect.

In addition to using the SAN fabric for reading and writing data, CFS offers storage checkpoints and rollback for backup and recovery.

Access to cluster storage in typical SF Oracle RAC configurations use CFS. Raw access to CVM volumes is also possible but not part of a common configuration.

CFS architecture

SF Oracle RAC uses CFS to manage a file system in a large database environment. Since CFS is an extension of VxFS, it operates in a similar fashion and caches metadata and data in memory (typically called buffer cache or vnode cache). CFS uses a distributed locking mechanism called Global Lock Manager (GLM) to ensure all nodes have a consistent view of the file system. GLM provides metadata and cache coherency across multiple nodes by coordinating access to file system metadata, such as inodes and free lists. The role of GLM is set on a per-file system basis to enable load balancing.

CFS involves a primary/secondary architecture. One of the nodes in the cluster is the primary node for a file system. Though any node can initiate an operation to create, delete, or resize data, the GLM master node carries out the actual operation. After creating a file, the GLM master node grants locks for data coherency across nodes. For example, if a node tries to modify a block in a file, it must obtain an exclusive lock to ensure other nodes that may have the same file cached have this cached copy invalidated.

SF Oracle RAC configurations minimize the use of GLM locking. Oracle RAC accesses the file system through the ODM interface and handles its own locking; only Oracle (and not GLM) buffers data and coordinates write operations to files. A single point of locking and buffering ensures maximum performance. GLM locking is only involved when metadata for a file changes, such as during create and resize operations.

CFS communication

CFS uses port f for GLM lock and metadata communication. SF Oracle RAC configurations minimize the use of GLM locking except when metadata for a file changes.

Oracle Disk Manager

SF Oracle RAC requires Oracle Disk Manager (ODM), a standard API published by Oracle for support of database I/O. Veritas provides a library for Oracle to use as its I/O library.

ODM architecture

When the Veritas ODM library is linked, Oracle is able to bypass all caching and locks at the file system layer and to communicate directly with raw volumes. The SF Oracle RAC implementation of ODM generates performance equivalent

to performance with raw devices while the storage uses easy-to-manage file systems.

All ODM features can operate in a cluster environment. Nodes communicate with each other before performing any operation that could potentially affect another node. For example, before creating a new data file with a specific name, ODM checks with other nodes to see if the file name is already in use.

Veritas ODM performance enhancements

Veritas ODM enables the performance benefits provided by Oracle Disk Manager:

- Locking for data integrity.
- Few system calls and context switches.
- Increased I/O parallelism.
- Efficient file creation and disk allocation.

Databases using file systems typically incur additional overhead:

- Extra CPU and memory usage to read data from underlying disks to the file system cache. This scenario requires copying data from the file system cache to the Oracle cache.
- File locking that allows for only a single writer at a time. Allowing Oracle to perform locking allows for finer granularity of locking at the row level.
- File systems generally go through a standard Sync I/O library when performing I/O. Oracle can make use of Kernel Async I/O libraries (KAIO) with raw devices to improve performance.

ODM communication

ODM uses port d to communicate with other ODM instances to support the file management features of Oracle Managed Files (OMF). OMF enables DBAs to set `init.ora` parameters for db datafile, controlfile, and logfile names and for those structures to be named automatically. OMF allows for the automatic deletion of physical data files when DBAs remove tablespaces.

Veritas Cluster Server

VCS directs SF Oracle RAC operations by controlling the startup and shutdown of components layers and providing monitoring and notification of failure.

In a typical SF Oracle RAC configuration, the RAC service groups for VCS run as “parallel” service groups rather than “failover” service groups; in the event of a

failure, VCS does not attempt to migrate a failed service group. Instead, the software enables you to configure the group to restart on failure.

VCS architecture

The High Availability Daemon (HAD) is the main VCS daemon running on each node. HAD tracks changes in the cluster configuration and monitors resource status by communicating over GAB and LLT. HAD manages all application services using agents, which are installed programs to manage resources (specific hardware or software entities).

The VCS architecture is modular for extensibility and efficiency; HAD does not need to know how to start up Oracle or any other application under VCS control. Instead, you can add agents to manage different resources with no effect on the engine (HAD). Agents only communicate with HAD on the local node, and HAD communicates status with HAD processes on other nodes. Because agents do not need to communicate across systems, VCS is able to minimize traffic on the cluster interconnect.

SF Oracle RAC provides specific agents for VCS to manage CVM, CFS, and Oracle agents.

VCS communication

SF Oracle RAC uses port `h` for HAD communication. Agents communicate with HAD on the local node about resources, and HAD distributes its view of resources on that node to other nodes through port `h`. HAD also receives information from other cluster members to update its own view of the cluster.

Cluster configuration files

VCS uses two configuration files in a default configuration:

- The `main.cf` file defines the entire cluster, including the cluster name, systems in the cluster, and definitions of service groups and resources, in addition to service group and resource dependencies.
- The `types.cf` file defines the resource types.

Additional files similar to `types.cf` may be present if you add agents. For example, SF Oracle RAC includes additional resource types files, such as `OracleTypes.cf` and `PrivNIC.cf`.

RAC extensions

Oracle RAC relies on several support services provided by VCS. Key features include Veritas Cluster Server Membership Manager (VCSMM) and Cluster Inter-Process Communication (VCSIPC), and LLT Multiplexer (LMX).

Veritas Cluster Server membership manager

To protect data integrity by coordinating locking between RAC instances, Oracle must know which instances actively access a database. Oracle provides an API called `skgxn` (system kernel generic interface node membership) to obtain information on membership. SF Oracle RAC implements this API as a library linked to Oracle after you install Oracle RAC. Oracle uses the linked `skgxn` library to make `ioctl` calls to VCSMM, which in turn obtains membership information for clusters and instances by communicating with GAB on port o.

Veritas Cluster Server inter-process communication

To coordinate access to a single database by multiple instances, Oracle uses extensive communications between nodes and instances. Oracle uses Inter-Process Communications (VCSIPC) for Global Enqueue Service locking traffic and Global Cache Service cache fusion. SF Oracle RAC uses LLT to support VCSIPC in a cluster and leverages its high-performance and fault-resilient capabilities.

Oracle has an API for VCSIPC, System Kernel Generic Interface Inter-Process Communications (`skgxp`), that isolates Oracle from the underlying transport mechanism. As Oracle conducts communication between processes, it does not need to know how data moves between systems; the cluster implementer can create the highest performance for internode communications without Oracle reconfiguration.

LLT multiplexer

Oracle instances use the `skgxp` library for interprocess communication. This interface enables Oracle to send communications between processes on instances.

SF Oracle RAC provides a library dynamically linked to Oracle at installation time to implement the `skgxp` functionality. This module communicates with the LLT Multiplexer (LMX) using `ioctl` calls.

The LMX module is a kernel module designed to receive communications from the `skgxp` module and pass them on to the correct process on the correct instance on other nodes. The LMX module “multiplexes” communications between multiple processes on other nodes. LMX leverages all features of LLT, including load balancing and fault resilience.

About I/O fencing

I/O Fencing is a feature that prevents data corruption in the event of a communication breakdown in a cluster.

Preventing data corruption with I/O fencing

To provide high availability, the cluster must be capable of taking corrective action when a node fails. In this situation, SF Oracle RAC configures its components to reflect the altered membership.

Problems arise when the mechanism that detects the failure breaks down because symptoms appear identical to those of a failed node. For example, if a system in a two-node cluster fails, the system stops sending heartbeats over the private interconnects and the remaining node takes corrective action. However, the failure of private interconnects (instead of the actual nodes) would present identical symptoms and cause each node to determine its peer has departed. This situation typically results in data corruption because both nodes attempt to take control of data storage in an uncoordinated manner.

In addition to a broken set of private networks, other scenarios can generate this situation. If a system is so busy that it appears to stop responding or “hang,” the other nodes could declare it as dead. This declaration may also occur for nodes using hardware that supports a “break” and “resume” function. When a node drops to PROM level with a break and subsequently resumes operations, the other nodes may declare the system dead even though the system later returns and begins write operations.

SF Oracle RAC uses a technology called I/O fencing to remove the risk associated with split brain. I/O fencing allows write access for members of the active cluster and blocks access to storage from non-members; even a node that is alive is unable to cause damage.

SCSI-3 persistent reservations

SCSI-3 Persistent Reservations (SCSI-3 PR) are required for I/O fencing and resolve the issues of using SCSI reservations in a clustered SAN environment. SCSI-3 PR enables access for multiple nodes to a device and simultaneously blocks access for other nodes.

SCSI-3 reservations are persistent across SCSI bus resets and support multiple paths from a host to a disk. In contrast, only one host can use SCSI-2 reservations with one path. If the need arises to block access to a device because of data integrity concerns, only one host and one path remain active. The requirements for larger clusters, with multiple nodes reading and writing to storage in a controlled manner, make SCSI-2 reservations obsolete.

SCSI-3 PR uses a concept of registration and reservation. Each system registers its own “key” with a SCSI-3 device. Multiple systems registering keys form a membership and establish a reservation, typically set to “Write Exclusive Registrants Only.” The WERO setting enables only registered systems to perform write operations. For a given disk, only one reservation can exist amidst numerous registrations.

With SCSI-3 PR technology, blocking write access is as simple as removing a registration from a device. Only registered members can “eject” the registration of another member. A member wishing to eject another member issues a “preempt and abort” command. Ejecting a node is final and atomic; an ejected node cannot eject another node. In SF Oracle RAC, a node registers the same key for all paths to the device. A single preempt and abort command ejects a node from all paths to the storage device.

I/O fencing components

Fencing in SF Oracle RAC involves coordinator disks and data disks. Each component has a unique purpose and uses different physical disk devices. The fencing driver, known as `vx fencing`, directs CVM as necessary to carry out actual fencing operations at the disk group level.

Data disks

Data disks are standard disk devices for data storage and are either physical disks or RAID Logical Units (LUNs). These disks must support SCSI-3 PR and are part of standard VxVM or CVM disk groups.

CVM is responsible for fencing data disks on a disk group basis. Disks added to a disk group are automatically fenced, as are new paths discovered to a device.

Coordinator disks

Coordinator disks are three standard disks or LUNs set aside for I/O fencing during cluster reconfiguration. Coordinator disks do not serve any other storage purpose in the SF Oracle RAC configuration. Users cannot store data on these disks or include the disks in a disk group for user data. The coordinator disks can be any three disks that support SCSI-3 PR. Coordinator disks cannot be special devices that array vendors use. For example, you cannot use EMC gatekeeper devices as coordinator disks.

Symantec recommends using the smallest possible LUNs for coordinator disks. Because coordinator disks do not store any data, cluster nodes need only register with them and do not need to reserve them.

These disks provide a lock mechanism to determine which nodes get to fence off data drives from other nodes. A node must eject a peer from the coordinator

disks before it can fence the peer from the data drives. This concept of racing for control of the coordinator disks to gain the ability to fence data disks is key to understanding prevention of split brain through fencing.

Dynamic Multipathing devices with I/O fencing

Dynamic Multipathing allows coordinator disks to take advantage of the path failover and dynamic adding and removal capabilities of DMP.

On HP-UX 11i v3, you must use DMP devices for I/O fencing. Note the following changes in HP-UX 11.31:

- Provides native multipathing support
- Does not provide access to individual paths through the device file entries

The metanode interface that HP-UX provides does not meet the SCSI-3 PR requirements for the I/O fencing feature. You can configure coordinator disks to use Veritas Volume Manager Dynamic Multipathing (DMP) feature.

For more information on using DMP, see the *Veritas Volume Manager Administrator's Guide*.

See “[Updating /etc/vxfenmode file](#)” on page 109.

I/O fencing operations

I/O fencing, provided by the kernel-based fencing module (`vxfen`), performs identically on node failures and communications failures. When the fencing module on a node is informed of a change in cluster membership by the GAB module, it immediately begins the fencing operation. The node attempts to eject the key for departed nodes from the coordinator disks using the `preempt` and `abort` command. When the node successfully ejects the departed nodes from the coordinator disks, it ejects the departed nodes from the data disks. In a split brain scenario, both sides of the split would race for control of the coordinator disks. The side winning the majority of the coordinator disks wins the race and fences the loser. The loser then panics and reboots the system.

I/O fencing communication

The `vxfen` driver connects to GAB port `b` to intercept cluster membership changes (reconfiguration messages). During a membership change, the fencing driver determines which systems are members of the cluster to allow access to shared disks.

After completing fencing operations, the driver passes reconfiguration messages to higher modules. CVM handles fencing of data drives for shared disk groups. After a node successfully joins the GAB cluster and the driver determines that a preexisting split brain does not exist, CVM can import all

shared disk groups. The CVM master coordinates the order of import and the key for each disk group. As each slave joins the cluster, it accepts the CVM list of disk groups and keys, and adds its proper digit to the first byte of the key. Each slave then registers the keys with all drives in the disk groups.

Additional features of SF Oracle RAC

Additional SF Oracle RAC features include:

- The ability to transition from a local high-availability cluster into a disaster recovery environment. This environment ensures maximum data protection and availability in the event of large-scale disasters and involves the global clustering feature of VCS (GCO) and replication with Veritas Volume Replicator (VVR).
- The ability to back up and recover data at the volume and file system levels using Veritas Storage Checkpoints and Veritas Database Flashsnap.
- The ability to evaluate or troubleshoot I/O performance with Veritas Storage Mapping. You can access mapping information that allows for a detailed understanding of the storage hierarchy in which files reside.

Planning SF Oracle RAC installation and configuration

This chapter contains the following topics:

- [Important pre-installation information](#)
- [About SF Oracle RAC component features](#)
- [Installation requirements](#)
- [Typical SF Oracle RAC cluster setup](#)
- [About CVM and CFS in an SF Oracle RAC environment](#)
- [Overview of SF Oracle RAC installation and configuration tasks](#)

Important pre-installation information

Before you install SF Oracle RAC, make sure you have reviewed the following:

- Current compatibility list in the Veritas Technical Support website to confirm the compatibility of your hardware
<http://entsupport.symantec.com/docs/283161>
- TechNote for late-breaking and new information on updates, patches, and software issues regarding this release
<http://entsupport.symantec.com/docs/293075>
- *Veritas Storage Foundation for Oracle RAC Release Notes* for 5.0 HP-UX 11i v3

About SF Oracle RAC component features

Review the description of the optional features and decide the features that you want to configure with SF Oracle RAC:

- [Symantec Product Authentication Service](#)
- [Veritas Cluster Management Console](#)
- [Notification for VCS events](#)
- [Global Clusters](#)
- [Veritas Volume Replicator](#)

Note: To configure the optional features of the SF Oracle RAC components, make sure to install all depots when the installation program prompts you.

Symantec Product Authentication Service

The Symantec Product Authentication Service is a common Veritas feature that validates identities based on existing network operating system domains (such as NIS and NT) or private domains. The authentication service protects communication channels among Symantec application clients and services through message integrity and confidentiality services.

Before you install the authentication service, refer to the *Symantec Product Authentication Service Installation Guide* at the following location on the Veritas software disc:

`authentication_service/docs/vxat_install.pdf`.

Symantec Product Authentication Service secures communication using digital certificates for authentication and SSL to encrypt communication over the public network. You can configure SF Oracle RAC to use the Authentication Service to secure communication between the following:

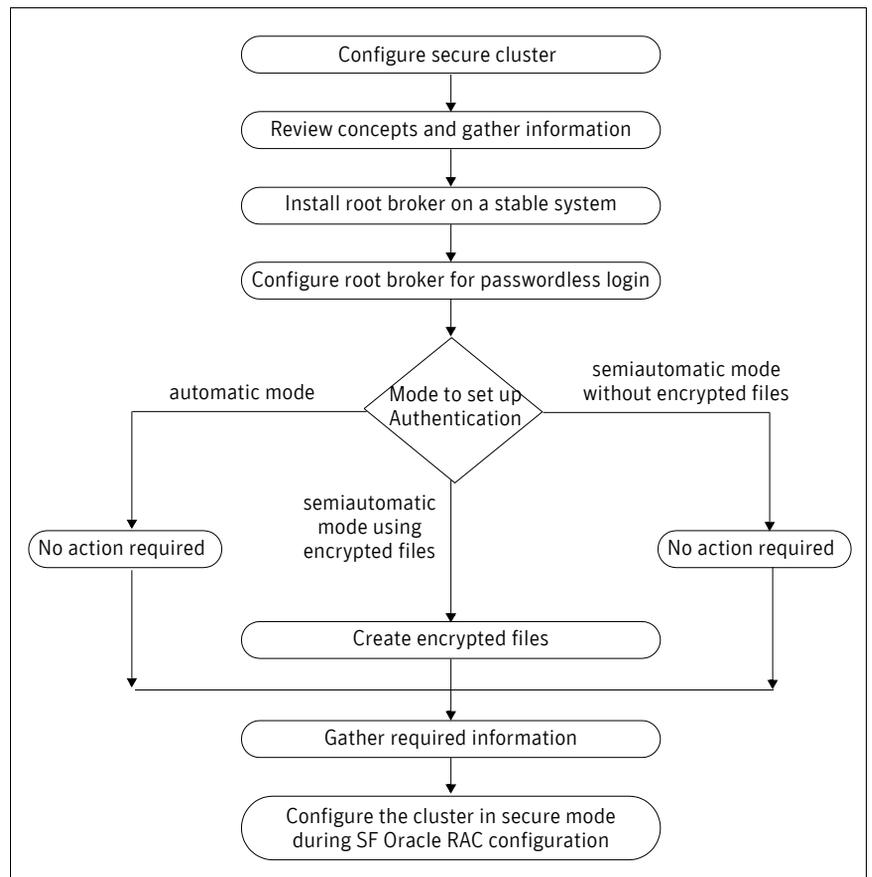
- Cluster nodes and clients, including the VCS Java and the Web consoles
You can set up Authentication Service for the cluster during the SF Oracle RAC installation and configuration process. If you want to enable Authentication Service after installation, refer to the *Veritas Cluster Server User's Guide*.
See [“Configuring the cluster in secure mode”](#) on page 88.
- Veritas Cluster Management Console Management Server and the centrally managed SF Oracle RAC clusters
See [“Veritas Cluster Management Console”](#) on page 41.

To configure the cluster in secure mode, SF Oracle RAC requires you to configure a system in your enterprise as root broker and all nodes in the cluster as authentication brokers.

- **Root broker**
 A root broker serves as the main registration and certification authority; it has a self-signed certificate and can authenticate other brokers. The root broker is only used during initial creation of an authentication broker.
- **Authentication brokers**
 Authentication brokers serve as intermediate registration and certification authorities. Authentication brokers have certificates that are signed by the root. Each node in SF Oracle RAC cluster serves as an authentication broker.

Figure 2-5 depicts the flow of configuring SF Oracle RAC in secure mode.

Figure 2-5 Secure SF Oracle RAC cluster configuration flowchart



If you decide to enable the Authentication Service, the root broker administrator must perform the following preparatory tasks:

- Install the root broker on another stable system.
The root broker is the main registration and certification authority and can serve multiple clusters. Symantec recommends that you install a single root broker on a utility computer such as an email server or domain controller, which can be highly available.
See [“Installing root broker for Symantec Product Authentication Service”](#) on page 50.
- Configure the root broker system for a passwordless login when you want to use the automatic mode.

The `installsfrac` program provides the following modes to configure Symantec Product Authentication Service:

- In the automatic mode, the installer configures Authentication Service automatically without any user intervention.
You must provide the name of the root broker system.
- In the semiautomatic modes, the installer provides you an option to use encrypted files or answer the installer prompts to enable security. The semiautomatic mode requires the root broker administrator to set up the basic authentication environment and create principals for authentication brokers. You must complete the following preparatory tasks to configure security in the semiautomatic mode:

- With encrypted files
- The root broker administrator must create an encrypted file for each node in the cluster.
See [“Creating encrypted files for Symantec Product Authentication Service”](#) on page 51.
 - The root broker administrator must provide the encrypted files in a media or make it available on a shared location that you can access.
 - You must copy the encrypted files to a directory in the installation node. Make a note of the path of this encrypted files.

- Without encrypted files
 - You must gather the following information from the root broker administrator:
 - Root broker name
 - Root broker domain name
 - Root broker port (Default is 2821)
 - Authentication broker principal name for each node
 - Authentication broker password for each Authentication broker
 - The root broker administrator must provide the root_hash file in a media or make it available on a shared location that you can access.
 - You must copy the root_hash file to a directory in the installation node. Make a note of the path of this root_hash file.

Refer to the *Symantec Product Authentication Service Administrator's Guide* for more information.

Note: Make sure that the system clocks of the Rook Broker and Authentication Brokers systems are in sync.

Veritas Cluster Management Console

The Veritas Cluster Management Console is a management interface that enables you to monitor and administer clusters from a web console. You can configure Cluster Management Console to manage single clusters, multiple clusters, or both. To configure Cluster Management Console, do the following:

- To manage multiple clusters, you must have Cluster Management Console management server setup. If the prerequisite is met, then you can manage multiple clusters using direct connection or cluster connector. See Veritas Cluster Management Console documentation. Depending on the connection mode you would use, do the following:

Direct connection If you can use direct connection to communicate with the management server, the clusters require no further preparation during VCS installation and configuration.

After configuring VCS, you can start the Cluster Management Console from the management server and configure the management server to connect to clusters using direct connection.

Cluster connector If a firewall separates the management server and cluster nodes, you need to install a component called the cluster connector on each cluster node.

The cluster connector enables communication with clusters through firewalls and provides buffering for cluster data. If the console goes offline and then comes back online, it can retrieve data collected during the offline period from the cluster connector buffer.

You can configure cluster connector during or after the VCS installation and configuration. Clusters using cluster connector connect to the management server automatically.

See [“Installing and configuring SF Oracle RAC components”](#) on page 77.

- To manage a single cluster, you must choose the option to install the Cluster Management Console during VCS installation and configuration.

See [“Installing and configuring SF Oracle RAC components”](#) on page 77.

See the *Veritas Cluster Server User’s Guide*.

Supported browsers for the Cluster Management Console

Veritas Cluster Management Console is supported on the following browsers:

- Microsoft Internet Explorer 6.0 or later
- Firefox 1.5 or later

Veritas Cluster Management requires the Macromedia Flash Plugin v8.0.

Notification for VCS events

You have the option to configure SMTP email notification and SNMP trap notification of VCS events by the VCS Notifier component. Refer to the *Veritas Cluster Server User’s Guide* for more information on SMTP and SNMP notification.

Global Clusters

Global clusters provide the ability to fail over applications between geographically distributed clusters when disaster occurs. You require a separate license to configure global clusters. You must add this license during the installation.

If you choose to configure global clusters, the installer enables you to choose whether or not to use the same NIC, virtual IP address, and netmask as are configured for the ClusterService group, which are the defaults. If you choose

not to use the same networking information, you must specify appropriate values for the NIC, virtual IP address, and netmask when you are prompted.

Veritas Volume Replicator

Veritas Volume Replicator is an optional, separately-licensable feature of SF Oracle RAC. Volume Replicator is a fully integrated component of Veritas Volume Manager that replicates data to remote locations over any standard IP network to provide continuous data availability.

Installation requirements

Make sure each node on which you want to install or upgrade SF Oracle RAC meets the installation requirements.

- [Hardware requirements](#)
- [Supported software](#)
- [Supported operating systems](#)
- [Supported HP-UX operating environments](#)
- [Typical SF Oracle RAC cluster setup](#)

Hardware requirements

Make sure that you have the correct equipment to install SF Oracle RAC.

Table 2-2 Hardware requirements

Item	Description
SF Oracle RAC systems	From two to eight HP-UX (Itanium or PA-RISC) systems connected to the public network running HP-UX 11i v3. Symantec recommends that each system have two or more CPUs at 2GHz or higher.
DVD drive	One drive that is accessible to all nodes in the cluster.

Table 2-2 Hardware requirements

Item	Description
Disks	<p>Typical SF Oracle RAC configurations require that shared disks support applications that migrate between systems in the cluster.</p> <p>The SF Oracle RAC I/O fencing feature requires that all disks used as data disks or as coordinator disks must support SCSI-3 Persistent Reservations (PR).</p> <p>Note: On HP-UX 11i v3, you must use DMP devices for I/O fencing. The coordinator disk does not store data, so configure the disk as the smallest possible LUN on a disk array to avoid wasting space. See “Performing basic system checks” on page 85.</p>
Disk space	See “Disk space (local)” on page 44.
RAM	Symantec recommends 2 GB or more of physical memory for each SF Oracle RAC system.
Swap space	Symantec recommends at least double of main memory.
Network links	<p>Two or more 100BaseT or Gigabit Ethernet links directly linking each node to the other node to form a private network that handles direct inter-system communication.</p> <p>Syantec recommends switches for the private network links. These links must be of the same type; you cannot mix 100BaseT and Gigabit.</p>
Fibre channel or SCSI host bus adapters	SF Oracle RAC requires at least one built-in SCSI adapter per system to access the operating system disks, and at least one additional SCSI or Fibre Channel Host Bus Adapter per system for shared data disks.

Disk space (local)

Confirm that your system has enough free disk space to install SF Oracle RAC. Each node in the cluster requires at least the following local disk space for the SF Oracle RAC depots:

Directory	Required depots
/opt	1.5 GB
/usr	225 MB
/var	32 MB
/	230 MB

Directory	Required depots
/tmp	512 MB
/var/tmp	600 MB
Total	3.1 GB

Supported software

Software versions that SF 5.0 for Oracle RAC supports include:

Oracle RAC	■ Oracle 10g Release 2
HP-UX operating system	See “ Supported operating systems ” on page 45.
VCS, VxVM, VxFS, VVR	Use only versions of VCS, VxVM, VxFS, and VVR provided on the software disc. Remove other versions before you install the software from the SF Oracle RAC product disc.

To verify the latest information on support for Oracle database versions, see the Veritas Technical Support TechNote:

<http://entsupport.symantec.com/docs/280186>

Supported operating systems

Within a cluster, all nodes must use the same operating system version.

- HP-UX 11i version 3.0.
If necessary, upgrade the operating system to this version.
See “[Supported HP-UX operating environments](#)” on page 45.
- For each platform, Symantec recommends applying any latest HP-UX operating system patches available from HP.

For Veritas Enterprise Administrator (VEA) functionality to work reliably on Itanium platform, you must have HP-UX patch PHSS_36311 installed prior to installing SF Oracle RAC. Check the *Veritas Storage Foundation Release Notes* for any additional information before you install and configure VEA.

Supported HP-UX operating environments

HP-UX 11i Operating Environments (OEs) are bundles of key applications available to users to help them install a complete operating system in a

single-pass process. Symantec supports SF Oracle RAC for HP-UX on the following OEs:

HPUX11i-OE	HP-UX 11i v3 Foundation OE
HPUX11i-OE-Ent	HP-UX 11i v3 Enterprise OE
HPUX11i-OE-MC	HP-UX 11i v3 Mission Critical OE

Identifying the operating environment

To identify the OE currently installed

- ◆ Run the `swlist` command to identify the OE currently installed on your system.

The output of this command includes a line that identifies the installed OE. For example:

```
# swlist | grep HPUX
HPUX11i-OE-MC          B.11.31          HP-UX Mission
                      Critical Operating Environment Component
HPUXBaseAux           B.11.31          HP-UX Base OS
                      Auxiliary
HPUXBaseOS            B.11.31          HP-UX Base OS
OnlineDiag            B.11.31          HPUX 11.31
                      Support Tools Bundle, Feb 2007
```

The output of this command includes a line that identifies the installed OE. For example:

```
HPUX11i-OE-MC          B.11.31          HP-UX Mission
                      Critical Operating Environment Component
```

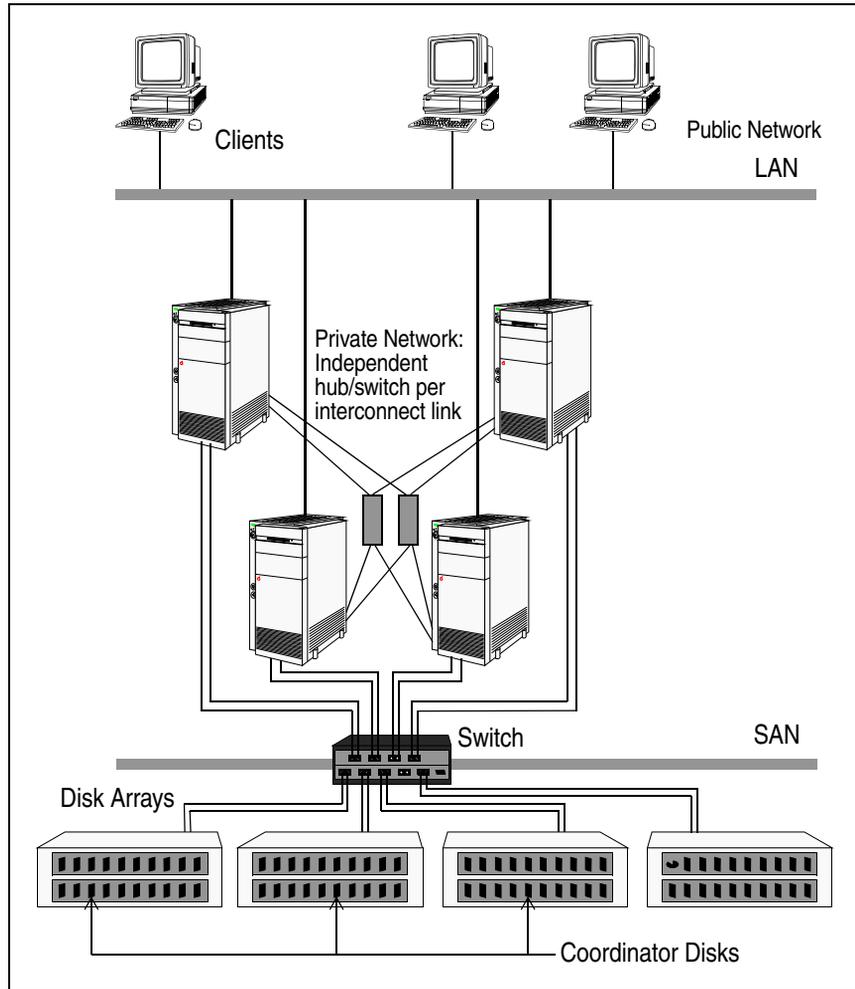
Typical SF Oracle RAC cluster setup

Figure 2-6 depicts a high-level view of an SF Oracle RAC configuration for a four-node cluster.

From a high-level, if you install SF Oracle RAC with Oracle 10g and create a database, the SF Oracle RAC cluster typically has the following characteristics:

- Nodes connected by at least two VCS private network links using 100 Base T or Gigabit Ethernet controllers on each system.
 For maximum performance, Symantec recommends the use of switches over hubs. In either case, use a minimum of two switches or hubs to provide necessary redundancy.
 If multiple links are present on a single switch, such as cases where three or four links are configured, a separate VLAN must be constructed for each link. Symantec does not support the use of multiple links on a single hub.
- Nodes connected to shared storage devices through Fibre Channel switch. Symantec does not support the use of shared SCSI with the SF Oracle RAC product. For a complete list of supported Fibre Channel storage devices, see the current hardware compatibility list on the Symantec Support Web site. See http://www.symantec.com/business/support/assistance_care.jsp.
- Oracle RAC database is configured on the shared storage that is available to each node. The shared storage could be cluster file system or raw volumes. All shared storage, including coordinator disks, must support SCSI-3 PR.
- VCS is configured to enable agents to direct and manage the resources required by Oracle RAC; these resources run in parallel on each node.

Figure 2-6 Example four-node cluster running SF Oracle RAC



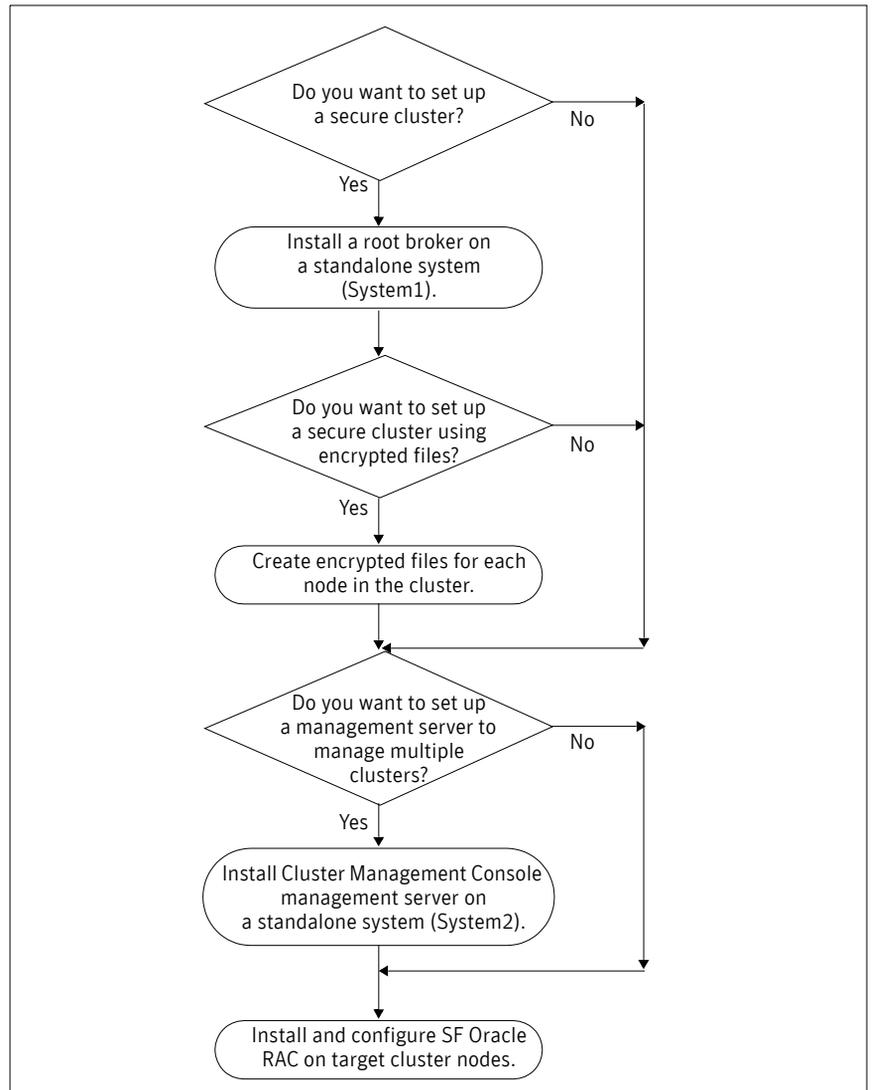
Preparing SF Oracle RAC cluster setup for optional features

After planning the SF Oracle RAC features that you want to configure, you must prepare to configure these features.

See [“About SF Oracle RAC component features”](#) on page 38.

[Figure 2-7](#) represents the major tasks and decisions required to install and configure SF Oracle RAC.

Figure 2-7 Workflow for fresh install of SF 5.0 for Oracle RAC



Complete the following preparatory tasks based on the SF Oracle RAC features you want to configure:

- [“Installing root broker for Symantec Product Authentication Service”](#) on page 50
- [“Creating encrypted files for Symantec Product Authentication Service”](#) on page 51

Installing root broker for Symantec Product Authentication Service

Install the root broker only if you plan on using Symantec Product Authentication Service. The root broker administrator must install and configure the root broker before you configure the Authentication Service for SF Oracle RAC. Symantec recommends that you install the root broker on a stable system that is outside the cluster. You can install the root broker on an AIX, HP-UX, Linux, or Solaris system. See *Symantec Product Authentication Service Installation Guide* for more information. You can configure the Authentication Service during or after SF Oracle RAC installation.

See [“Symantec Product Authentication Service”](#) on page 38.

To install the root broker

- 1 Change to the directory where you can start the `installsfrac` program:

```
# cd cluster_server
```
- 2 Start the Root Broker installation program:

```
# ./installsfrac -security
```
- 3 Select to install the Root Broker from the three choices that the installer presents:

```
[3] Install Symantec Product Authentication Service Root Broker.
```
- 4 Enter the name of the system where you want to install the Root Broker.
Enter the system name on which to install Symantec Product Authentication Service: **venus**
- 5 Review the output as the installer:
 - checks to make sure that the SF Oracle RAC supports the operating system
 - verifies that you are installing from the global zone (only on Solaris)
 - checks if the system already runs the security depot
- 6 Review the output as the `installsfrac` program checks for the installed depots on the system.
The `installsfrac` program lists the depots that will be installed on the system. Press Enter to continue.

- 7 Review the output as the installer installs the root broker on the system.
- 8 Enter **y** when the installer prompts you to configure the Symantec Product Authentication Service.
- 9 Enter a password for the root broker. Make sure the password contains a minimum of five characters.
- 10 Enter a password for the authentication broker. Make sure the password contains a minimum of five characters.
- 11 Press Enter to start the Authentication Server processes.

```
Do you want to start Symantec Product Authentication Service
processes now? [y,n,q] y
```
- 12 Review the output as the installer starts the Authentication Service.
- 13 If you plan to configure the Authentication Service during SF Oracle RAC installation, choose to configure the cluster in secure mode when the installer prompts you.
See [“Configuring SF Oracle RAC components”](#) on page 87.

Creating encrypted files for Symantec Product Authentication Service

Create encrypted files only if you plan on choosing the semiautomatic mode that uses an encrypted file to configure the Authentication Service. The encrypted files must be created by the administrator on the root broker node. The administrator must create encrypted files for each node that would be a part of the cluster before you configure the Authentication Service for SF Oracle RAC. See *Veritas Cluster Server User’s Guide* for more information. You can configure the Authentication Service during or after SF Oracle RAC installation.

See [“Symantec Product Authentication Service”](#) on page 38.

The example procedure assumes venus as the root broker node. The example procedure creates encrypted files for nodes galaxy and nebula that would form the SF Oracle RAC cluster rac_cluster101.

To create encrypted files

- 1 Determine the root broker domain name. Enter the following command on the root broker system:

```
venus> # vssat showalltrustedcreds
```

For example, the domain name would resemble
“Domain Name: root@venus.symantecexample.com” in the output.
- 2 For each node in the cluster, make sure that you have created an account on root broker system.
For example, to verify on node galaxy:

```
venus> # vssat showprpl --pdrtype root \
--domain root@venus.symantecexample.com --prplname galaxy
```

- If the output displays the principal account on root broker for the authentication broker on the node, then delete the existing principal accounts. For example:

```
venus> # vssat deleteprpl --pdrtype root \
--domain root@venus.symantecexample.com \
--prplname galaxy --silent
```

- If the output displays an error similar to “Failed To Get Attributes For Principal,” then the account for given authentication broker is not created on this root broker. Proceed to [step 3](#).

- 3 Create a principal account for each authentication broker in the cluster. For example:

```
venus> # vssat addprpl --pdrtype root --domain \
root@venus.symantecexample.com --prplname galaxy \
--password password --prpltype service
```

You must use this password that you create in the input file for the encrypted file.

- 4 Make a note of the following information that is required for the input file for the encrypted file.

- hash - The root hash string that consists of 40 characters, as shown by the command:

```
venus> # vssat showbrokerhash
```

- identity - Authentication broker identity
The value that you provide for `--prplname` in [step 3](#) (for example, galaxy).
- password - Authentication broker password
The value that you provide for `--password` in [step 3](#).
- root_domain - Domain name of the root broker system
The value that you determined in [step 1](#).
- broker_admin_password - Authentication broker password for Administrator account on the node
Provide a password of at least five characters long.

- 5 For each node in the cluster, create the input file for the encrypted file. The installer presents the format of the input file for the encrypted file when you proceed to configure the Authentication Service using encrypted file. For example, the input file for authentication broker on galaxy would resemble:

```
[setuptrust]
broker=venus.symantecexample.com
hash=758a33dbd6fae751630058ace3dedb54e562fe98
securitylevel=high
```

```
[configab]
identity=galaxy
password=password
root_domain=vx:root@venus.symantecexample.com
root_broker=venus.symantecexample.com:2821
broker_admin_password=ab_admin_password
start_broker=true
enable_pbx=false
```

- 6 Back up these input files that you created for the authentication broker on each node in the cluster.

Note that for security purposes, the command to create the output file for the encrypted file deletes the input file.

- 7 For each node in the cluster, create the output file for the encrypted file from the root broker system using the following command.

```
RootBroker> # vssat createpkg --in /path/to/blob/input/file.txt
--out /path/to/encrypted/blob/file.txt --host_ctx AB-hostname
```

For example:

```
venus> # vssat createpkg --in /tmp/galaxy.blob.in \
--out /tmp/galaxy.blob.out --host_ctx galaxy
```

Note that this command creates a encrypted file even if you provide wrong password for “password=” entry, but the encrypted file will fail to install on authentication broker node.

- 8 After you complete creating output files for the encrypted file, you must copy these files to the installer node.
- 9 After you have created the encrypted file, you can start the SF Oracle RAC installation and choose to configure the cluster in secure mode. See “[Configuring SF Oracle RAC components](#)” on page 87.

About CVM and CFS in an SF Oracle RAC environment

Before installing SF Oracle RAC, you can review concepts on CVM and CFS to better understand the overall setup and plan your SF Oracle RAC configuration.

- [About CVM](#)
- [About CFS](#)
- [Coordinating CVM and CFS configurations](#)
- [About shared disk groups](#)
- [About raw volumes versus CFS for data files](#)

About CVM

Review CVM configuration differences from VxVM and CVM recovery operations.

See “[Cluster Volume Manager](#)” on page 27.

CVM configuration differences

CVM configuration differs from VxVM configuration in these areas:

- Configuration commands occur on the master node.
- Disk groups are created (could be private) and imported as shared disk groups.
- Disk groups are activated per node.
- Shared disk groups are automatically imported when CVM starts.

CVM recovery

When a node leaves a cluster, it can leave some mirrors in an inconsistent state. The membership change is communicated through GAB to the `vxconfigd` daemon, which automatically calls the `vxrecover` utility with the `-c` option when necessary.

CVM supports both the FastResync option and dirty region logging (DRL) as optional features to improve resynchronization performance. FastResync improves performance when reorganizing volumes (moving, splitting, and joining disk groups). This is useful when performing off-host processing. DRL speeds up resynchronization after a node failure.

Special considerations exist when using the DRL in an SF Oracle RAC environment. As in a non-clustered environment, the DRL in clusters exists on a log subdisk in a mirrored volume. The size of the DRL in clusters is typically larger than in non-clustered systems. The log size depends on the volume size and the number of nodes. The `vxassist` command automatically imports a sufficiently large DRL.

You can reimport a private disk group as a shared disk group but the DRL for any mirrored volume in the disk group is probably too small to accommodate maps for all the cluster nodes. Adding nodes to the cluster can also result in too small a log size. In this situation, VxVM marks the log invalid and performs full volume recovery instead of using DRL.

About CFS

Review CFS File System benefits, CFS configuration differences from VxFS and CFS recovery operations.

See “[Cluster File System](#)” on page 28.

CFS file system benefits

Many features available in VxFS do not come into play in an SF Oracle RAC environment because ODM handles such features. CFS adds such features as high availability, consistency and scalability, and centralized management to VxFS. Using CFS in an SF Oracle RAC environment provides the following benefits:

- Increased manageability, including easy creation and expansion of files
 Without a file system, you must provide Oracle with fixed-size partitions. With CFS, you can grow file systems dynamically to meet future requirements.
- Less prone to user error
 Raw partitions are not visible and administrators can compromise them by mistakenly putting file systems over the partitions. Nothing exists in Oracle to prevent you from making such a mistake.
- Data center consistency
 If you have raw partitions, you are limited to a RAC-specific backup strategy. CFS enables you to implement your backup strategy across the data center.

CFS configuration differences

The first node to mount a CFS file system as shared becomes the primary node for that file system. All other nodes are “secondaries” for that file system.

Use the `fsclustadm` command from any node to view which node is primary and set the CFS primary node for a specific file system.

Mount the cluster file system individually from each node. The `-o cluster` option of the `mount` command mounts the file system in shared mode, which means you can mount the file system simultaneously on mount points on multiple nodes.

When using the `fsadm` utility for online administration functions on VxFS file systems, including file system resizing, defragmentation, directory reorganization, and querying or changing the `largefiles` flag, run `fsadm` from the primary node. This command fails from secondaries.

CFS recovery

The `vxfsckd` daemon is responsible for ensuring file system consistency when a node crashes that was a primary node for a shared file system. If the local node is a secondary node for a given file system and a reconfiguration occurs in which

this node becomes the primary node, the kernel requests `vxfsckd` on the new primary node to initiate a replay of the intent log of the underlying volume. The `vxfsckd` daemon forks a special call to `fsck` that ignores the volume reservation protection normally respected by `fsck` and other VxFS utilities. The `vxfsckd` can check several volumes at once if the node takes on the primary role for multiple file systems.

After a secondary node crash, no action is required to recover file system integrity. As with any crash on a file system, internal consistency of application data for applications running at the time of the crash is the responsibility of the applications.

Coordinating CVM and CFS configurations

After installing SF Oracle RAC, a VCS cluster attribute (`HacliUserLevel`) is set to give root the ability to run commands on remote systems by way of the cluster interconnect. CFS takes advantage of this mechanism to enable you to perform file system operations requiring the primary node be initiated on secondary nodes and carried out on the primary node transparently.

If you reset this attribute, be aware of which node is the primary for certain file system operations and perform those tasks from that node. Unlike a non-RAC environment, you cannot run a sequence of VxVM and VxFS commands, such as resizing a volume and a file system, on the same node unless it is both the CVM master and CFS primary node.

About shared disk groups

Make sure you review the following general information when dealing with disk groups and volumes. Refer to the *Veritas Volume Manager Administrator's Guide* for complete details on creating and managing shared disk groups.

Viewing information on a disk group

To display information about a specific disk group, type:

```
vx dg list disk_group
```

Checking the connectivity policy on a shared disk group

By default, the connectivity policy for a shared disk group is set to “global.” This setting protects against possible data corruption and causes all nodes in the cluster to detach from the disk group when any node reports a disk failure.

The output of the `vx dg list shared_disk_group` command includes the following line:

```
detach-policy: global
```

To change the connectivity policy for a disk group from “local” to “global,” type:

```
# vxedit set diskdetpolicy=global shared_disk_group
```

Determining whether a node is CVM master or slave

On one node (nebula), determine whether the node is the master or slave:

```
# vxdctl -c mode
```

On nebula, which is the slave, the output shows:

```
mode: enabled: cluster active - SLAVE
master: galaxy
```

On galaxy, which is the master, the output shows:

```
mode: enabled: cluster active - MASTER
master:galaxy
```

Enabling write access to volumes in the disk groups

By default, the activation mode for shared disk groups is inactive (set to `off`). To create databases on the shared volumes, enable the write access to the volumes:

On the CVM master node, enter:

```
vx dg -s import shared_disk_group
vx vol -g shared_disk_group startall
vx dg -g shared_disk_group set activation=sw
```

On the slave nodes, enter:

```
vx dg -g shared_disk_group set activation=sw
```

Refer to the description of disk group activation modes in the *Veritas Volume Manager Administrator’s Guide* for more information.

Deporting and importing shared disk groups

Shared disk groups in an SF Oracle RAC environment are configured for “Autoimport” at the time of CVM startup. If the user manually deports the shared disk group on the CVM master, the disk group is deported on all nodes. To reimport the disk group, the user must import the disk group as a shared group from the CVM master.

To deport a shared disk group, use the following command on the CVM master:

```
vx dg deport shared_disk_group
```

To import a shared disk group, use the following command on the CVM master:

```
vx dg -s import shared_disk_group
```

To import a disk group as a standalone disk group, deport it from the CVM master and use the following command on any node:

```
vx dg -C import shared_disk_group
```

To reimport a disk group as a shared disk group, deport it from the standalone node and use the following command on the CVM master node:

```
vx dg -C -s import shared_disk_group
```

Reviewing limitations of shared disk groups

The cluster functionality of VxVM (CVM) does not support RAID-5 volumes or task monitoring for shared disk groups in a cluster. These features can function in private disk groups attached to specific nodes of a cluster. Online relayout is available provided it does not involve RAID-5 volumes.

The boot disk group (usually aliased as `bootdg`) is a private group that cannot be shared in a cluster.

CVM only provides access to raw device; it does not support shared access to file systems in shared volumes unless you install and configure the appropriate software, such as Veritas Cluster File System (CFS). If a shared disk group contains unsupported objects, deport the group and reimport it as a private group on any node. Reorganize the volumes into layouts supported for shared disk groups, and then deport and reimport the group as a shared one.

About raw volumes versus CFS for data files

Keep these points in mind about raw volumes and CFS for data files:

- If you use file-system-based data files, the file systems containing these files must be located on shared disks. Create the same file system mount point on each node.
- If you use raw devices, such as VxVM volumes, set the permissions for the volumes to be owned permanently by the database account.

For example, type:

```
# vxedit -g dgname set group=oracle owner=oracle mode 660 \  
/dev/vx/rdisk/dgname/volume_name
```

VxVM sets volume permissions on import. The VxVM volume, and any file system that is created in it, must be owned by the Oracle database account.

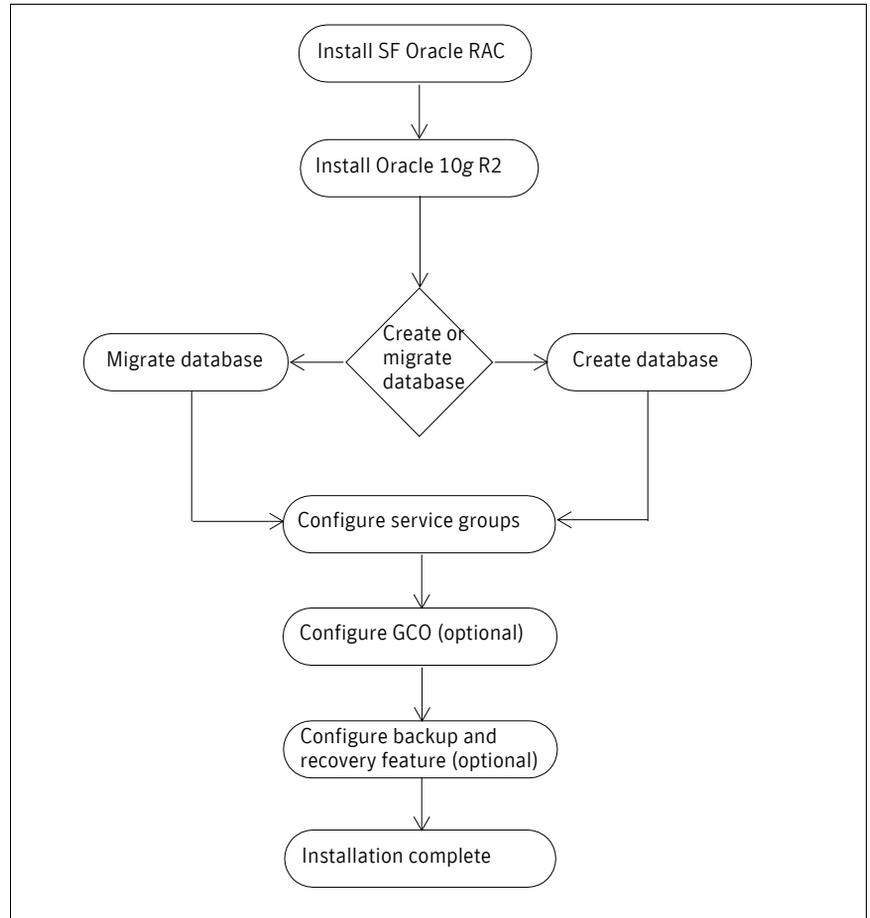
Overview of SF Oracle RAC installation and configuration tasks

Phases involved in installing and configuring SF 5.0 for Oracle RAC include:

- [Preparing to install and configure SF Oracle RAC](#)
- [Installing SF Oracle RAC and configuring its components](#)
- [Installing Oracle RAC and creating Oracle RAC database](#)
- [Setting up VCS to manage RAC resources](#)
- [Setting up disaster recovery in SF Oracle RAC environment \(optional\)](#)
- [Setting up backup and recovery feature for SF Oracle RAC \(optional\)](#)

Figure 2-8 describes a high-level flow of the SF 5.0 for Oracle RAC installation and configuration process.

Figure 2-8 SF Oracle RAC installation and configuration flow chart



Preparing to install and configure SF Oracle RAC

Before installing or upgrading SF Oracle RAC, you must:

- Make sure you meet the installation requirements.
See “[Installation requirements](#)” on page 43.
- Set up the basic hardware.
Details about supported hardware are on the Veritas support web site:
<http://entsupport.symantec.com/docs/283161>

Decide the optional features you want to set up and plan your configuration.
See [“Typical SF Oracle RAC cluster setup”](#) on page 47.

- Perform the SF Oracle RAC pre-installation and pre-configuration tasks.
Gather the required information to install and configure SF Oracle RAC.
See [“Preparing to install and configure SF Oracle RAC”](#) on page 65.

Installing SF Oracle RAC and configuring its components

Install SF Oracle RAC on clusters of up to eight nodes; note that a global cluster environment only allows a maximum of two nodes per cluster. In addition to installing SF 5.0 for Oracle RAC, you can also upgrade an existing cluster to SF 5.0 for Oracle RAC.

See [“Installing and configuring SF Oracle RAC components”](#) on page 77.

See [“Upgrading SF Oracle RAC”](#) on page 117.

Installing and configuring SF Oracle RAC involves:

Installing SF Oracle RAC Use the Veritas product installer or the `installsfrac` program.

On each node, the interactive installer installs depots for:

- Veritas Cluster Server (VCS)
- Veritas Volume Manager (VxVM)
- Veritas File System (VxFS)
- Veritas Cluster Server Agent for Oracle
- Veritas Volume Replicator (VVR)
- Other SF Oracle RAC modules

Performing system checks to configure SF Oracle RAC Use the Veritas product installer or `installsfrac -configure` option of `installsfrac` program.

The installer guides you to perform basic system checks.

Configuring SF Oracle RAC stack Use the Veritas product installer or `installsfrac -configure` option of `installsfrac` program.

After you perform basic system checks, you can configure the SF Oracle RAC components:

- VCS
- CVM, Veritas Volume Manager enabled for clusters
- CFS, Veritas File System enabled for clusters
- VVR (optional)

The installer also starts the SF Oracle RAC processes.

- | | |
|------------------------|--|
| Setting up I/O fencing | <p>Manually configure I/O fencing feature for SF Oracle RAC:</p> <ul style="list-style-type: none"> ■ Set up coordinator disks for the I/O fencing feature into a disk group. ■ Set the UseFence=SCSI3 attribute in the configuration file. ■ Shut down and restart the node. |
|------------------------|--|

Installing Oracle RAC and creating Oracle RAC database

After installing and configuring components of Storage Foundation for Oracle RAC, proceed to install Oracle RAC. SF 5.0 for Oracle RAC supports Oracle 10g R2.

- Prepare to install Oracle RAC
See [“Preparing to install Oracle 10g RAC”](#) on page 129.
- Install Oracle RAC.
See [“Installing Oracle 10g RAC”](#) on page 147.
- Create a raw database on raw volumes within a VxVM disk group or on a Veritas cluster file system.
Numerous procedures exist for creating a database. If you decide to use the Oracle dbca utility, review the procedure to create a database.
See [“Creating a starter database”](#) on page 381.

Setting up VCS to manage RAC resources

SF Oracle RAC provides the capability to completely automate the RAC environment. This capability ranges from enabling automatic control of the entire database environment to having VCS mount cluster file systems or enable CVM and CFS daemons. The user or DBA is free to choose the level of control and automation.

VCS uses the main.cf configuration file to manage resources in the cluster. The SF Oracle RAC installation process creates a basic VCS configuration file. After installing Oracle and creating the database, you can modify the main.cf file on one of the cluster nodes to reflect the new resources and their configuration.

You can configure VCS service groups using the configuration wizard or manually.

See [“Configuring VCS service groups for Oracle 10g”](#) on page 161.

See [“Sample VCS configuration files for SF Oracle RAC”](#) on page 355.

Setting up disaster recovery in SF Oracle RAC environment (optional)

You can create a global cluster environment with SF Oracle RAC and volume replication capability. VCS provides the Global Cluster Option (GCO) for wide-area failover and disaster recovery, and Veritas Volume Replicator provides the volume replication capability.

After installing SF Oracle RAC on each node in the cluster, you can choose to set up a global cluster environment for disaster recovery. The general process for setting up a global cluster involves the following tasks on the secondary site:

- Creating a cluster on a secondary site with hardware set up
- Installing SF Oracle RAC
- Installing Oracle RAC
- Configuring VCS service groups

You do not need to create a database for the secondary site, the database will be replicated from the primary site.

See [“Preparing for global clustering”](#) on page 199.

See [“Configuring global clustering”](#) on page 207.

Setting up backup and recovery feature for SF Oracle RAC (optional)

You can configure the following SF Oracle RAC optional features to back up and recover data at the volume and file system levels:

- Veritas Storage Checkpoint
Allows efficient backup and recovery of Oracle RAC databases. This feature is available with SF Oracle RAC as part of the Veritas File System.
See [“Using Storage Checkpoints and Storage Rollback”](#) on page 243.
- Veritas Database FlashSnap
Allows you to create a point-in-time copy of an Oracle RAC database for backup and off-host processing.
See [“Using FlashSnap for backup and recovery”](#) on page 271.
- Veritas Storage Mapping
Allows you to evaluate or troubleshoot I/O performance. You can access mapping information that allows for a detailed understanding of the storage hierarchy in which files reside.
See [“Investigating I/O performance for SF Oracle RAC: Storage Mapping”](#) on page 325.

SF Oracle RAC installation and upgrade

- [Chapter 3, “Preparing to install and configure SF Oracle RAC” on page 65](#)
- [Chapter 4, “Installing and configuring SF Oracle RAC components” on page 77](#)
- [Chapter 5, “Upgrading SF Oracle RAC” on page 117](#)

Preparing to install and configure SF Oracle RAC

This chapter contains the following topics:

- [About preparing to install and configure SF 5.0 for Oracle RAC](#)
- [Preparing to install SF 5.0 for Oracle RAC](#)
- [Gathering information to install and configure SF Oracle RAC](#)

About preparing to install and configure SF 5.0 for Oracle RAC

Before you perform the pre-installation tasks, make sure you reviewed the installation requirements, set up the basic hardware, and planned your SF Oracle RAC setup.

See [“Planning SF Oracle RAC installation and configuration”](#) on page 37.

Preparing to install SF 5.0 for Oracle RAC

Perform the following tasks before proceeding to install SF 5.0 for Oracle RAC:

- [Obtaining SF Oracle RAC license keys](#)
- [Synchronizing time settings on cluster nodes](#)
- [Setting up inter-system communication](#)
- [Setting up shared storage](#)
- [Setting the PATH variable](#)
- [Setting the MANPATH variable](#)

- [Mounting the product disc](#)
- [Verifying the systems before installation](#)

Obtaining SF Oracle RAC license keys

SF Oracle RAC includes a License Key certificate. The certificate specifies the product keys and the number of product licenses purchased. A single key lets you install the product on the number and type of systems for which you purchased the license. A key may enable the operation of more products than are specified on the certificate; however, you are legally limited to the number of product licenses purchased.

The product installation procedure includes instructions on how to activate the key. If you encounter problems while licensing this product, visit the Symantec licensing support website at:

<http://www.veritas.com/buy/vLicense/vLicenseHome.jhtml>

See “Installing SF 5.0 for Oracle RAC” on page 81.

The VRTSvlic depot enables product licensing. After the VRTSvlic is installed, the following commands and their manual pages are available on the system:

<code>vxlicinst</code>	Installs a license key for a Symantec product
<code>vxlicrep</code>	Displays currently installed licenses
<code>vxlictest</code>	Retrieves features and their descriptions encoded in a license key

Even though other products are included on the enclosed software discs, you can only install the Symantec software products for which you have purchased a license.

Synchronizing time settings on cluster nodes

If you plan to configure the SF Oracle RAC in secure mode, Symantec recommends all cluster nodes have the same time. If you do not run the Network Time Protocol (NTP) daemon, make sure to synchronize the time settings on each node.

Setting up inter-system communication

When you install SF Oracle RAC using the `installsfrac` program, to install and configure the entire cluster at one time, make sure that communication between systems exists. By default, the installer uses `ssh` for inter-system communication. You must grant permissions for the system where you run

installsfrac program to issue `ssh` or `remsh` commands as root on all systems in the cluster.

If `ssh` is used for communication, you must configure it in a way such that it operates without requests for passwords or passphrases.

If `remsh` is used for communication, on each node, place a "+" character in the first line of the `.rhosts` file to give remote access to the system running the install program. You can limit the remote access to specific nodes. Refer to the manual page for the `rhosts` file for more information. Remove the remote `remsh` access permissions after the installation and disk verification process.

Refer to the *Veritas Storage Foundation and High Availability Solutions Getting Started Guide* for more information on setting up `ssh` or `remsh` communication.

Mounting the product disc

You must have superuser (root) privileges to load the SF Oracle RAC software.

You can unmount the product disc after completing your SF Oracle RAC installation and configuration.

To mount the product disc using mount command

- 1 Log in as superuser to a cluster node or a remote node in the same subnet as the cluster nodes.

- 2 Insert the product disc in the appropriate drive on your local system.

- 3 Determine the block device file for the disc drive:

```
# ioscand -fnC disk
```

Make a note of the device file as it applies to your system.

- 4 Create a directory in which to mount the software disc and mount the disc using the appropriate drive name. For example:

```
# mkdir -p /cdrom
# mount /dev/vx/rdmp/c0t0d0 /cdrom
```

- 5 Verify that the disc is mounted:

```
# mount
```

Setting the PATH variable

You can find the installation and other commands located in various directories. Add these directories to your `PATH` environment variable on each system.

To set the PATH variable

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

```
Bourne Shell (sh or ksh) # PATH=/usr/sbin:/sbin:/usr/bin:\
                          /usr/lib/vxvm/bin:/opt/VRTSvxfs/cfs/bin:\
                          /opt/VRTSvcs/bin:/opt/VRTS/bin:\
                          /opt/VRTSvcs/rac/bin:/opt/VRTSob/bin:\
                          $PATH; export PATH
```

```
C Shell (csh)             # setenv PATH /usr/sbin:/sbin:/usr/bin:\
                          /usr/lib/vxvm/bin:/opt/VRTSvxfs/cfs/bin:\
                          /opt/VRTSvcs/bin:/opt/VRTS/bin:\
                          /opt/VRTSvcs/rac/bin:/opt/VRTSob/bin:$PATH
```

Do not define paths for the root user to a cluster file system in the *LD_LIBRARY_PATH* variable. You can define *\$ORACLE_HOME/lib* in *LD_LIBRARY_PATH* for the oracle user.

The path defined as */opt/VRTSob/bin* is optional unless you choose to install Veritas Enterprise Administrator.

Setting the MANPATH variable

Set the MANPATH variable to enable viewing manual pages.

To set the MANPATH variable

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

```
Bourne Shell (sh or ksh) # MANPATH=/usr/share/man:/opt/VRTS/man;\
                          export MANPATH
```

```
C Shell (csh)           # setenv MANPATH /usr/share/man:\
                          /opt/VRTS/man
```

Setting up shared storage

You need to set up shared storage so that it is visible to the SCSI layer from all the nodes in the cluster. The shared storage that you add for use with SF Oracle RAC software must support SCSI-3 persistent reservations, a functionality that enables the use of I/O fencing.

See [“About I/O fencing”](#) on page 33.

See [“Checking shared disks for I/O fencing”](#) on page 101.

Removing pre-existing license keys

If you have pre-existing license keys, clean those keys.

See [“Obtaining SF Oracle RAC license keys”](#) on page 66.

To remove pre-existing license keys

- 1 View license key files currently installed on a node:

```
# cd /opt/VRTS/bin
# ./vxlicrep
```

The output lists the license keys and information about their respective products.

- 2 Navigate to the directory containing the license key files and list the files.

```
# cd /etc/vx/licenses/lic
# ls -a
```

- 3 Move the license key files to another location for future reference.

For example, in the directory containing the license key files, create a sub-directory and move the files to that new location:

```
# mkdir OLD
# mv * OLD
```

Verifying the systems before installation

Before beginning the installation of SF Oracle RAC software, you can check the readiness of the systems where you plan to install SF Oracle RAC. The command to start the preinstallation check is:

```
installsfrac -precheck system1 system2 ...
```

To verify the systems

- 1 Navigate to the folder containing the installsfrac program.

```
# cd /cdrom/storage_foundation_for_oracle_rac
```

- 2 Start the pre-installation check:

```
# ./installsfrac -precheck galaxy nebula
```

The program proceeds in a non-interactive mode, examining the systems for licenses, depots, disk space, and system-to-system communications. The program displays the results of the check and saves the results of the check in a log file.

Gathering information to install and configure SF Oracle RAC

The SF Oracle RAC installation and configuration program prompts you for information about some SF Oracle RAC components. The program provides default values for some information, which you can choose to use. Keep the following required information at hand.

[Table 3-3](#) lists the required information for installing SF Oracle RAC and configuring each SF Oracle RAC component.

Table 3-3 Required information to install and configure SF Oracle RAC

Required information	Description
To install SF Oracle RAC depots	
System names on which to install SF Oracle RAC	Example: galaxy, nebula
License keys	<p>License keys could be one of the following types:</p> <ul style="list-style-type: none"> ■ Valid license keys for each system in the cluster ■ Valid site license key ■ Valid demo license key <p>If you want to configure global clusters and enable disaster recovery, you must enter appropriate license keys.</p> <p>See “Obtaining SF Oracle RAC license keys” on page 66.</p>
Do you want to install required SF Oracle RAC depots or all SF Oracle RAC depots?	<p>Install only the required depots if you do not want to configure any optional components or features.</p> <p>Default option is to install all depots.</p>
To configure Veritas Cluster Server component	
Name of the cluster	<p>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 (_).</p> <p>Example: rac_cluster101</p>

Table 3-3 Required information to install and configure SF Oracle RAC

Required information	Description
Unique ID number for the cluster	Number in the range of 0-65535. Within the site containing the cluster, each cluster must have a unique ID. Example: 101
Device names of the NICs used by the private networks among systems	Do not enter the network interface card that is used for the public network, which is typically lan0. Example: 1an1, 1an2
To configure SF Oracle RAC clusters in secure mode (optional)	
Which mode do you want to choose to configure Authentication Service?	The installer provides you three modes to configure Authentication Service in the SF Oracle RAC clusters. <ul style="list-style-type: none"> ■ automatic mode ■ semiautomatic mode using encrypted files ■ semiautomatic mode without using encrypted files Default option is automatic mode. See “Symantec Product Authentication Service” on page 38.
Host name of the Symantec Product Authentication Service Root Broker System	Example: venus
To add SF Oracle RAC users (required if you did not configure the cluster in secure mode)	
User name	Example: smith
User password	Enter the password at the prompt.
User privilege	Users have three levels of privileges: A=Administrator, O=Operator, or G=Guest. Example: A
To configure the Cluster Management Console cluster connector (optional)	
Management Server network address for Cluster Management Console	Example: mgmtserver1.symantecexample.com

Table 3-3 Required information to install and configure SF Oracle RAC

Required information	Description
Cluster Management Console service account password	You must have set this account password while installing the management server.
Root hash of the management server	You can use <code>vssat showbrokerhash</code> command and copy the root hash of the management server.

Table 3-3 Required information to install and configure SF Oracle RAC

Required information	Description
To configure the Cluster Management Console (optional)	
Name of the public NIC for each node in the cluster	The device name for the NIC that provides public network access. Example: 1an0
Virtual IP address of the NIC for Cluster Management Console	This virtual IP address becomes a resource for use by the ClusterService group that includes the Cluster Management Console. The "Cluster Virtual IP address" can fail over to another cluster system, making the Web Console highly available. Example: 10.10.12.1
Netmask for the virtual IP address	The subnet used with the virtual address. Example: 255.255.240.0
NetworkHosts IP addresses to check the connection	This IP address is used to check the adapter connections. Example: 10.10.12.12
To configure SMTP email notification (optional)	
Domain-based address of the SMTP server	The SMTP server sends notification email about the events within the cluster. Example: smtp.symantecexample.com
Email address of each SMTP recipient to be notified	Example: john@symantecexample.com
Minimum severity of events for SMTP email notification	Events have four levels of severity: I=Information, W=Warning, E=Error, and S=SevereError. Example: E
To configure SNMP trap notification (optional)	
Port number for the SNMP trap daemon	Default port number is 162.
Machine name for each SNMP console	Example: neptune
Minimum severity of events for SNMP trap notification	Events have four levels of severity: I=Information, W=Warning, E=Error, and S=SevereError. Example: E

Table 3-3 Required information to install and configure SF Oracle RAC

Required information	Description
To configure global cluster option (optional)	
Name of the public NIC	You may use the same NIC that you configured for the ClusterService group. Otherwise, you must specify appropriate values for the NIC when you are prompted. Example: 1an0
Virtual IP address of the NIC	You may use the same virtual IP address that you configured for the ClusterService group. Otherwise, you must specify appropriate values for the virtual IP address when you are prompted. Example: 10.10.12.1
Netmask for the virtual IP address	You may use the same netmask as configured for the ClusterService group. Otherwise, you must specify appropriate values for the netmask when you are prompted. Example: 255.255.240.0
NetworkHosts IP addresses	You may use the same NetworkHosts IP address as configured for the ClusterService group. Otherwise, you must specify appropriate values for the netmask when you are prompted. Example: 10.10.12.2
To configure Cluster Volume Manager	
CVM cluster reconfiguration timeout in seconds	Default is 200.
Do you want to configure enclosure-based naming scheme?	Dynamic Multipathing (DMP) is a prerequisite for enclosure-based naming schemes.
To configure Veritas Volume Replicator (optional)	
Frequency of VVR statistics collection	Default is 10 seconds.
Number of days to preserve the collected statistics	Default is 3 days.

Table 3-3 Required information to install and configure SF Oracle RAC

Required information	Description
To configure I/O fencing	
DMP nodes names for each disk in the coordinator disk group	Example: <code>/dev/vx/rdmp</code>

Installing and configuring SF Oracle RAC components

This chapter contains the following topics:

- [About installing and configuring SF Oracle RAC](#)
- [Invoking SF Oracle RAC installation program](#)
- [Installing SF 5.0 for Oracle RAC](#)
- [Invoking SF Oracle RAC configuration program](#)
- [Performing basic system checks](#)
- [Configuring SF Oracle RAC components](#)
- [Setting up I/O fencing for SF Oracle RAC](#)
- [Unmounting the product disc](#)
- [Verifying SF Oracle RAC installation using VCS configuration file](#)

About installing and configuring SF Oracle RAC

You can install SF 5.0 for Oracle RAC on clusters of up to eight nodes. Use one of the following tools to install and configure SF Oracle RAC:

Veritas product installer	<p>Offers a high-level approach to installing or configuring multiple Veritas products.</p> <p>You can use the Veritas product installer to install Symantec Product Authentication Service, Veritas Cluster Management Console, and Veritas Centralized Management Server. For more information on these product and features, refer to the respective product documentation.</p>
installsfrac program	<p>Offers a direct approach to specifically installing and configuring SF Oracle RAC.</p> <p>The <code>installsfrac</code> and <code>installsfrac -installonly</code> commands only install the depots. After you complete installing the depots, you must configure SF Oracle RAC using <code>installsfrac -configure</code> command.</p>

Note: If you obtained SF Oracle RAC from an electronic download site, you must use the `installsfrac` program instead of the Veritas product installer.

After installing the SF Oracle RAC depots, you must start the SF Oracle RAC configuration program through the Veritas product installer or the `installsfrac` program. The configuration program provides the following options:

- **Check system for SFRAC** - Provides guidelines to verify private interconnects for LLT and to verify the disks intended for shared storage support
- **Configure SFRAC** - Asks you a set of questions to configure SF Oracle RAC components, and provides guidelines to configure I/O fencing
- **Prepare to install Oracle** - Provides guidelines and prompts you to complete pre-installation tasks for Oracle in an SF Oracle RAC environment. See [“Preparing to install Oracle 10g RAC”](#) on page 129.
- **Install or relink Oracle** - Launches Oracle Installer to install Oracle RAC, and links Oracle with the Veritas IPC library. See [“Installing Oracle 10g RAC”](#) on page 147.

Overview of installation and configuration tasks

Table 4-4 lists the tasks to complete SF Oracle RAC installation and configuration.

Table 4-4 SF Oracle RAC installation and configuration tasks

Task	Procedural reference
Invoking SF Oracle RAC installation program	<ul style="list-style-type: none"> ■ Mounting the product disc ■ Starting software installation ■ Specifying systems for installation
Licensing SF Oracle RAC and adding SF Oracle RAC packages to each cluster node	<ul style="list-style-type: none"> ■ Licensing SF Oracle RAC ■ Installing SF Oracle RAC depots ■ Rebooting the nodes after installation
Invoking SF Oracle RAC configuration program	<ul style="list-style-type: none"> ■ Mounting the product disc ■ Starting software configuration ■ Specifying systems for configuration ■ Choosing the configuration task
Performing basic system checks	<ul style="list-style-type: none"> ■ Preparing private interconnects for LLT ■ Checking shared disks for I/O fencing
Configuring SF Oracle RAC components	<ul style="list-style-type: none"> ■ Configuring the Veritas Cluster Server and optional features ■ Setting permissions for database administration ■ Configuring the cluster volume manager ■ Configuring VVR on each node ■ Starting the VAILAgent ■ Creating SF Oracle RAC configuration files ■ Configuring standalone hosts for Storage Foundation ■ Starting SF Oracle RAC
Configuring I/O fencing	<ul style="list-style-type: none"> ■ Verifying the SF Oracle RAC configuration ■ Initializing disks ■ Setting up coordinator disk groups ■ Starting I/O fencing ■ Modifying VCS configuration to enable I/O fencing ■ Starting VCS, CVM, and CFS

The example procedure installs and configures SF Oracle RAC on two nodes: galaxy and nebula. For the sample installation, the cluster's name is rac_cluster101 and the cluster's ID is 101. The example installation chooses all optional features.

Invoking SF Oracle RAC installation program

Mount the product disc and launch the installation program. At the end of each product installation, the installer creates a new directory with three files:

- A log file containing any system commands executed, and their output.
- A response file that can be used in conjunction with the `-responsefile` option of the installer.
- A summary file containing the output of the install scripts.

Mounting the product disc

Make sure you have superuser (root) privileges to load the SF Oracle RAC software. You can use the `mount` command to mount the disc.

See [“Mounting the product disc”](#) on page 67.

Starting software installation

You can install SF Oracle RAC using the Veritas product installer or the `installsfrac` program.

To install SF Oracle RAC using the product installer

- 1 Confirm that you are logged in as the superuser.
- 2 Navigate to the folder containing the installer program.
`# cd /cdrom`
- 3 Start the installer.
`# ./installer`
The installer starts the product installation program with a copyright message and specifies the directory where the logs are created.
- 4 From the opening Selection Menu, choose: “I” for “Install/Upgrade a Product.”
- 5 From the displayed list of products to install, choose: **Veritas Storage Foundation for Oracle RAC.**

To install SF Oracle RAC using the installsfrac program

- 1 Confirm that you are logged in as the superuser.
- 2 Navigate to the folder containing the installsfrac program.

```
# cd /cdrom/storage_foundation_for_oracle_rac
```
- 3 Start the installsfrac program.

```
# ./installsfrac
```

The installer begins with a copyright message and specifies the directory where the logs are created.

Specifying systems for installation

The installer prompts for the system names on which you want to install and then performs an initial system check.

To specify system names for installation

- 1 Review the installer note on the VxVM that is bundled with HP-UX. The installer replaces any previous VxVM version with VxVM 5.0.
- 2 Confirm that you want to proceed with the installation at the prompt.
- 3 Enter the names of the systems where you want to install SF Oracle RAC.
Enter the system names separated by spaces on which to install SF Oracle RAC: **galaxy nebula**
- 4 Review the output as the installer checks that the local node running the script can communicate with remote nodes and checks whether a previous version of SF Oracle RAC is installed.
If a previous version of SF Oracle RAC is installed, the installer provides an option to upgrade.
See [“Upgrading SF Oracle RAC”](#) on page 117.

Installing SF 5.0 for Oracle RAC

SF Oracle RAC is a licensed product. During the installation, you must also add licenses for other SF Oracle RAC features that require licensing. After the program installs the depots, you must reboot the nodes and then proceed to configure SF Oracle RAC.

Licensing SF Oracle RAC

The installer checks whether SF Oracle RAC license keys are currently in place on each system. If license keys are not installed, the installer prompts you for the license keys.

To license SF Oracle RAC

- 1 Review the output as the utility checks system licensing and installs the licensing depot.
 - If the VRTSvlic depot is not present, the utility installs it on each node after checking for sufficient disk space.
 - If a previous version of the VRTSvlic depot is present, the utility replaces the depot with the current version.

- 2 Enter the license key for Veritas Storage Foundation for Oracle RAC as the installer prompts for each node.

```
Enter a SFRAC license key for galaxy: [?] XXXX-XXXX-XXXX-XXXX-XXXX-XXXX-XXX  
XXXX-XXXX-XXXX-XXXX-XXXX-XXXX-XXX successfully registered on  
galaxy  
SFRAC license registered on galaxy
```

- 3 Enter keys for additional product features such as VVR, if you want to set up a Global Cluster environment.

```
Do you want to enter another license key for galaxy? [y,n,q,?]  
(n) y
```

```
Enter a SFRAC license key for galaxy: [?] XXXX-XXXX-XXXX-XXXX-XXXX-XXXX-XXX  
XXXX-XXXX-XXXX-XXXX-XXXX-XXXX-XXX successfully registered on  
galaxy
```

```
Do you want to enter another license key for galaxy? [y,n,q,?]  
(n)
```

- 4 Enter the license keys for other nodes.

```
SFRAC is not licensed on nebula  
.  
.  
Do you want to enter another license key for nebula? [y,n,q,?]  
(n)  
SFRAC licensing completed successfully.
```

Installing SF Oracle RAC depots

The installer verifies for any previously installed depots and then based on your choice installs all the SF Oracle RAC depots or only the required depots.

To install SF Oracle RAC depots

- 1 Review the output as the installer checks the depots that are already installed.
- 2 Choose the SF Oracle RAC depots to be installed. Do one of the following:
 - Enter **1** to install only the required SF Oracle RAC depots.

- Enter **2** to install all the SF Oracle RAC depots. (default option)
You must install the optional depots for any optional feature you plan to use.
- 3 View the list of depots that the installer would install on each node.
- 4 Verify that the installation process completes successfully.
- 5 Review the output at the end of the installation and note the location of the summary and log files for future reference.
The installer also notes that the VCS cvm service group remains in a `FAULTED` state until you configure the coordinator disks.
See [“Setting up coordinator disk groups”](#) on page 107.

Rebooting the nodes after installation

Some of the depots would require a reboot before you can proceed to configuring SF Oracle RAC. When you need to reboot any of the nodes, make sure SF Oracle RAC is properly shut down by always using the `shutdown` command with the `-r` option instead of the `reboot` command.

To reboot the nodes after installation

- ◆ When the installer prompts you to reboot the nodes before configuring SF Oracle RAC, review the installer output and follow the instructions.
Execute the `/usr/sbin/shutdown -r now` command to reboot the nodes.

Invoking SF Oracle RAC configuration program

Mount the product disc and launch the configuration program. At the end of each product configuration, the installer creates a new directory with three files:

- A log file containing any system commands executed, and their output.
- A response file that can be used in conjunction with the `-responsefile` option of the installer.
- A summary file containing the output of the install scripts.

Mounting the product disc

Make sure you have superuser (root) privileges to load the SF Oracle RAC software. You can use the `mount` command to mount the disc.

See [“Mounting the product disc”](#) on page 67.

Starting software configuration

After mounting the product disc, you can configure SF Oracle RAC using the Veritas product installer or the `installsfrac` program.

To configure SF Oracle RAC using the product installer

- 1 Confirm that you are logged in as the superuser.
- 2 Navigate to the folder containing the installer program.
`# cd /cdrom`
- 3 Start the installer.
`# ./installer`
The installer starts the product installation program with a copyright message and specifies the directory where the logs are created.
- 4 From the opening Selection Menu, choose: **C** for “Configure an Installed Product.”
- 5 From the displayed list of products to install, choose: **Veritas Storage Foundation for Oracle RAC**.

To configure SF Oracle RAC using the `installsfrac` program

- 1 Confirm that you are logged in as the superuser.
- 2 Navigate to the folder containing the `installsfrac` program.
`# cd /cdrom/storage_foundation_for_oracle_rac`
- 3 Start the `installsfrac` program with the `-configure` option.
`# ./installsfrac -configure`
The installer begins with a copyright message and specifies the directory where the logs are created.

Specifying systems for configuration

The configuration program prompts for the system names on which you want to configure and then performs an initial system check for communication.

To specify system names for configuration

- 1 Confirm that you want to proceed with the configuration at the prompt.
- 2 Enter the names of the nodes where you want to configure the software.
Enter the system names separated by spaces on which to configure
SFRAC: **galaxy nebula**
- 3 Review the output as the program checks that the local node running the script can communicate with remote nodes and checks whether SF 5.0 for Oracle RAC is installed successfully.

If SF 5.0 for Oracle RAC is not installed, the program exits.

Choosing the configuration task

The configuration program provides you menu options from which you can choose the configuration task. The menu options that are marked ****INSTRUCTIONS ONLY**** require you to:

- Follow the instructions in the order mentioned
- Keep the shells on all systems open with superuser privileges

To choose the configuration task

- 1 After the program verifies the license keys for SF Oracle RAC, review the program options.
- 2 Make sure that the systems are ready for configuration. Select **Check systems for SFRAC** to perform LLT and I/O fencing checks. See [“Performing basic system checks”](#) on page 85.
- 3 After performing the system checks, select **Configure SFRAC** and proceed to setting up the stack.
- 4 Select **Configure VCS, CVM, and CFS** to configure the SF Oracle RAC components.
Review the output as the program stops various response files and processes. Answer the prompts to configure the different components and their optional features.
See [“Configuring SF Oracle RAC components”](#) on page 87.
- 5 After configuring these components, review the output as the configuration program starts SF Oracle RAC and exits.
Note that you must launch the configuration program again to perform any other task.
- 6 Select **Configure I/O fencing** if you completed configuring SF Oracle RAC components. Review the program guidelines for setting up the I/O fencing feature of SF Oracle RAC.
See [“Setting up I/O fencing for SF Oracle RAC”](#) on page 104.

Performing basic system checks

Choose **Check systems for SFRAC** from the configuration program’s menu to perform the basic system checks for SF Oracle RAC.

You can follow the SF Oracle RAC configuration program instructions to check the private interconnects for LLT at this point. However, you must check the shared disks for I/O fencing after you configure SF Oracle RAC components.

- Do not choose “Perform all the following tasks” option.
You must check the I/O fencing disks after you configure SF Oracle RAC components.
- If you choose **Check LLT links**, the program lists the prerequisites for the LLT links and prompts you to check each LLT link on all cluster nodes.
See “[Preparing private interconnects for LLT](#)” on page 86.
- Choose **Check I/O fencing disks** after you configure SF Oracle RAC components.
Follow the program guidelines to check I/O fencing disks on all nodes.
See “[Checking shared disks for I/O fencing](#)” on page 101.

Preparing private interconnects for LLT

You must attach each network interface that is configured for LLT to a separate and distinct physical network.

Optimizing LLT media speed settings on private NICs

For optimal LLT communication among the cluster nodes, the interface cards on each node must use the same media speed settings. Also, the settings for switches or hubs used for the interconnects must match that of the interface cards. Incorrect settings can cause poor network performance or even network failure.

Guidelines for setting the media speed of the LLT interconnects

- If you have hubs or switches for LLT interconnects, Symantec recommends using the Autonegotiation media speed setting on each Ethernet card on each node.
- If you have hubs or switches for LLT interconnects and you do not use the Autonegotiation media speed setting, set the hub or switch port to the same setting as that used on the cards on each node.
- If you use directly connected Ethernet links (using crossover cables), set the media speed to the highest value common to both cards, typically 100 Full-Duplex.
- Symantec does not recommend using dissimilar network cards for private links.

For details on setting the media speeds for specific devices, consult the device's documentation.

Configuring SF Oracle RAC components

Choose **Configure SFRAC > Configure VCS, CVM and CFS** from the configuration program's menu to configure the SF Oracle RAC components.

Make sure that you gathered the required information to configure the SF Oracle RAC components. Also, make sure that you performed the pre-configuration tasks if you want to configure the cluster in secure mode, centralized cluster management, or centralized Storage Foundation management.

See "[Preparing to install and configure SF Oracle RAC](#)" on page 65.

Answer the configuration program prompts to configure the SF Oracle RAC components such as VCS, CVM, and CFS.

Configuring the Veritas Cluster Server and optional features

You can configure the Veritas Cluster Server and its optional features at this time. You must have chosen to install all SF Oracle RAC depots to configure the optional VCS features. Refer to the *Veritas Cluster Server Installation Guide* for more information.

Tasks involved in configuring the cluster server include:

- [Configuring the cluster](#)
- [Configuring the cluster in secure mode](#)
- [Adding SF Oracle RAC users](#)
- [Configuring cluster connector](#)
- [Configuring the Cluster Management Console](#)
- [Configuring SMTP email notification](#)
- [Configuring SNMP trap notification](#)
- [Configuring global clusters](#)

Configuring the cluster

Enter a cluster name and ID to perform the basic cluster configuration.

To configure the cluster

- 1 After reviewing the instructions on how to respond to the configuration questions, enter the unique cluster name and ID.

```
Enter the unique cluster name: [?] rac_cluster101  
Enter the unique Cluster ID number between 0-65535: [b,?] 101
```

- 2 Review the NICs available on the first system as the installer discovers and reports them.
- 3 Enter the details for private heartbeat links.
You must not enter the network interface card that is used for the public network (typically, lan0).
- 4 Specify whether you use the same NICs on all nodes.
 - If you use the same NICs for private heartbeat links on all nodes, make sure the same NICs are available on each system and enter **y**.
 - If you use NICs with different device names on some of the nodes, enter **n**.

```
Are you using the same NICs for private heartbeat links on all systems? [y,n,q,b,?] (y)
```
- 5 Review the output as the installer checks that the interface cards on each node use the same media speed settings.
- 6 Verify the information that the program summarizes and confirm the information.

Configuring the cluster in secure mode

Before you configure a cluster in a secure mode, make sure you have installed a root broker on another stable system. Also, make sure you meet the requirements for automatic or semiautomatic mode of configuration.

See [“Symantec Product Authentication Service”](#) on page 38.

To configure the cluster in secure mode

- 1 Choose whether to configure SF Oracle RAC to use Symantec Product Authentication Service.
 - If you want to configure Authentication Service, make sure that you have installed the root broker, and answer **y**.
 - If you decline to configure Authentication Service, answer **n** and proceed to adding SF Oracle RAC users.
See [“Adding SF Oracle RAC users”](#) on page 90.

```
Would you like to configure SF Oracle RAC to use Symantec Product Authentication Service? [y,n,q] (n) y
```
- 2 Select one of the options to configure security.
 - 1) Configure security automatically
 - 2) Configure security semiautomatically using encrypted file
 - 3) Configure security semiautomatically answering prompts

Select the mode you want to use [1-3,q,?] (1)

Based on the mode of configuration you want to use, enter one of the following:

Option	Tasks
1. Automatic configuration	<p>Enter the name of the root broker system when prompted.</p> <p>Requires remote access to the root broker.</p> <p>Review the output as the installer verifies communication with the root broker system, checks vxatd process and version, and checks security domain.</p>
2. Semi-automatic using encrypted files	<p>Enter the path of the file for each node when prompted.</p>
3. Semi-automatic entering authentication information at installer prompts	<p>Enter the following root broker information as the installer prompts you:</p> <pre> Enter root Broker name: venus.symantecexample.com Enter root broker FQDN: [b] (symantecexample.com) symantecexample.com Enter root broker domain: [b] (root@venus.symantecexample.com) root@venus.symantecexample.com Enter root broker port: [b] (2821) 2821 Enter path to the locally accessible root hash [b] (/var/tmp/installvcs-1Lcljr/root_hash) /root/root_hash </pre> <p>Enter the following authentication broker information as the installer prompts you for each node:</p> <pre> Enter authentication broker principal name on galaxy [b] (galaxy.symantecexample.com) galaxy.symantecexample.com Enter authentication broker password on galaxy: Enter authentication broker principal name on nebula [b] (nebula.symantecexample.com) nebula.symantecexample.com Enter authentication broker password on nebula: </pre>

- 3 After configuring the cluster in secure mode, proceed to configure the Cluster Management Console cluster connector.
See [“Configuring cluster connector”](#) on page 90.

Adding SF Oracle RAC users

If you configured the cluster in secure mode, you need not add SF Oracle RAC users now. Proceed to configure the Cluster Management Console cluster connector. Otherwise, you can add SF Oracle RAC users at this time.

See “[Configuring the cluster in secure mode](#)” on page 88.

See “[Configuring cluster connector](#)” on page 90.

To add SF Oracle RAC users

- 1 Review the required information to add SF Oracle RAC users.
- 2 Reset the password for the Admin user, if necessary.
- 3 To add a user, enter **y** at the prompt.
- 4 Enter the user’s name, password, and level of privileges.
Enter the user name: [?] **smith**
Enter New Password: *********

Enter Again: *********
Enter the privilege for user smith (A=Administrator, O=Operator, G=Guest): [?] **a**
- 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 cluster connector

If you configured the Cluster Management Console management server to centrally manage this cluster, you can now configure cluster connector for the buffering feature. If a firewall exists between the management server and this cluster, then you must configure cluster connector to enable centralized management. Make sure you meet the prerequisites to configure cluster connector.

See “[Veritas Cluster Management Console](#)” on page 41.

To configure cluster connector

- 1 Review the information to configure Cluster Management Console.
- 2 Choose whether to configure cluster connector or not. Do one of the following:
 - To configure cluster connector on the systems, press Enter.
Do you want this cluster to be managed by a management server? Enter 'y' if you have set up a management server.
[y,n,q] (y) **y**

- To skip configuring cluster connector and advance to configuring Cluster Management Console for local cluster management, enter **n**. See “[Configuring the Cluster Management Console](#)” on page 91.
- 3 Review the required information to configure cluster connector.
 - 4 Enter the Management Server network address for Cluster Management Console.

```
Enter the network address used by the management server [?]
(galaxy) mgmtserver1.symantecexample.com
```
 - 5 Verify and confirm the management server information.
 - 6 Enter the following information that is required to securely communicate with the management server.
 - Password for the service account that is created during the management service installation
 - Hash of Cluster Management Console management server's root broker
 - 7 Verify and confirm the information.

Configuring the Cluster Management Console

If you want to locally manage this cluster, then you must configure the Cluster Management Console. Note that this cluster can also be a part of the centrally managed clusters.

See “[Veritas Cluster Management Console](#)” on page 41.

To configure the Cluster Management Console

- 1 Review the required information to configure the Cluster Management Console.
- 2 Choose whether to configure the Cluster Management Console or not. Do one of the following:
 - To configure the Cluster Management Console on the systems, press **Enter**.

```
Do you want to configure the Cluster Management Console
[y,n,q] (y)
```
 - To skip configuring the Cluster Management Console and advance to configuring SMTP, enter **n**.
See “[Configuring SMTP email notification](#)” on page 92.
- 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 north: lan0
Enter the NIC for Cluster Management Console to use on north:
[b,?] (lan0)
```

- 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 lan0 to be the public NIC used by all systems [y,n,q,b,?] (y)
```

- 5 Enter the virtual IP address for the Cluster Management Console.

```
Enter the Virtual IP address for Cluster Management Console:
[b,?] 10.10.12.1
```

- 6 Confirm the default netmask or enter another one:

```
Enter the netmask for IP 10.10.12.1: [b,?] (255.255.240.0)
```

- 7 Enter the NetworkHosts IP addresses, separated by spaces, for checking the connections.

```
Enter the NetworkHosts IP addresses, separated by spaces: [b,?]
10.10.12.2
```

- 8 Verify and confirm the Cluster Management Console information.

```
Cluster Management Console verification:
```

```
NIC: lan0
IP: 10.10.12.1
Netmask: 255.255.240.0
NetworkHosts: 10.10.12.2
```

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

Configuring SMTP email notification

You can choose to configure SF Oracle RAC to send event notifications to SMTP e-mail services. You need to provide the SMTP server name and e-mail addresses of people to be notified. Note that it is also possible to configure notification after installation.

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. Do one of the following:

- To configure SMTP notification, press **Enter**.

```
Do you want to configure SMTP notification? [y,n,q] (y) y
```

- To skip configuring SMTP notification and advance to configuring SNMP notification, enter **n**.
See “Configuring SNMP trap notification” on page 93.
- 3 Provide information to configure SMTP notification.
 - Enter the SMTP server’s host name.
Enter the domain-based hostname of the SMTP server
(example: smtp.yourcompany.com): [b,?]
smtp.symantecexample.com
 - Enter the email address of each recipient.
Enter the full email address of the SMTP recipient
(example: user@yourcompany.com): [b,?]
ozzie@symantecexample.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,?] **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,?]
harriet@symantecexample.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,?] **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.

Configuring SNMP trap notification

You can choose to configure SF Oracle RAC 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 it is also possible to configure notification after installation.

To configure the SNMP trap notification

- 1 Review the required information to configure the SNMP notification feature of SF Oracle RAC.

- 2 Specify whether you want to configure the SNMP notification. Do one of the following:
 - To configure SNMP notification, press **Enter**.
Do you want to configure SNMP notification? [y,n,q] (y)
 - To skip configuring SNMP notification and advance to configuring global clustering option, enter **n**.
See “[Configuring global clusters](#)” on page 94.
- 3 Provide information to configure SNMP trap notification.
 - Enter the SNMP trap daemon port.
Enter the SNMP trap daemon port: [b,?] (162)
 - Enter the SNMP console system name.
Enter the SNMP console system name: [b,?] **saturn**
 - 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 saturn [I=Information, W=Warning, E=Error, S=SevereError]: [b,?] **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,?] **jupiter**
Enter the minimum severity of events for which SNMP traps should be sent to jupiter [I=Information, W=Warning, E=Error, S=SevereError]: [b,?] **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.

Configuring global clusters

You can configure global clusters to link clusters at separate locations and enable wide-area failover and disaster recovery. Note that you must have entered a valid license key for SF Oracle RAC global clusters.

See “[Preparing for global clustering](#)” on page 199.

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. Do one of the following:
 - To configure global cluster option, press **Enter**.

```
Do you want to configure the Global Cluster Option? [y,n,q]
(y)
```

- To skip configuring global cluster option and advance to setting permissions for database administration, enter **n**.
See “[Setting permissions for database administration](#)” on page 95.

3 Provide information to configure the Global Cluster option.

If you configured Cluster Management Console to manage this cluster locally, the installer discovers and displays the virtual IP address and netmask used by the Cluster Management Console. You can use the same virtual IP address and netmask.

See “[Configuring the Cluster Management Console](#)” on page 91.

Do one of the following:

- If you want to use the default values, press **Enter**.
- If you do not want to use the default value, enter another IP address. The installer prompts you for a NIC, value for the netmask, and value for the network hosts.

```
Enter the Virtual IP address for Global Cluster Option:
[b,?] (10.10.12.1)
```

4 Verify and confirm the configuration of the global cluster.

Setting permissions for database administration

After SF Oracle RAC is installed, the default settings allow only the superuser to access the /opt/VRTSdbed folder.

- If you already have an Oracle user and group, and you want database administrators (DBAs) to access SF Oracle RAC components, you must set the required permissions.
- If you do not have an Oracle user and group, defer setting the database administration permission and advance to configuring the cluster volume manager. You can configure the database permissions when setting up the database repository.

To set permissions for database administration

- 1 Review the required information to set up the permissions for database administration.
- 2 Specify whether you want to add single user access, group access, or both on each of the nodes as the installer prompts.
 - Provide information if you want to add single user access.

```
Do you want to add single user access on galaxy [y,n,g,?] (y)
Enter login account name for DBA user: dba
```

- Provide information if you want to add group access.
Do you want to add group access on galaxy [y,n,q,?] (y)
Enter group name for DBA users: **oper**

Configuring the cluster volume manager

Cluster volume manager configuration tasks include:

- [Specifying the CVM timeout interval](#)
- [Setting up naming scheme](#)
- [Setting up default disk group](#)

Specifying the CVM timeout interval

By default, the Cluster Volume Manager (CVM) component is configured with a reconfiguration timeout value of 200 seconds. When you are prompted, accept the default, which is appropriate for most clusters.

To specify the CVM timeout interval

- ◆ Enter the timeout interval for CVM:
Enter Cluster Volume Manager cluster reconfiguration timeout
(sec): (200)
Refer to the Veritas Volume Manager documentation for more information on CVM.

Setting up naming scheme

Disks on HP-UX systems typically use device names such as `/dev/rdisk/c#t#d#` to identify disks on the system. It is possible to use the VxVM enclosure-based naming scheme, which allows disk arrays to be more readily recognizable. Dynamic Multipathing (DMP) is a prerequisite for enclosure-based naming schemes. Refer to the Veritas Volume Manager documentation for details on this scheme.

To set up the naming scheme

- 1 If you want to set up the enclosure-based naming scheme, enter **y**.
Do you want to set up the enclosure-based naming scheme?
[y,n,q,?] (n)
- 2 Specify if you want to use the new naming scheme for all eligible systems.
Do you want to use the enclosure-based naming scheme for all of the eligible systems? [y,n,q,?] (y)

Setting up default disk group

If applicable, set up the default disk group. Because some VxVM commands require that a disk group be specified, the installer enables you to register the name of a default VxVM disk group on each eligible node. Note that you can create the default disk group later.

- 1 If you want to set up a default disk group, enter **y**.

```
Do you want to set up a system wide default disk group?
[y,n,q,?] (y) y
```

```
Which disk group? [<group>,list,q,?] xyz_dg
```

- 2 If you specified setting up a default disk group, review the setup output.

```
Volume Manager default disk group setup and daemon startup
```

```
Setting default diskgroup to xyz_dg on galaxy ..... Done
Starting vxrelocd on galaxy ..... Started
Starting vxcached on galaxy ..... Started
Starting vxconfigbackupd on galaxy ..... Started
.
.
```

Configuring VVR on each node

If you added license for Veritas Volume Replicator during installation, you can now accept the default settings or modify the settings for VVR. The installer prompts you for the information on each node.

Setting up VVR ports

The installer identifies the default ports that would be used VVR. You can also assign different ports at this point. Note that the port settings must be identical for systems that will be part of the same Replicated Data Set. They must also be identical for all the systems in a cluster.

To set up VVR ports

- 1 Review and accept the default port values that the configuration program displays.

```
Following are the default ports that will be used by VVR on
galaxy:
```

```
.
.
```

```
Do you want to change any of the VVR ports on galaxy? [y,n,q]
(n)
```

- 2 If you want to change any of the VVR ports on the system, enter **y**.

```
Do you want to change any of the VVR ports on galaxy? [y,n,q]
(n) y
```

- 3 Follow the instructions to change the port values. Note the following points:
 - The port settings must be identical for systems that will be part of the same Replicated Data Set.
 - The port settings must also be identical for all the systems in a cluster.

Configuring VVR statistics collector

The VVR administrative daemon vradmind collects and maintains various statistics, which are helpful in solving VVR performance issues. You can tune the collection using a few tunables:

frequency	for gathering the statistics default = 10 seconds
number of days	for which the collected statistics should be preserved, after which the earlier statistics are automatically deleted default = 3 days

The installation program provides an option to change the default settings.

To configure VVR statistics collector

- 1 Enter **y** at the prompt to change the default setting.
- 2 Enter the values when the installer prompts.

Configuring VVR tunables

As an advanced user, you can modify the VVR tunable parameters. Refer to Veritas Volume Replicator documentation for more details.

To configure VVR tunables

- 1 Enter **y** to view or modify the VVR tunables.
- 2 Review the output to find whether the configuration is successful.

Starting the VAILAgent

You must start the VAILAgent to access array discovery service for deep mapping.

The storage mapping feature available with SF Oracle RAC provides “deep” mapping information for supported storage arrays. Veritas Array Integration Layer (VAIL) software interfaces third-party hardware storage arrays with Veritas storage software.

See [“About arrays for Storage Mapping and statistics”](#) on page 337.

After starting the VAILAgent agent, this service for deep mapping becomes accessible across the domain. Refer to the Veritas Volume Manager documentation for more information.

To start the VAILAgent

- ◆ Press Enter to start the VAILAgent on the target system.

```
Would you like to start VAILAgent (Array discovery service for  
Deep mapping) on the target system (once started, this  
service will be accessible across the domain)? [y,n,q,?] (y)
```

Creating SF Oracle RAC configuration files

After gathering the configuration information for SF Oracle RAC components, the configuration program creates the configuration files.

If you chose to configure the cluster in secure mode, the installer also configures the Symantec Product Authentication Service. Depending on the mode you chose to set up Authentication Service, the installer creates security principal or executes the encrypted file to create security principal on each node in the cluster. The installer creates the VxSS service group, creates Authentication Server credentials on each node in the cluster, and Web credentials for SF Oracle RAC users, and sets up trust with the root broker. Then, the installer proceeds to start SF Oracle RAC in secure mode.

- 1 Review the output as the configuration program creates security principal, starts VCS, creates VCS configuration files, and copies the files to each node.
- 2 When the configuration prompts you, confirm the fully qualified host names of the cluster nodes.

```
Is the fully qualified hostname of system "galaxy" =  
"galaxy.example.com"? [y,n,q] (y)  
Is the fully qualified hostname of system "nebula" =  
"nebula.example.com"? [y,n,q] (y)
```

- 3 Review the output as the program verifies communication with the remote nodes.

Configuring standalone hosts for Storage Foundation

When the configuration program prompts you to enable Storage Foundation Management Server Management, enter **n** to configure the standalone hosts for Storage Foundation.

Warning: If you press Enter to accept the default option, you must restart the configuration program. SF Management Server is not available on the Storage Foundation and High Availability Solutions release. Make sure to enter **n** at the prompt and configure standalone hosts.

See “[About Veritas Storage Foundation Management Server](#)” on page 100.

To configure standalone hosts for Storage Foundation

- 1 Review the output on the ways to manage the Storage Foundation hosts.
- 2 Enter **n** at the prompt to configure standalone hosts.
Enable Storage Foundation Management Server Management? [y,n,q]
(y) **n**
- 3 Review the output as the configuration program configures to manage the Storage Foundation hosts in standalone mode.

About Veritas Storage Foundation Management Server

Veritas Storage Foundation Management Server by Symantec (SF Management Server) ties together the Storage Foundation product offerings to ensure that the hosts in your data center use storage as efficiently as possible. You can use it to centrally monitor, visualize, and manage Storage Foundation hosts and generate reports about the hosts and the storage resources they consume.

The central console seamlessly integrates a wide range of management tasks like as monitoring and reporting. SF Management Server also offers customizable policy-based management that helps you automate:

- notification
- recovery
- other user-definable actions

You are prompted to set up an optional SF Management Server managed host during SF Oracle RAC installation. SF Management Server is not available on the Storage Foundation and High Availability Solutions release and must be obtained separately. For information on ordering SF Management Server, visit:

www.symantec.com/enterprise/sfms

Refer to the Storage Foundation Management Server documentation for details on preparing the SF Oracle RAC environment to enable centrally managed Storage Foundation hosts.

Starting SF Oracle RAC

The installer starts SF Oracle RAC, SF Oracle RAC processes, and configures the agents that you opted during the configuration phase.

To start SF Oracle RAC

- 1 Confirm that you want to start the SF Oracle RAC processes.
- 2 Review the output as the installer starts SF Oracle RAC and its 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. See [“Setting up I/O fencing for SF Oracle RAC”](#) on page 104.
- 3 Review the output as the program starts the VxVM daemons and configures the SF Oracle RAC agents.
- 4 Review the output as the configuration program sets the default disk group on each node. You must have a specified a default disk group for VxVM during the configuration,
- 5 Review the output at the end of the configuration and note the location of the summary and log files for future reference.

Checking shared disks for I/O fencing

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

Data disks	Stores shared data
Coordinator disks	Act as a global lock during membership changes. Coordinator disks are small LUNs (typically three per cluster)

Note: On HP-UX 11i v3, you must use DMP devices for I/O fencing.

See [“Setting up shared storage”](#) on page 68.

Perform the following checks for I/O fencing disks:

- Identify three SCSI-3 PR compliant shared disks as coordinator disks. List the disks on each node and pick three disks as coordinator disks. For example, execute the following commands to list the disks:

```
# ioscan -nfc disk
# insf -e
```

- Test the shared disks using the `vxfcntlsthdw` script.
See [“Testing the shared disks for SCSI-3”](#) on page 102.

Testing the shared disks for SCSI-3

Use the `vxfcntlsthdw` utility to test the shared storage arrays support SCSI-3 persistent reservations and I/O fencing. Review the guidelines to run `vxfcntlsthdw` program, verify that the systems see the same disk, and proceed to test the disks. Make sure to test disks serving as coordinator disks.

See [“Setting up coordinator disk groups”](#) on page 107.

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

See [“Verifying data storage arrays using the vxfcntlsthdw utility”](#) on page 388.

Reviewing guidelines on vxfcntlsthdw

- Verify the connection of the shared storage for data to two of the nodes on which you installed SF Oracle RAC.

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

- The two nodes must have `ssh` (default) or `remsh` communication. If you use `remsh`, launch the `vxfcntlsthdw` utility with the `-n` option.
See [“Setting up inter-system communication”](#) on page 66.
After completing the testing process, remove permissions for communication and restore public network connections.
See [“Removing permissions for communication”](#) on page 112.
- To ensure both nodes are connected to the same disk during the testing, use the `vxfenadm -i diskpath` command to verify the disk serial number.
See [“Verifying the nodes see the same disk”](#) on page 102.

Verifying the nodes see the same disk

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 to verify that the same serial number for the LUN is returned on all paths to the LUN.

For example, an EMC array is accessible by the `/dev/vx/rdmp/c2t13d0` path on node A and by the `/dev/vx/rdmp/c2t11d0` path on node B.

From node A, enter:

```
# vxfenadm -i /dev/vx/rdmp/c2t13d0
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/vx/rdmp/c2t11d0 path.

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

```
# vxfenadm -i /dev/vx/rdmp/c2t0d2
Vendor id      : HITACHI
Product id     : OPEN-3      -HP
Revision       : 0117
Serial Number  : 0401EB6F0002
```

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

Testing the disks using vxfentsthdw script

This procedure uses the /dev/vx/rdmp/c2t13d0 disk in the steps.

If the utility does not show a message stating a disk is ready, verification has failed. Failure of verification can be the result of an improperly configured disk array. It can also be caused by 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/vx/rdmp/c2t13d0 is ready to be configured for I/O
Fencing on node galaxy
```

See [“Adding or removing coordinator disks”](#) on page 348.

To test disks using vxfentsthdw script

- 1 Make sure system-to-system communication is functioning properly.
See [“Setting up inter-system communication”](#) on page 66.
- 2 From one node, start the utility. Do one of the following:
 - If you use `ssh` for communication:


```
# /opt/VRTSvcs/vxfen/bin/vxfentsthdw
```
 - If you use `remsh` for communication:


```
# /opt/VRTSvcs/vxfen/bin/vxfentsthdw -n
```
- 3 After reviewing the overview and warning that the tests overwrite data on the disks, confirm to continue the process and enter the node names.

```
***** 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: galaxy  
Enter the second node of the cluster: nebula
```

- 4 Enter the names of the disks you are checking. For each node, the disk may be known by the same name:

```
Enter the disk name to be checked for SCSI-3 PGR on node galaxy  
in the format: /dev/vx/rdmp/cxtxdx  
/dev/vx/rdmp/c2t13d0
```



```
Enter the disk name to be checked for SCSI-3 PGR on node nebula  
in the format: /dev/vx/rdmp/cxtxdx  
Make sure it's the same disk as seen by nodes galaxy and nebula  
/dev/vx/rdmp/c2t13d0
```

If the serial numbers of the disks are not identical, then the test terminates.
- 5 Review the output as the utility performs the checks and report its activities.
- 6 If a disk is ready for I/O fencing on each node, the utility reports success:

```
ALL tests on the disk /dev/vx/rdmp/c2t13d0 have PASSED.  
The disk is now ready to be configured for I/O Fencing on node  
galaxy.  
ALL tests on the disk /dev/vx/rdmp/c2t13d0 have PASSED.  
The disk is now ready to be configured for I/O Fencing on node  
nebula.  
Removing test keys and temporary files, if any ...  
.  
.
```
- 7 Run the `vxfcntlshdw` utility for each disk you intend to verify.

Setting up I/O fencing for SF Oracle RAC

Choose **Configure SFRAC > Configure I/O fencing** from the configuration program's menu to configure the I/O fencing feature. The program presents guidelines to set up I/O fencing in SF Oracle RAC. Make sure you checked the I/O fencing disks before you set up I/O fencing.

See [“Checking shared disks for I/O fencing”](#) on page 101.

I/O fencing requires coordinator disks that are configured in a disk group and accessible to each node in the cluster. These disks enables the `vxfen` driver to resolve potential split-brain conditions and prevent data corruption. Review the description of I/O fencing and the role of coordinator disks.

See [“About I/O fencing”](#) on page 33.

Tasks involved in setting up I/O fencing include:

- [Verifying the SF Oracle RAC configuration](#)
- [Initializing disks](#)
- [Setting up coordinator disk groups](#)

- Stopping SF Oracle RAC on all nodes
- Configuring `/etc/vxfendg` disk group for I/O fencing
- Updating `/etc/vxfenmode` file
- Starting I/O fencing
- Modifying VCS configuration to enable I/O fencing
- Starting SF Oracle RAC on all nodes
- Verifying I/O fencing configuration
- Removing permissions for communication

Verifying the SF Oracle RAC configuration

After installing SF Oracle RAC and configuring VCS, CVM, and CFS, verify the installation using the `gabconfig -a` command.

The output of the `gabconfig -a` command displays the nodes that have membership with the modules installed and configured. The first line indicates that each node (0, 1, 2, and 3) has membership with the GAB utility that uses “Port a.” [Table 4-5](#) lists the different ports that the software configures for different functions.

Table 4-5 GAB port description

Port	Function
a	GAB
b	I/O fencing
d	Oracle Disk Manager (ODM)
f	Cluster File System (CFS)
h	Veritas Cluster Server (VCS: High Availability Daemon)
o	VCSMM driver
v	Cluster Volume Manager (CVM)
w	<code>vxconfigd</code> (module for CVM)

To verify the SF Oracle RAC configuration

- ◆ Run the `gabconfig -a` command on all nodes.

For example:

```
# gabconfig -a
GAB Port Memberships
```

```
-----  
Port a gen   ada401 membership 0123  
Port b gen   ada40d membership 0123  
Port d gen   ada409 membership 0123  
Port f gen   ada41c membership 0123  
Port h gen   ada40f membership 0123  
Port o gen   ada406 membership 0123  
Port v gen   ada416 membership 0123  
Port w gen   ada418 membership 0123
```

Initializing disks

Install the driver and HBA card. Refer to the documentation from the vendor for instructions.

After you physically add shared disks to the nodes, you must initialize them as VxVM disks and verify that all the nodes see the same disk. Use the example procedure; see the *Veritas Volume Manager Administrator's Guide* for more information on adding and configuring disks.

To initialize disks

- 1 Make the new disks recognizable. On each node, enter:

```
# ioscan -nfc disk  
# insf -e
```

Warning: The HP-UX man page for the `insf` command instructs you to run the command in single-user mode only. Running `insf -e` in multi-user mode is acceptable assuming no other user is accessing any of the device files. This command can change the mode, owner, or group of an existing special (device) file, or unlink and recreate a file. Special files that are currently open may be left in an indeterminate state.

- 2 If the ASL for the array you are adding is not installed, obtain and install it on each node before proceeding. The ASL for the supported storage device you are adding is available from the disk array vendor or Veritas technical support.
- 3 Verify that the Array Support Library (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:

```
# vxddladm listsupport all  
LIB_NAME                ARRAY_TYPE  VID        PID  
-----  
.  
libvxemc.sl             A/A        EMC        SYMMETRIX  
.
```

- 4 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 adding and configuring disks.
- 5 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 *Veritas Volume Managers Administrator's Guide*.
 - Use the `vxdisksetup` command to initialize a disk as a VxVM disk.

```
vxdisksetup -i device_name format=cdsdisk
```

The example specifies the CDS format:

```
# vxdisksetup -i c2t13d0 format=cdsdisk
```

Setting up coordinator disk groups

I/O fencing requires coordinator disks that are configured in a disk group and accessible to each node in the cluster. These disks enables the `vxfen` driver to resolve potential split-brain conditions and prevent data corruption. Make sure to meet the requirements for coordinator disks and then create the coordinator disk group.

Requirements for coordinator disks

After adding and initializing disks for use as coordinator disks, make sure coordinator disks meet the following requirements:

- You must have three coordinator disks.
- Each of the coordinator disks must use a physically separate disk or LUN.
- Each of the coordinator disks should exist on a different disk array, if possible.
- You must initialize each disk as a VxVM disk.
- The coordinator disks must support SCSI-3 persistent reservations. See [“Testing the shared disks for SCSI-3”](#) on page 102.
- The coordinator disks must exist in a disk group (for example, `vxfencoorddg`). See [“Creating the coordinator disk group”](#) on page 108.
- Symantec recommends using hardware-based mirroring for coordinator disks.

Creating the coordinator disk group

From one node, create a disk group named `vxfencoorddg`. This group must contain three disks or LUNs. Refer to the *Veritas Volume Manager Administrator's Guide* for details on creating disk groups.

You must also set the coordinator attribute for disk groups. The `vxfen` driver uses this attribute to prevent the reassignment of coordinator disks to other disk groups.

Note that if you create a coordinator disk group as a regular disk group, you can turn on the coordinator attribute in Volume Manager. You do this with a `vx dg set coordinator=on` command.

The example procedure assumes that the disks have the device names `c1t1d0`, `c2t1d0`, and `c3t1d0`.

To create the coordinator disk group

- 1 On one node, create the disk group by specifying the device name of one of the disks:

```
# vx dg init vxfencoorddg c1t1d0
```
- 2 Add the other two disks to the disk group:

```
# vx dg -g vxfencoorddg adddisk c2t1d0  
# vx dg -g vxfencoorddg adddisk c3t1d0
```
- 3 Set the option `coordinator=on` to set the coordinator attribute:

```
# vx dg -g vxfencoorddg set coordinator=on
```

Stopping SF Oracle RAC on all nodes

Before configuring the coordinator disk for use, you must stop SF Oracle RAC on all nodes.

To stop SF Oracle RAC on all nodes

- ◆ On one node, enter:

```
# hastop -all
```

Configuring /etc/vxfendg disk group for I/O fencing

After setting up the coordinator disk group, configure it for use.

To configure the disk group for fencing

- 1 Deport the disk group:

```
# vx dg deport vxfencoorddg
```
- 2 Import the disk group with the `-t` option to avoid automatically importing it when the nodes restart:

```
# vxdg -t import vxfencoorddg
```

- 3 Deport the disk group. Deporting the disk group prevents the coordinator disks from serving other purposes:

```
# vxdg deport vxfencoorddg
```

- 4 On all nodes, type:

```
# echo "vxfencoorddg" > /etc/vxfendg
```

Do not use spaces between the quotes in the “vxfencoorddg” text.

This command creates the `/etc/vxfendg` file, which includes the name of the coordinator disk group.

Based on the contents of the `/etc/vxfendg` and `/etc/vxfenmode` files, the rc script creates the `/etc/vxfentab` file for use by the `vxfen` driver when the system starts. The rc script also invokes the `vxfenconfig` command, which configures the `vxfen` driver to start and use the coordinator disks listed in `/etc/vxfentab`. The `/etc/vxfentab` file is a generated file; do not modify this file.

Example `/etc/vxfentab` file

The `/etc/vxfentab` file gets created when you start the I/O fencing driver.

See “[Starting I/O fencing](#)” on page 109.

An example of the `/etc/vxfentab` file on one node resembles:

```
/dev/vx/rdmp/c1t1d0
/dev/vx/rdmp/c2t1d0
/dev/vx/rdmp/c3t1d0
```

In some cases you must remove disks from or add disks to an existing coordinator disk group.

See “[Adding or removing coordinator disks](#)” on page 348.

Updating `/etc/vxfenmode` file

You must update the `/etc/vxfenmode` file to use `scsi3` mode. You must configure the `vxfen` module to use DMP devices. You must use the `dmp` SCSI-3 disk policy on all the nodes.

To update `/etc/vxfenmode` file

- ◆ On all cluster nodes, type:

```
# cp /etc/vxfen.d/vxfenmode_scsi3_dmp /etc/vxfenmode
```

Starting I/O fencing

You now need to start I/O fencing on each node. After you configure the SF Oracle RAC components, the configuration program starts I/O fencing in the

disabled mode. So, you need to restart the driver for the new configuration to take effect.

To stop I/O fencing on a node

- ◆ Stop the I/O fencing driver.


```
# /sbin/init.d/vxfen stop
```

To start I/O fencing on a node

- ◆ Start the I/O fencing driver.


```
# /sbin/init.d/vxfen start
```

Modifying VCS configuration to enable I/O fencing

After adding coordinator disks and configuring I/O fencing, add the UseFence = SCSI3 cluster attribute to the VCS configuration file, `/etc/VRTSvcs/conf/config/main.cf`. This entry gives the user the flexibility to disable or enable I/O fencing by modifying the UseFence attribute.

To modify VCS configuration to enable I/O fencing

- 1 Save the existing configuration:


```
# haconf -dump -makero
```
- 2 Stop VCS on all nodes:


```
# hastop -all
```
- 3 Make a backup copy of the main.cf file:


```
# cd /etc/VRTSvcs/conf/config
# cp main.cf main.orig
```
- 4 On one node, use vi or another text editor to edit the `main.cf` file. Modify the list of cluster attributes by adding the UseFence attribute and assigning its value of SCSI3.


```
cluster rac_cluster101
UserNames = { admin = "cDRpdxPmHpzS." }
Administrators = { admin }
HacliUserLevel = COMMANDROOT
CounterInterval = 5
UseFence = SCSI3
)
```
- 5 Save and close the file.
- 6 Verify the syntax of the file `/etc/VRTSvcs/conf/config/main.cf`:


```
# hacf -verify /etc/VRTSvcs/conf/config
```
- 7 Using `rcp` or another utility, copy the VCS configuration file from a node (for example, galaxy) to the remaining cluster nodes. For example, on each remaining node, enter:

```
# rcp galaxy:/etc/VRTSvcs/conf/config/main.cf \
/etc/VRTSvcs/conf/config
```

Starting SF Oracle RAC on all nodes

You must start SF Oracle RAC on each node to bring up the cluster configuration with I/O fencing enabled.

Starting VCS, CVM, and CFS

You must start VCS, CVM, and CFS on all nodes in the cluster.

To start VCS, CVM, and CFS on a node

- ◆ With the configuration file in place on each system, start VCS, CVM, and CFS:

```
# hastart
```

Verifying GAB port membership

After setting up I/O fencing and starting VCS, CVM, and CFS on each node, verify GAB port membership.

See [Table 4-5, “GAB port description,”](#) on page 105.

To verify GAB port membership

- ◆ Run the `gabconfig -a` command.

For example:

```
galaxy# gabconfig -a
GAB Port Memberships
=====
Port a gen   ada401 membership 0123
Port b gen   ada40d membership 0123
Port d gen   ada409 membership 0123
Port f gen   ada41c membership 0123
Port h gen   ada40f membership 0123
Port o gen   ada406 membership 0123
Port v gen   ada416 membership 0123
Port w gen   ada418 membership 0123
```

Verifying the CVM group is online

Make sure the `cvm` group is in the `ONLINE` state.

To verify CVM group

- ◆ On all nodes, type:

```
# hagrps -state cvm
```

Verifying I/O fencing configuration

Verify from the `vxfenadm` output that the SCSI-3 disk policy reflects the configuration in the `/etc/vxfenmode` file.

To verify I/O fencing configuration

- ◆ On one of the nodes, type:

```
# vxfenadm -d

I/O Fencing Cluster Information:
=====

Fencing Protocol Version: 201
Fencing Mode: SCSI3
Fencing SCSI3 Disk Policy: dmp
Cluster Members:

    * 0 (galaxy)
    * 1 (nebula)

RFSM State Information:
node  0 in state  8 (running)
node  1 in state  8 (running)
```

Removing permissions for communication

After completing the installation of SF Oracle RAC and verification of disk support for I/O fencing, remove the temporary `remsh` access permissions 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 at this time.

Unmounting the product disc

After verifying the successful installation and configuration of all required packages, unmount the disc.

To unmount the product disc using `umount`

- ◆ Run the `umount` command.

For example:

```
# umount /cdrom
```

Where `/cdrom` is the location of the disc mount point.

Verifying SF Oracle RAC installation using VCS configuration file

You can verify the SF Oracle RAC installation and configuration by examining the VCS configuration file, `main.cf`, in the directory `/etc/VRTSvcs/conf/config`.

About VCS configuration file after SF Oracle RAC installation

Note the following information about the VCS configuration file after the SF Oracle RAC installation:

- The “include” statements list types files for VCS (`types.cf`), CFS (`CFSTypes.cf`), and CVM (`CVMTypes.cf`). These files are in the `/etc/VRTSvcs/conf/config` directory. The types file for the Oracle enterprise agent (`OracleTypes.cf`) is also located in `/etc/VRTSvcs/conf/config`. These files define the agents that control the resources in the cluster.
 - The VCS types include all agents bundled with VCS. Refer to the *Veritas Cluster Server Bundled Agents Reference Guide* for information about VCS agents.
 - The CFS types include `CFSMount` and `CFSfsckd`.
The `CFSMount` agent mounts and unmounts the shared volume file systems.
The `CFSfsckd` types are defined for cluster file system daemons and do not require user configuration.
 - The CVM types include the `CVMCluster`, `CVMVxconfigd`, and `CVMVolDg`.
The `CVMCluster` agent, which is automatically configured during installation, starts CVM in the cluster by autoimporting shared disk groups, controls node membership in the cluster, and defines how nodes communicate the state of volumes.
See “[SF Oracle RAC agents](#)” on page 369.
The `CVMVxconfigd` agent starts and monitors the `vxconfigd` daemon; this daemon maintains disk and disk group configurations, communicates configuration changes to the kernel, and modifies configuration information stored on disks.
The `CVMVolDg` agent starts the volumes in a shared disk group, sets activation modes, and monitors specified critical volumes.
- The Oracle enterprise agent types includes definitions for the Oracle agent and the `Netlsnr` agent. The Oracle agent monitors the resources for an Oracle database, and the `Netlsnr` agent manages the resources for the listener process.

- 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. See [“Modifying VCS configuration to enable I/O fencing”](#) on page 110.
- The `main.cf` now includes the `cvm` service group. This group includes definitions for monitoring the CFS and CVM resources. The `CVMCluster` agent resource definition indicates 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.

Sample `main.cf` file after SF Oracle RAC installation

The configuration file created on your node is located in `/etc/VRTSvcs/conf/config/main.cf`. Review the sample VCS configuration file that is created after SF Oracle RAC installation and before Oracle installation.

```
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"
include "OracleTypes.cf"

cluster rac_cluster101 (
    UserNames = { admin = bopHo }
    Administrators = { admin }
    UseFence = SCSI3
)

system galaxy (
)

system nebula (
)

group cvm (
    SystemList = { galaxy = 0, nebula = 1 }
    AutoFailOver = 0
    Parallel = 1
    AutoStartList = { galaxy, nebula }
)

CFSfsckd vxfsckd (
)

CVMCluster cvm_clus (
    CVMClustName = rac_cluster101
    CVMNodeId = { galaxy = 0, nebula = 1 }
    CVMTransport = gab
```

```
CVMTimeout = 200
)

CVMVxconfigd cvm_vxconfigd (
  Critical = 0
  CVMVxconfigdArgs = { syslog }
)

vxfscsd requires cvm_clus
cvm_clus requires cvm_vxconfigd
```


Upgrading SF Oracle RAC

This chapter contains the following topics:

- [About upgrading to SF 5.0 for Oracle RAC on HP-UX 11i v3](#)
- [Upgrading SF Oracle RAC](#)

About upgrading to SF 5.0 for Oracle RAC on HP-UX 11i v3

You can upgrade from SF Oracle RAC version 4.1 or 5.0 on HP-UX 11i v2 to SF Oracle RAC version 5.0 on HP-UX 11i v3.

The upgrade process involves a combination of the following:

- SF Oracle RAC upgrade
- HP-UX upgrade
- Oracle upgrade

Note: If you use SF Oracle RAC 3.5 or 3.5 Update versions, then you must first upgrade SF Oracle RAC to one of the versions that support the upgrade path. Then, you can upgrade SF Oracle RAC to 5.0 on HP-UX 11.31.

Table 5-6 presents the possible scenarios for upgrading to SF 5.0 for Oracle RAC.

Table 5-6 SF 5.0 for Oracle RAC upgrade scenarios

Upgrade scenarios	Upgrade From			Upgrade To		
	SF Oracle RAC Version	Oracle Version	HP-UX Version	SF Oracle RAC Version	Oracle Version	HP-UX Version
SF Oracle RAC, Oracle, and HP-UX	4.1 4.1 MP1 5.0 5.0 MP1	9i	11i v2 (11.23)	5.0 for HP-UX 11i v3	10g R2	11i v3
	4.1 MP1 5.0 5.0 MP1	10g R1	11i v2 (11.23)	5.0 for HP-UX 11i v3	10g R2	11i v3
SF Oracle RAC and HP-UX	4.1 MP1 PP1 5.0 5.0 MP1	10g R2	11i v2 (11.23)	5.0 for HP-UX 11i v3	10gR2	11i v3

Depending on the upgrade scenario, upgrading SF Oracle RAC involves the following tasks:

- SF Oracle RAC, Oracle, and HP-UX ■ Upgrade SF Oracle RAC and HP-UX
 See [“Upgrading SF Oracle RAC to version 5.0 for HP-UX 11i v3”](#) on page 119.
- Upgrade and migrate Oracle software
 See [“Upgrading and migrating to Oracle 10g R2 software”](#) on page 157.
- SF Oracle RAC and HP-UX ■ Upgrade SF Oracle RAC and HP-UX
 See [“Upgrading SF Oracle RAC to version 5.0 for HP-UX 11i v3”](#) on page 119.

Upgrading SF Oracle RAC

Review the upgrade matrix for a detailed list of the supported upgrade paths.
See “[About upgrading to SF 5.0 for Oracle RAC on HP-UX 11i v3](#)” on page 117.

Upgrading SF Oracle RAC to version 5.0 for HP-UX 11i v3

Upgrade tasks include:

- [Preparing to upgrade SF Oracle RAC](#)
- [Upgrading to HP-UX 11i Version 3](#)
- [Upgrading SF Oracle RAC](#)
- [Configuring SF Oracle RAC after upgrading to 5.0](#)
- [Performing post-upgrade tasks](#)

Note: The SF Oracle RAC upgrade procedure necessitates a reboot after installing the SF 5.0 for Oracle RAC depots.

Preparing to upgrade SF Oracle RAC

Perform the following preparatory tasks before upgrading SF Oracle RAC to version 5.0.

To prepare to upgrade SF Oracle RAC

- 1 Log in as superuser to one of the nodes in the cluster and make a backup of the VCS configuration file, main.cf.

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

- 2 If you are using Oracle9i, stop GSD daemon.
As oracle user, stop gsd processes on each node.

- To determine if gsd processes are running, type:

```
$ $ORACLE_HOME/bin/gsdctl stat
```

- If gsd processes are running, stop them:

```
$ $ORACLE_HOME/bin/gsdctl stop
```

- 3 Take the resources offline.

- For Oracle9i

On each node in the cluster, take the Oracle and Netlsnr resources in the VCS configuration offline. For example:

```
# hares -offline VRT -sys galaxy  
# hares -offline rac -sys galaxy
```

```
# hares -offline VRT -sys nebula
# hares -offline rac -sys nebula
# hares -offline LISTENER -sys galaxy
# hares -offline LISTENER -sys nebula
```

These commands stop the Oracle instances running on the specified systems.

- For Oracle 10g

On each node in the cluster, take the Oracle and cssd resources offline. For example:

To take the Oracle resources offline, run the following commands:

```
# hares -offline VRT -sys galaxy
# hares -offline rac -sys galaxy
# hares -offline VRT -sys nebula
# hares -offline rac -sys nebula
```

To take the cssd resources offline, run the following commands:

```
# hares -offline cssd-resource1 -sys galaxy
# hares -offline cssd-resource1 -sys nebula
```

- 4 If you are using Oracle 10g R1 or R2, then comment the CRS_HOME entry in the /etc/fstab file.

- 5 Stop all applications that use the CFS mounts that are not under VCS control.

To stop the processes using the CFS mount point, use the application-specific commands.

- 6 Unmount Cluster File Systems.

On each node, unmount any CFS file systems that are not under VCS control.

- Determine the file systems to unmount by checking the /etc/mnttab file. For example:

```
# cat /etc/mnttab | grep vxfs | grep cluster
```

The output shows each line of the /etc/mnttab file that contains an entry for a VxFS file system mounted in cluster mode.

- To unmount each of the file systems listed in the output, specify the appropriate mount point:

```
# umount mount_point
```

- 7 Freeze the VCS service groups. Run the following commands:

```
# haconf -makerw
# hagrps -freeze servicegroup -persistent
# haconf -dump -makero
```

- 8 Stop VCS on all nodes:

```
# hastop -all -force
```

- 9 If the cluster-wide attribute “UseFence” is set to SCSI3, then reset the value to NONE in the /etc/VRTSvcs/conf/config/main.cf file.

- 10 On each node, edit the `/etc/vxfenmode` file to configure I/O fencing in disabled mode.

```
# cat /etc/vxfenmode
vxfen_mode=disabled
```

Note: To upgrade from SF 4.1 for Oracle RAC on HP-UX 11iv2, create `/etc/vxfenmode` file and populate it as above.

- 11 On each node, change `LLT_START=0` in the file `/etc/rc.config.d/lltconf`.

- 12 On each node, remove the following device files:

```
# rm -f /dev/llt
# rm -f /dev/gab*
# rm -f /dev/vxfen
# rm -f /dev/lmx
# rm -f /dev/vcsmm
```

- 13 To upgrade from SF 5.0 for Oracle RAC on HP-UX 11iv2, you must convert Rootable VxVM volumes to LVM Rootable.

- On each node, execute the following command to convert Rootable VxVM volumes to LVM Rootable on disk `c1t1d0`.

```
# /etc/vx/bin/vxres_lvmroot -v -b c1t1d0
```

After executing the command, the primary boot device is set to newly cloned LVM root disk.
- Reboot the node:

```
# /usr/sbin/shutdown -r now
```

Upgrading to HP-UX 11i Version 3

Refer to the HP documentation for upgrading to HP-UX 11i Version 3.

Note: During the OS upgrade, the following errors may be reported that you can ignore. The OS upgrade will complete successfully.

```
ERROR: A later revision (one with a higher revision
number) of fileset "VRTSvxvm.VXVM-KRN,r=4.1.032" has
already been installed.
```

```
ERROR: A later revision (one with a higher revision
number) of fileset "VRTSvxvm.VXVM-RUN,r=4.1.032" has
already been installed.
```

Upgrading SF Oracle RAC

Invoke the `installsrac` program to upgrade to SF 5.0 for Oracle RAC.

To upgrade SF Oracle RAC

- 1 Mount the software disc.
See “[Mounting the product disc](#)” on page 67.
- 2 Make sure you performed other pre-installation tasks before you upgrade SF Oracle RAC.
See “[Preparing to install SF 5.0 for Oracle RAC](#)” on page 65.
- 3 On the node where the disk is mounted, browse to the directory containing the installer program.
- 4 Start the product upgrade.

```
# ./installsfrac
```

After viewing a copyright notice, review the report after the program examines the configuration files and discovers the existing cluster configuration.
- 5 To upgrade to SF 5.0 for Oracle RAC, press **Enter**.
Do you want to upgrade to version 5.0 on these systems using the current configuration? [y,n,q,?] (y) **y**
- 6 Confirm that you want to continue with the installation at the installer prompt. Note that the installation program will replace any previous version of VxVM that was bundled with HP-UX with VxVM 5.0.
- 7 Review the output as the program performs the following:
 - Checks system licensing and installs the licensing depot
 - Checks the installed depots on each node
 - Lists the SF Oracle RAC depots that the program will install or upgrade
- 8 Confirm that you want to upgrade SF Oracle RAC. Note that the installer will stop all the currently running SF Oracle RAC processes.
- 9 Review the output as the installer stops the CVM and CFS agents, and updates the gabtab files on each node in the cluster.
- 10 Confirm that you are ready to upgrade SF Oracle RAC. Note that the installation program makes configuration updates and stops the cluster to upgrade the SF Oracle RAC depots.
- 11 Review the output as the program does the following:
 - Backs up the configuration files
 - Freezes all the VCS service groups
 - Updates the types.cf files
Refer to the *Veritas Cluster Server Installation Guide* for more information on the `types.cf` updates.
 - Updates gabtab files on each node

- 12 Review the output as the program stops SF Oracle RAC processes and shuts down SF Oracle RAC on all nodes in the cluster.
- 13 Review the output as the program uninstalls the depots from the previous version of SF Oracle RAC.
- 14 Review the output as the program installs SF 5.0 for Oracle RAC depots.
- 15 Note the location of the summary and log files for future reference.
- 16 When the installer prompts you to reboot the nodes before configuring SF Oracle RAC, review the installer output and follow the instructions.
Execute the `/usr/sbin/shutdown -r now` command to reboot the nodes.

Configuring SF Oracle RAC after upgrading to 5.0

After upgrading SF Oracle RAC to version 5.0, you can now choose to use the previous configuration of SF Oracle RAC. Note that, if you want to take advantage of the new features in SF 5.0 for Oracle RAC, you must reconfigure SF Oracle RAC using the configuration program.

To configure SF Oracle RAC after upgrading to version 5.0

- 1 Confirm that the product disc is mounted and that you are logged in as the superuser.
- 2 Start the configuration program.
The program specifies the directory where the logs are created and begins with a copyright message.
- 3 Enter the names of the nodes where you want to configure the software.
Enter the system names separated by spaces on which to configure SFRAC: **galaxy nebula**
- 4 Review the output as the program checks that the local node running the script can communicate with remote nodes and checks whether SF 5.0 for Oracle RAC is installed successfully.
- 5 Confirm that you want to use the previous configuration of SF Oracle RAC as the installer prompts.
- 6 Review the output as the program checks the system licensing.
- 7 Select **Configure SFRAC > Configure VCS, CVM, and CFS** and confirm at the installer prompt that you want to continue with the configuration.
- 8 Review the output as the program stops various SF Oracle RAC processes on each node.

- 9 If you want the database administrators (DBAs) to access SF Oracle RAC components, then you must set the required permissions.
The default settings allows only the superuser to access the `/opt/VRTSdbed` folder.
See [“Setting permissions for database administration”](#) on page 95.
- 10 Confirm the fully qualified host names of the nodes in the cluster as the installer prompts you.
- 11 Review the output as the installer verifies communication with the nodes in the cluster.
- 12 If you want to configure the management console for Storage Foundation, you can do that now.
See [“Configuring standalone hosts for Storage Foundation”](#) on page 99.
- 13 Review the output as the installer starts the SF Oracle RAC processes.
- 14 Review the output as the installer starts the VxVM daemons.
- 15 Note the location of the upgrade log files and summary files.
If you did not configure authentication broker before upgrading, the authentication broker will remain unconfigured after the upgrade. So, the `vxatd` process cannot be started. The installation logs will contain this message which you can ignore.

Performing post-upgrade tasks

After reconfiguring the upgraded SF Oracle RAC, you must perform the following post-upgrade tasks.

To perform post-upgrade tasks

- 1 Before upgrading HP-UX to 11i v3, if you converted the rootable VxVM volumes to LVM, convert them back to VxVM rootable.
 - On each node, execute the following command to convert LVM Rootable volumes back to rootable VxVM volumes.

```
# /etc/vx/bin/vxcp_lvmroot -v -b c1t0d0
```


After executing the command, the primary boot device is set to the newly cloned VxVM root device.
 - Reboot the node:

```
# /usr/sbin/shutdown -r now
```
- 2 Verify the output of the `gabconfig -a` command to ensure that SF Oracle RAC is configured correctly.
See [“Verifying the SF Oracle RAC configuration”](#) on page 105.
- 3 Link the SF Oracle RAC libraries to Oracle.

See [“Relinking the SF Oracle RAC libraries to Oracle 10g”](#) on page 154.

- 4 Stop VCS on all nodes:

```
# haconf -dump -makero
# hastop -all -force
```
- 5 Configure I/O fencing to use the dmp scsi disk policy.
See [“Setting up I/O fencing for SF Oracle RAC”](#) on page 104.
- 6 Set the clusterwide attribute "UseFence" to use SCSI3. Add the following line to the `/etc/VRTSvcs/conf/config/main.cf` file:

```
UseFence=SCSI3
```
- 7 If you upgraded SF Oracle RAC with Oracle 10g R1 or R2, then uncomment the `CRS_HOME` entry in the `/etc/fstab` file that you commented before the upgrade.
- 8 If `PrivNIC.cf` is not included in the `main.cf` file after the upgrade, then add the following line in the `main.cf` file:

```
include "PrivNIC.cf"
```
- 9 Start the VCS engine on each system:

```
# hstart
```
- 10 Verify that the `CVMVolDg` resource type has the `CVMDGAction` attribute:

```
# /opt/VRTS/bin/haattr -display CVMVolDg | \
grep -w CVMDGAction
```
- 11 If the `CVMVolDg` resource type does not have the `CVMDGAction` attribute, then add the attribute.
 - Add the `CVMDGAction` attribute:

```
# /opt/VRTS/bin/haconf -makerw
# /opt/VRTS/bin/haattr -add CVMVolDg CVMDGAction -string
# /opt/VRTS/bin/hatype -modify CVMVolDg ArgList \
CVMDiskGroup CVMDGAction CVMVolume CVMActivation
# /opt/VRTS/bin/haconf -dump -makero
```
 - Verify that the `CVMVolDg` resource type has the `CVMDGAction` attribute:

```
# /opt/VRTS/bin/haattr -display CVMVolDg | \
grep -w CVMDGAction
```
- 12 Import any VxVM disk groups that are not under VCS control.
- 13 Mount any VxFS file systems that are not under VCS control.
- 14 Unfreeze the VCS service groups. Run the following commands:

```
# haconf -makerw
# hagr -unfreeze servicegroup -persistent
# haconf -dump -makero
```
- 15 Stop and restart the VCS on all nodes:

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

```
# /opt/VRTS/bin/hastart
```

- 16 Verify all VCS groups are now online and the database instances are up and running.

Upgrading CVM protocol and disk group version

To take advantage of the new features in release 5.0, you must upgrade the Veritas Cluster Volume Manager (CVM) protocol version (70), and upgrade to the latest disk group version (140).

To verify the cluster protocol version

- ◆ Verify the cluster protocol version, enter the following command:

```
# /opt/VRTS/bin/vxdctl protocolversion
```

If the cluster protocol version is less than 70, then it needs to be upgraded to 70 for SF 5.0 for Oracle RAC.

To upgrade the cluster protocol version

- ◆ From the CVM master node run:

```
# /opt/VRTS/bin/vxdctl upgrade
```

To upgrade the disk group version

- ◆ Upgrade the disk group version to 140 by entering the following command on the master node:

```
# vxdg -T 140 upgrade disk_group_name
```

Setting up SF Oracle RAC with Oracle 10g R2

- [Chapter 6, “Preparing to install Oracle 10g RAC” on page 129](#)
- [Chapter 7, “Installing Oracle 10g RAC” on page 147](#)
- [Chapter 8, “Upgrading and migrating to Oracle 10g R2 software” on page 157](#)
- [Chapter 9, “Configuring VCS service groups for Oracle 10g” on page 161](#)
- [Chapter 10, “Adding and removing SF Oracle RAC nodes for Oracle 10g” on page 175](#)
- [Chapter 11, “Uninstalling SF Oracle RAC with Oracle 10g” on page 187](#)

Preparing to install Oracle 10g RAC

This chapter contains the following topics:

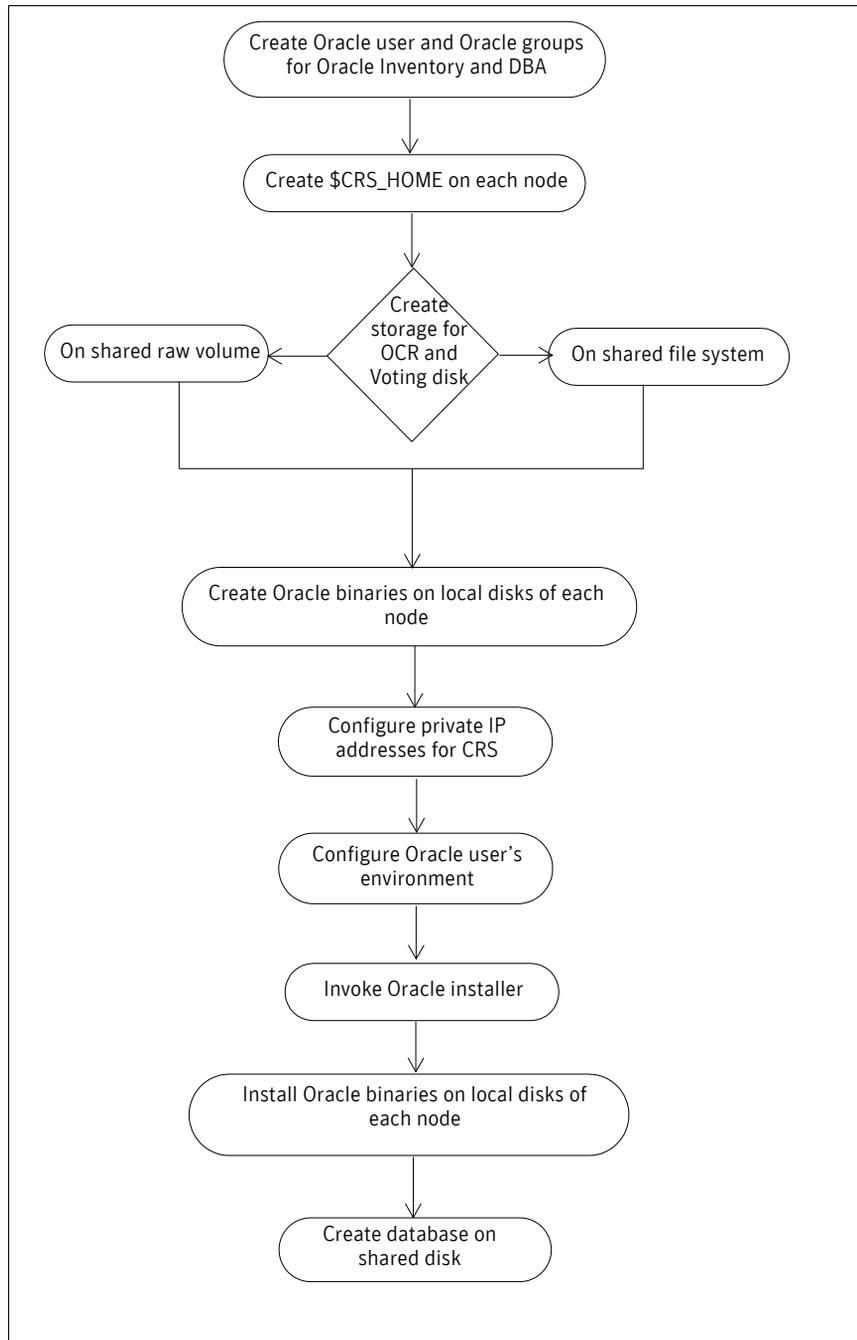
- [About preparing to install Oracle 10g RAC](#)
- [About Oracle 10g RAC in a Veritas SF Oracle RAC environment](#)
- [Identifying storage for Oracle 10g RAC components](#)
- [Performing the Oracle 10g preinstallation tasks](#)

About preparing to install Oracle 10g RAC

After setting up Veritas Storage Foundation for Oracle RAC, prepare to install Oracle 10g software. You can install the Oracle 10g software locally on each node. Make sure to review the Oracle installation manuals before installing Oracle 10g.

[Figure 6-9](#) describes a high-level flow of the Oracle 10g RAC installation process.

Figure 6-9 Oracle 10g RAC installation flow chart



About Oracle 10g RAC in a Veritas SF Oracle RAC environment

Review the information on Cluster Ready Services, the Oracle Cluster Registry, application resources, and the voting disk. Refer to the Oracle RAC documentation for more information.

Cluster Ready Services

Oracle clusterware, also known as Cluster Ready Services (CRS), manages Oracle cluster-related functions including membership, group services, global resource management, and databases. CRS is required for every Oracle 10g RAC instance and is only applicable to Oracle 10g.

CRS has the following major components:

- A cluster interconnect that allows for cluster communications.
- A private virtual IP address for cluster communications over the interconnect.
- A public virtual IP address for client connections.
- Shared storage accessible by each node that needs to run the application.

Oracle Cluster Registry

The Oracle Cluster Registry (OCR) contains cluster and database configuration and state information for Oracle 10g RAC and CRS. This is roughly analogous to the main.cf file and in-memory configuration in VCS. However, only one process performs I/O to the OCR file on disk.

The information maintained in the OCR includes:

- The list of nodes
- The mapping of database instances to nodes
- CRS application resource profiles
- Resource profiles that define the properties of resources under CRS control
- Rules that define dependencies between the CRS resources
- The current state of the cluster

The OCR data exists on a shared raw volume or a cluster file system that is accessible to each node. This requires approximately 100 MB of disk space.

You can specify redundant storage for the OCR to protect against failures. CRS faults nodes if the OCR is not accessible because of corruption or disk failure. Oracle automatically backs up OCR data. You can also export the OCR contents

before making configuration changes in CRS. This way, if you encounter configuration problems and are unable to restart CRS, you can restore the original contents.

Application Resources

CRS application resources are similar to VCS resources. Each component controlled by CRS is defined by an application resource, including databases, instances, services, and node applications.

Unlike VCS, CRS uses separate resources for components that run in parallel on multiple nodes.

Resource Profiles

Resources are defined by profiles, which are similar to the attributes that define VCS resources. The OCR contains application resource profiles, dependencies, and status information.

CRS Node Applications

CRS node applications are the resources CRS uses to manage RAC databases in a high availability environment. The installation of CRS creates a set of node applications by default. CRS uses these application resources to manage Oracle components, such as databases, listeners, and virtual IP addresses.

During installation, a `nodeapps` object is created as a container for all these application resources. This enables starting and stopping application resources as a group action.

Voting Disk

The voting disk is a heartbeat mechanism used by CRS to maintain cluster node membership. Voting disk data exists on a shared raw volume or a cluster file system that is accessible to each node. This requires approximately 20 MB of disk space.

The `ocssd` processes of CRS provides cluster node membership and group membership information to RAC instances. On each node, `ocssd` processes write a heartbeat to the voting disk every second. If a node is unable to access the voting disk, CRS determines the cluster is in a split brain condition and panics the node.

Because the voting disk is a critical component and single point of failure in a CRS cluster, the storage for the voting disk uses either the Oracle redundancy feature or third-party mirroring and multipathing features to prevent interruptions to the cluster. Placing the storage objects for the voting disk under

VCS control ensures that the voting disk is in place before the ocssd process is started (through resource dependencies) and provides notification in the event the voting disk fails.

Identifying storage for Oracle 10g RAC components

The SF Oracle RAC configuration program provides options to perform Oracle 10g preinstallation tasks, install Oracle 10g CRS, install Oracle 10g binaries, and link Veritas libraries with Oracle.

The configuration program prompts you for information while installing Oracle 10g in an SF Oracle RAC environment. The program provides default values for some information, which you can choose to use. The examples assume a two-node cluster, with nodes galaxy and nebula. Keep the following information at hand.

Getting the information to create Oracle user and group id

Keep the following information at hand to create the primary and secondary groups for Oracle.

- To create Oracle user and Oracle Inventory group as the primary group:

Oracle user name	oracle Default is oracle.
Oracle user id	100
Oracle user home directory	/app/oracle/orahome
Oracle group name	oinstall Default is oinstall.
Oracle group id	1000

- To create OS DBA group as the secondary group:

Oracle user name	oracle Default is oracle.
Oracle user id	100
Oracle user home directory	/app/oracle/orahome
Oracle group name	dba Default is oinstall.

Mount point for shared file system	/ocrvote
Directory for OCR files	ocr
Directory for Voting Disk files	vote

Identifying storage for Oracle 10g binaries

- To install Oracle 10g binaries on local disks of each node:

Local disk on each node	c4t0d2
Disk group for each local disk	<ul style="list-style-type: none"> ■ On galaxy: orabindg_galaxy ■ On nebula: orabindg_nebula
Volume for each local disk group	<ul style="list-style-type: none"> ■ On galaxy: orabinvol_galaxy ■ On nebula: orabinvol_nebula
File system on each local volume	<ul style="list-style-type: none"> ■ On galaxy: /dev/vx/rdisk/orabindg_galaxy/orabinvol_galaxy ■ On nebula: /dev/vx/rdisk/orabindg_nebula/orabinvol_nebula
Mount point for each local file system	/app/oracle/orahome
Directory to install Oracle 10g binaries on each local disk	/app/oracle/orahome

Identifying storage for CRS component

- To configure private IP addresses for CRS on each node:

Private IP address for each node	<ul style="list-style-type: none"> ■ On galaxy: 192.168.12.1 ■ On nebula: 192.168.12.2
NIC for network on each node	<ul style="list-style-type: none"> ■ On galaxy: 1an1 ■ On nebula: 1an2
Netmask for the cluster	255.255.240.0

Getting the information to configure Oracle user's environment

Keep the following information in hand to configure Oracle user's environment:

Value for DISPLAY environment variable	10.20.12.150:0.0 You can provide the host name or IP address of the system that you want to use to display the Oracle installer.
Value for ORACLE_BASE environment variable	/app/oracle

Getting the path information of the installer

Keep the following information in hand to invoke Oracle installer:

Absolute path of CRS installer	/var/Oracle10g/clusterware/Disk1
Absolute path of Oracle installer	/var/Oracle10g/database/Disk1

Performing the Oracle 10g preinstallation tasks

The Oracle installation requires the following preparation:

- [Creating Oracle user and group id](#)
- [Creating disk groups, volumes, and mount points for Oracle](#)
- [Configuring private IP addresses for CRS](#)
- [Editing the VCS configuration for additional CRS resources](#)
- [Creating public virtual IP addresses for Oracle](#)
- [Starting VCS, CVM, and CFS on all nodes](#)
- [Verifying GAB port membership](#)

You can choose to perform all preinstallation tasks or only specific tasks using the SF Oracle RAC configuration program. Note that you must manually perform some of the preinstallation tasks.

To perform the Oracle 10g preinstallation tasks

- 1 Launch the SF Oracle RAC configuration program:
`# installsfrac -configure`
- 2 Enter the system names.
- 3 From the configuration program menu, select **Prepare to install Oracle**.

- 4 Select the preinstallation task from the configuration menu.
- 5 If you select **Perform all Oracle preinstallation tasks**, the configuration program takes you through the other menu options:
 - Create Oracle user and group id
See [“Creating Oracle user and group id”](#) on page 137.
 - Create disk groups, volumes, and mount points for Oracle
See [“Creating disk groups, volumes, and mount points for Oracle”](#) on page 138.
 - Configure private IP addresses for CRS
See [“Configuring private IP addresses for CRS”](#) on page 141.

Creating Oracle user and group id

You must create the groups oinstall (the Oracle Inventory group) and dba, and the user oracle, assigning the primary group for oracle to be oinstall and the secondary group for oracle to be dba. Refer to *Oracle Real Application Clusters Installation and Configuration Guide* for more information.

You can create Oracle users and groups using the SF Oracle RAC configuration program. To create different groups, you must launch the configuration program and perform the procedure again.

To create an Oracle user and group id on all nodes

- 1 Select **Create Oracle user and group id** if you want to specifically perform this task.
- 2 Provide the required information.
The configuration program assigns the same values on all nodes. The text within parentheses is the default, which you may select by pressing Enter.


```
Enter Oracle UNIX user name: (oracle)
Enter Oracle UNIX user id: 100
Enter Oracle UNIX user home dir: /home/oracle
Enter Oracle UNIX group name: (oinstall)
Enter Oracle UNIX group id: 1000
User oracle does not exist on any node. Do you want to create it
with the information provided [y,n,q] (y)
```
- 3 Create secondary groups for Oracle user as required.
- 4 Create a password for the oracle user on each node.
passwd oracle
- 5 Review as the configuration program verifies the ssh or remsh permission for Oracle user on all the nodes.
Oracle installation requires ssh or remsh permission to be set for the Oracle user.

- 6 If the `ssh` or `remsh` verification fails on any nodes, enable `ssh` or `remsh` access in those nodes.

Creating disk groups, volumes, and mount points for Oracle

To create disk groups, volumes, and mount points for Oracle

- 1 Select **Create disk groups, volumes, and mount points for Oracle** if you want to specifically perform this task.
- 2 Review the displayed guidelines and proceed to [step 3](#).

For Oracle 10g, create the following on shared storage:

1. CVM volume or a file on a CFS filesystem for OCR.
2. CVM volume or a file on a CFS filesystem for VOTE-disk.
3. CVM volumes or a directory on a CFS filesystem for storing the database.

Oracle and CRS home directories (where the binaries are installed) can be on a local filesystem or on CFS.

Remember to change the permission for all of the above to `oracle user/group`.

- 3 Before you install the Oracle Cluster Ready Services (CRS) and Oracle 10g binaries, you must create storage space for these installations. Create storage space for:

The home directory (`$CRS_HOME`) for CRS binaries. On each node, create a volume and file system for `$CRS_HOME`.

See [“Preparing `\$CRS_HOME` on each node”](#) on page 139.

The CRS files for Oracle Cluster Registry (OCR) and the Vote-disk.

The CRS files can reside in volumes on raw device, or in directories in a cluster file system.

See [“Creating volumes or directories for OCR and Vote-disk”](#) on page 139.

- The home directory (`$ORACLE_HOME`) for Oracle binaries. Symantec recommends you to install the Oracle binaries on local storage on each node.
- The CVM volumes or a directory on CFS for database file storage.

See [“Preparing to store the Oracle binary and data files”](#) on page 141.

Preparing \$CRS_HOME on each node

On each system in the SF Oracle RAC cluster, create a directory for CRS_HOME. The minimum disk space required is 2 GB. The example procedure uses galaxy as one of the nodes.

To prepare \$CRS_HOME on each node

- 1 Make sure you are logged in as superuser on one system.
- 2 On one node, create a private disk group on local storage. For example:

```
# vxdg init crsdg_galaxy c4t0d0
```
- 3 Create the volume in the group for the CRS_HOME. The volume should be a minimum of 2 GB:

```
# vxassist -g crsdg_galaxy make crsvol_galaxy 2000M
```
- 4 Create a VxFS file system on which to install CRS. For example:

```
# mkfs -F vxfs /dev/vx/rdsk/crsdg_galaxy/crsvol_galaxy
```
- 5 Create the mount point for the file system.

```
# mkdir /app/crshome
```
- 6 Mount the file system, using the device file for the block device:

```
# mount -F vxfs /dev/vx/dsk/crsdg_galaxy/crsvol_galaxy \
/app/crshome
```
- 7 Add an entry to the /etc/fstab file. For example:

```
/dev/vx/dsk/crsdg_galaxy/crsvol_galaxy /app/crshome vxfs
delaylog 0 2
```
- 8 Set the CRS_HOME directory for the oracle user as /app/crshome.
- 9 Assign ownership of the CRS_HOME directory to oracle and the group oinstall:

```
# chown -R oracle:oinstall /app/crshome
```
- 10 Repeat [step 1](#) through [step 9](#) on each of the other nodes in the cluster.

Creating volumes or directories for OCR and Vote-disk

The OCR and Vote-disk must be shared among all nodes in a cluster. You can do one of the following:

- Create OCR and Vote-disk volumes on shared raw volumes.
 See [“To create OCR and Vote-disk volumes on raw volumes”](#) on page 140.
- Create OCR and Vote-disk directories in a cluster file system.
 See [“To create OCR and Vote-disk directories on CFS”](#) on page 140.

Whether you create shared volumes or shared file system directories, you can add them in the VCS configuration to make them highly available. Symantec recommends to create OCR and Vote-disk on shared raw volumes.

To create OCR and Vote-disk volumes on raw volumes

- 1 Log in as superuser.
- 2 On the master node, create a shared disk group:

```
# vxdbg -s init ocrvotedg c4t0d1
```
- 3 Create volumes in the shared group for OCR and Vote-disk:

```
# vxassist -g ocrvotedg make ocrvotevol 200M
```
- 4 Assign ownership of the volumes using the vxedit(1M) command:

```
vxedit -g disk_group set group=group user=user mode=660 volume
```

For example:

```
# vxedit -g ocrvotedg set group=oinstall user=oracle mode=660 ocrvotevol
```
- 5 Start the volume:

```
# vxvol -g ocrvotedg startall
```

To create OCR and Vote-disk directories on CFS

- 1 Log in as superuser.
- 2 On the master node, create a shared disk group:

```
# vxdbg -s init ocrvotedg c4t0d1
```
- 3 Create volumes in the shared group for OCR and Vote-disk:

```
# vxassist -g ocrvotedg make ocrvotevol 200M
```
- 4 Assign ownership of the volumes using the vxedit(1M) command:

```
vxedit -g disk_group set group=group user=user mode=660 volume
```

For example:

```
# vxedit -g ocrvotedg set group=oinstall user=oracle mode=660 ocrvotevol
```
- 5 Start the volume:

```
# vxvol -g ocrvotedg startall
```
- 6 Set the activation mode for the disk group on all cluster nodes:

```
# vxdbg -g ocrvotedg set activation=sw
```
- 7 Create the file system:

```
# mkfs -F vxfs /dev/vx/rdisk/ocrvotedg/ocrvotevol
```
- 8 Create a mount point for the Vote-disk and OCR files on all the nodes.

```
# mkdir /ocrvote
```
- 9 On all nodes, mount the file system:

```
# mount -F vxfs -o cluster /dev/vx/dsk/ocrvotedg/ocrvotevol\ /ocrvote
```
- 10 Create a directory for the Vote-disk and OCR files.

```
# cd /ocrvote  
# mkdir ocr
```

```
# mkdir vote
```

- 11 Set "oracle" to be the owner of the file system, and set "755" as the permissions:

```
# chown -R oracle:oinstall /ocrvote
# chmod 755 /ocrvote
```

Preparing to store the Oracle binary and data files

Symantec recommends you to install the Oracle 10g Database Binaries locally on each cluster node. You must create the disk group, volume, and mount points on each node that would be a part of the cluster.

To create storage space for Oracle files on local disks

Perform the following steps in each node in the cluster.

- 1 Log in as superuser on one node.
- 2 On one node, create a disk group in the available local storage.


```
# vxdg init orabindg_galaxy c4t0d2
```
- 3 Create the volume in the group. For the Oracle 10g binaries, make the volume 7 GB.


```
# vxassist -g orabindg_galaxy make orabinvol_galaxy 7168M
```
- 4 Create a VxFS file system on orbin_vol to install the Oracle 10g binaries. For example:


```
# mkfs -F vxfs /dev/vx/rdisk/orabindg_galaxy/orabinvol_galaxy
```
- 5 Create the mount point for the file system.


```
# mkdir /app/oracle/orahome
```
- 6 Mount the file system using the device file for the block device.


```
# mount -F vxfs /dev/vx/dsk/orabindg_galaxy/orabinvol_galaxy \
/app/oracle/orahome
```
- 7 Edit the /etc/fstab file and list the new file system. For example:


```
/dev/vx/dsk/orabindg_galaxy/orabinvol_galaxy /app/oracle/
orahome vxfs delaylog 0 1
```
- 8 Assign ownership of the Oracle directory to oracle:


```
# chown -R oracle:oinstall /app/oracle/orahome
# chmod 775 /app/oracle/orahome
```
- 9 Repeat [step 1](#) through [step 8](#) on the other nodes.

Configuring private IP addresses for CRS

The CRS daemon requires a private IP address on each node to enable communications and heartbeating.

Determine a private NIC device for which LLT is configured. Look at the file / etc/llttab. For example, if lan0 is used as an LLT interconnect on one system, you can configure an available IP address for it. The private IP addresses of all nodes should be on the same physical network in the same IP subnet.

To configure private IP addresses for CRS

- 1 Select **Configure private IP addresses for CRS** if you want to specifically perform this task.

- 2 Enter the private IP address for the system and interface for each host.

```
Enter the private IP for galaxy: [b] 192.168.12.1
Enter the NIC for network 1 for galaxy (x if done): [b] lan1
Enter the NIC for network 2 for galaxy (x if done): [b] lan2
Enter the NIC for network 3 for galaxy (x if done): [b] x
```

```
Enter the private IP for nebula: [b] 196.168.12.2
Enter the NIC for network 1 for nebula (x if done): [b] (lan1)
Enter the NIC for network 2 for nebula (x if done): [b] (lan2)
Enter the NIC for network 3 for nebula (x if done): [b] x
```

- 3 Enter the netmask for the cluster.

```
Enter the netmask for private network: [b] 255.255.240.0
```

- 4 Review and confirm the private IP address information for CRS.

CRS private IP configuration information verification

```
System Name: galaxy
CRS IP address: 192.168.12.1
Interfaces: lan1 lan2
```

```
System Name: nebula
CRS IP address: 192.168.12.2
Interfaces: lan1 lan2
```

```
Netmask: 255.255.240.0
```

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

After confirming the values, the utility adds a new section in the VCS configuration file (main.cf) for the PrivNIC resource in the CVM group. Here is a sample segment of an updated PrivNIC information in the main.cf:

```
.
.
group cvm (
  SystemList = { galaxy = 0, nebula = 1 }
  AutoFailOver = 0
  Parallel = 1
  AutoStartList = { galaxy, nebula }
)
.
.
```

```
PrivNIC ora_priv (
  Critical = 0
  Device = { lan1 = 0, lan2 = 1}
  Address@galaxy = "192.168.12.1"
  Address@nebula = "192.168.12.2"
  NetMask = "255.255.240.0"
)
.
.
```

See [“About PrivNIC agent”](#) on page 376.

- 5 On each system, add the configured private IP addresses of all nodes to the `/etc/hosts` file, mapping them to symbolic names such as `galaxy-priv`.
- 6 From each system, try pinging each of the other nodes, using the symbolic system name associated with the private NIC IP address.

Editing the VCS configuration for additional CRS resources

In addition to automatically updating the PrivNIC resource, you must manually modify the CVM service group in the `main.cf` to add or modify other resources for CRS:

- Add the CVMVolDg resource for the volume created for the cluster file system. The resource needs to contain the OCR and Vote-disk directories. You can use the same CVMVolDg resource for multiple OCR and Vote-disk volumes created on raw device.
- Add the resource for the shared file system containing the OCR and Vote-disk directories. This process does not apply to separate volumes for the OCR and Vote-disk.
- Modify the Application resource for the Oracle `cssd` process within the CVM service group. This resource is required for monitoring the `cssd` process. If the `cssd` process is not configured in the Application resource, the commands used to offline the shared mounts (such as `hastop -local`, `hastop -all`, or `hagrpl -offline`) cause the nodes to automatically reboot.

To modify the `main.cf` for additional CRS resources

- 1 Log into one of the nodes in the cluster.
- 2 Save the existing configuration to disk and make the configuration read-only while making the changes:


```
# haconf -dump -makero
```
- 3 To ensure VCS is not running while you edit the `main.cf`, use the `hastop` command to stop the VCS engine on all nodes and leave the resources available:

```
# hastop -all -force
```

- 4 Make a backup copy of the main.cf file:

```
# cd /etc/VRTSvcs/conf/config  
# cp main.cf main.orig
```

- 5 Use vi or another text editor to edit the main.cf file:

- 6 Add the CVMVoIDG resource for the volume. In this example, only one volume exists because the OCR and Vote-disk are directories in a cluster file system:

```
CVMVoIdg ocrvote_voldg (  
    Critical = 0  
    CVMdiskGroup = ocrvotedg  
    CVMVolume = { ocrvotevol }  
    CVMActivation = sw  
)  
.  
.
```

- 7 If you created a cluster file system for the OCR and Vote-disk directories, add the CFSSMount resources for the cluster file system:

```
CFSSMount ocrvote_mnt (  
    Critical = 0  
    MountPoint = "/ocrvote"  
    BlockDevice = "/dev/vx/dsk/ocrvotedg/ocrvotevol"  
)  
.  
.
```

- 8 Revise the dependencies section for the CVM group.

For example, you must change the dependencies that resembles:

```
cvm_clus requires cvm_vxconfigd  
vxfsckd requires cvm_clus
```

To:

```
ocrvote_mnt requires ocrvote_voldg  
ocrvote_mnt requires vxfsckd  
ocrvote_voldg requires cvm_clus
```

- 9 Modify the CVM service group to monitor the cssd program that uses the Application resource.

```
Application cssd (  
    Critical = 0  
    StartProgram = "/opt/VRTSvcs/rac/bin/cssd-online"  
    StopProgram = "/opt/VRTSvcs/rac/bin/cssd-offline"  
    MonitorProgram = "/opt/VRTSvcs/rac/bin/cssd-monitor"  
    CleanProgram = "/opt/VRTSvcs/rac/bin/cssd-clean"  
)
```

- 10 Save and close the main.cf file.

- 11 Verify the syntax of the file /etc/VRTSvcs/conf/config/main.cf:

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

12 Copy the saved `main.cf` file to all nodes in the cluster.

Creating public virtual IP addresses for Oracle

Before starting the Oracle installation, you must create virtual IP addresses for each node. An IP address and an associated host name should be registered in the domain name service (DNS) for each public network interface.

To create public virtual IP addresses for Oracle

- 1 On each cluster node, select a public NIC device and configure a virtual IP address.

For example:

```
# ifconfig lan1 plumb
# ifconfig lan1 inet 10.10.11.1 netmask 255.255.240.0
# ifconfig lan1 up
```

- 2 On each node, add the public virtual IP addresses to the `/etc/hosts` file, mapping them to their respective host names, such as `galaxy` and `nebula`.

Starting VCS, CVM, and CFS on all nodes

With the configuration file in place on each node, verify that the resources you configured come online.

To start VCS, CVM, and CFS on all nodes

- ◆ Start VCS, CVM, and CFS from each node:

```
# hstart
```

Verifying GAB port membership

After starting VCS, CVM, and CFS on each node, verify GAB port membership by running the `gabconfig -a` command.

For example:

```
galaxy# gabconfig -a
GAB Port Memberships
=====
Port a gen   ada401 membership 0123
Port b gen   ada40d membership 0123
Port d gen   ada409 membership 0123
Port f gen   ada41c membership 0123
Port h gen   ada40f membership 0123
Port o gen   ada406 membership 0123
Port v gen   ada416 membership 0123
Port w gen   ada418 membership 0123
```


Installing Oracle 10g RAC

This chapter contains the following topics:

- [About installing Oracle 10g RAC](#)
- [Setting Oracle environment variables and invoking the Oracle installer](#)
- [Installing CRS](#)
- [Installing Oracle 10g Binaries](#)
- [Completing Oracle 10g post-installation tasks](#)

About installing Oracle 10g RAC

After completing the Oracle 10g pre-installation tasks, proceed to install Oracle 10g software locally on each node. SF Oracle RAC supports Oracle 10g R2. Make sure to review the Oracle installation manuals.

When you install Oracle, do not create a database or install a pre-created database. For example, when you choose to use the **Enterprise Edition** or the **Standard Edition** displayed by the Oracle installer, select the **Software only** option to install the binaries without creating a database. If you choose **Custom installation**, do *not* create a database when prompted. Create the database after installing Oracle and relinking the SF Oracle RAC libraries to Oracle.

You can also upgrade to Oracle 10g from Oracle9i. Refer to the Oracle documentation or access the Oracle Technology Network for specific information.

See [“Migrating the database”](#) on page 158.

See [“Applying Oracle patchsets”](#) on page 158.

Setting Oracle environment variables and invoking the Oracle installer

Review the procedure to set the environment variables using the SF Oracle RAC configuration program. You must perform the procedure before installing Oracle 10g CRS and before installing Oracle 10g binaries. Depending on what you want to install, do one of the following:

- If you want to install Oracle 10g CRS, use the procedure to invoke Oracle 10g CRS.
- If you want to install Oracle 10g binaries, invoke the Oracle 10g binaries installation.

To set the Oracle environment variables and invoke the Oracle installer

- 1 If you want to pass any arguments to Oracle installer, then set up the `OUI_ARGS` environment variable. You must perform this step before you start the SF Oracle RAC configuration program.
Refer to the Oracle documentation for details on the arguments that you can pass to the Oracle installer.

For example:

```
# export OUI_ARGS=-ignoreSysPrereqs
```

- 2 Launch the SF Oracle RAC configuration program.

```
# installsfrac -configure
```
- 3 Select **Install or relink Oracle** from the menu.
- 4 Review the latest SF Oracle RAC support matrix as the installer instructs.
- 5 Choose **Oracle 10gR2** as the version of the Oracle software.
- 6 Depending on what you want to install, select one of the following:

CRS **Install Oracle Clusterware (CRS)**

Oracle binaries **Install Oracle RDBMS server**

- 7 Set the `DISPLAY` environment variable that is required for the Oracle Universal Installer.

```
Enter DISPLAY environment variable: 10.20.12.150:0.0
```

- 8 Enter the Oracle user name.

```
Enter Oracle UNIX user name: (oracle) oracle
```

- 9 Enter the Oracle group name. For example:

```
CRS                    Enter Oracle UNIX group name: (oinstall)
```

```
Oracle binaries      Enter Oracle UNIX group name: (oinstall) dba
```

- 10** Enter the absolute path of the Oracle CRS or database software install image.

This is the path to the software disc where the installer resides.

```
CRS                  Enter absolute path of CRS install image: [b] /  
var/Oracle10g/clusterware/Disk1
```

```
Oracle binaries     Enter absolute path of Database install image: /  
var/Oracle10g/database/Disk1
```

- 11** Enter the Oracle base directory.

```
Enter Oracle base directory: /app/oracle
```

- 12** Enter the absolute path of the CRS home directory.

```
Enter absolute path of CRS Home directory: [b] /app/crshome
```

- 13** If the CRS home directory you specified does not exist, the SF Oracle RAC configuration program provides guidelines to create one.

```
Do you wish to create this directory? [y,n,q,b] (n) y
```

```
mkdir -p /app/crshome  
chown -R oracle:oinstall /app/crshome  
chmod -R 744 /app/crshome  
Run above command on all systems? [y,n,q,b] (n) y
```

Now, enter the absolute path. See [step 12](#).

- 14** Review the output as the SF Oracle RAC configuration program verifies the CRS health status.

```
Verifying CRS health status  
CRS check succeeded on galaxy.  
CRS check succeeded on nebula.
```

Proceed to install CRS when the program launches the Oracle 10g CRS utility runInstaller.

See [“Installing CRS”](#) on page 150.

- 15** If you chose **Install Oracle RDBMS server**, enter the absolute path of the database home directory.

```
Enter absolute path of Database Home directory: [b] /app/oracle/  
orahome
```

Proceed to install Oracle RAC software when the program launches the Oracle Universal Installer.

See [“Installing Oracle 10g Binaries”](#) on page 152.

- 16** If the database home directory you specified does not exist, the SF Oracle RAC configuration program provides guidelines to create one.

```
Do you wish to create this directory? [y,n,q,b] (n) y
```

```
mkdir -p /app/oracle/orahome
chown -R oracle:oinstall /app/oracle/orahome
chmod -R 744 /app/oracle/orahome
Run above command on all systems? [y,n,q,b] (n) y
Now, enter the absolute path. See step 15.
```

- 17 Proceed with the Oracle 10g installation.
See [“Installing Oracle 10g RAC”](#) on page 147.

Installing CRS

If you chose **Install Oracle Clusterware (CRS)** in the SF Oracle RAC configuration program, then after you set the Oracle 10g environment variables, the program launches the Oracle utility to install CRS.

Make sure any required HP-UX patches listed in the Oracle documentation are installed before you install the CRS software. You must install CRS in the `$CRS_HOME` location on each node that you created.

See [“Preparing \\$CRS_HOME on each node”](#) on page 139.

Note: When the Oracle 10g CRS installation prompts you to run the `root.sh` script, make sure to patch the `init.cssd` script before running `root.sh`.

To install CRS

- 1 Make sure you set the environment variables using the configuration program.
See [“Setting Oracle environment variables and invoking the Oracle installer”](#) on page 148.
- 2 When the Oracle installer prompts you to run the `/oracle/oraInventory/orainstRoot.sh` script, make sure the script exists on each node before proceeding.
- 3 If the `/oracle/oraInventory/orainstRoot.sh` script does not exist on each node, copy it from the first node to each of the other cluster nodes.
- 4 Run the `/oracle/oraInventory/orainstRoot.sh` script on each node.

- 5 As you run the installer, be prepared with the following information required for the installation and configuration of the CRS component:

Specify File Locations dialog box	<ul style="list-style-type: none">■ The installation destination (\$CRS_HOME) on each node. For example, /oracrs.■ The path to \$CRS_HOME.
Cluster Configuration dialog box	<p>The names of the nodes and their associated host names.</p> <p>In the example CRS installation, the nodes and their associated host names are galaxy and nebula.</p>
Private Interconnect Enforcement dialog box	<p>The private NIC IP addresses you have set up for each node.</p> <p>For example, lan1 on subnet 1.0.0.0.</p>
Oracle Cluster Registry dialog box	<p>The name of a file in the OCR directory or raw volume.</p> <p>For example, a file named ocr_file in the directory /ocrvote/ocr.</p>
Voting Disk dialog box	<p>The name of a file in the Vote-disk directory or raw volume.</p> <p>For example, a file named vote_file in the directory /ocrvote/vote.</p>

- 6 When you arrive at the Install screen, click **Install**.
- 7 Follow the wizard instructions to install CRS.
- 8 In the Setup Privileges Notice dialog box, click **OK**.
The installer prompts you to run the `root.sh` script.
- 9 Before you run the `root.sh` script, you need to add the `init.cssd.patch`.
- Open another window on the system where you are running the installer.
 - Log in as superuser.
 - Change to the directory where the patch is to be copied:

```
# cd $CRS_HOME/css/admin  
# cp /opt/VRTSvcs/rac/patch/init.cssd-10gR2.patch .
```
 - Run the following command to install the patch:

```
# patch < init.cssd.patch init.cssd
```
 - Run the `root.sh` script. For example:

```
# cd $CRS_HOME  
# ./root.sh
```

This starts the CRS daemons on the node where you enter the command.

- 10 Repeat [step 9](#) on each cluster node.

Installing Oracle 10g Binaries

Review the procedure to install the Oracle 10g software in an SF Oracle RAC environment. Symantec recommends you to install Oracle binaries locally on each node. Make sure to review the Oracle installation manuals.

See [“Preparing to store the Oracle binary and data files”](#) on page 141.

Make sure any required HP-UX patches listed in the Oracle documentation are installed before you install the Oracle software. Follow the Oracle Universal Installer wizard instructions and complete the installation. Refer to Oracle documentation for specific information on how to install Oracle.

To install Oracle 10g

- 1 If you chose **Install Oracle RDBMS server** in the SF Oracle RAC configuration program, then after you set the Oracle 10g environment variables, the program launches the Oracle installer.
- 2 As you run the installer, be prepared with the following information required for the installation and configuration of the Oracle 10g binaries:

Specify File Locations dialog box ■ The installation destination (\$ORACLE_HOME). Depending on where you install Oracle 10g binaries, this location is either on shared storage or an identical location on each of the local cluster nodes. For example, /oracle/VRT.

■ The path to \$ORACLE_HOME.

Specify Hardware Cluster Installation Mode dialog box The names of the cluster nodes and their associated host names.

In the example Oracle installation, the nodes and their associated host names are galaxy and nebula.

- 3 When you come to the Select Database Configuration dialog box, choose *not* to have a starter database created. Click **Do not create a starter database**. Symantec recommends you to create a database later.
- 4 When you arrive at the Install screen, click **Install**.

- 5 Follow the wizard instructions to install Oracle 10g binaries.
- 6 In the Setup Privileges Notice dialog box, click **OK**.
The installer prompts you to run the `root.sh` script.
- 7 Run the `root.sh` script on each node. For example:

```
# cd $ORACLE_HOME
# ./root.sh
```
- 8 When you start the script, a VIP Configuration Assistant window appears.
Enter the virtual IP address you are configuring for the node on which you run `root.sh`.
See “[Creating public virtual IP addresses for Oracle](#)” on page 145.
- 9 Complete with the Oracle 10g installation.
- 10 Review the output as the SF Oracle RAC configuration program links the SF Oracle RAC libraries to Oracle.
- 11 After installing Oracle 10g software, perform the post-installation tasks.
See “[Completing Oracle 10g post-installation tasks](#)” on page 153.

Completing Oracle 10g post-installation tasks

After installing Oracle 10g software, complete the post-installation tasks that include:

- “[Adding Oracle 10g patches](#)” on page 153
- “[Relinking the SF Oracle RAC libraries to Oracle 10g](#)” on page 154
- “[Creating Oracle 10g database](#)” on page 156
- “[Configuring Oracle 10g service group in a VCS configuration](#)” on page 156

Adding Oracle 10g patches

Before installing Oracle 10g Patch software,

- Review the latest information on supported Oracle patches:
<http://entsupport.symantec.com/docs/280186>
- You must have installed Oracle 10g R2 (10.2.0.1) software.
- Review the *Patch Set Notes* that accompany the patch set for instructions on installing the patch set and performing the post-installation operations.

Review the procedure to add Oracle 10g patches to your node if you have installed Oracle, but have not yet configured Oracle in your cluster.

To install Oracle 10g patches

- 1 Log in as oracle user.
- 2 On one node, create the directory where you intend to copy the Oracle 10g patch software. For example:

```
$ mkdir /oracle/patch  
$ cd /oracle/patch
```
- 3 Copy all files included with the downloaded Oracle 10g patch software to the /oracle/patch.
When you uncompress and untar the downloaded ZIP file, the software resides in a Disk1 directory.
- 4 On each node, edit the .rhosts file to provide the other node access to the local system during the installation.
Place a "+" character in the first line of the file. You can remove this permission after the patch installation is complete.
- 5 Set the DISPLAY variable. Do one of the following:

```
Bourne Shell (sh or ksh)  $ DISPLAY=host:0.0;export DISPLAY  
C Shell (csh or tcsh)   $ setenv DISPLAY host:0.0
```
- 6 Run the Oracle 10g runInstaller utility.

```
$ $ORACLE/patch/Disk1/runInstaller
```
- 7 Select products.xml from the /oracle/patch/Disk1/stage directory.
- 8 In the Node Selection dialog box, select all nodes for installation to install the patches on local disks of each node.
If you are installing Oracle 10g patches on a local file system, install the software by invoking the installer on each node, one node at a time.
- 9 Proceed with the installation.
- 10 After installing the Oracle 10g patches, you must perform the following tasks:
 - Preparatory tasks to relink the SF Oracle RAC libraries
 - Relink the SF Oracle RAC libraries to OracleSee [“Relinking the SF Oracle RAC libraries to Oracle 10g”](#) on page 154.

Relinking the SF Oracle RAC libraries to Oracle 10g

If you added or upgraded the Oracle patches, you must relink the SF Oracle RAC libraries to Oracle. You must link Oracle with the Veritas IPC library and enable

Oracle to use the Veritas ODM and cluster membership (VCSMM) libraries. You can relink the libraries using the SF Oracle RAC configuration program.

Note: Symantec recommends you to relink the SF Oracle RAC libraries *only* after completing all the required patch additions.

To prepare to relink the SF Oracle RAC libraries

- 1 Back up the following files, if present:

```
$ORACLE_HOME/lib/libskgxpu.a  
$ORACLE_HOME/lib/libskgxpg.a  
$ORACLE_HOME/lib/libskgxpu.so
```

You can restore these backed up files if the linking gets cancelled or interrupted during execution.

- 2 Remove the following files, if present:

```
$ORACLE_HOME/lib/libskgxpu.oracle.a  
$ORACLE_HOME/lib/libskgxpg.oracle.a  
$ORACLE_HOME/lib/libskgxpu.oracle.so
```

To relink the SF Oracle RAC libraries to Oracle 10g

- 1 Launch the SF Oracle RAC configuration program.
`# installsfrac -configure`
- 2 Select **Install or relink Oracle** from the menu.
- 3 Choose **Oracle 10gR2** depending on the version of the Oracle software.
- 4 Select **Relink Oracle**.
- 5 Enter the following information as the configuration program prompts you:
Enter Oracle UNIX user name: [b] (oracle)
Enter Oracle UNIX group name: [b] (oinstall)
Enter Oracle base directory: [b] (/oracle)
Enter absolute path of CRS Home directory: [b] (/crs_home)
Enter absolute path of Database Home directory: [b] (/oracle/
VRT)
- 6 Confirm the Oracle environment information that you entered.
- 7 Review the output as the SF Oracle RAC configuration program relinks the libraries.
The configuration program links Oracle with Veritas libraries on each cluster node.

Verifying whether Oracle uses Veritas libraries

After relinking the libraries, you must verify whether Oracle uses Veritas libraries.

To verify whether Oracle uses Veritas libraries

- 1 On each node, start the Oracle service group:

```
# hagrps -online oracle_database_grp -sys system_name
```
- 2 After starting Oracle instances, confirm that Oracle uses the Veritas libraries.

On each node in the cluster, examine the Oracle alert file, alert_\${ORACLE_SID}.log, for the following lines:

```
Oracle instance running with ODM: Veritas 5.0 ODM Library,  
Version 5.0  
cluster interconnect IPC version:Veritas IPC xxxxxxxxx
```

Creating Oracle 10g database

Refer to the *Oracle Real Application Clusters Installation and Configuration Guide* for instructions on how to install the Oracle 10g database. Create the Oracle database on shared storage. Use your own tools or review the guidelines on using the Oracle dbca (Database Creation Assistant) tool. to create a database on shared raw VxVM volumes or shared VxFS file systems.

See [“Creating a starter database for Oracle 10g”](#) on page 381.

Upgrading the databases

If you currently have Oracle databases running and want to migrate them to the latest Oracle patch level, refer to the README.html file downloaded with the patch. The file is located in the /oracle/patch directory.

Configuring Oracle 10g service group in a VCS configuration

After installing Oracle 10g and creating a database, proceed to modify the VCS configuration file. Review the sample VCS configurations and details on configuring service groups in an Oracle 10g environment.

See [“Configuring VCS service groups for Oracle 10g”](#) on page 161.

Upgrading and migrating to Oracle 10g R2 software

This chapter contains the following topics:

- [Before you upgrade and migrate to Oracle 10g R2](#)
- [Migrating the database](#)
- [Applying Oracle patchsets](#)

Before you upgrade and migrate to Oracle 10g R2

The migrating procedure assumes that you have installed Oracle9i or Oracle 10g R1 on the cluster nodes.

Before you upgrade and migrate the Oracle9i or Oracle 10g R1 software, do the following:

- Review the configuration of VCS service groups.
See [“Configuring VCS service groups for Oracle 10g”](#) on page 161.
- Review the example main.cf file.
See [“Sample VCS configuration files for SF Oracle RAC”](#) on page 355.
- Install SF 5.0 for Oracle RAC if you have not already installed it.
See [“Installing and configuring SF Oracle RAC components”](#) on page 77.
- Configure Oracle 10g prerequisites.
See [“Preparing to install Oracle 10g RAC”](#) on page 129.
- Install CRS for Oracle 10g R2.
See [“Installing CRS”](#) on page 150.
- Install Oracle 10g R2 binaries.
See [“Installing Oracle 10g Binaries”](#) on page 152.

Migrating the database

Follow any of the methods to migrate the database:

- Using the database upgrade assistant utility
- Performing a manual upgrade
- Data copying using export/import utilities

For the details on the methods to migrate to Oracle 10g R2, refer to the *Oracle Database Upgrade Guide 10g Release 1(10.2)*, Part Number B14238-01.

After upgrading the database, ensure that Oracle is linked to the Veritas libraries for Oracle.

See [“Relinking the SF Oracle RAC libraries to Oracle 10g”](#) on page 154.

Applying Oracle patchsets

Storage Foundation for Oracle RAC supports:

- Oracle 10g 10.2.0.1 or greater

To apply an Oracle 10g patchset

- 1 Log in as the oracle user.
- 2 Apply the patchset.

Note: Refer to the Oracle 10g patchset Installation and Configuration Guide for instructions on applying patchsets.

When applying a patchset to Oracle CRS, make sure to patch `init.cssd` before running `root<xxx>.sh` script as follows:

- a Change to the directory to which the patch is to be copied:

```
# cd /oracrs/install/patch<xxx>/css/admin
# cp /opt/VRTSvcs/rac/patch/init.cssd-10gR2.patch.
```
- b Run the following command to install the patch:

```
# patch < init.cssd-10gR2.patch init.cssd
```

- 3 Perform post-upgrade relinking.

See [“Relinking the SF Oracle RAC libraries to Oracle 10g”](#) on page 154.

After applying an Oracle patchset, the Oracle libraries must be relinked. This post-upgrade reconfiguration allows Oracle to use the Veritas RAC libraries rather than the native Oracle libraries. This is a required step in the Oracle RAC installation process as Oracle will not function properly in an SF Oracle RAC environment without the Veritas libraries.

- 4 When Oracle 10g installation is complete, remove the temporary `rsh` access permissions you have set for the systems in the cluster. For example, if you added a “+” character in the first line of the `.rhosts` file, remove that line. For the complete procedure, see the *Veritas Cluster Server Installation Guide for HP-UX*.
- 5 Check if Oracle is using the Veritas ODM library. After starting Oracle instances, you can confirm that Oracle is using the Veritas ODM libraries by examining the Oracle alert file, `alert_$(ORACLE_SID).log`. Look for the line that reads:

```
Oracle instance running with ODM: Veritas xxx ODM Library,  
Version x.x
```
- 6 Create an Oracle database on shared storage. Use your own tools to create a starter database.
See [“Creating a starter database”](#) on page 381.

Configuring VCS service groups for Oracle 10g

This chapter contains the following topics:

- [About VCS service group for Oracle 10g dependencies](#)
- [Configuring CVM and Oracle service groups](#)
- [Location of VCS log files](#)

About VCS service group for Oracle 10g dependencies

Review the information on how to set up VCS to automate the Oracle 10g RAC environment and how VCS manages resources within a cluster.

VCS service group dependencies are based on whether you use the VCS Oracle agent or not. [Figure 9-10](#) and [Figure 9-11](#) illustrate the dependencies.

- In a configuration without the VCS Oracle agent, CRS controls the database. See [Figure 9-10, "Configuration without the Oracle agent."](#)
- In a configuration with the VCS Oracle agent, VCS controls the Oracle database. An online local firm dependency exists between the Oracle group and the CVM group. For more details on service group dependencies, refer to the *Veritas Cluster Server User's Guide*. See [Figure 9-11, "Configuration with the Oracle agent."](#)

Figure 9-10 Configuration without the Oracle agent

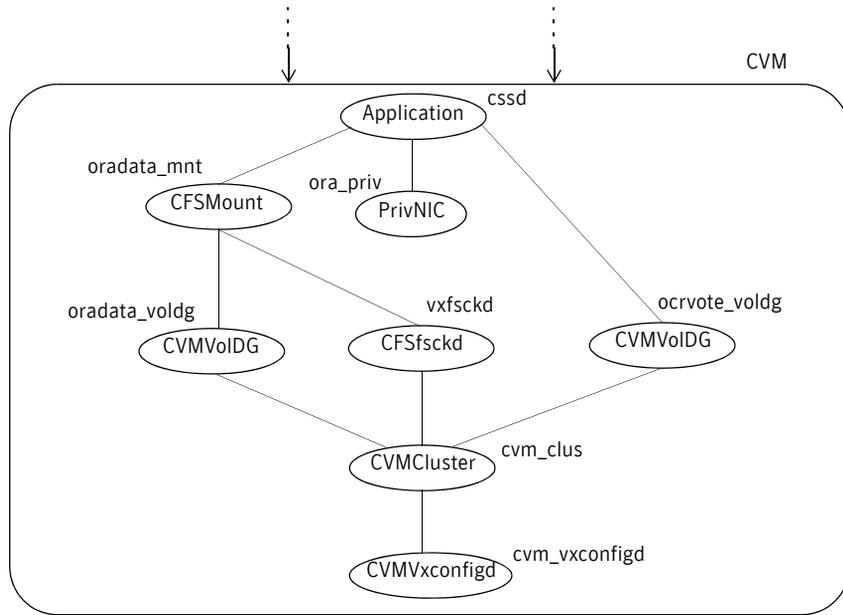
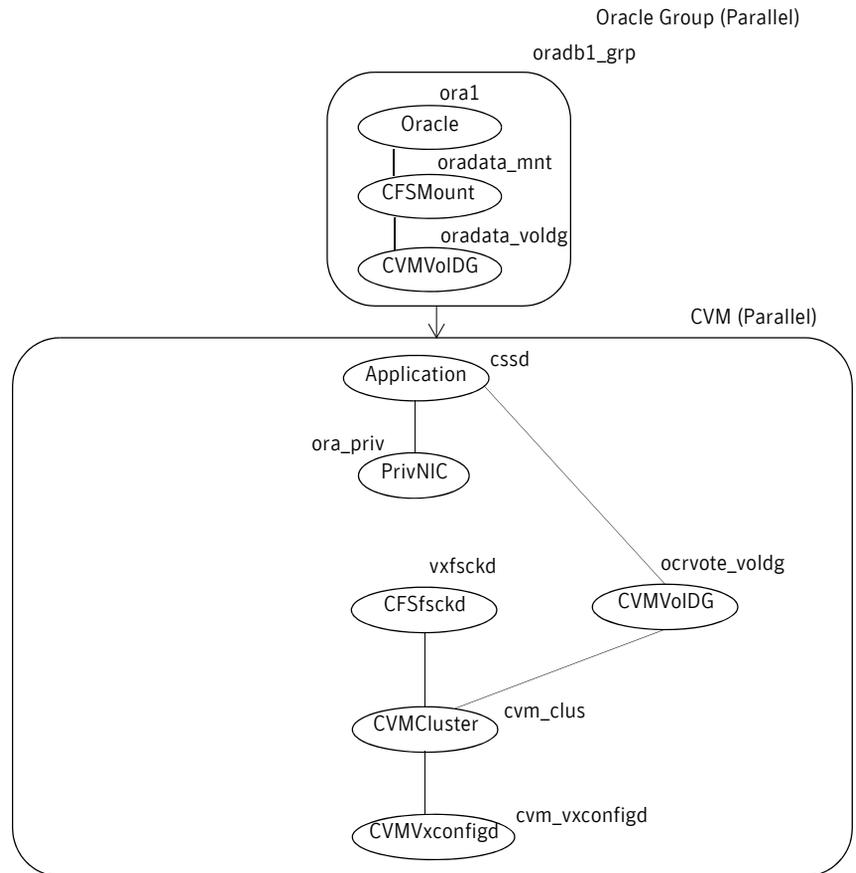


Figure 9-11 Configuration with the Oracle agent



CRS and Oracle agent might attempt to start the instance at the same time if the database mount is available. To prevent automatic database starting, you must change the Management policy for the database (automatic, manual) to MANUAL using SRVCTL command. The command will change AUTO_START attribute of CRS database and instance resources.

To prevent automatic database starting

- ◆ Depending on the status of the database:
 - If the database is already registered and running:
 - To change it to manual execute the following commands as Oracle user:
 - \$ **srvctl stop database -d db-name**
 - \$ **srvctl modify database -d db-name -y manual**

- If the database is not registered, use the following command to register database:

```
$ srvctl add database -d db-name -p \  
location-of-parameter-file -y manual
```

Configuring CVM and Oracle service groups

The CVM and Oracle service groups can be configured using the following two methods:

- By editing the VCS configuration file, main.cf, to define the service groups
See [“Configuring CVM service group for Oracle 10g manually”](#) on page 164.
- By using a configuration wizard for Oracle RAC
See [“Configuring the service groups using the wizard”](#) on page 165.

Configuring CVM service group for Oracle 10g manually

This section describes how to manually edit the main.cf file to configure the CVM and Oracle service groups.

To configure CVM service group for Oracle 10g manually

- 1 Log in to one system as superuser.
- 2 Save your existing configuration to prevent any changes while you modify main.cf:

```
# haconf -dump -makero
```
- 3 Ensure VCS is not running while you edit main.cf by using the `hastop` command to stop the VCS engine on all systems and leave the resources available:

```
# hastop -all -force
```
- 4 Make a backup copy of the main.cf file:

```
# cd /etc/VRTSvcs/conf/config  
# cp main.cf main.orig
```
- 5 Using vi or another text editor, edit the main.cf file, modifying the cvm service group and creating Oracle service groups using the sample main.cf files as a guideline.
See [“Sample main.cf for Oracle 10g without Oracle agent”](#) on page 356.
See [“Sample main.cf for Oracle 10g with Oracle agent”](#) on page 357.

Saving and checking the configuration

When you finish configuring the CVM and Oracle service groups by editing the main.cf file, verify the new configuration.

To save and check the configuration

- 1 Save and close the `main.cf` file.
- 2 Verify the syntax of the file `/etc/VRTSvcs/conf/config/main.cf`:

```
# cd /etc/VRTSvcs/conf/config
# hacf -verify .
```
- 3 Start the VCS engine on one system:

```
# hastart
```
- 4 Type the command `hastatus`:

```
# hastatus
```
- 5 When “LOCAL_BUILD” is listed in the message column, start VCS on the other system:

```
# hastart
```
- 6 Verify that the service group resources are brought online. On one system, enter:

```
# hagrps -display
```

Modifying the VCS Configuration

For additional information and instructions on modifying the VCS configuration by editing the `main.cf` file, refer to the *Veritas Cluster Server User's Guide*.

Configuring the service groups using the wizard

You can use a configuration wizard to configure the VCS service groups for Storage Foundation for Oracle RAC environment. The wizard enables you to modify the CVM service group to include the CRS resources. Note that the wizard for Oracle 10g does *not* create the service group for the Oracle database. To monitor the Oracle database using the Oracle agent provided by VCS, you must edit the `main.cf` manually after you finish running the wizard. See [“Configuring CVM service group for Oracle 10g manually”](#) on page 164 for details.

Before starting the wizard

Before starting the Wizard, you can verify that your Oracle installation can be configured. Review the requirements listed below. Also, you need to provide the wizard information as it proceeds. Make sure you have that information at hand.

Prerequisites

- Oracle RAC instances and listeners must be running on all cluster nodes.
- The database files of all instances must be on a cluster file system.

Note: The Wizard does not support using the same file system for the Oracle binary and Oracle datafiles.

- The OCR file and VOTE file location must be on a raw volume or a cluster file system.
- Each Oracle instance must be associated with a listener.

Note: The RAC configuration wizard requires that for the default listener, the listener parameter file, listener.ora, must reside in \$ORACLE_HOME/network/admin. No such restriction applies for non-default listeners.

- The IP addresses and host names specified in the files listener.ora and tnsnames.ora must be the same.
- Virtual IPs required for CRS must be up.

Information Required From the User

- RAC database instances to be configured
- NICs for Private NIC resource
- Registry and vote disk location for CRS

Establishing graphical access for the wizard

The configuration wizard requires graphical access to the VCS systems where you want to configure service groups. If your VCS systems do not have monitors, or if you want to run the wizards from a remote HP system, do the following:

To establish graphical access from a remote system

- 1 From the remote system, (*jupiter*, for example), run `xhost +`
`# xhost +`
- 2 Complete one of the following operations (depending on your shell):
 - If you are running `ksh`, run this step on one of the systems where the wizard will run (for example, *jupiter*):
`# export DISPLAY=jupiter:0.0`
 - If you are running `csh`, run this step
`# setenv DISPLAY jupiter:0.0`
- 3 Verify the `DISPLAY` environment variable is updated:
`# echo $DISPLAY`
`jupiter:0.0`

- 4 Make sure to set the JRE_HOME variable to /opt/VRTSjre/jre1.4. If VRTSjre1.4 is not installed, the hawizard exits after displaying an error message.

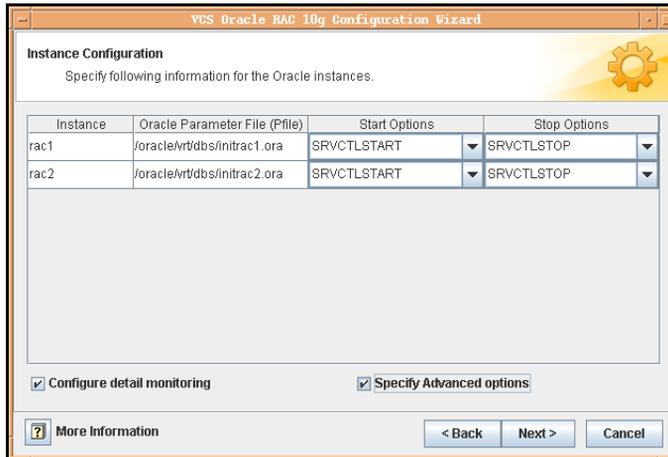
Creating service groups using the configuration wizard

Start the configuration wizard for Oracle RAC at the command-line.

To create service groups using the configuration wizard

- 1 Log on to one of your VCS systems as superuser.
- 2 Start the configuration wizard.
`# /opt/VRTSvcs/bin/hawizard rac`
- 3 Read the information on the Welcome screen.
 - If your configuration does not meet the requirements, click **Cancel** to stop the wizard. Start the wizard again after taking the necessary steps to meet the requirements.
 - If your configuration meets the requirements, click **Next**. The wizard begins discovering the current Oracle RAC information before proceeding.
 If the wizard does not find all databases and listeners running on all nodes in the cluster, it halts with an error, indicating the problem. Click **Cancel**, and start the wizard again after you correct the problem.
- 4 In the Wizard Options dialog box, select the **Create RAC Service Group** option.
- 5 Enter a name for the RAC service group in the **Service group name** box and click **Next**.
- 6 In the Database Selection dialog box, select a database and click **Next**.

- 7 In the Instance Configuration dialog box, specify information for all instances of the database you selected.



Specify the following information for each Oracle instance that is displayed and click **Next**:

Oracle Parameter File (Pfile) Verify the location of the Oracle Parameter File. Edit the information if necessary.

Start Options Choose the Start options, if desired. Default is STARTUP_FORCE.

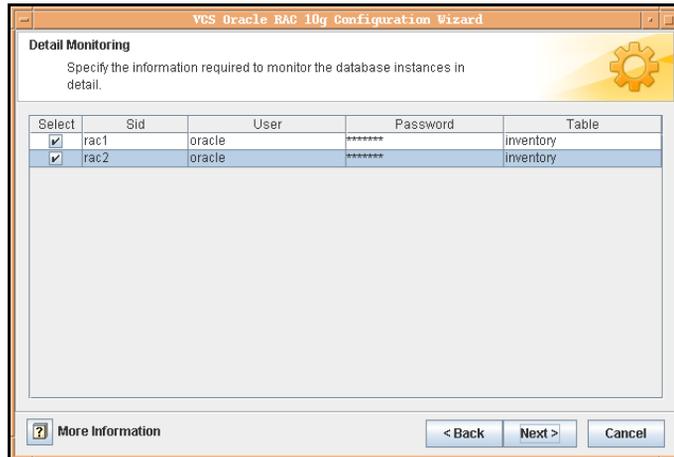
Stop Options Choose the Stop options, if desired. Default is IMMEDIATE.

Configure detail monitoring Select the check box if you want to monitor the database in detail.

If you want to enable Detail Monitoring, be sure you have previously set up the database table, user, and password for the agent to use during monitoring.

Specify Advanced Options Select the check box to enter advanced configuration information for the database instances.

- 8 If you chose to monitor the database in detail, the Detail Monitoring dialog box is displayed.



Specify the following information for the database instances that you want the agent to monitor in detail and click **Next**:

Select Select the check box corresponding to the database to be monitored in detail.

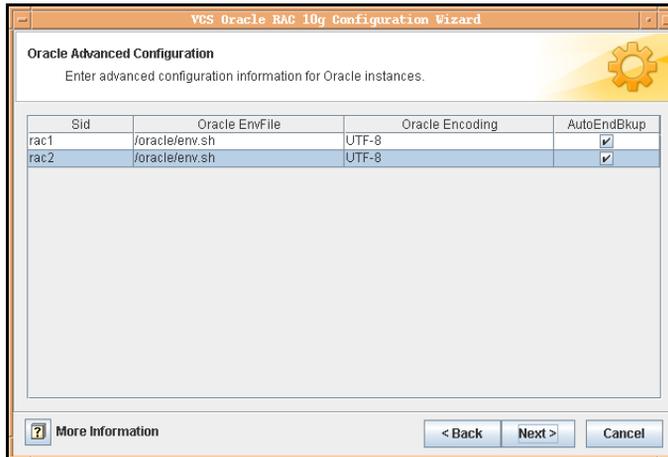
User Enter a valid user name for the database that the Oracle agent uses to log in to monitor the health of the database.

Password Enter a valid password for the database user.

Note: Do not encrypt passwords when entering them through the Agent Configuration Wizard; the wizard takes care of encrypting passwords.

Table Enter the name of a table that will be queried to validate the status of the database.

- 9 If you chose to specify advanced options, the Oracle Advanced Configuration dialog box is displayed.



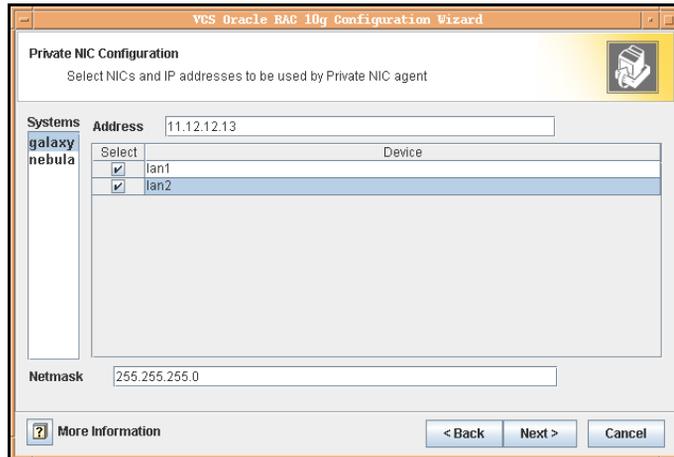
Specify the following information for the Oracle instances that you want to configure advanced attributes and click **Next**:

- Oracle EnvFile** Enter the location of the Oracle Envfile, the source file used by the agent entry point scripts.
- Oracle Encoding** Enter the operating system encoding that corresponds to Oracle encoding for the displayed Oracle output.
The encoding value must match the encoding value used by the Netlsnr configuration.
- AutoEndBkup** Select the check box, if desired.
Specifies that data files in the database are taken out of the backup mode when instance is brought online.

Refer to the *Veritas Cluster Server Agent for Oracle Installation and Configuration Guide* for a complete description of these attributes.

- 10 In the Monitor option Configuration dialog box, specify the monitor option for the Oracle instances, and click **Next**.
The default monitor option is **Process check**.

- 11 In the Private NIC Configuration dialog box, specify the NIC and IP address for Private NIC agent.



Specify the following information for each node in the cluster and click **Next**:

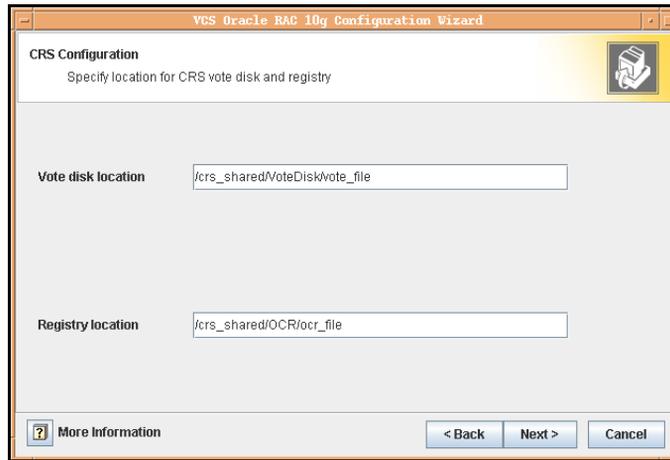
- Address** Enter the private IP address that is used by Oracle 10g CRS.
- Select** Select the checkbox against the network cards in the **Device** column. This NIC will be used by the PrivNIC agent.
- Netmask** Enter the netmask.

- 12 In the CRS Configuration dialog box, specify the location for CRS vote disk and OCR registry.

Enter the cluster file system or raw volume location for the CRS vote disk and registry. Example vote disk location:

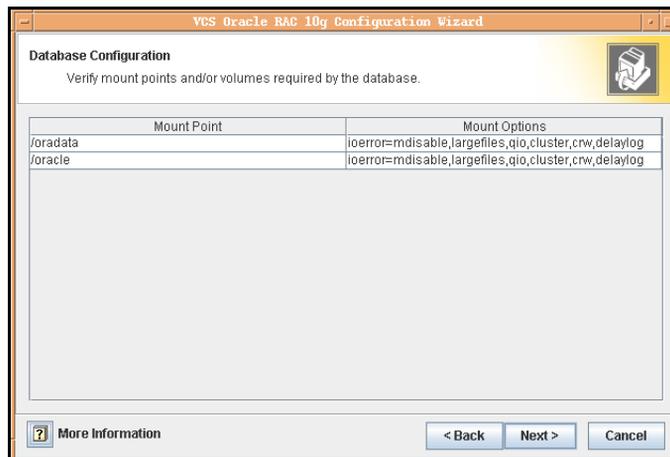
- /ora_crs/VOTE-disk/vote_file (if you are using a cluster file system)

- /dev/vx/rdisk/crs_oradg/crsvol (if you are using raw volumes)



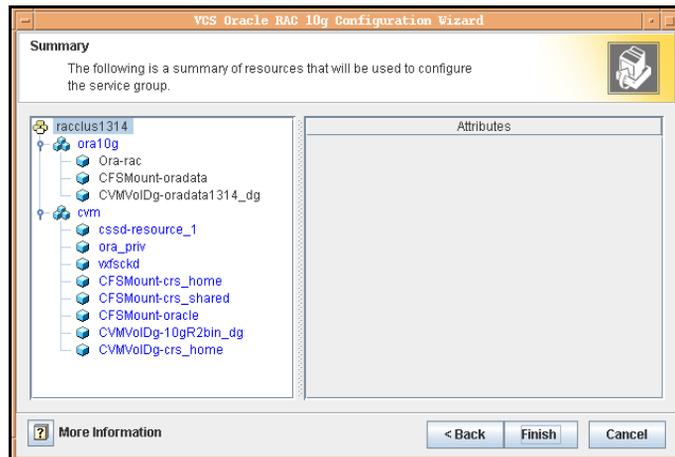
- 13 In the Database Configuration dialog box, verify the mount point of the database that the wizard displays. Confirm or modify the mount options displayed and click **Next**.

Note that the wizard discovers the mount point if the database is installed on a cluster file system. If the database exists on raw volumes, the wizard discovers the volumes.



- 14 In the Service Group Summary dialog, review your configuration.

Click on a resource to view its attributes and their configured values in the **Attributes** box.



- Click a resource within the service group to display its attributes and their values.
 For example, if you click on the name of the cssd application resource, cssd-resource, the wizard displays details of the cssd application resource.
 Attributes for the CFSMount resource show dependencies.
 The NetLsnr resource is configured as part of the CVM service group. The CVM service group also contains other resources, which may not be displayed by the wizard because the wizard does not control them.
 - Change names of resources, if desired; the wizard assigns unique names to resources based on their respective name rules.
 To edit a resource name, select the resource name and click on it, press Enter after editing each attribute.
- 15 Review your configuration and click **Finish**.
 The wizard starts running commands to create the Oracle RAC service group. Various messages indicate the status of these commands.
 - 16 In the Completing the Oracle Configuration wizard dialog box, select the **Bring the service group online** check box to bring the service group online on the local system.
 - 17 Click **Close**.
 The wizard creates the Oracle RAC service group in your cluster and adds the Netlsnr resource to the CVM configuration.

Location of VCS log files

On all cluster nodes, look at the following log files for any errors or status messages:

```
/var/VRTSvcs/log/engine_A.log
```

When large amounts of data are written, multiple log files may be required. For example, `engine_B.log`, `engine_C.log`, and so on, may be required. The `engine_A.log` contains the most recent data.

Adding and removing SF Oracle RAC nodes *for* Oracle 10g

This chapter contains the following topics:

- [Adding a node to an Oracle 10g cluster](#)
- [Removing a node from an Oracle 10g cluster](#)

Adding a node to an Oracle 10g cluster

An SF Oracle RAC cluster can have as many as eight nodes; SF Oracle RAC enables you to add and remove nodes. The example procedure shows how to add a node to a two-node cluster. The example procedure adds a node saturn to an existing cluster rac_cluster101 with two nodes galaxy and nebula.

Tasks involved in adding a node to your existing Storage Foundation for Oracle RAC cluster include:

- [“Checking system requirements for new node”](#) on page 176
- [“Physically adding a new system to the cluster”](#) on page 176
- [“Installing SF Oracle RAC on the new system”](#) on page 176
- [“Running vxinstall to start VxVM”](#) on page 178
- [“Configuring LLT, GAB, VCSMM, ODM, and VXFEN drivers”](#) on page 179
- [“Configuring the new node”](#) on page 180
- [“Using the Oracle add node procedure”](#) on page 182

Checking system requirements for new node

Verify the new nodes joining the cluster meet all of the requirements for installing and using SF Oracle RAC.

- The new node must have the identical operating system and patch level as the existing nodes.
- Use a text window of 80 columns (minimum) by 24 lines (minimum) Symantec recommends 80 columns by 24 lines for the optimum display of the `installsfrac` program.
- Verify the `/etc/fstab` file contains only valid entries, each of which specifies a file system to mount.

Physically adding a new system to the cluster

The new node must have the identical operating system and patch level as the existing nodes. When you physically add the new node to the cluster, make sure the node is connected to the same shared storage devices as the existing nodes and has private network connections to two independent switches for the cluster. Refer to the *Veritas Cluster Server Installation Guide*.

After installing SF Oracle RAC on the new node and starting VxVM, the new node can access the same shared storage devices (including coordinator disks). If the new node does not see the same disks as the existing nodes, the new node cannot join the cluster, as indicated by a CVM error on the console.

Installing SF Oracle RAC on the new system

Tasks involved in installing SF Oracle RAC on the new node include:

- [Mounting the software disc](#)
- [Checking the new node](#)
- [Installing SF Oracle RAC on new node](#)

Mounting the software disc

Make sure you have superuser (root) privileges to load the SF Oracle RAC software. You can use the `mount` command to mount the disc.

See “[Mounting the product disc](#)” on page 67.

To navigate to the folder containing the `installsfrac` program

- ◆ Navigate to the `/cdrom` folder containing the `installsfrac` script:

```
# cd /cdrom/storage_foundation_for_oracle_rac
```

Checking the new node

Before installing the SF Oracle RAC software on the new node, you must verify that the new node meets the installation requirements.

To check the new node for installation

- 1 Run the `installsfrac` program with the `precheck` option to verify the current operating system level, patch level, licenses, and disk space are adequate for a successful installation:

```
# ./installsfrac -precheck saturn
```

 The `precheck` function of the utility proceeds without user interaction.
- 2 Upon completion, review as the utility displays the results of the verification and saves the results in a log file.
- 3 If the `precheck` function indicates a requirement for licensing, add the license when running the installation utility.
 The `precheck` function may prompt you for other requirements.
- 4 If the verification is successful, proceed to run `installsfrac` with the `-installonly` option.

Installing SF Oracle RAC on new node

You must install SF Oracle RAC using the `-installonly` option and later manually configure the new node joining the cluster.

To install SF Oracle RAC on new node without configuration

- 1 On the new node, use the `-installonly` option to install SF Oracle RAC without performing configuration:

```
# ./installsfrac -installonly
```

 The new node will use the configuration from the existing cluster nodes.
- 2 Enter the name of the new node.
- 3 After the script performs initial checks, confirm to start the installation.
- 4 Review the output as the script checks system licensing and installs the licensing depot.
- 5 Enter the license key as the installer prompts.

```
Enter a SFRAC license key for saturn: [?] XXXX-XXXX-XXXX-XXXX-XX  

XXXX-XXXX-XXXX-XXXX-XX successfully registered on saturn  

SFRAC license registered on saturn
```
- 6 Enter keys for additional product features such as VVR, if you want to set up a Global Cluster environment.

```
Do you want to enter another license key for saturn? [y,n,q,?]  

(n)
```

- 7 Review the output as the script lists the depots and patches to install and checks whether any of them are present on the node.
- 8 After the script installs the packages and patches, note the location of the summary, log, and response files in the output.
- 9 Ignore the message advising you to run `installsfrac -configure`.
- 10 Restart the new node.
`# /usr/sbin/shutdown -r now`

Running vxinstall to start VxVM

To start VxVM using vxinstall

- 1 To start Veritas Volume Manager on the new node, use the vxinstall utility:
`# vxinstall`
- 2 VxVM uses license keys to control access. As you run the utility, answer “n” when prompted about licensing; you installed the appropriate license when you ran the installsfrac utility:

VxVM uses license keys to control access. If you have not yet installed a VxVM license key on your system, you will need to do so if you want to use the full functionality of the product.

Licensing information:

System host ID: 2653658338

Host type: 9000/800/rp3410

Are you prepared to enter a license key [y,n,q,?] (default: n)

Do you want to use enclosure based names for all disks ?

[y,n,q,?] (default: n)

Populating VxVM DMP device directories

V-5-1-0 vxvm:vxconfigd: NOTICE: Generating /etc/vx/array.info

-

The Volume Daemon has been enabled for transactions.

Starting the relocation daemon, vxrelocd.

Starting the cache daemon, vxcached.

Starting the diskgroup config backup daemon, vxconfigbackupd.

Starting the dg monitoring daemon, vxvrsecdgd.

Do you want to setup a system wide default disk group?

[y,n,q,?] (default: y) n

- 3 Decline to set up a disk group for the node:

Do you want to setup a system wide default disk group? [y,n,q,?]

(default: y) **n**

The installation is successfully completed.

- 4 Verify the daemons are up and running:

```
# vxdisk list
```

 Make sure the output displays the shared disks without errors.

Configuring LLT, GAB, VCSMM, ODM, and VXFEN drivers

To configure LLT, GAB, VCSMM, ODM, and VXFEN drivers

- 1 On the new node, set the shared memory parameter using SAM. The value of the shared memory parameter becomes effective when the system restarts. Refer to the *Oracle 10g Installation Guide* for details.

- 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 galaxy
1 nebula
2 saturn
```

- 3 Copy the `/etc/llthosts` file from one of the existing nodes to the new node. The `/etc/llthosts` file must be identical on all cluster nodes.

- 4 Create the `/etc/llttab` file on the new node. For example:

```
set-node saturn
set-cluster 101
link lan1 /dev/lan:1 - ether --
link lan2 /dev/lan:2 - 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 *must* be the same line on *all* nodes.

- 5 Use vi or another text editor to create the `/etc/gabtab` file on the new node. The file resembles:

```
/sbin/gabconfig -c -nN
```

Where *N* represents the number of nodes in the cluster. For a three-node cluster, the value of *N* is 3.

- 6 Edit the `/etc/gabtab` file on each of the existing nodes, changing the content to match the file on the new node.

- 7 If you are adding a node to a single-node SF Oracle RAC cluster, then do the following:

- Configure I/O fencing on both the original node and the new node. See [“Checking shared disks for I/O fencing”](#) on page 101.
- Set up the `/etc/vcsmmtab` files on the new node by copying them from one of the existing nodes:

```
# scp galaxy:/etc/vcsmmtab /etc
```

- Proceed to step [step 9](#).

- 8 Set up the `/etc/vcsmmtab` and `/etc/vxfendg` files on the new node by copying them from one of the existing nodes:

```
# scp galaxy:/etc/vcsmmtab /etc
# scp galaxy:/etc/vxfendg /etc
# scp galaxy:/etc/vxfenmode /etc
```

- 9 Start LLT and GAB on the new node:

```
# /sbin/init.d/llt start
# /sbin/init.d/gab start
```

- 10 On the new node, start the VXFEN, VCSMM, and LMX drivers. Use the following commands in this order:

```
# /sbin/init.d/vxfen start
# /sbin/init.d/vcsmm start
# /sbin/init.d/lmx start
```

- 11 On the new node, start the ODM driver. Use the following commands in this order:

```
# kcmodule vxgms=loaded
# kcmodule odm=loaded
# /sbin/init.d/odm stop
# /sbin/init.d/odm start
```

- 12 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
```

Configuring the new node

You must manually configure the new node after installing the SF Oracle RAC depots.

To configure the new node

- 1 Create a local group and local user for Oracle in the new node. The user and group ids for oracle must be the same as those on the other cluster nodes.

```
# groupadd -g 1000 oinstall
# groupadd -g 1001 dba
# groupadd -g 1002 oper

# useradd -u 999 -g oinstall -G dba,oper -d /oracle oracle
```

- 2 Create a password for the Oracle user on each node:

```
# passwd oracle
```

- 3 Make sure that the user Oracle can `remsh` between the nodes without having to use or confirm a password. If not, in the Oracle user account directory (for example, `/oracle`), create a `.rhosts` file that includes plus characters (“+ +”) in the first line.
- 4 Create mount point directories for the Oracle base and CRS file system having the OCR and Vote Disk files.
 - Create the Oracle base mount point.


```
# mkdir /oracle
```
 - Create the mount point for the CRS file system having the OCR and Vote Disk files.


```
# mkdir /ora_crs
```
 - Create the mount point for the Oracle Database.


```
# mkdir /rac_db
```
- 5 Create the local CRS volume and the file system.
 See “[Preparing \\$CRS_HOME on each node](#)” on page 139.
- 6 Plumb the private NIC and the virtual public IP addresses.
 See “[Configuring private IP addresses for CRS](#)” on page 141.
 See “[Creating public virtual IP addresses for Oracle](#)” on page 145.
- 7 Modify the VCS configuration as illustrated in the following example. Execute the following commands from one of the existing nodes. For example, `galaxy` or `nebula`.
 - Enable write access to the VCS configuration:


```
# haconf -makerw
```
 - Add the new node to the cluster:


```
# hasys -add saturn
```
 - If the ClusterService service group is configured, add the new node to its system list and specify a failover priority.
 For example, add a node `saturn` with failover priority “2”:


```
# hagrps -modify ClusterService SystemList -add saturn 2
```
 - If the ClusterService service group is configured, add the new node to its AutoStartList:


```
# hagrps -modify ClusterService AutoStartList \  
galaxy nebula saturn
```
 - Configure the service groups on the new node (`saturn`). For example, to add the CVM service group:
 - Add the node to the SystemList of the CVM service group.


```
# hagrps -modify cvm SystemList -add saturn 2
```
 - Add the node to the AutoStartList of the CVM service group by specifying the remaining nodes in the desired order:


```
# hagrps -modify cvm AutoStartList galaxy nebula saturn
```

- Add the node to the CVMCluster resource by adding it to the CVMNodeId attribute key list:
hares -modify cvm_clus CVMNodeId -add saturn 2
 - Add the resource for the PrivNIC network interfaces.
hares -modify ora_priv Device lan1 0 lan2 1 -sys saturn
 - Add the resource for the PrivNIC IP address.
hares -modify ora_priv Address "10.10.19.23" -sys saturn
 - Add the new node information for other VCS resources that exist.
 - Save the new configuration to disk:
haconf -dump -makero
- 8 On each of the existing nodes, enable the nodes to recognize the new node.
/opt/VRTS/bin/vxclustadm -m vcs -t gab reinit
- 9 Restart the new node.
/usr/sbin/shutdown -r now
As the new node boots, VCS propagates the configuration from the existing cluster nodes to the new node. All the configuration files located in the /etc/VRTSvcs/conf/config directory, including main.cf, CVMTypes.cf, CFSTypes.cf, and OracleTypes.cf are identical on each node.
- 10 At this point, GAB membership shows membership for all the nodes. Following ports must be up on all the nodes:
gabconfig -a
GAB Port Memberships
=====
- ```
Port a gen df205 membership 012
Port b gen df20e membership 012
Port d gen df20f membership 012
Port f gen df219 membership 012
Port h gen df211 membership 012
Port o gen df208 membership 012
Port v gen df215 membership 012
Port w gen df217 membership 012
```

## Using the Oracle add node procedure

For the Oracle procedure for adding a node, refer to:

Metalink Article 270512.1, Adding a Node to a 10g RAC Cluster

In this procedure, Oracle copies the CRS\_HOME and ORACLE\_HOME from an existing node in the cluster.

After performing the Oracle add node procedure, you must relink Oracle binaries to the appropriate Veritas libraries for Oracle 10g R2.

See [“Relinking the SF Oracle RAC libraries to Oracle 10g”](#) on page 154.

## Removing a node from an Oracle 10g cluster

You can remove one or more nodes from a Veritas Storage Foundation for Oracle RAC cluster by using the following procedures:

- [“Using the Oracle remove node procedure”](#) on page 183
- [“Removing SF 5.0 for Oracle RAC”](#) on page 183

### Using the Oracle remove node procedure

For the Oracle procedure for removing a node, refer to:

Metalink Article 269320.1, Removing a Node from a 10g RAC Cluster

### Removing SF 5.0 for Oracle RAC

- 1 Stop VCS on the node on which you want to remove SF Oracle RAC.  
`# /opt/VRTSvcs/bin/hastop - local`
- 2 Remove SF 5.0 for Oracle RAC.  
 See [“Removing SF Oracle RAC using the uninstallsfrac utility”](#) on page 183.
- 3 Edit VCS configuration files on existing nodes.  
 See [“Editing VCS configuration files on existing nodes”](#) on page 184.

### Removing SF Oracle RAC using the uninstallsfrac utility

Run the uninstallsfrac script from any node in the cluster. Prior to invoking the uninstallsfrac script, you must take all service groups offline and shut down VCS.

The example procedure involves removing SF Oracle RAC from the node saturn.

#### To remove SF Oracle RAC using the uninstallsfrac utility

- 1 Take all the Oracle parallel groups offline.  
`# hagrps -offline oracle_group -sys nebula`
- 2 Take the CRS resources offline:  
`# hares -offline cssd_resource -sys nebula`  
 If CRS is not under VCS control, enter the following command:  
`# /sbin/init.d/init.crs stop`
- 3 As superuser, start the uninstallation.  
`# cd /opt/VRTS/install`  
`# ./uninstallsfrac`
- 4 Decline to uninstall SF Oracle RAC from *all* systems.  
 VCS configuration files exist on this system with the following information:

```
Cluster Name: rac_cluster101
Cluster ID Number: 101
Systems: galaxy nebula saturn
Service Groups: cvm oradb1_grp
```

```
Do you want to uninstall SFRAC from these systems? [y,n,q] (y) n
```

---

**Warning:** Be sure to answer "n". Otherwise, the utility begins the process of uninstalling SF Oracle RAC from all systems.

---

- 5 Specify the name of the system on which you are uninstalling SF Oracle RAC.

```
Enter the system names separated by spaces on which to uninstall
SFRAC: saturn
```

- 6 Review the output as the script checks for packages currently installed on your system. The utility also checks for dependencies between packages to determine the packages it can safely uninstall and in which order.

- 7 Confirm to uninstall SF Oracle RAC.

```
All SFRAC processes that are currently running will be stopped.
```

```
Are you sure you want to uninstall SFRAC? [y,n,q] (y)
```

- 8 Review the output as the script stops processes and drivers running on each node.

- 9 After the script uninstalls the packages, make a note of the location of the summary and log files that the uninstaller creates.

- 10 Reboot the node:

```
/usr/sbin/shutdown -r now
```

## Editing VCS configuration files on existing nodes

After running `uninstallsfrac` program, modify the configuration files on the existing nodes to remove references to the removed nodes.

### Editing `/etc/llhosts`

On the each of the existing nodes, use `vi` or another editor to edit `/etc/llhosts`. Remove lines corresponding to the removed nodes; for example, if `saturn` is the node removed from the cluster, remove the line "2 saturn" from the file:

```
0 galaxy
1 nebula
2 saturn
```

Change to:

```
0 galaxy
1 nebula
```

### Editing /etc/gabtab

In /etc/gabtab, change the command contained in the file to reflect the number of cluster nodes after the removal of a node:

```
/sbin/gabconfig -c -nN
```

N is the number of remaining nodes. For example, a file with two remaining nodes appears as:

```
/sbin/gabconfig -c -n2
```

### Modifying the VCS configuration to remove the node

Modify the VCS configuration by editing /etc/VRTSvcs/conf/config/main.cf directly, using the VCS Cluster Manager, or using the command line, as illustrated in the following example. Refer to the *Veritas Cluster Server User's Guide* for details on configuring VCS.

Complete the procedure on one of the existing nodes.

#### To modify the VCS configuration

- 1 As superuser, enable write access to the configuration:  

```
haconf -makerw
```
- 2 Remove the node from the AutoStartList of the cvm service group by specifying the remaining nodes in the desired order:  

```
hagrps -modify cvm AutoStartList galaxy nebula
```
- 3 Remove the node from the SystemList of the cvm service group:  

```
hagrps -modify cvm SystemList -delete saturn
```
- 4 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 2](#) and [step 3](#) for each of them.
- 5 Delete the node from the CVMCluster resource by removing it from the CVMNodeId attribute key list:  

```
hares -modify cvm_clus CVMNodeId -delete saturn
```
- 6 After deleting the removed node from all service groups in the configuration, delete the node from the cluster system list:  

```
hasys -delete saturn
```
- 7 Save the new configuration to disk:  

```
haconf -dump -makero
```



# Uninstalling SF Oracle RAC with Oracle 10g

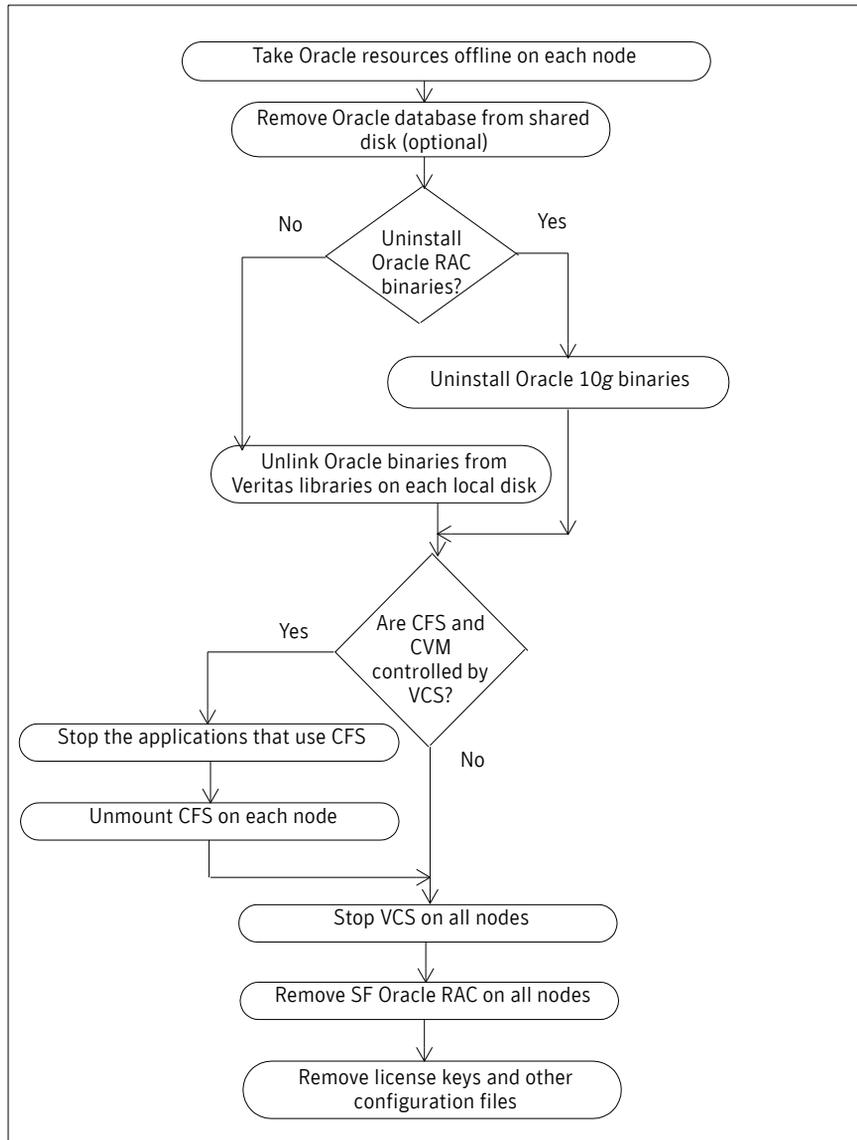
This chapter contains the following topics:

- [About uninstalling SF 5.0 for Oracle RAC from Oracle 10g cluster](#)
- [Preparing to uninstall SF 5.0 for Oracle RAC from Oracle 10g cluster](#)
- [Removing SF 5.0 for Oracle RAC from Oracle 10g cluster](#)

## About uninstalling SF 5.0 for Oracle RAC from Oracle 10g cluster

You can remove the SF 5.0 for Oracle RAC software from the nodes in a cluster. [Figure 11-12](#) describes a high-level flow of uninstalling SF Oracle RAC from a cluster having Oracle 10g instances. At the completion of the uninstallation activity, you can continue to run Oracle using the single-instance binary that is generated when you unlink the Veritas binaries from Oracle.

Figure 11-12 SF Oracle RAC with Oracle 10g uninstallation flowchart



# Preparing to uninstall SF 5.0 for Oracle RAC from Oracle 10g cluster

Perform the following steps before you uninstall SF Oracle RAC from a cluster having Oracle 10g instances:

- [Taking the Oracle resources offline](#)
- [Removing the Oracle database \(Optional\)](#)
- [Uninstalling Oracle 10g \(optional\)](#)
- [Unlinking the Oracle Binary From Veritas Libraries](#)
- [Stopping applications using CFS \(outside of VCS control\)](#)
- [Unmounting CFS file systems \(outside of VCS control\)](#)
- [Stopping VCS](#)

## Taking the Oracle resources offline

If *only* the Oracle 10g instances are under VCS control, you must take the Oracle resources offline.

### To take the Oracle resources offline

- 1 Log in as superuser.
- 2 On each node, take the Oracle resources in the VCS configuration file (main.cf) offline.
 

```
hagrps -offline oracle_group -sys node_name
```

 For example:
 

```
/opt/VRTSvcs/bin/hagrps -offline Oracle1 -sys galaxy
/opt/VRTSvcs/bin/hagrps -offline Oracle1 -sys nebula
```

 These commands stop the Oracle resources under VCS control running on the specified nodes.
- 3 On each node, take the CSSD resources in the VCS configuration file (main.cf) offline.
 

```
hares -offline cssd_resource -sys node_name
```

 For example:
 

```
hares -offline cssd-resource1 -sys galaxy
hares -offline cssd-resource1 -sys nebula
```

 These commands stop the CSSD resources under VCS control running on the specified nodes.
- 4 Verify the state of the Oracle and CVM service groups.
 

```
/opt/VRTSvcs/bin/hagrps -state
```

| #Group | Attribute | System | Value |
|--------|-----------|--------|-------|
|--------|-----------|--------|-------|

|         |       |        |         |
|---------|-------|--------|---------|
| Oracle1 | State | galaxy | OFFLINE |
| Oracle1 | State | nebula | OFFLINE |
| cvm     | State | galaxy | ONLINE  |
| cvm     | State | nebula | ONLINE  |

## Removing the Oracle database (Optional)

You can remove the Oracle database after safely relocating the data as necessary. Refer to the Oracle documentation for instructions on removing the database.

## Uninstalling Oracle 10g (optional)

You can uninstall Oracle 10g or choose to run Oracle after uninstalling SF Oracle RAC. If you do not want to uninstall Oracle, then you must unlink the Oracle binaries from the Veritas libraries.

See [“Unlinking the Oracle Binary From Veritas Libraries”](#) on page 190.

If you choose to uninstall Oracle 10g, use the Oracle runInstaller utility.

- If Oracle 10g is installed locally on each node, run the utility on each node.
- If Oracle 10g is installed on a cluster file system, run the utility once.

### To uninstall Oracle 10g

- 1 Log in as oracle user.
- 2 Set the DISPLAY variable. Depending on the shell you use, execute the following command:

```
Bourne Shell (sh or ksh) $ DISPLAY=host:0.0;export DISPLAY
```

```
C Shell (csh or tcsh) $ setenv DISPLAY host:0.0
```

- 3 Run the Oracle 10g runInstaller utility:  
\$ /cdrom/Disk1/runInstaller
- 4 Select the option to uninstall the Oracle 10g software.  
Refer to the Oracle 10g documentation for details on this utility.
- 5 If necessary, remove Oracle 10g from the other nodes.

## Unlinking the Oracle Binary From Veritas Libraries

If you uninstalled Oracle 10g, proceed to stop VCS. Otherwise, unlink the Veritas libraries. You must perform the procedure in each node of the cluster.

See [“Stopping VCS”](#) on page 192.

**To unlink the Veritas libraries**

- 1 Log in as oracle user.
- 2 Using vi or another text editor, open the init\$ORACLE\_SID.ora file.
- 3 Set the cluster\_database parameter to FALSE in the init\$ORACLE\_SID.ora file.

```
cluster_database=FALSE
```

- 4 Navigate to the following directory:

```
$ cd $ORACLE_HOME/lib
```

- 5 Link the original ODM library file to remove the Veritas ODM function.

```
PA systems $ ln -s $ORACLE_HOME/lib/libodmd10.sl \
 $ORACLE_HOME/lib/libodm10.sl
```

```
IA systems $ ln -s $ORACLE_HOME/lib/libodmd10.so \
 $ORACLE_HOME/lib/libodm10.so
```

- 6 To convert from Oracle RAC binary to Oracle single-instance binary, enter:

```
$ cd $ORACLE_HOME/rdbms/lib
$ make -f ins_rdbms.mk rac_off
$ make -f ins_rdbms.mk ioracle
```

- 7 Perform [step 1](#) through [step 6](#) on each of the other nodes.

- 8 To verify the state of the Oracle RAC configuration, type:

```
$ ldd $ORACLE_HOME/bin/oracle | egrep "11tdb|vcsmm"
```

- 9 Review the output of the command that you executed to verify the state:
  - If the command does not generate output, then Oracle single-instance binary is configured successfully.
  - If the command generates the following output, then Oracle RAC is still enabled.

```
PA systems /usr/lib/pa20_64/libvcsmm.1 => /usr/lib/
 pa20_64/libvcsmm.1
 /usr/lib/pa20_64/libl1tdb.1 => /usr/lib/
 pa20_64/libl1tdb.1
```

```
IA systems /usr/lib/hpux64/libvcsmm.1 => /usr/lib/hpux64/
 libvcsmm.1
 /usr/lib/hpux64/libl1tdb.so.1 => /usr/lib/
 hpux64/libl1tdb.so.1
```

You might have to redo the unlinking task to configure the Oracle single-instance binary.

## Stopping applications using CFS (outside of VCS control)

Stop all applications using the CFS mounts *not* under VCS control.

### To stop applications that use CFS

- 1 Ensure that no processes are using the CFS mount point.  
# `fuser -c mount_point`
- 2 Stop any processes using a CFS mount point.  
# `fuser -ck mount_point`

## Unmounting CFS file systems (outside of VCS control)

Unmount any CFS file systems that are *not* under VCS control on all nodes.

### To unmount CFS file systems

- 1 Determine the file systems to unmount by checking the `/etc/mnttab` file.  
For example:  
# `cat /etc/mnttab | grep vxfs | grep cluster`  
The output shows each line of the `/etc/mnttab` file that contains an entry for a VxFS file system mounted in the cluster mode.
- 2 By specifying the mount point for the file system, unmount each file system listed in the output:  
# `umount mount_point`

## Stopping VCS

Stop VCS to take the service groups on all nodes offline:

### To stop VCS

- 1 Log in as superuser on one cluster node.
- 2 Stop VCS on all nodes.  
# `/opt/VRTSvcs/bin/hastop -all`

## Removing SF 5.0 for Oracle RAC from Oracle 10g cluster

You can remove the SF 5.0 for Oracle RAC depots from all nodes in the SF Oracle RAC cluster using the `uninstallsfrac` program. Note that the uninstallation program removes all SF Oracle RAC depots regardless of the Oracle version used.

If you need to uninstall SF Oracle RAC after an incomplete installation, or if the `uninstallsfrac` program is not available in `/opt/VRTS/install`, you may need to use the `uninstallsfrac` program on the SF 5.0 for Oracle RAC disc.

Removing SF Oracle RAC depots involve the following tasks:

- [Removing the SF Oracle RAC depots](#)
- [Removing other configuration files \(optional\)](#)

## Removing the SF Oracle RAC depots

The installer provides you an option to remove Veritas Volume Manager and Veritas Volume Replicator depots. Note that uninstallation program can remove these depots only if both the boot disk is not under VxVM control and there are no open volumes.

### To remove the SF Oracle RAC depots

- 1 Do one of the following to begin uninstalling:
  - If you can execute commands as superuser on the remote nodes in the cluster using `ssh` or `remsh`, run `uninstallsfrac` program on one node to uninstall SF Oracle RAC on all nodes in the cluster.
  - If you cannot execute commands as superuser on remote nodes in the cluster using `ssh` or `remsh`, you must run `uninstallsfrac` program on each node in the cluster.
- 2 Navigate to the folder containing the `uninstallsfrac` program.  

```
cd /opt/VRTS/install
```
- 3 Start the `uninstallsfrac` program.  

```
./uninstallsfrac
```

The program specifies the directory where the logs are created and begins with a copyright message.
- 4 Indicate whether or not you want to remove VxVM and VVR depots from all nodes in the cluster; enter "y" only if the root disk is outside of VM control.
- 5 If you invoke the uninstallation program from a remote system in the same subnet, enter the system names where you want the uninstallation to take place.  

```
Enter the system names separated by spaces on which to uninstall
SFRAC: galaxy nebula
```
- 6 If you invoke the uninstallation program from a node in the SF Oracle RAC cluster, review the cluster information and confirm to uninstall SF Oracle RAC.  

```
VCS configuration files exist on this system with the following
information:
```

```
Cluster Name: rac_cluster101
Cluster ID Number: 101
Systems: galaxy nebula
Service Groups: cvm
```

```
Do you want to uninstall SFRAC from these systems? [y,n,q] (y)
```

- 7 Review the output as the uninstallation program checks the operating system on each system, verifies system-to-system communication, and verifies the system licensing.
- 8 Review the output as the uninstallation program checks for Storage Foundation for Oracle RAC depots currently installed on the nodes. This process involves identifying system uninstall requirements and dependencies between depots to determine the safety and order of uninstalling depots.
- 9 Confirm to uninstall SF Oracle RAC:  
All SFRAC processes that are currently running will be stopped.  
  
Are you sure you want to uninstall SFRAC? [y,n,q] (y)
- 10 Make sure that the user VxFS file systems that are currently mounted on each node are unmounted before uninstalling SF Oracle RAC.  
All user VxFS filesystems that are currently mounted on galaxy should be unmounted before uninstalling.  
See [“Unmounting CFS file systems \(outside of VCS control\)”](#) on page 192.
- 11 Make sure that each node in the cluster does not have any attached volumes.
- 12 Review the output as the uninstallation program stops agents and performs verifications on each node to proceed with the uninstallation.
- 13 Review the output as the uninstallation program stops the SF Oracle RAC processes.
- 14 Review the output as the uninstallation program uninstalls the SF Oracle RAC depots.
- 15 If necessary, review the summary and log files of uninstallation activities.
- 16 Reboot the nodes to completely uninstall SF Oracle RAC:  

```
./usr/sbin/shutdown -r now
```

## Removing other configuration files (optional)

You can remove other SF Oracle RAC configuration files.

### To remove the configuration files

- ◆ Execute the following commands to remove the configuration files.

```
rm /etc/vcsmmtab
rm /etc/vxfentab
rm /etc/vxfendg
rm /etc/llttab
rm /etc/gabtab
rm /etc/llthosts
```



# Disaster Recovery with SF Oracle RAC

- [Chapter 12, “Preparing for global clustering” on page 199](#)
- [Chapter 13, “Configuring global clustering” on page 207](#)



# Preparing for global clustering

This chapter contains the following topics:

- [About preparing for global clustering](#)
- [About global clustering](#)
- [Setting up a secondary site](#)

## About preparing for global clustering

You can set up SF Oracle RAC for ultimate data protection with global clustering and replication features. VCS provides an option to configure global clusters for wide-area failover and disaster recovery, and Veritas Volume Replicator provides the volume replication capability.

If you are currently administering an Oracle database in a cluster running SF Oracle RAC, you can configure a global clustering environment with a new cluster on a secondary site. SF 5.0 for Oracle RAC supports up to four nodes each on two clusters. You will need to set up a secondary site and modify the existing cluster configuration to support replication.

Preparing for global clustering requires:

- Creating a cluster on a secondary site with hardware set up
- Installing SF Oracle RAC
- Installing Oracle
- Configuring VCS service groups

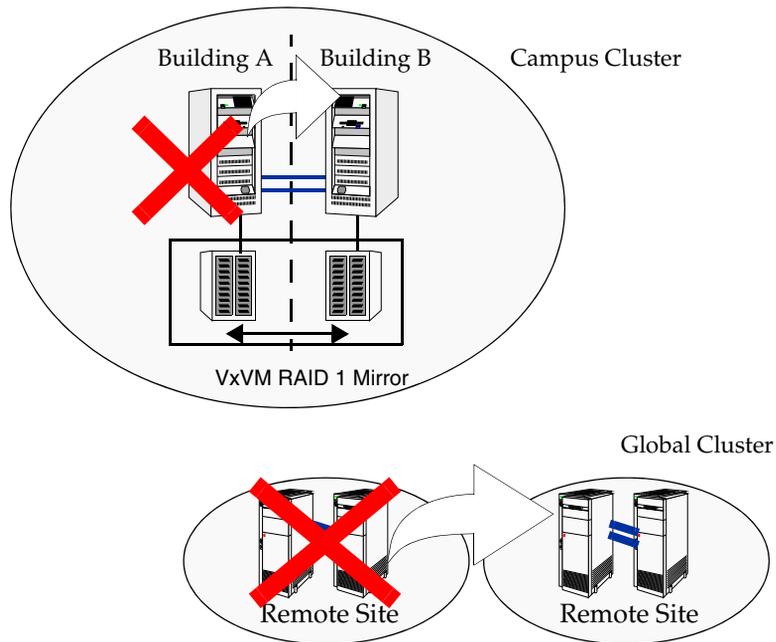
You do not need to create a database for the secondary site, since it will be replicated from the primary site.

## About global clustering

The global clustering feature of VCS enables wide-area failover and disaster recovery. Local clustering provides local failover for each site or building. [Figure 12-13](#) depicts campus clusters and global clusters. Campus and replicated cluster configurations offer some degree of protection against disasters within limited geographic regions. These types of configurations do not provide protection against outages caused by large-scale disasters such as major floods, hurricanes, and earthquakes that affect an entire city or region. An entire cluster could be affected by such outages.

Global clustering ensures data availability during large-scale disasters. This type of clustering involves migrating applications between clusters over a considerable distance.

**Figure 12-13** Campus clusters and global clusters

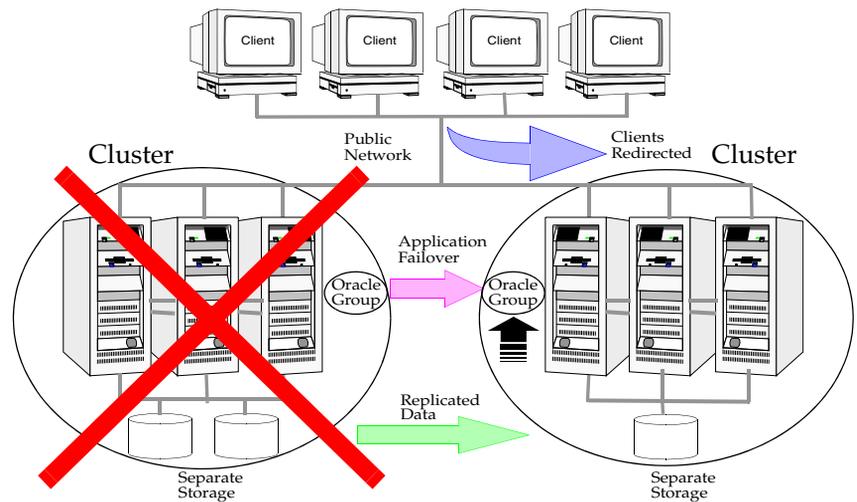


## Global clustering for Oracle RAC

VCS enables you to link clusters running at separate locations and switch service groups across clusters. Global clustering provides complete protection against failure of a cluster. [Figure 12-14](#) depicts a sample global cluster setup.

To understand how global clusters work, review the example of an Oracle RAC database configured using global clustering. Oracle RAC is installed and configured in cluster A and cluster B. Oracle data is located on shared disks within each cluster and is replicated across clusters to ensure data concurrency. The VCS service groups for Oracle are online on a node in cluster A and are configured to fail over on cluster A and cluster B.

**Figure 12-14** Sample global cluster setup



VCS continuously monitors and communicates events between clusters. Cluster-to-cluster communications ensures that the global cluster environment is aware of the state of global service group at all times.

In the event of a local node or application failure, the Oracle RAC service groups become available on other cluster nodes. If cluster A fails, the service groups can fail over to a remote cluster B. VCS also redirects clients when the application is online at the new location. Refer to the *Veritas Cluster Server User's Guide* for complete details on VCS global clusters.

## Replication in a shared disk environment

Veritas Volume Replicator (VVR) enables you to replicate data volumes on a shared disk group in SF Oracle RAC. In this environment, a two-node cluster on the primary site exists with a shared disk group. A two-node or single-node cluster exists on the secondary site; the disk group need not be a shared disk group.

The VVR feature is provided with SF Oracle RAC. Review the configuring instructions for VVR.

See “[Configuring global clustering](#)” on page 207.

Refer to the *Veritas Volume Replicator Administrator’s Guide* for complete details of VVR in a shared disk environment.

## Setting up a secondary site

Setting up SF Oracle RAC in a global cluster environment to prepare for replication requires preparing two sites:

- [Obtaining and installing license keys for VVR](#)
- [Installing SF Oracle RAC on the secondary site](#)
- [Installing Oracle on the secondary site](#)
- [Configuring VCS service groups for the secondary site](#)

## Obtaining and installing license keys for VVR

Make sure you have licenses for the following products:

- Veritas Storage Foundation for Oracle RAC
- Veritas Volume Replicator (VVR)

See “[Obtaining SF Oracle RAC license keys](#)” on page 66.

To install GCO/VCS and CVR license keys

- 1 Install the VVR license:  

```
vxlicinst -k XXXX-XXXX-XXXX-XXXX-XXXX-XXX
```
- 2 Make sure that VVR is enabled:  

```
vxlicrep | grep -e "CVM_FULL" -e "VVR"
```

## Installing SF Oracle RAC on the secondary site

Important requirements for global clustering:

- Cluster names on the primary and secondary sites must be unique.
- Node and resource names must be unique within a cluster but not across clusters.
- Make sure you use the same oracle group and user IDs for both primary and secondary clusters.

You can set up a multi-node or single-node cluster on the secondary site. The only difference between setting up a single node-cluster and a multi-node cluster is the fencing configuration.

See [“Installing SF Oracle RAC on a multi-node cluster”](#) on page 203.

See [“Installing SF Oracle RAC on a single-node cluster”](#) on page 203.

## Installing SF Oracle RAC on a multi-node cluster

If you install SF Oracle RAC on a multi-node cluster on the secondary site, you can enable I/O fencing feature of SF Oracle RAC.

### To install SF Oracle RAC on a multi-node cluster

- 1 Prepare to install SF Oracle RAC by planning your installation, reviewing the requirements, and performing pre-installation tasks.  
See [“Planning SF Oracle RAC installation and configuration”](#) on page 37.  
See [“Preparing to install and configure SF Oracle RAC”](#) on page 65.
- 2 Install SF Oracle RAC.
  - Global clustering requires different names for the clusters on the primary and secondary sites.
  - To install a multi-node cluster:  
See [“Installing and configuring SF Oracle RAC components”](#) on page 77.
- 3 Set up storage on the secondary site and configure I/O fencing.
  - Verify that the shared storage on the secondary site supports SCSI-3 reservations.
  - Set up the coordinator disks.
  - Configure I/O fencing for multi-node clusters.
  - Update the main.cf file on the secondary site.  
See [“Checking shared disks for I/O fencing”](#) on page 101.
- 4 Shut down and restart all nodes on the secondary site.
- 5 After starting the cluster nodes, check that the cvm group is online:  

```
hagr -state cvm
```

## Installing SF Oracle RAC on a single-node cluster

For a single-node cluster, after installing and configuring SF Oracle RAC, do not enable I/O fencing. You must not create a coordinator disk group and edit the configuration files to enable fencing. Fencing will run in disabled mode.

If you add more nodes later to the existing single-node cluster on the secondary site, then you must enable I/O fencing.

See [“Adding a node to an Oracle 10g cluster”](#) on page 175.

#### To install SF Oracle RAC on a one-node cluster

- 1 Prepare to install SF Oracle RAC by planning your installation, reviewing the requirements, and performing pre-installation tasks.  
See [“Planning SF Oracle RAC installation and configuration”](#) on page 37.  
See [“Preparing to install and configure SF Oracle RAC”](#) on page 65.
- 2 Install SF Oracle RAC.
  - Global clustering requires different names for the clusters on the primary and secondary sites.
  - To install a single-node cluster on the secondary site:  
See [“Installing and configuring SF Oracle RAC components”](#) on page 77.
- 3 Set up storage on the secondary site.
- 4 Shut down and restart the node on the secondary site.
- 5 After starting the cluster node, check the cvm group is online:  

```
hagrps -state cvm
```

## Installing Oracle on the secondary site

Make sure you have Oracle pre-configured and installed and configured on the secondary site.

#### To install Oracle on the secondary site

- 1 Pre-configure Oracle.  
See [“Preparing to install Oracle 10g RAC”](#) on page 129.
- 2 Install Oracle on the secondary site.
  - Use the same name for the database disk group and the database volume on the secondary site as the one on the primary site.
  - Set the same capacity for the database disk group and the database volume on the secondary site as the one on the primary site.
  - Do not create a database on the secondary site. The disk group will contain the database replicated from the primary site.  
See [“Installing Oracle 10g RAC”](#) on page 147.
- 3 On the secondary site, set up the disk group and volumes for the Oracle database, but do not create a database.  
The database will be replicated from the primary site.

- 4 Make sure that the database disk group and database volume on the secondary site have the same name and size as that on the primary site.

## Configuring VCS service groups for the secondary site

You must configure VCS service groups for Oracle on the secondary site.

Each cluster requires a virtual IP address associated with the cluster. The VCS installation and creation of the ClusterService group typically involves defining this IP address. If you did not configure the ClusterService group when you installed SF Oracle RAC on the secondary site, configure it when you configure global clustering.

See [“Configuring global clustering”](#) on page 287.

### To configure VCS service groups for the secondary site

- ◆ After installing Oracle on the secondary site, configure VCS service groups. See [“Configuring VCS service groups for Oracle 10g”](#) on page 161.



# Configuring global clustering

This chapter contains the following topics:

- [About configuring global clustering](#)
- [Preparing clusters for replication](#)
- [Setting up replication](#)
- [Configuring VCS to replicate the database volume](#)
- [Migration and takeover of primary replication role](#)

## About configuring global clustering

After setting up a secondary cluster running SF Oracle RAC, you can configure a global cluster environment. You must modify both cluster configurations to operate with replication in the global cluster environment.

Configuring SF Oracle RAC for global clusters requires:

- Setting up both clusters as part of a global cluster environment
- Setting up replication on both clusters
- Starting replication of the database
- Configuring VCS on the primary site for replication
- Configuring VCS on the secondary site for replication

## Preparing clusters for replication

Before configuring clusters for global clustering, make sure both clusters have:

- SF Oracle RAC and Oracle installed and configured
- License keys installed for CVR
- VVR and global clustering enabled

The example procedure assumes rac\_cluster101 as the local cluster with nodes galaxy and nebula, and rac\_cluster102 as the remote cluster on the secondary site with nodes mercury and jupiter. Preparing clusters for replication in both sites requires:

- [Adding the VVR resource types to the VCS configuration](#)
- [Configuring global clustering](#)
- [Defining the remote cluster and heartbeat Cluster Objects](#)

## Adding the VVR resource types to the VCS configuration

After SF Oracle RAC is installed on two clusters and fencing is configured, add the VVR types to the VCS configuration.

### To add VVR types to the VCS configuration on each cluster

- 1 On the first cluster, make sure that CVM is up.  
For example:

```
LOCAL_CLUSTER_NAME: rac_cluster101 (galaxy,nebula)
REMOTE_CLUSTER_NAME: rac_cluster102 (mercury,jupiter)
```
- 2 Make sure you have installed CVR license keys with "VVR" enabled.  
See ["Obtaining and installing license keys for VVR"](#) on page 202.
- 3 On each node, run the script to add definitions for VVR types:

```
cd /etc/VRTSvcs/conf/sample_vvr
./addVVRTypes.sh
```

## Configuring global clustering

You can configure global clustering in one of the following ways:

- Running the global clustering configuration wizard  
See ["Modifying the global clustering configuration using the wizard"](#) on page 209.
- Editing the VCS configuration file  
See ["Modifying the global clustering configuration using the main.cf"](#) on page 210.

Before you configure global clustering, review the following requirements:

- Cluster names on the primary and secondary sites must be unique.

- Node and resource names must be unique within a cluster but not across clusters.
- Each cluster requires a virtual IP address associated with the cluster. The VCS installation and creation of the ClusterService group typically involves defining this IP address. If you did not configure the ClusterService group when you installed SF Oracle RAC, configure it when you configure global clustering.
- One WAN (Wide Area Network) heartbeat must travel between clusters, assuming each cluster has the means to monitor the health of the remote cluster. Configure the heartbeat resource manually.
- All oracle user and group IDs must be the same on all nodes.
- The Oracle database, which VVR replicates from the storage on the primary site to the secondary site, must be defined in a global group having the same name on each cluster. Each resource in the group may differ from cluster to cluster, but clients redirected to a remote cluster after a wide-area failover must see the same application as the one in the primary cluster.

See the *Veritas Cluster Server User's Guide* for complete details on global clustering.

## Modifying the global clustering configuration using the wizard

The global clustering wizard completes the following tasks:

- Validates the ability of the current configuration to support a global cluster environment.
- Creates the components that enable the separate clusters, each of which contains a different set of GAB memberships, to connect and operate as a single unit.
- Creates the ClusterService group, or updates an existing ClusterService group.

Run the global clustering configuration wizard on each of the clusters; you must have the global clustering license in place on each node in the cluster.

### To use the global clustering wizard

- 1 On a node in the primary site, start the global clustering configuration wizard:  

```
/opt/VRTSvcs/bin/gcoconfig
```
- 2 After discovering the NIC devices on the local node, specify or confirm the device for the cluster joining the global cluster environment.

- 3 Indicate whether the NIC you entered is for all cluster nodes. If you enter **n**, enter the names of NICs on each node.
- 4 Enter or confirm the virtual IP address for the local cluster.
- 5 When the wizard discovers the net mask associated with the virtual IP address, accept the discovered value or enter another value.  
With NIC and IP address values configured, the wizard creates a ClusterService group or updates an existing one. After modifying the VCS configuration file, the wizard brings the group online.
- 6 Perform through [step 1](#) through [step 5](#) on the secondary cluster.

## Modifying the global clustering configuration using the main.cf

Edit the main.cf file to specify the virtual IP address for the local cluster and define the ClusterService group for the local cluster.

The example global clustering configuration shows the rac\_cluster101 cluster on the primary site. The additions to the configuration appear in bold text.

```
include "types.cf"
include "CFSTypes.cf"
include "CVMTTypes.cf"
include "OracleTypes.cf"
include "VVRTypes.cf"

cluster rac_cluster101 (
 UserNames = { admin = "cDRpdxPmHpzS." }
 ClusterAddress = "10.10.10.101"
 Administrators = { admin }
 CounterInterval = 5
 UseFence = SCSI3
)

system galaxy (
)

system nebula (
)

group ClusterService (
 SystemList = { galaxy = 0, nebula = 0 }
 AutoStartList = { galaxy, nebula }
 OnlineRetryLimit = 3
 OnlineRetryInterval = 120
)

Application wac (
```

```

 StartProgram = "/opt/VRTSvcs/bin/wacstart"
 StopProgram = "/opt/VRTSvcs/bin/wacstop"
 MonitorProcesses = "/opt/VRTSvcs/bin/wac" }
 RestartLimit = 3
)

IP gcoip (
 Device =lan0
 Address = "10.10.10.101"
 NetMask = "255.255.240.0"
)

NIC csgnic (
 Device =lan0
 NetworkHosts = { "10.10.12.2", "10.10.12.3" }
)

gcoip requires csgnic
wac requires gcoip
.
.
.

group cvm (
.
.
.

```

## Defining the remote cluster and heartbeat Cluster Objects

After configuring global clustering, add the `remotecluster` cluster object to define the IP address of the cluster on the secondary site, and the `heartbeat` object to define the cluster-to-cluster heartbeat.

Heartbeats monitor the health of remote clusters. VCS can communicate with the remote cluster only after you set up the heartbeat resource on both clusters.

### To define the remote cluster and heartbeat

- 1 On the primary site, enable write access to the configuration:  
# **haconf -makerw**
- 2 Define the `remotecluster` and its virtual IP address.  
In this example, the remote cluster is `rac_cluster102` and its IP address is 10.11.10.102:  
# **haclus -add rac\_cluster102 10.11.10.102**
- 3 Complete [step 1](#) and [step 2](#) on the secondary site using the name and IP address of the primary cluster.  
In this example, the primary cluster is `rac_cluster101` and its IP address is 10.10.10.101:

- 4 On the primary site, add the heartbeat object for the cluster.  
 In this example, the heartbeat method is ICMP ping.

```
hahb -add Icmp
```

- 5 Define the following attributes for the heartbeat resource:
  - ClusterList attribute lists the remote cluster.
  - Arguments attribute enables you to define the virtual IP address for the remote cluster.

For example:

```
hahb -modify Icmp ClusterList rac_cluster102
hahb -modify Icmp Arguments 10.11.10.102 -clus rac_cluster102
```

- 6 Save the configuration and change the access to read-only on the local cluster:

```
haconf -dump -makero
```

- 7 Complete [step 4](#) through [step 6](#) on the secondary site using appropriate values to define the cluster on the primary site and its IP as the remote cluster for the secondary cluster.

- 8 Verify Clusters status with `hastatus -sum` command on two systems. And `haclus -list` should show the two clusters.

```
hastatus -sum
haclus -list
rac_cluster101
rac_cluster102
```

After completing the steps, review the example additions to the `main.cf` file on the primary site:

```
.
.
remotecluster rac_cluster102 (
Cluster Address = "10.11.10.102"
)
heartbeat Icmp (
ClusterList = { rac_cluster102 }
Arguments @rac_cluster102 = { "10.11.10.102" }
)

system galaxy (
)
.
.
```

Review the example additions to the `main.cf` file on the secondary site:

```
.
.
remotecluster rac_cluster101 (
Cluster Address = "10.11.10.101"
)
.
```

```
heartbeat Icmp (
 ClusterList = { rac_cluster101 }
 Arguments @rac_cluster101 = { "10.11.10.101" }
)

system mercury (
)
.
.
```

See the *Veritas Cluster Server User's Guide* for details for configuring the required and optional attributes of the heartbeat object.

## Setting up replication

Setting up replication in a global cluster environment involves the following tasks:

- Creating the Storage Replicator Log in the disk group for the database  
See [“Creating the SRL volume on the primary site”](#) on page 213.
- Creating the RVG on the primary site.  
See [“Setting up replication objects on the primary site”](#) on page 214.
- Setting up replication objects on the secondary site.  
See [“Configuring replication for the secondary site”](#) on page 215.

## Creating the SRL volume on the primary site

Create the Storage Replicator Log (SRL), a volume in the Replicated Volume Group (RVG). The RVG also holds the data volumes for replication.

- The data volume on the secondary site has the same name and the same size as the data volume on the primary site.  
See [“Setting up a secondary site”](#) on page 202.
- The SRL on the secondary site has the same name and same size as the SRL on the primary site.
- The data volume and SRL volume should exist on the same disk group. If possible, create SRLs on disks without other volumes.
- Mirror SRLs and data volumes in the absence of hardware-based mirroring.

After determining the size of the SRL volume, create the volume in the shared disk group for the Oracle database. If hardware-based mirroring does not exist in your setup, use the `nmirror` option to mirror the volume. In this example, the Oracle database is in the `oradatadg` shared disk group on the primary site.

**To create the SRL volume on the primary site**

- 1 On the primary site, determine the size of the SRL volume based on the configuration and use.  
See the Veritas Volume Replicator documentation for details.
- 2 Determine whether a node is the master or the slave:  

```
vxctl -c mode
```
- 3 From the master node, issue the following command:  

```
vxassist -g oradatadg make rac1_srl 1500M nmirror=2 c4t1d1 c4t1d2
```

  
Make sure that the data disk has a minimum of 500M of free space after creating the SRL volume.
- 4 Start the SRL volume by starting all volumes in the disk group:  

```
vxvol -g oradatadg startall
```

## Setting up replication objects on the primary site

Before creating the Replicated Volume Group (RVG) on the primary site, make sure the replication objects are active and online.

**To review the status of replication objects on the primary site**

- 1 Verify the volumes you intend to include in the group are active.
- 2 Review the output of the `hagrp -state cvm` command.
- 3 Check that the cvm group is online.

**To create the RVG**

- ◆ Create the primary RVG using the `vradmin` command:  

```
vradmin -g disk_group createpri rvg_name data_volume srl_volume
```

  
where:
  - `disk_group` is the name of the disk group containing the database
  - `rvg_name` is the name for the RVG
  - `data_volume` is the volume that VVR replicates
  - `srl_volume` is the volume for the SRL
 For example, to create the `rac1_rvg` RVG, enter:  

```
vradmin -g oradatadg createpri rac1_rvg rac1_vol rac1_srl
```

  
The command creates the RVG on the primary site and adds a Data Change Map (DCM) for each data volume. In this case, a DCM exists for `rac1_vol`.

## Configuring replication for the secondary site

To create objects for replication on the secondary site, use the `vradmin` command with the `addsec` option. Setting up replication on the secondary site involves:

- Creating a disk group on the storage with the same name as the equivalent disk group on the primary site if you have not already done so. See [“Installing Oracle on the secondary site”](#) on page 204.
- Creating volumes for the database and SRL on the secondary site. See [“Creating the data and SRL volumes on the secondary site”](#) on page 215.
- Editing the `/etc/vx/vras/.rdg` file on the secondary site. See [“Editing the /etc/vx/vras/.rdg files”](#) on page 216.
- Resolvable virtual IP addresses that set network RLINK connections as host names of the primary and secondary sites. See [“Setting up IP addresses for RLINKs on each cluster”](#) on page 216.
- Creating the replication objects on the secondary site. See [“Setting up disk group on secondary site for replication”](#) on page 217.

### Creating the data and SRL volumes on the secondary site

Note the following when creating volumes for the data and SRL:

- The sizes and names of the volumes must reflect the sizes and names of the corresponding volumes in the primary site.
- Create the data and SRL volumes on different disks in the disk group. Use the `vxdisk -g diskgroup list` command to list the disks in the disk group.
- Mirror the volumes.

#### To create the data and SRL volumes on the secondary site

- 1 In the disk group created for the Oracle database, create a volume for data; in this case, the `rac_vol1` volume on the primary site is 6.6 GB:  

```
vxassist -g oradatadg make rac_vol1 6600M nmirror=2 c4t2d1
c4t2d2
```
- 2 Create the volume for the SRL, using the same name and size of the equivalent volume on the primary site. Create the volume on different disks from the disks for the database volume, but on the same disk group that has the data volume.  

```
vxassist -g oradatadg make rac1_srl 1500M nmirror=2 c4t2d3
c4t2d4
```

## Editing the `/etc/vx/vras/.rdg` files

Editing the `/etc/vx/vras/.rdg` file on the secondary site enables VVR to replicate the disk group from the primary site to the secondary site. On each node, VVR uses the `/etc/vx/vras/.rdg` file to check the authorization to replicate the RVG on the primary site to the secondary site. The file on each node in the secondary site must contain the primary disk group ID, and likewise, the file on each primary system must contain the secondary disk group ID.

### To edit the `/etc/vx/vras/.rdg` files

- 1 On a node in the primary site, display the primary disk group ID:  

```
vxprint -l diskgroup
.....
```
- 2 On each node in the secondary site, edit the `/etc/vx/vras/.rdg` file and enter the primary disk group ID on a single line.
- 3 On each cluster node of the primary cluster, edit the file and enter the primary disk group ID on a single line.

## Setting up IP addresses for RLINKs on each cluster

Creating objects with the `vradm` command requires resolvable virtual IP addresses that set network RLINK connections as host names of the primary and secondary sites.

### To set up IP addresses for RLINKs on each cluster

- 1 For each RVG running on each cluster, set up a virtual IP address. These IP addresses are part of the RLINK.
  - Set up the virtual IP address on one node in the primary cluster:  
The example assumes that the public network interface is `lan0:1`, the virtual IP address is `10.10.9.101`, and the net mask is `255.255.240.0` for the cluster on the primary site.  

```
ifconfig lan0:1 plumb
ifconfig lan0:1 inet 10.10.9.101 netmask 255.255.240.0
ifconfig lan0:1 up
```
  - Set up the virtual IP address on one node in the secondary cluster:  
Use the same commands with appropriate values for the interface, IP address, and net mask on the secondary site.  
The example assumes the interface is `lan0:1`, virtual IP address is `10.11.9.102`, and the net mask is `255.255.240.0` on the secondary site.
- 2 Define the virtual IP addresses to correspond to a virtual cluster host name on the primary site and a virtual cluster host name on the secondary site.

Update the `/etc/hosts` file on all the nodes on both the primary and secondary sites.

The examples assume `rac_clus101_priv` has IP address 10.10.9.101 and `rac_clus102_priv` has IP address 10.11.9.102.

- 3 Use the `ping` command to verify the links are functional.

## Setting up disk group on secondary site for replication

Create the replication objects on the secondary site from the master node on the primary site, using the `vradm` command.

### To set up the disk group on the secondary site for replication

- 1 Issue the command in the following format from the cluster on the primary site:

```
vradm -g dg_pri addsec rvg_pri pri_host sec_host
```

where:

- *dg\_pri* is the disk group on the primary site that VVR will replicate. For example: `oradatadg`
- *rvg\_pri* is the RVG on the primary site. For example: `rac1_rvg`
- *pri\_host* is the virtual IP address or resolvable virtual host name of the cluster on the primary site. For example: `10.10.9.101` or `rac_clus101_priv`
- *sec\_host* is the virtual IP address or resolvable virtual host name of the cluster on the secondary site. For example: `10.11.9.102` or `rac_clus102_priv`

For example, the command to add the cluster on the primary site to the RDS is:

```
vradm -g oradatadg addsec rac1_rvg rac_clus101_priv
rac_clus102_priv
```

- 2 On the secondary site, the command does the following:
  - Creates an RVG within the specified disk group using the same name as the one for the primary site
  - Associates the data and SRL volumes that have the same names as the ones on the primary site with the specified RVG
  - Adds a data change map (DCM) for the data volume
  - Creates cluster RLINKS for the primary and secondary sites with the default names; for example, the “primary” RLINK created for this example is `rlk_rac_clus102_priv_rac1_rvg` and the “secondary” RLINK created is `rlk_rac_clus101_priv_rac1_rvg`.
- 3 Verify the list of RVGs in the RDS by executing the following command.

```
vradmin -g oradg -l printrvg
For example:
Reeplicated Data Set: rac1_rvg
Primary:
HostName: 10.180.88.187 <localhost>
RvgName: rac1_rvg
DgName: oradatadg
datavol_cnt: 1
vset_cnt: 0
srl: rac1_srl
RLinks:
name=rlk_10.11.9.102_ rac1_rvg, detached=on,
synchronous=off
Secondary:
HostName: 10.190.99.197
RvgName: rac1_rvg
DgName: oradatadg
datavol_cnt: 1
vset_cnt: 0
srl: rac1_srl
RLinks:
name=rlk_10.10.9.101_ rac1_rvg, detached=on,
synchronous=off
```

---

**Note:** After you start the replication, the value of detached flag changes the status from on to off.

---

## Starting replication of Oracle database volume

When you have both the primary and secondary sites set up for replication, you can start replication from the primary site to the secondary site.

Start with the default replication settings:

- Mode of replication: synchronous=off
- Latency Protection: latencyprot=off
- SRL overflow protection: srlprot\_autodcm
- Packet size: packet\_size=8400
- Network protocol: protocol=UDP

Method of initial synchronization:

- Automatic synchronization
- Full synchronization with Checkpoint

For for guidelines on modifying these settings and information on choosing the method of replication for the initial synchronization, see the *Veritas Volume Replicator Administrator's Guide*.

## Starting replication using automatic synchronization

### To start replication using automatic synchronization

- ◆ From the primary site, automatically synchronize the RVG on the secondary site:

```
vradmin -g disk_group -a startrep pri_rvg sec_host
```

where:

- *disk\_group* is the disk group on the primary site that VVR will replicate
- *pri\_rvg* is the name of the RVG on the primary site
- *sec\_host* is the virtual host name for the secondary site

For example:

```
vradmin -g oradatadg -a startrep rac1_rvg rac_clus102_priv
```

Because the cluster on the secondary site uses only one host name, the command does not require the *sec\_host* argument. The command starts replication or the transfer of data from the primary site to the secondary site over the network.

## Starting replication using full synchronization with Checkpoint

### To start replication using full synchronization with Checkpoint

- 1 From the primary site, synchronize the RVG on the secondary site with full synchronization (using the `-c checkpoint` option):

```
vradmin -g disk_group -full -c ckpt_name syncrvg pri_rvg sec_host
```

where:

- *disk\_group* is the disk group on the primary site that VVR will replicate
- *ckpt\_name* is the name of the checkpoint on the primary site
- *pri\_rvg* is the name of the RVG on the primary site
- *sec\_host* is the virtual host name for the secondary site

For example:

```
vradmin -g oradatadg -c rac1_ckpt syncrvg rac1_rvg
rac_clus102_priv
```

- 2 To start replication after full synchronization, enter:

```
vradmin -g oradatadg -c rac1_ckpt startrep rac1_rvg
rac_clus102_priv
```

## Verifying replication status

You must verify that replication is functioning properly.

### To verify replication status

- 1 Use the `vxprint` command on the primary site:

```
vxprint -g diskgroup -l rlink_name
```

- 2 Review the `flags` output for the status. The output may appear as connected and consistent. For example:

```
vxprint -g oradatadg -l rlk_10.182.13.221_oradatadg
Rlink: rlk_10.182.13.221_oradatadg
info: timeout=500 packet_size=8400 rid=0.1078
 latency_high_mark=10000 latency_low_mark=9950
 bandwidth_limit=none
state: state=ACTIVE
 synchronous=off latencyprot=off srlprot=autodcm
.
.
protocol: UDP/IP
checkpoint: rac1_ckpt
flags: write enabled attached consistent connected
asynchronous
```

## Configuring VCS to replicate the database volume

After configuring both clusters for global clustering and setting up the Oracle database for replication, configure VCS to provide high availability for the database. Specifically, configure VCS agents to control the resources in the clusters, including resources for replication.

Review the sample `main.cf` files that illustrate the changes to the VCS configuration when you set up the existing Oracle database for replication.

See [“Sample main.cf for Oracle 10g for CVM/VVR primary site”](#) on page 360.

See [“Sample main.cf for Oracle 10g for CVM/VVR secondary site”](#) on page 363.

## Modifying the VCS configuration for replication

Resources that must be configured or modified for replication:

- Log owner group
- RVG group
- Oracle database service group
- RVGSharedPri resource
- The CVMVolDg resource from the existing Oracle database service group.

## About VCS resources for replication

Review the information on the groups and resources required for replication.

### Log owner group

A log owner group including the RVGLogowner resources must be created. The RVGLogowner resources are used by the RLINKs for the RVG, and the RVGLogowner resource, for which the RVG and its associated disk group are defined as attributes. The RVG log owner service group has an online local firm dependency on the service group containing the RVG.

### RVG group

An RVG group including the RVGShared resource replication objects must be created. Define the RVGShared resource and CVMVolDg resource together within a parallel service group. The group is defined as parallel because it may be online at the same time on all cluster nodes. The CVMVolDg resource does not have volumes specified for the CVMVolume attribute; the volumes are contained in the RVG resource. The CVMVolume attribute for the CVMVolDg resource is empty because all volumes in the RVG are defined by the RVG attribute of the RVGShared resource. The RVG service group has an online local firm dependency on the CVM service group.

VCS uses RVGLogowner agent to control the RVGLogowner resource, and the RVGShared agent to control the RVGShared resource. Refer to the *Veritas Cluster Server Agents for Veritas Volume Replicator Configuration Guide*.

### Oracle database service group

The RVGSharedPri resource must be added to the existing Oracle database service group. The CVMVolDg resource must be removed from the existing Oracle database service group.

The existing Oracle database service group is a parallel group consisting of the Oracle database resource, CVMVolDg resource, and CFMount resource (if the database resides in a cluster file system). Define the Oracle service group as a global group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute. Refer to the *Veritas Cluster Server Agents for Veritas Volume Replicator Configuration Guide* for more information on replication resources.

See [“About the CVMVolDg agent”](#) on page 373.

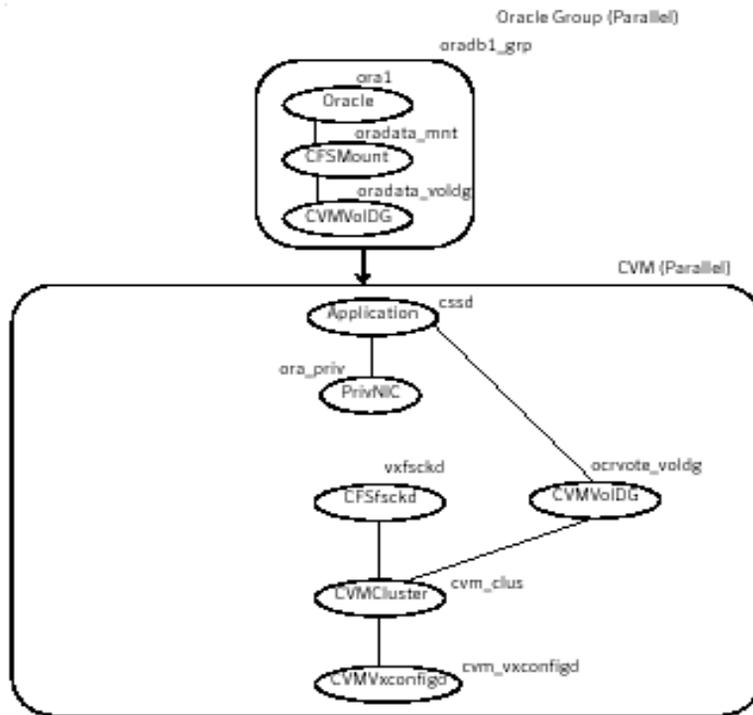
## Configuration before and after modification

Review the illustrations that show the changes to the VCS configuration when you set up the existing Oracle database for replication.

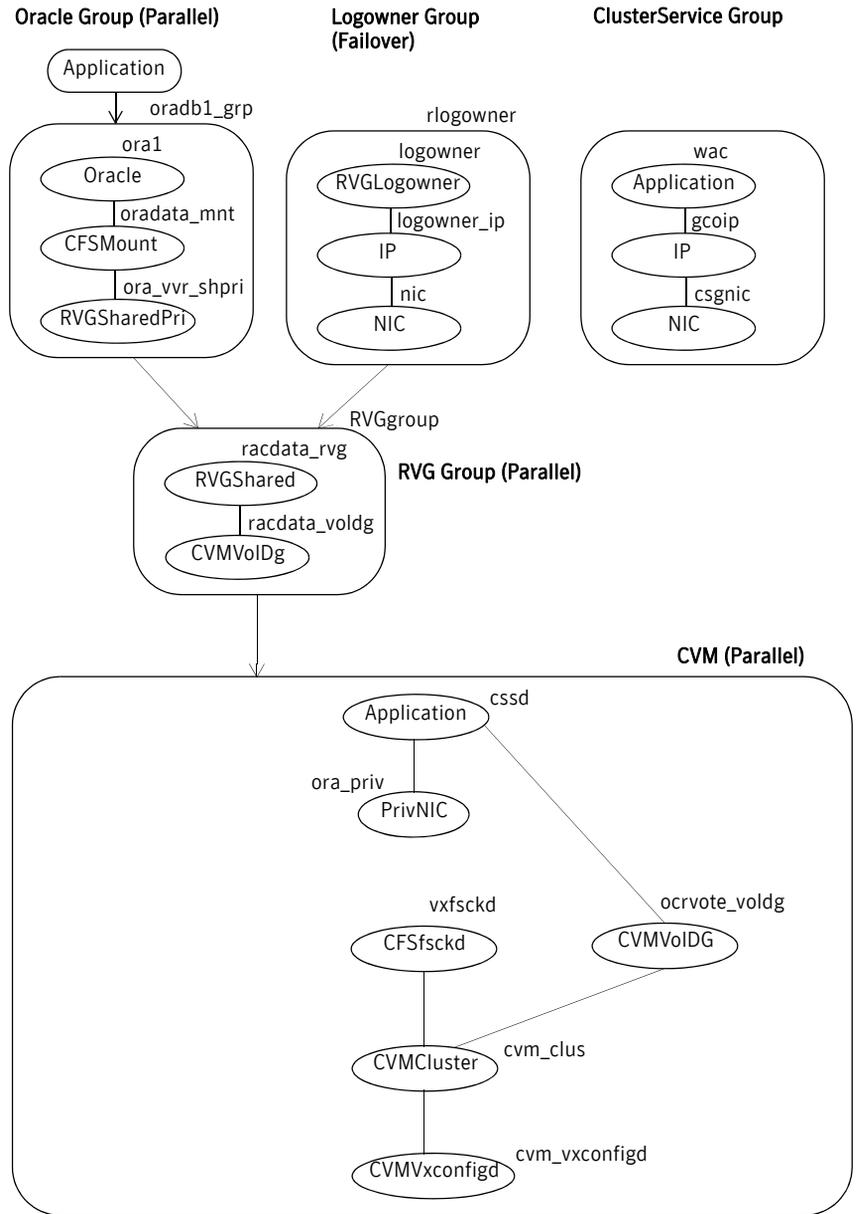
- Configuration before modification:

- See [Figure 13-15, "Illustration of dependencies: Configuration before modification for replication \(Oracle 10g\)."](#)
- Configuration after modification:
  - See [Figure 13-16, "Illustration of dependencies: Configuration after modification for replication \(Oracle 10g\)."](#)

**Figure 13-15** Illustration of dependencies: Configuration before modification for replication (Oracle 10g)



**Figure 13-16** Illustration of dependencies: Configuration after modification for replication (Oracle 10g)



## Modifying the VCS configuration on the primary site

Highlights of the procedure to modify the existing VCS configuration on the primary site include:

- Configure two service groups:
  - A log owner group including the RVGLogowner resource.
  - An RVG group including the RVGShared resource replication objects.
- Add the RVGSharedPri resource to the existing Oracle database service group and define this group as a global group by setting the ClusterList and ClusterFailOverPolicy attributes.
- Move the CVMVolDg resource from the existing Oracle database service group to the newly created RVGShared service group.

### To modify VCS on the primary site

- 1 Log into one of the nodes on the primary cluster.
- 2 Save the existing configuration to disk and make the configuration read-only while you are make the changes:

```
haconf -dump -makero
```
- 3 Make sure VCS is not running while you edit `main.cf` by stopping the VCS engine on all nodes and leave the resources available:

```
hastop -all -force
```
- 4 Make a backup copy of the `main.cf` file:

```
cd /etc/VRTSvcs/conf/config
cp main.cf main.orig
```
- 5 Use `vi` or another text editor to edit the `main.cf` file.
- 6 Add a failover service group using the appropriate values for your cluster and nodes. Include:
  - RVGLogowner resource. The node on which the group is online functions as the log owner (node connected to the second cluster for the purpose of replicating data).

- IP resource
- NIC resources

Example RVGLogowner service group:

```
group rlogowner (
 SystemList = { galaxy = 0, nebula = 1 }
 AutoStartList = { galaxy, nebula }
)

IP logowner_ip (
 Device =lan0
```

```

Address = "10.10.9.101"
NetMask = "255.255.240.0"
)

NIC nic (
 Device =lan0
 NetworkType = ether
)

RVGLogowner logowner (
 RVG = rac1_rvg
 DiskGroup = oradatadg
)
requires group RVGgroup online local firm
logowner requires logowner_ip
logowner_ip requires nic

```

- 7** Add the RVG service group using the appropriate values for your cluster and nodes.

Example RVGgroup service group:

```

group RVGgroup (
 SystemList = { galaxy = 0, nebula = 1 }
 Parallel = 1
 AutoStartList = { galaxy, nebula }
)

RVGShared racdata_rvg (
 RVG = rac1_rvg
 DiskGroup = oradatadg
)
CVMVoldg racdata_voldg(
 CVMDiskGroup = oradatadg
 CVMActivation = sw
)
requires group cvm online local firm
racdata_rvg requires racdata_voldg

```

- 8** Modify the Oracle service group using the appropriate values for your cluster and nodes:
- Define the Oracle service group as a *global* group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute (see the bolded attribute in the example that follows).
  - Add the ClusterFailOverPolicy cluster attribute. Symantec recommends using the Manual value. See the bolded attribute in the example.
  - Add the RVGSharedPri resource to the group configuration.

- Remove the CVMVolDg resource, if it has been configured in your previous configuration. This resource is now part of the RVG service group.
- Specify the service group to depend (online, local, firm) on the RVG service group.
- Remove the existing dependency of the Oracle Database service group on the CVM service group. Remove the line:  
requires group CVM online local firm
- Remove the existing dependency between the CFSSMount for the oracle database and the CVMVolDg for the Oracle database. Remove the line:  
oradata\_mnt requires oradata\_voldg

Example Oracle database service group configured for replication:

```
group oradb1_grp
SystemList = { galaxy = 0, nebula = 1 }
ClusterList = { rac_cluster101 = 0, rac_cluster102 = 1 }
Parallel = 1
ClusterFailOverPolicy = Manual
Authority = 1
AutoStartList = { galaxy, nebula }
)

CFSSMount oradata_mnt
MountPoint = "/oradata"
BlockDevice = "/dev/vx/dsk/oradatadg/racdb_vol"
)

RVGSharedPri ora_vvr_shpri (
 RvgResourceName = racdata_rvg
 OnlineRetryLimit = 0
)

Oracle rac_db (
Sid @galaxy = vrts1
Sid @nebula = vrts2
Owner = Oracle
Home = "/oracle/orahome/dbs"
Pfile @galaxy = "/oracle/orahome/dbs/initvrts1.ora"
Pfile @nebula = "/oracle/orahome/dbs/initvrts2.ora"
StartupOpt = SRVCTLSTART
ShutdownOpt = SRVCTLSTOP
MonScript = "./bin/Oracle/SqlTest.pl"
)
requires group RVGgroup online local firm
oradata_mnt requires ora_vvr_shpri
rac_db requires oradata_mnt
```

- 9 Save and close the main.cf file.

- 10 Verify the syntax of the `/etc/VRTSvcs/conf/config/main.cf` file:  

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

## Modifying the VCS configuration on the secondary site

Highlights of the procedure to modify the existing VCS configuration on the secondary site include:

- Add the log owner and RVG service groups.
- Add a service group to manage the Oracle database and the supporting resources.
- Define the replication objects and agents, such that the cluster at the secondary site can function as a companion to the primary cluster.

The steps are similar to what you performed on the primary site.

### To modify VCS on the secondary site

- 1 Log into one of the nodes on the secondary site as superuser.
- 2 Save the existing configuration to disk and make the configuration read-only while you are make the changes:  

```
hacnf -dump -makero
```
- 3 Ensure VCS is not running while you edit the `main.cf` by stopping the VCS engine on all systems and leave the resources available:  

```
hstop -all -force
```
- 4 Make a backup copy of the `main.cf` file:  

```
cd /etc/VRTSvcs/conf/config
cp main.cf main.orig
```
- 5 Use `vi` or another text editor to edit the `main.cf` file.
- 6 Edit the CVM group on the secondary site.  
Review the sample configuration file after the SF Oracle RAC installation to see the CVM configuration.  
See [“Verifying SF Oracle RAC installation using VCS configuration file”](#) on page 113.  
In our example, the secondary site has `rac_cluster102` consisting of the nodes `mercury` and `jupiter`. To modify the CVM service group on the secondary site, use the CVM group on the primary site as your guide.
- 7 Add a failover service group using the appropriate values for your cluster and nodes. Include:
  - `RVGLogowner` resource. The node on which the group is online functions as the log owner (node connected to the second cluster for the purpose of replicating data).

- IP resource
- NIC resources

Example RVGLogowner service group:

```
group rlogowner (
 SystemList = { galaxy = 0, nebula = 1 }
 AutoStartList = { galaxy, nebula }
)

 IP logowner_ip (
 Device =lan0
 Address = "10.11.9.102"
 NetMask = "255.255.240.0"
)

 NIC nic (
 Device =lan0
 NetworkType = ether
)

 RVGLogowner logowner (
 RVG = rac1_rvg
 DiskGroup = oradatadg
)

requires group RVGgroup online local firm
logowner requires logowner_ip
logowner_ip requires nic
```

- 8 Add the RVG service group using the appropriate values for your cluster and nodes.

Example RVGgroup service group:

```
group RVGgroup (
 SystemList = { mercury = 0, jupiter = 1 }
 Parallel = 1
 AutoStartList = { mercury, jupiter }
)

RVGShared racdata_rvg (
 RVG = rac1_rvg
 DiskGroup = oradatadg
)

 CVMVoldg racdata_voldg
 CVMDiskGroup = oradatadg
 CVMActivation = sw
)

requires group cvm online local firm
racdata_rvg requires racdata_voldg
```

- 9 Add an Oracle service group. Use the Oracle service group on the primary site as a model for the Oracle service group on the secondary site.
  - Define the Oracle service group as a *global* group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute.
  - Assign this global group the same name as the group on the primary site; for example, oradb1\_grp.
  - Include the ClusterList and ClusterFailOverPolicy cluster attributes. Symantec recommends using the Manual value.
  - Add the RVGSharedPri resource to the group configuration.
  - Remove the CVMVolDg resource, if it has been configured in your previous configuration. This resource is now part of the RVG service group.
  - Specify the service group to depend (online, local, firm) on the RVG service group.

Below is an example of the Oracle group on the secondary site:

```
.
group oradb1_grp
 SystemList = { mercury = 0, jupiter = 1 }
 ClusterList = { rac_cluster101 = 0, rac_cluster102 = 1 }
 Parallel = 1
 OnlineRegrInterval = 300
 ClusterFailOverPolicy = Manual
 Authority = 1
 AutoStartList = { mercury, jupiter }
)

 CFSMount oradata_mnt
 MountPoint = "/oradata"
 BlockDevice = "/dev/vx/dsk/oradatadg/racdb_vol"
)

RVGSharedPri ora_vvr_shpri (
 RvgResourceName = racdata_rvg
 OnlineRetryLimit = 0
)

Oracle rac_db (
 Sid @mercury = vrts1
 Sid @jupiter = vrts2
 Owner = Oracle
 Home = "/oracle/orahome/dbs"
 Pfile @mercury = "/oracle/orahome/dbs/initvrts1.ora"
 Pfile @jupiter = "/oracle/orahome/dbs/initvrts2.ora"
 StartUpOpt = SRVCTLSTART
 ShutDownOpt = SRVCTLSTOP
 MonScript = "./bin/Oracle/SqlTest.pl"
```

```
)

requires group RVGgroup online local firm
oradata_mnt requires ora_vvr_shpri
rac_db requires oradata_mnt
```

- 10 Save and close the main.cf file.
- 11 Verify the syntax of the `/etc/VRTSvcs/conf/config/main.cf` file:  

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

## Starting VCS on all nodes in both clusters

Start VCS on both clusters.

### To start VCS on all node in both clusters

- 1 From the primary site, start the VCS engine on one node:  

```
hstart
```
- 2 Type:  

```
hastatus
```
- 3 When LOCAL\_BUILD or RUNNING is listed in the message column, start VCS on the other node:  

```
hstart
```
- 4 Verify that VCS brings all resources online. On one node, enter:  

```
hagr -display
```

The Oracle, RVG, and CVM groups are online on both nodes of the primary site. The RVGLogOwner group is online on one node of the cluster. If either the RVG group or the RVGLogOwner group is partially online, manually bring the groups online using the `hagr -online` command. This information applies to the secondary site, except for the Oracle group which must be offline.
- 5 On the secondary site, start VCS from one node:  

```
hstart
```
- 6 Type:  

```
hastatus
```
- 7 When LOCAL\_BUILD or RUNNING is listed in the message column, start VCS on the other node:  

```
hstart
```
- 8 Verify the service groups and their resources that are brought online. On one node, enter:  

```
hagr -display
```

The Oracle service group is offline on the secondary site, but the CVM, RVG log owner, and RVG groups are online.

# Migration and takeover of primary replication role

Migration refers to the planned transfer of the role of primary replication host from one cluster to a remote cluster. This transfer enables the application on the remote cluster to actively use the replicated data. The former primary cluster becomes free for maintenance or other activity.

Takeover occurs when an unplanned event (such as a disaster) causes a failure, making it necessary for the applications using the replicated data to be brought online on the remote cluster.

## Migrating the role of primary site to the remote site

After configuring the replication objects within VCS, you can use VCS commands to migrate the role of the cluster on the primary site to the remote cluster. In the procedure below, VCS takes the replicated Oracle RAC database service group, `oradb1_grp`, offline on the primary site and brings it online on the secondary site; the secondary site now assumes the role of the primary site.

---

**Note:** The `hagrp -switch` command cannot migrate a parallel group within a cluster or between clusters in a global cluster environment.

---

- 1 From the primary site, take the Oracle service group offline on all nodes.  

```
hagrp -offline oradb1_grp -any
```

 Wait for VCS to take all Oracle service groups offline on the primary site.
- 2 Verify that the RLINK between the primary and secondary is up to date. Use the `vxrlink -g` command with the `status` option and specify the RLINK for the primary cluster (`rlk_rac_clus102_priv_rac1_rvg`, in this example). You can use the command from any node on the primary cluster. For example:

```
vxrlink -g oradatadg status rlk_rac_clus102_priv_rac1_rvg
```

- 3 On Secondary make sure that CRS is up, add listener resource using `netca`. Make changes to `tnsnames.ora`.

Example `tnsnames.ora` (Here `vrts` is the database name)

```
LISTENERS_VRTS =
 (ADDRESS_LIST =
 (ADDRESS = (PROTOCOL = TCP) (HOST = mercury-vip) (PORT =
1521))
 (ADDRESS = (PROTOCOL = TCP) (HOST = jupiter-vip) (PORT =
1521))
)

VRTS2 =
 (DESCRIPTION =
```

**Migration and takeover of primary replication role**

```

 (ADDRESS = (PROTOCOL = TCP)(HOST = jupiter-vip)(PORT =
1521))
 (CONNECT_DATA =
 (SERVER = DEDICATED)
 (SERVICE_NAME = vrts)
 (INSTANCE_NAME = vrts2)
)
)

VRTS1 =
 (DESCRIPTION =
 (ADDRESS = (PROTOCOL = TCP)(HOST = mercury-vip)(PORT =
1521))
 (CONNECT_DATA =
 (SERVER = DEDICATED)
 (SERVICE_NAME = vrts)
 (INSTANCE_NAME = vrts1)
)
)

VRTS =
 (DESCRIPTION =
 (ADDRESS = (PROTOCOL = TCP)(HOST = mercury-vip)(PORT =
1521))
 (ADDRESS = (PROTOCOL = TCP)(HOST = jupiter-vip)(PORT =
1521))
 (LOAD_BALANCE = yes)
 (CONNECT_DATA =
 (SERVER = DEDICATED)
 (SERVICE_NAME = vrts)
)
)

```

**4 Register the database using `srvctl` command. (On Secondary)**

```
srvctl add database -d <database_name> -o <oracle_home> -p
<spfile-on-sharedisk>
```

To prevent automatic database instance restart, change Management policy for the database (automatic, manual) to MANUAL using `srvctl` command:

```
srvctl modify database -d <database_name> -y manual
```

**5 Register the instances using `srvctl` command. Execute the following command on each node on the secondary site:**

```
srvctl add instance -d <database_name> -i <instance_name> -n
<node-name>
```

**6 Create directories `adump`, `bdump`, `cdump`, `dpdump`, `hdump`, `udump` in `$ORACLE_HOME/admin/<db_name>`.****7 Create pfile using spfile as follows on each node.**

```
echo "SPFILE=<location of spfile on shared volume>" >>
$ORACLE_HOME/dbs/init<database_name><instance-no>.ora
```

- 8 On the secondary site, bring the Oracle service group online on all nodes:  

```
hagrps -online oradb1_grp -any
```
- 9 Verify CRS resources by executing the command `crs_stat -t`. All resources must be online.
- 10 Make sure that all CRS resources are online, and switch back the group `oradb1_grp` to the primary site.
  - Issue the following command on the remote site:  

```
hagrps -offline oradb1_grp -any
```
  - Make sure that `oradb1_grp` is offline on the primary site. Then, execute the following command on the primary site to bring `oradb1_grp` online:  

```
hagrps -online oradb1_grp -any
```

## Taking over the primary role by the remote cluster

Takeover occurs when the remote cluster on the secondary site starts the application that uses replicated data. This situation may occur if the secondary site perceives the primary site as dead, or when the primary site becomes inaccessible (perhaps for a known reason). See the *Veritas Volume Replicator Administrator's Guide* for detailed description of concepts of taking over the primary role.

Before enabling the secondary site to take over the primary role, the administrator on the secondary site must “declare” the type of failure at the remote (primary, in this case) site. Designate the failure type using one of the options for the `haclus` command, are discussed in the following sections.

### disaster

When the cluster on the primary site is inaccessible and appears dead, the administrator declares the failure type as *disaster*. For example, fire may destroy a data center, including the primary site and all data in the volumes. After making this declaration, the administrator can bring the service group online on the secondary site, which now has the role as “primary” site.

### outage

When the administrator of a secondary site knows the primary site is inaccessible for a known reason, such as a temporary power outage, the administrator may declare the failure as an *outage*. Typically, an administrator expects the primary site to return to its original state.

After the declaration for an outage occurs, the `RVGSharedPri` agent enables DCM logging while the secondary site maintains the primary replication role. After the original primary site becomes alive and returns to its original state,

DCM logging makes it possible to use fast fail back resynchronization when data is resynchronized to the original cluster.

Before attempting to resynchronize the data using the fast fail back option from the current primary site to the original primary site, take the precaution at the original primary site to make a snapshot of the original data. This action provides a valid copy of data at the original primary site for use in the case the current primary site fails before the resynchronization is complete.

See “[replica](#)” on page 234.

## disconnect

When both clusters are functioning properly and the heartbeat link between the clusters fails, a split-brain condition exists. In this case, the administrator can declare the failure as *disconnect*, meaning no attempt will occur to take over the role of the primary site at the secondary site. This declaration is merely advisory, generating a message in the VCS log indicating the failure results from a network outage rather than a server outage.

## replica

In the rare case where the current primary site becomes inaccessible while data is resynchronized from that site to the original primary site using the fast fail back method, the administrator at the original primary site may resort to using a data snapshot (if it exists) taken before the start of the fast fail back operation. In this case, the failure type is designated as *replica*.

## Example of takeover for an outage

### To take over after an outage

- 1 From any node of the secondary site, issue the `haclus` command:  

```
haclus -declare outage -clus rac_cluster101
```
- 2 After declaring the state of the remote cluster, bring the Oracle service group online on the secondary site. For example:  

```
hagrps -online -force oradb1_grp -any
```

## Example of resynchronization after an outage

### To resynchronize after an outage

- 1 On the original primary site, create a snapshot of the RVG before resynchronizing it in case the current primary site fails during the resynchronization. Assuming the disk group is `oradatadg` and the RVG is `rac1_rvg`, type:

```
vxrvg -g oradatadg -F snapshot rac1_rvg
```

See the *Veritas Volume Replicator Administrator's Guide* for details on RVG snapshots.

- 2 Resynchronize the RVG. From the CVM master node of the current primary site, issue the `hares` command and the `-action` option with the `fbsync` action token to resynchronize the `RVGSharedPri` resource. For example:
 

```
hares -action ora_vvr_shpri fbsync -sys mercury
```

 To determine which node is the CVM master node, type:
 

```
vxdctl -c mode
```
- 3 Perform one of the following commands, depending on whether the resynchronization of data from the current primary site to the original primary site is successful:
  - If the resynchronization of data is successful, use the `vxrvg` command with the `snapback` option to reattach the snapshot volumes on the original primary site to the original volumes in the specified RVG:
 

```
vxrvg -g oradatadg snapback rac1_rvg
```
  - A failed attempt at the resynchronization of data (for example, a disaster hits the primary RVG when resynchronization is in progress) could generate inconsistent data. You can restore the contents of the RVG data volumes from the snapshot taken in [step 1](#):
 

```
vxrvg -g oradatadg snaprestore rac1_rvg
```

**Migration and takeover of primary replication role**

# Backup and recovery for SF Oracle RAC

- [Chapter 14, “Configuring the repository database for Oracle” on page 239](#)
- [Chapter 15, “Using Storage Checkpoints and Storage Rollback” on page 243](#)
- [Chapter 16, “Using FlashSnap for backup and recovery” on page 271](#)



# Configuring the repository database for Oracle

This chapter contains the following topics:

- [About configuring the repository database for Oracle](#)
- [Creating and configuring the repository database for Oracle](#)

## About configuring the repository database for Oracle

After installing SF Oracle RAC, you can create and configure the repository database using the `sfua_db_config` script. This repository database configuration enable you to use SF Oracle RAC features such as Checkpoint, flashsnap and storage mapping. The script detects that the system is running in an HA configuration and automatically configures the repository database.

## Creating and configuring the repository database for Oracle

Before running the `sfua_db_config` script, review the following requirements:

- Make sure a disk group exists with at least one volume, which should not be shared. A VxFS file system must be created on the disk group.
- The volume must be mounted.
- Obtain an unique virtual IP address for public NIC interface.
- Obtain a device name for the public NIC interface (for example: `lan0`).
- Obtain a subnet mask for the public NIC interface.

---

**Note:** The volume is used to store the repository database.

---

**Table 14-1** indicates the options available for the sfua\_db\_config script.

**Table 14-1** sfua\_db\_config options

Options	Description
-ssh	Use this option in a high availability (HA) configuration. The option indicates that ssh and scp are to be used for communication between systems. Either ssh or rsh should be preconfigured so that you can execute the commands without being prompted for passwords or confirmations.
-o dropdb	Drops the repository database.
-o unconfig_cluster	Use this option in a high availability (HA) configuration. Unconfigures the repository database from the VCS cluster.
-o dbstatus	Verifies the status of the database and database server.
-o stopserver	Stops the database server.
-o startserver	Starts the database server.
-o serverstatus	Reports the database server status.
-o stopdb	Detaches the repository database from the database server.
-o startdb	Attaches the repository database to the database server.

**To create and configure the repository database**

- 1 Run the sfua\_db\_config script:

```
/opt/VRTSdbcom/bin/sfua_db_config
```

- 2 The following is an example of configuring SF Oracle RAC:

```
/opt/VRTSdbcom/bin/sfua_db_config
```

```
Welcome to the SFORA configuration script.
This script creates repository for standalone and HA
configuration. Please create a Veritas File System on a Veritas
Volume and mount it, before starting configuration using this
script. This mount point will be used to store repository. The
following is required to configure SFORA repository for HA
solution:
```

- \* A mount point of already mounted Veritas Volume on a shared storage, with Veritas File system.
- \* A public NIC used by each system in the cluster.
- \* A Virtual IP address and netmask.

```

Are you ready to configure SFORA repository (y/n/q) [y]?
Enter Veritas filesystem mount point for SFORA repository:
/sfua_rep

Discovering public NIC on galaxylan0 lan1

Enter the NIC for system galaxy for HA Repository configuration:
[lan0] lan0

Discovering public NIC on nebulalan0 lan1

Enter the NIC for system nebula for HA Repository configuration:
[lan0] lan0

Enter the Virtual IP address for repository failover:
10.182.186.249

Enter the netmask for public NIC interface: [255.255.240.0]

Following information will be used for SFORA HA configuration:

Public IP address: 10.182.186.249
Subnet mask: 255.255.240.0
Public interface: galaxy -> lan0,nebula -> lan0
Mount point: /sfua_rep
Volume Name for mount point: dbed_rep
Diskgroup for mount point: sfua_rep
Is this correct (y/n/q) [y]? y

Adding repository information in VCS (HA) configuration...

Added repository information successfully in VCS (HA)
configuration.

Repository configuration completed successfully for HA
environment.

```

- 3 If you are upgrading, migrate your old repository information into the new repository. If you are installing or upgrading Veritas Storage Foundation for Oracle, run the `dbed_update` command.

## Setting administrative permissions

To allow database administrators to administer a database using SF Oracle RAC, you must change permission settings. During SF Oracle RAC installation, you are asked if you want to allow database administrators access. If you did not change permissions installation, you can do so at a later time.

The default settings at installation time for the `/opt/VRTSdbed` directory allow only the root login to access the directory.

**To enable access for users other than root**

- 1 To enable the user “oracle” access to the `/opt/VRTSdbed` directory, use the `chown` and `chmod` commands, as follows:  

```
chown oracle /opt/VRTSdbed
chmod 500 /opt/VRTSdbed
```
- 2 To allow users in the group “dba” access to the `/opt/VRTSdbed` directory, use the `chgrp` and `chmod` commands, as follows:  

```
chgrp dba /opt/VRTSdbed
chmod 550 /opt/VRTSdbed
```

# Using Storage Checkpoints and Storage Rollback

This chapter contains the following topics:

- [About Storage Checkpoints and Storage Rollback in SF Oracle RAC](#)
- [Using Storage Checkpoints and Storage Rollback for backup and restore](#)
- [Determining Space Requirements for Storage Checkpoints](#)
- [Performance of Storage Checkpoints](#)
- [Backing up and recovering the database using Storage Checkpoints](#)
- [Guidelines for Oracle Recovery](#)
- [Using the Storage Checkpoint Command Line Interface \(CLI\)](#)
- [Examples of using the Command Line Interface](#)

## About Storage Checkpoints and Storage Rollback in SF Oracle RAC

Veritas Storage Checkpoints enable efficient backup and recovery of Oracle databases. The Veritas Storage Checkpoint feature is available with SF Oracle RAC as part of the Veritas File System package and is used for the efficient backup and recovery of Oracle databases. Storage Checkpoints can also be mounted, allowing regular file system operations to be performed or secondary databases to be started. Review the information on Storage Checkpoints and Storage Rollback and how to use these technologies through Storage Foundation for Oracle RAC.

## Using Storage Checkpoints and Storage Rollback for backup and restore

SF Oracle RAC provides a Storage Checkpoint facility that is similar to the snapshot file system mechanism; however, a Storage Checkpoint persists after a system reboot. A Storage Checkpoint creates an exact image of a database instantly and provides a consistent image of the database from the point in time the Storage Checkpoint was created. The Storage Checkpoint image is managed and available through the Veritas Storage Foundation command line interface (CLI).

For more information on creating Storage Checkpoints with the CLI:

See “[Creating Storage Checkpoints using `dbed\_ckptcreate`](#)” on page 257.

A direct application of the Storage Checkpoint facility is Storage Rollback. Because each Storage Checkpoint is a consistent, point-in-time image of a file system, Storage Rollback is the restore facility for these on-disk backups. Storage Rollback rolls back changed blocks contained in a Storage Checkpoint into the primary file system for restoring the database faster.

For more information on Storage Checkpoints and Storage Rollback, see the *Veritas File System Administrator’s Guide*.

### About Storage Checkpoints and Storage Rollback

A Storage Checkpoint is a disk and I/O efficient snapshot technology for creating a “clone” of a currently mounted file system (the *primary* file system). Like a snapshot file system, a Storage Checkpoint appears as an exact image of the snapped file system at the time the Storage Checkpoint was made. However, unlike a snapshot file system that uses separate disk space, all Storage Checkpoints share the same free space pool where the primary file system resides unless a Storage Checkpoint allocation policy is assigned. A Storage Checkpoint can be mounted as read-only or read-write, allowing access to the files as if it were a regular file system. A Storage Checkpoint is created using the `dbed_ckptcreate` command.

Initially, a Storage Checkpoint contains no data—it contains only the inode list and the block map of the primary fileset. This block map points to the actual data on the primary file system. Because only the inode list and block map are needed and no data is copied, creating a Storage Checkpoint takes only a few seconds and very little space.

A Storage Checkpoint initially satisfies read requests by finding the data on the primary file system, using its block map copy, and returning the data to the requesting process. When a write operation changes a data block *n* in the primary file system, the old data is first copied to the Storage Checkpoint, and

then the primary file system is updated with the new data. The Storage Checkpoint maintains the exact view of the primary file system at the time the Storage Checkpoint was taken. Subsequent writes to block *n* on the primary file system do not result in additional copies to the Storage Checkpoint because the old data only needs to be saved once. As data blocks are changed on the primary file system, the Storage Checkpoint gradually fills with the original data copied from the primary file system, and less and less of the block map in the Storage Checkpoint points back to blocks on the primary file system.

You can set a quota to limit how much space a file system will give to all storage checkpoints, to prevent the checkpoints from consuming all free space. See the command `dbed_ckptquota` for more information.

Storage Rollback restores a database, a tablespace, or datafiles on the primary file systems to the point-in-time image created during a Storage Checkpoint. Storage Rollback is accomplished by copying the “before” images from the appropriate Storage Checkpoint back to the primary file system. As with Storage Checkpoints, Storage Rollback restores at the block level, rather than at the file level. Storage Rollback is executed using the `dbed_ckptrollback` command.

---

**Note:** Whenever you change the structure of the database (for example, by adding or deleting datafiles, converting PFILE to SPFILE, or converting SPFILE to PFILE), you must run `dbed_update`.

---

Mountable Storage Checkpoints can be used for a wide range of application solutions, including backup, investigations into data integrity, staging upgrades or database modifications, and data replication solutions.

If you mount a Storage Checkpoint as read-write, the command will not allow you to roll back to this Storage Checkpoint. This ensures that any Storage Checkpoint data that has been modified incorrectly cannot be a source of any database corruption. When a Storage Checkpoint is mounted as read-write, the `dbed_ckptmount` command creates a “shadow” Storage Checkpoint of and mounts this “shadow” Storage Checkpoint as read-write. This allows the database to still be rolled back to the original Storage Checkpoint.

For more information on mountable Storage Checkpoints:

See [“Mounting Storage Checkpoints using `dbed\_ckptmount`”](#) on page 263.

## Determining Space Requirements for Storage Checkpoints

To support Block-level Incremental (BLI) Backup and storage rollback, the file systems need extra disk space to store the Storage Checkpoints. The extra space

needed depends on how the Storage Checkpoints are used. Storage Checkpoints that are used to keep track of the block changes contain only file system block maps, and therefore require very little additional space (less than 1 percent of the file system size).

If the database is online while the backup is running, the additional space required by each file system for Storage Checkpoints depends on the duration of the backup and the database workload. If workload is light during the backup or the backup window is relatively short (for example, for incremental backups), for most database configurations, an additional 10 percent of the file system size will be sufficient. If the database has a busy workload while a full backup is running, the file systems may require more space.

To support Storage Checkpoints and storage rollback, VxFS needs to keep track of the original block contents when the Storage Checkpoints were created. The additional space needed is proportional to the number of blocks that have been changed since a Storage Checkpoint was taken. The number of blocks changed may not be identical to the number of changes. For example, if a data block has been changed many times, only the first change requires a new block to be allocated to store the original block content. Subsequent changes to the same block require no overhead or block allocation.

If a file system that has Storage Checkpoints runs out of space, by default VxFS removes the oldest Storage Checkpoint automatically instead of returning an ENOSPC error code (UNIX errno 28- No space left on device), which can cause the Oracle instance to fail. Removing Storage Checkpoints automatically ensures the expected I/O semantics, but at the same time, eliminates a key recovery mechanism.

When restoring a file system that has data-full Storage Checkpoints from tape or other offline media, you need extra free space on the file system. The extra space is needed to accommodate the copy-on-write algorithm needed for preserving the consistent image of the Storage Checkpoints. The amount of free space required depends on the size of the restore and the number of Storage Checkpoints on the file system.

If you are restoring the entire file system, in most cases, you no longer need the existing Storage Checkpoint. You can simply re-make the file system using the mkfs command, and then restore the file system from tape or other offline media.

If you are restoring some of the files in the file system, you should first remove the data-full Storage Checkpoints that are no longer needed. If you have very limited free space on the file system, you may have to remove all data-full Storage Checkpoints in order for the restore to succeed.

To avoid unnecessary Storage Checkpoint removal, instead of using a low quota limit use the SFDB utility to set up a Monitoring Agent to monitor file system space usage. When file system space usage exceeds a preset threshold value

(say, 95 percent full), the Monitoring Agent alerts the system administrator and optionally grows the volume and the file system. Automatic notifications to the system administrator on the status of space usage and file system resizing are available through electronic mail, the `syslogd(1M)` program, or by logging messages to a simple log file.

Always reserve free disk space for growing volumes and file systems. You can also preallocate sufficient space for each file system when the file system is first created or manually grow the file system and logical volume where the file system resides.

See the `vxassist(1)` and `fsadm_vxfs(1)` manual pages for more information.

## Performance of Storage Checkpoints

Veritas File System attempts to optimize the read and write access performance on both the Storage Checkpoint and the primary file system. Reads from a Storage Checkpoint typically perform at nearly the throughput of reads from a normal VxFS file system, allowing backups to proceed at the full speed of the VxFS file system.

Writes to the primary file system are typically affected by the Storage Checkpoints because the initial write to a data block requires a read of the old data, a write of the data to the Storage Checkpoint, and finally, the write of the new data to the primary file system. Having multiple Storage Checkpoints on the same file system, however, will not make writes slower. Only the initial write to a block suffers this penalty, allowing operations like writes to the intent log or inode updates to proceed at normal speed after the initial write.

The performance impact of Storage Checkpoints on a database is less when the database files are Direct I/O files. A performance degradation of less than 5 percent in throughput has been observed in a typical OLTP workload when the Storage Checkpoints only keep track of changed information. For Storage Checkpoints that are used for storage rollback, higher performance degradation (approximately 10 to 20 percent) has been observed in an OLTP workload. The degradation should be lower in most decision-support or data-warehousing environments.

Reads from the Storage Checkpoint are impacted if the primary file system is busy, because the reads on the Storage Checkpoint are slowed by all of the disk I/O associated with the primary file system. Therefore, performing database backup when the database is less active is recommended.

## Backing up and recovering the database using Storage Checkpoints

Storage Checkpoints can be created by specifying one of the following options: online, offline, or instant. To create a Storage Checkpoint with the online option, the database should be online and you must enable ARCHIVELOG mode for the database. For the offline option, the database should be offline.

During the creation of the Storage Checkpoint, the tablespaces are placed in backup mode. Because it only takes a few seconds to take a Storage Checkpoint, the extra redo logs generated while the tablespaces are in online-backup mode are very small. You can roll back the entire database or individual tablespaces or datafiles to an online or offline Storage Checkpoint. After the rollback is complete, you may roll the database forward to restore the database if you have used an online Storage Checkpoint.

For the instant option, the database should be online and it can be running in either ARCHIVELOG or NOARCHIVELOG mode. You can only roll back the entire database to an instant Storage Checkpoint. Rolling back individual tablespaces or datafiles to an instant Storage Checkpoint is not possible. After the rollback is complete, you need to perform database recovery. Rolling the database forward is not supported; that is, you cannot apply archived redo logs.

To allow the easiest recovery, always keep ARCHIVELOG mode enabled, regardless of whether the database is online or offline when you create Storage Checkpoints.

### Verifying a Storage Checkpoint using the command line

After creating a Storage Checkpoint and before using it to back up or restore a database, you can verify that the Storage Checkpoint is free of errors.

Usage notes                      See the `dbed_ckptcreate(1M)` and `dbed_ckptmount(1M)` manual pages for more information.

See [“Creating Storage Checkpoints using `dbed\_ckptcreate`”](#) on page 257.

See [“Mounting Storage Checkpoints using `dbed\_ckptmount`”](#) on page 263.

#### To verify that a Storage Checkpoint is error-free using the command line

- 1 Create and mount a Storage Checkpoint:

```
$ /opt/VRTS/bin/dbed_ckptcreate -S PROD -H /oracle/product \
-o online
```

```
Storage Checkpoint Checkpoint_903937870 created.
```

```
$ mkdir /tmp/ckpt_ro
```

```
$ /opt/VRTS/bin/dbed_ckptmount -S PROD -c Checkpoint_903937870 \
-m /tmp/ckpt_ro
```

If the specified mount point directory does not exist, then `dbed_ckptmount` creates it before mounting the Storage Checkpoint, as long as the Oracle DBA user has permission to create it.

## 2 Examine the contents of the Storage Checkpoint:

```
$ ls -l /tmp/ckpt_ro/dbvol_82/dbinst1
drwxr-xr-x3 oracle dba 1024 Nov 11 2000 .
drwxr-xr-x3 oracle dba 512 Nov 16 11:00 ..
-rw-r--r--1 oracle dba 209747968 Nov 16 10:58 .tstmp
-rw-r--r--1 oracle dba 209747968 Nov 16 10:58 .tstab
lrwxrwxrwx1 oracle dba 18 Nov 11 2000 tstmp ->
.tstmp::cdev:vxfs:
lrwxrwxrwx1 oracle dba 18 Nov 11 2000 tstab ->
.tstab::cdev:vxfs:
```

## 3 Run dbv tool against Quick I/O file tstmp:

```
$ dbv file=/tmp/ckpt_ro/db01/tstmp
DBVERIFY: Release 9.2.0.2.0 - Production on Mon Mar 7 11:48:35 \
2005
```

Storage Checkpoints can only be used to restore from logical errors (for example, a human error). Because all the data blocks are on the same physical device, Storage Checkpoints cannot be used to restore files due to a media failure. A media failure requires a database restore from a tape backup or a copy of the database files kept on a separate medium. The combination of data redundancy (disk mirroring) and Storage Checkpoints is recommended for highly critical data to protect them from both physical media failure and logical errors.

## Backing up using a Storage Checkpoint

You can back up a database by creating a Storage Checkpoint using the `dbed_ckptcreate` command, mount the Storage Checkpoint as read-only using the `dbed_ckptmount` command, and then back it up using tools such as `tar` or `cpio`.

Usage notes See the `dbed_ckptcreate(1M)`, `dbed_ckptmount(1M)`, `tar(1)`, and `cpio(1)` manual pages for more information.

See [“Creating Storage Checkpoints using `dbed\_ckptcreate`”](#) on page 257.

See [“Mounting Storage Checkpoints using `dbed\_ckptmount`”](#) on page 263.

In the example procedure, all the database datafiles reside on one VxFS file system named `/db01`.

### To back up a frozen database image using the command line

- 1 Create a Storage Checkpoint using the `dbed_ckptcreate` command:  

```
$ /opt/VRTS/bin/dbed_ckptcreate -S PROD -H /oracle/product \
-o online
Storage Checkpoint Checkpoint_903937870 created.
```
- 2 Mount the Storage Checkpoint using the `dbed_ckptmount` command:  

```
$ /opt/VRTS/bin/dbed_ckptmount -S PROD -c Checkpoint_903937870 \
-m /tmp/ckpt_ro
```

If the specified mount point directory does not exist, then `dbed_ckptmount` creates it before mounting the Storage Checkpoint, as long as the Oracle DBA user has permission to create it.
- 3 Use `tar` to back up the Storage Checkpoint:  

```
$ cd /tmp/ckpt_ro
$ ls
db01
$ tar cvf /tmp/PROD_db01_903937870.tar ./db01
```

## Recovering a database using a Storage Checkpoint

Because Storage Checkpoints record the before images of blocks that have changed, you can use them to do a file-system-based storage rollback to the exact time when the Storage Checkpoint was taken. You can consider Storage Checkpoints as backups that are online, and you can use them to roll back an entire database, a tablespace, or a single database file. Rolling back to or restoring from any Storage Checkpoint is generally very fast because only the changed data blocks need to be restored.

---

**Note:** Some database changes made after a Storage Checkpoint was taken may make it impossible to perform an incomplete recovery of the databases after Storage Rollback of an online or offline Storage Checkpoint using the current control files. For example, you cannot perform incomplete recovery of the database to the point right before the control files have recorded the addition or removal of datafiles. To provide recovery options, a backup copy of the control file for the database is saved under the `/etc/vx/vxdbed/$ORACLE_SID/checkpoint_dir/CKPT_NAME` directory immediately after a Storage Checkpoint is created. You can use this file to assist with database recovery, if necessary. If possible, both ASCII and binary versions of the control file will be left under the `/etc/vx/vxdbed/$ORACLE_SID/checkpoint_dir/CKPT_NAME` directory. The binary version will be compressed to conserve space. Use extreme caution when recovering your database using alternate control files.

---

Suppose a user deletes a table by mistake right after 4:00 p.m., and you want to recover the database to a state just before the mistake. You created a Storage Checkpoint (Checkpoint\_903937870) while the database was running at 11:00 a.m., and you have ARCHIVELOG mode enabled.

If the Oracle RAC database is running under VCS control, you must perform the following steps before you recover the database:

- 1 As superuser, temporarily freeze the VCS resource group for the database.  
`# hagr -freeze resource_group`
- 2 Shut down the primary database.
- 3 Run the `dbed_ckptrollback` command.
- 4 Unfreeze the resource group:  
`# hagr -unfreeze resource_group`

#### To recover the database using a Storage Checkpoint

- 1 Ensure that the affected datafiles, tablespaces, or database are offline, and use Storage Rollback to roll back any datafiles in the database that contained the table data from the Storage Checkpoint you created at 11:00 a.m.
- 2 Start up the database instance if it is down.
- 3 Re-apply archive logs to the point before the table was deleted to recover the database to 4:00 p.m. Use one of the following:  

```
recover database until cancel
recover database until change
recover database until time
```
- 4 Open the database with the following command:  

```
alter database open resetlogs
```
- 5 Delete the Storage Checkpoint you created at 11:00 a.m. and any other Storage Checkpoints created before that time.
- 6 Create a new Storage Checkpoint.

## Guidelines for Oracle Recovery

For optimal Oracle recovery, follow these guidelines:

- Back up all control files before storage rollback in case the subsequent Oracle recovery is not successful. Oracle recommends that you keep at least two copies of the control files for each Oracle database and that you store the copies on different disks. It is also a good idea to back up the control files before and after making structural changes to databases.

---

**Note:** The `dbed_ckptcreate` command automatically saves control file and log information when you create a Storage Checkpoint.

---

See “[Creating Storage Checkpoints using `dbed\_ckptcreate`](#)” on page 257.

- Make sure that the control files are *not* rolled back.  
A control file is a small binary file that describes the structure of the database and must be available to mount, open, and maintain the database. The control file stores all necessary database file information, log file information, the name of the database, the timestamp of database creation, and synchronization information, such as the Storage Checkpoint and log-sequence information needed for recovery. Rolling back the control file will result in an inconsistency between the physical database structure and the control file.

---

**Note:** If your intention is to roll back the database to recover from structural changes that you do not want to maintain, you may want to use the backup control file that was created by the `dbed_ckptcreate` command.

---

- Make sure that all archived redo logs are available.  
A database backup with online and archived logs is required for a complete database recovery. Query `V$ARCHIVED_LOG` to list all the archived log information and `V$ARCHIVE_DEST` to list the location of archive destinations.  
To restore the necessary archived redo log files, you can query `V$LOG_HISTORY` to list all the archived redo log history or query `V$RECOVERY_LOG` to list only the archived redo logs needed for recovery. The required archived redo log files can be restored to the destination specified in the `LOG_ARCHIVE_DEST` parameter or to an alternate location. If the archived redo logs were restored to an alternate location, use the `ALTER DATABASE RECOVER . . . FROM` statement during media recovery.
- After storage rollback, perform Oracle recovery, applying some or all of the archived redo logs.

---

**Note:** After rolling back the database (including control files and redo logs) to a Storage Checkpoint, you need to recover the Oracle database instance. Rolling the database forward is not supported; that is, you cannot apply archived redo logs.

---

- To perform a complete media recovery:  

```
SET AUTORECOVERY ON;
RECOVER DATABASE;
```

- To perform an incomplete media recovery, use one of the following:
  - `RECOVER DATABASE UNTIL CANCEL;`
  - `RECOVER DATABASE UNTIL TIME 'yyyy-mm-dd:hh:mm:ss';`  
(You can confirm the time of error by checking the `../bdump/alert*.log` file.)
  - `RECOVER DATABASE UNTIL TIME 'yyyy-mm-dd:hh:mm:ss' \`  
`using backup controlfile;`
  - `RECOVER DATABASE UNTIL CHANGE scn;`
- To open the database after an incomplete media recovery, use the following:  
`ALTER DATABASE OPEN RESETLOGS;`  
RESETLOGS resets the log sequence. The RESETLOGS option is required after an incomplete media recovery. After opening the database with the RESETLOGS option, remove the Storage Checkpoint you just rolled back to as well as any Storage Checkpoints that were taken before that one. These earlier Storage Checkpoints can no longer be used for storage rollback. After removing these Storage Checkpoints, be sure to create a new Storage Checkpoint.

---

**Caution:** Attempting to roll back to the same Storage Checkpoint more than once can result in data corruption. After rolling back, be sure to delete the Storage Checkpoint that you rolled back to and then create a new one.

---

See your Oracle documentation for complete information on recovery.

## Using the Storage Checkpoint Command Line Interface (CLI)

Veritas Storage Foundation for Oracle RAC provides a command line interface to many key operations. The command line interface lets you incorporate command operations into scripts and other administrative processes.

---

**Note:** The SF Oracle RAC command line interface depends on certain tablespace and container information that is collected and stored in a repository. Some CLI commands update the repository by default. It is also important to regularly ensure the repository is up-to-date by using the `dbed_update` command.

---

---

**Note:** For SF Oracle RAC database, when you issue the commands, replace \$ORACLE\_SID with \$ORACLE\_SID=instance\_name and provide the instance name on which the instance is running.

---

## Overview of Commands

SF Oracle RAC commands supported in the command line interface are located in the /opt/VRTS/bin directory. Online manual pages for these commands are located in the /opt/VRTS/man directory.

[Table 15-2](#) summarizes the commands available to you from the command line.

**Table 15-2** Veritas Storage Foundation for Oracle RAC Commands

Command	Description
dbed_update	Updates the repository in Veritas Storage Foundation for Oracle RAC See <a href="#">“Creating or updating the repository using dbed_update”</a> on page 256.
dbed_ckptcreate	Creates a Storage Checkpoint for an Oracle database See <a href="#">“Creating Storage Checkpoints using dbed_ckptcreate”</a> on page 257.
dbed_ckptdisplay	Displays Storage Checkpoints associated with an Oracle instance See <a href="#">“Displaying Storage Checkpoints using dbed_ckptdisplay”</a> on page 258.
dbed_ckptmount	Mounts a Storage Checkpoint for an Oracle instance See <a href="#">“Mounting Storage Checkpoints using dbed_ckptmount”</a> on page 263.
dbed_ckptumount	Unmounts a Storage Checkpoint for an Oracle instance See <a href="#">“Unmounting Storage Checkpoints using dbed_ckptumount”</a> on page 264.
dbed_ckptrollback	Rolls back an Oracle instance to a Storage Checkpoint point-in-time image See <a href="#">“Performing Storage Rollback using dbed_ckptrollback”</a> on page 264.

**Table 15-2** Veritas Storage Foundation for Oracle RAC Commands

Command	Description
<code>dbed_ckptremove</code>	Removes a Storage Checkpoint for an Oracle instance See <a href="#">“Removing Storage Checkpoints using dbed_ckptremove”</a> on page 266.
<code>dbed_clonedb</code>	Creates a copy of an Oracle database by cloning all existing database files and recreating the control file accordingly. This cloned database can only be started on the same host as the existing database as long as it uses a different SID. See <a href="#">“Cloning the Oracle instance using dbed_clonedb”</a> on page 266.

## Examples of using the Command Line Interface

Review the examples for using the Veritas Storage Foundation for Oracle RAC command line interface to perform administrative operations. For more detailed information about the commands and their syntax and available options, see the individual manual pages.

### Prerequisites

You must log in as the database administrator to use the following CLI commands.

- `dbed_update`
- `dbed_ckptcreate`
- `dbed_clonedb`
- `dbed_ckptdisplay`
- `dbed_ckptmount`
- `dbed_ckptumount`
- `dbed_ckptrollback`
- `dbed_ckptremove`

### Creating or updating the repository using `dbed_update`

You can use the `dbed_update` command to update the repository.

Any time you change the structure of the database (for example, by adding or deleting datafiles, converting PFILE to SPFILE, or converting SPFILE to PFILE), you must run `dbed_update`.

Before creating or updating the repository, the following conditions must be met:

- Prerequisites
- You must be logged on as the database administrator (typically, the user ID oracle).

## Usage notes

- The `dbed_update` command creates a repository in the `/etc/vx/vxdbed/$ORACLE_SID` directory where information used by SF Oracle RAC is kept. If the repository already exists, the command will refresh the information.
- The database must be up and running, and the `ORACLE_SID` and the `ORACLE_HOME` variable arguments must be specified with the `-S` and `-H` options, respectively.
- See the `dbed_update(1M)` manual page for more information.

## To update the repository

- ◆ Use the `dbed_update` command as follows:

```
$ /opt/VRTS/bin/dbed_update -S PROD -H /oracle/product/10g
```

## Creating Storage Checkpoints using `dbed_ckptcreate`

You can use the `dbed_ckptcreate` command to create a Storage Checkpoint from the command line.

Storage Checkpoints can be either online, offline, or instant. By default, Storage Checkpoints are offline. If online is specified, the database is put into hot-backup mode when the Storage Checkpoint is created. If offline is specified, the database is expected to be down. If instant is specified, the database must be online and a Storage Checkpoint will be taken for a “crash recovery”-type Storage Rollback.

Before creating a Storage Checkpoint, the following conditions must be met:

## Prerequisites

- You must be logged on as the database administrator (typically, the user ID oracle).
- For best recoverability, always keep ARCHIVELOG mode enabled when you create Storage Checkpoints.

## Usage notes

- `dbed_ckptcreate` stores Storage Checkpoint information under the following directory:  
`/etc/vx/vxdbed/$ORACLE_SID/  
checkpoint_dir`
- See the `dbed_ckptcreate(1M)` manual page for more information.

## To create Storage Checkpoints while the database is online

- ◆ Use the `dbed_ckptcreate` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptcreate -S PROD \
```

```
-H /oracle/product/10g -o online
Storage Checkpoint Checkpoint_971672042 created.
```

### To create Storage Checkpoints without updating the repository while the database is online

- ◆ Use the `dbed_ckptcreate` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptcreate -S PROD \
-H /oracle/product/10g -o online -n
Storage Checkpoint Checkpoint_971672043 created.
```

### To create Storage Checkpoints while the database is offline

- ◆ Use the `dbed_ckptcreate` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptcreate -S PROD \
-H /oracle/product/10g -o offline
Storage Checkpoint Checkpoint_971672044 created.
The default option is online.
```

### To create an instant Storage Checkpoints

- ◆ Ensure that the database is online and use the `dbed_ckptcreate` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptcreate -S PROD \
-H /oracle/product/10g -o instant
Storage Checkpoint Checkpoint_971672045 created.
```

### To assign a Storage Checkpoint allocation policy to a Storage Checkpoint

- ◆ Use the `dbed_ckptcreate` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptcreate -S PROD \
-H /oracle/product/10g -o online -p ckpt_data,ckpt_metadata
Storage Checkpoint Checkpoint_971672044 created.
```

## Displaying Storage Checkpoints using `dbed_ckptdisplay`

You can use the `dbed_ckptdisplay` command to display the Storage Checkpoints associated with an Oracle database from the command line.

You can also use it to display fileset quota values.

Before displaying Storage Checkpoints, the following conditions must be met:

- Prerequisites
  - You must be logged in as the database administrator (typically, the user ID `oracle`).

## Usage Notes

- In addition to displaying the Storage Checkpoints created by SF Oracle RAC, `dbed_ckptdisplay` also displays other Storage Checkpoints (for example, Storage Checkpoints created by the Capacity Planning Utility and NetBackup).
- The Status field identifies if the Storage Checkpoint is partial (P), complete (C), invalid (I), mounted (M), read-only (R), writable (W), or of type online (ON), offline (OF), instant (IN), or unknown (UN).
- Database FlashSnap commands are integrated with Storage Checkpoint functionality. It is possible to display and mount Storage Checkpoints carried over with snapshot volumes to other nodes. However limitations apply.  
See [“Mounting the snapshot volumes and backing up”](#) on page 311.
- See the `dbed_ckptdisplay(1M)` manual page for more information.

### To display Storage Checkpoints created by Veritas Storage Foundation for Oracle

- ◆ Use the `dbed_ckptdisplay` command as follows to display information for Storage Checkpoints created by SF Oracle RAC:

```
$ /opt/VRTS/bin/dbed_ckptdisplay -S PROD \
-H /oracle/product/10g
```

Storage Checkpoint	Creation Time	Status
Checkpoint_975876659	Sun Apr 3 12:50:59 2005	P+R+IN
Checkpoint_974424522_wr001	Thu May 16 17:28:42 2005	C+R+ON
Checkpoint_974424522	Thu May 16 17:28:42 2004	P+R+ON

### To display other Storage Checkpoints

- ◆ Use the `dbed_ckptdisplay` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptdisplay -S PROD \
-H /oracle/product/10g -o other
```

Storage Checkpoint	Creator	Location
NetBackup_incr_PROD_955133480	NBU	/db01
NetBackup_full_PROD_9551329 52	NBU	/db01

### To display other Storage Checkpoints without updating the repository

- ◆ Use the `dbed_ckptdisplay` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptdisplay -S PROD \
-H /oracle/product/10g -o other -n
```

Storage Checkpoint	Creator	Location
NetBackup_incr_PROD_955133480	NBU	/db01
NetBackup_full_PROD_9551329 52	NBU	/db01

### To display all Storage Checkpoints

- ◆ Use the `dbed_ckptdisplay` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptdisplay -S PROD \
-H /oracle/product/10g -o all
```

Storage Checkpoint	Creation Time	Status
Checkpoint_971672042	Sun May 15 13:55:53 2005	C+R+IN
Checkpoint_903937870	Fri May 13 22:51:10 2005	C+R+ON
Checkpoint_901426272	Wed May 11 16:17:52 2005	P+R+ON

Storage Checkpoint	Creator	Location
NetBackup_incr_PROD_955133480	NBU	/db01
NetBackup_full_PROD_9551329 52	NBU	/db01

### To display all Storage Checkpoints without updating the repository

- ◆ Use the `dbed_ckptdisplay` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptdisplay -S PROD \
```

```
-H /oracle/product/10g -o all -n
```

```
Storage Checkpoint Creation Time Status

Checkpoint_971672042 Sun May 15 13:55:53 2005 C+R+IN
Checkpoint_903937870 Fri May 13 22:51:10 2005 C+R+ON
Checkpoint_901426272 Wed May 11 16:17:52 2005 P+R+ON
Storage Checkpoint Creator Location

NetBackup_incr_PROD_955133480 NBU /db01
NetBackup_full_PROD_9551329 52 NBU /db01
```

### To display fileset quota values

- ◆ Use the `dbed_ckptdisplay` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptdisplay -S PROD -c \
Checkpoint_903937870 -Q
Checkpoint_903937870 Wed Mar 19 9:12:20 2005 C+R+ON
Filesystem HardLim SoftLim CurrentUse
/oradata1/indx1_1 100000 50000 2028
/oradata1/user1_1 100000 50000 2028
/oradata1/temp 150000 80000 2142
/oradata1/system1 150000 70000 3092
```

### Scheduling Storage Checkpoints using `dbed_ckptcreate` and `cron`

You can use the `dbed_ckptcreate` command to schedule Storage Checkpoint creation in a `cron` job or other administrative script.

Before scheduling Storage Checkpoints, the following conditions must be met:

- |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prerequisites | ■ You must be logged on as the database administrator (typically, the user ID oracle).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Usage notes   | <ul style="list-style-type: none"> <li>■ Create a new crontab file or edit an existing crontab file to include a Storage Checkpoint creation entry with the following space-delimited fields:<br/> <i>minute hour day_of_month month_of_year day_of_week \</i><br/> <i>/opt/VRTS/bin/dbed_ckptcreate</i><br/>           where:<br/> <i>minute</i> - numeric values from 0-59 or *<br/> <i>hour</i> - numeric values from 0-23 or *<br/> <i>day_of_month</i> - numeric values from 1-31 or *<br/> <i>month_of_year</i> - numeric values from 1-12 or *<br/> <i>day_of_week</i> - numeric values from 0-6, with 0=Sunday or *<br/>           Each of these variables can either be an asterisk (meaning all legal values) or a list of elements separated by commas. An element is either a number or two numbers separated by a hyphen (meaning an inclusive range).</li> <li>■ See the <code>dbed_ckptcreate(1M)</code>, <code>cron(1M)</code>, and <code>crontab(1)</code> manual pages for more information.</li> </ul> |

### Scheduling Storage Checkpoint creation in a cron job

- To create a Storage Checkpoint twice a day, at 5:00 a.m. and 7:00 p.m., every Monday through Friday, include the following entry in your crontab file:

```
0 5,19 * * 1-5 /opt/VRTS/bin/dbed_ckptcreate -S PROD \
-H /oracle/product/10g -o instant
```

- To create a Storage Checkpoint at 11:30 p.m., on the 1st and 15th day of each month, include the following entry in your crontab file:

```
30 23 1,15 * * /opt/VRTS/bin/dbed_ckptcreate -S PROD \
-H /oracle/product/10g -o instant
```

- To create a Storage Checkpoint at 1:00 a.m. every Sunday while the database is offline, include the following entry in your crontab file:

```
0 1 * * 0 /opt/VRTS/bin/dbed_ckptcreate -S PROD \
-H /oracle/product/10g -o offline
```

## Mounting Storage Checkpoints using dbed\_ckptmount

You can use the `dbed_ckptmount` command to mount a Storage Checkpoint for the database from the command line.

Before mounting Storage Checkpoints, the following conditions must be met:

- |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prerequisites | <ul style="list-style-type: none"> <li>■ You must be logged on as the database administrator (typically, the user ID oracle).</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Usage notes   | <ul style="list-style-type: none"> <li>■ The <code>dbed_ckptmount</code> command is used to mount a Storage Checkpoint into the file system namespace. Mounted Storage Checkpoints appear as any other file system on the machine and can be accessed using all normal file system based commands.</li> <li>■ Storage Checkpoints can be mounted as read-only or read-write. By default, Storage Checkpoints are mounted as read-only.</li> <li>■ If the <code>rw</code> (read-write) option is used, <code>_wrxxx</code>, where <code>xxx</code> is an integer, will be appended to the Storage Checkpoint name.</li> <li>■ If the specified mount point directory does not exist, then <code>dbed_ckptmount</code> creates it before mounting the Storage Checkpoint, as long as the Oracle database owner has permission to create it.</li> <li>■ Database FlashSnap commands are integrated with Storage Checkpoint functionality. It is possible to display and mount Storage Checkpoints carried over with snapshot volumes to other nodes. However limitations apply. See <a href="#">“Mounting the snapshot volumes and backing up”</a> on page 311.</li> <li>■ See the <code>dbed_ckptmount(1M)</code> manual page for more information.</li> </ul> |

### To mount Storage Checkpoints with the read/write option

- ◆ Use the `dbed_ckptmount` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptmount -S PROD -c Checkpoint_971672042 \
-m /tmp/ckpt_rw -o rw
```

```
Creating Storage Checkpoint on /tmp/ckpt_rw/share/oradata with
name Checkpoint_971672042_wr001
```

#### To mount Storage Checkpoints with the read-only option

- ◆ Use the `dbed_ckptmount` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptmount -S PROD -c Checkpoint_971672042 \
-m /tmp/ckpt_ro -o ro
```

## Unmounting Storage Checkpoints using `dbed_ckptumount`

You can use the `dbed_ckptumount` command to unmount a Storage Checkpoint for an Oracle database from the command line.

Before unmounting Storage Checkpoints, the following conditions must be met:

- |               |                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prerequisites | ■ You must be logged on as the database administrator (typically, the user ID oracle).                                                                                                                                                                                                                                                                                                                                           |
| Usage notes   | ■ The <code>dbed_ckptumount</code> command is used to unmount a mounted Storage Checkpoint from the file system namespace. Mounted Storage Checkpoints appear as any other file system on the machine and can be accessed using all normal file system based commands. When mounted Storage Checkpoints are not required, they can be unmounted.<br>■ See the <code>dbed_ckptumount(1M)</code> manual page for more information. |

#### To unmount Storage Checkpoints

- ◆ Use the `dbed_ckptumount` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptumount -S PROD \
-c Checkpoint_971672042_wr001
```

## Performing Storage Rollback using `dbed_ckptrollback`

You can use the `dbed_ckptrollback` command to rollback an Oracle database to a Storage Checkpoint.

Before performing a Storage Rollback, the following conditions must be met:

- |               |                                                                                        |
|---------------|----------------------------------------------------------------------------------------|
| Prerequisites | ■ You must be logged on as the database administrator (typically, the user ID oracle). |
|---------------|----------------------------------------------------------------------------------------|

- Usage notes
- The `dbed_ckptrollback` command rolls an Oracle database back to a specified Storage Checkpoint. You can perform a Storage Rollback for the entire database, a specific tablespace, or list of datafiles. Database rollback for the entire database requires that the database be inactive before Storage Rollback commences. The `dbed_ckptrollback` command will not commence if the Oracle database is active. However, to perform a Storage Rollback of a tablespace or datafile, only the tablespace or datafile to be rolled back must be offline (not the entire database).
  - You must run the `dbed_update` command after upgrading to Storage Foundation 5.0 for Oracle RAC from a previous release. This will allow you to roll back to a Storage Checkpoint that was created with an earlier version of this product.
  - See the `dbed_ckptrollback(1M)` manual page for more information.

### To roll back an Oracle database to a Storage Checkpoint

- ◆ Use the `dbed_ckptrollback` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptrollback -S PROD \
-H /oracle/product/10g -c Checkpoint_903937870
```

### To rollback a tablespace to a Storage Checkpoint

- ◆ Use the `dbed_ckptrollback` command with the `-T` option as follows:

```
$ /opt/VRTS/bin/dbed_ckptrollback -S PROD \
-H /oracle/product/10g -T DATA01 -c Checkpoint_903937870
```

If the Oracle database is running, you must take the tablespace offline before running this command. If the tablespace is online, the command will fail.

In the case of an instant Storage Checkpoint, rolling back a tablespace does not apply.

### To rollback datafiles to a Storage Checkpoint

- ◆ Use the `dbed_ckptrollback` command with the `-F` option as follows:

```
$ /opt/VRTS/bin/dbed_ckptrollback -S PROD \
-H /oracle/product/10g -F /share/oradata1/data01.dbf \
/share/oradata2/index01.dbf -c Checkpoint_903937870
```

If the Oracle database is running, you must take the datafile offline before running this command. If the datafile is online, the command will fail.

In the case of an instant Storage Checkpoint, rolling back datafiles does not apply.

## Removing Storage Checkpoints using `dbed_ckptremove`

You can use the `dbed_ckptremove` command to remove a Storage Checkpoint for an Oracle database at the command line.

Before removing Storage Checkpoints, the following conditions must be met:

- |               |                                                                                                                                                                                                                                                                                                                                                                                                          |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prerequisites | ■ You must be logged on as the database administrator (typically, the user ID oracle).                                                                                                                                                                                                                                                                                                                   |
| Usage notes   | ■ The <code>dbed_ckptremove</code> command is used to remove a Storage Checkpoint from the file system, or file systems, it is associated with. The Storage Checkpoint must have been created using the <code>dbed_ckptcreate(1M)</code> command.<br>■ You must unmount the Storage Checkpoint before you can remove it.<br>■ See the <code>dbed_ckptremove(1M)</code> manual page for more information. |

### To remove Storage Checkpoints

- ◆ Use the `dbed_ckptremove` command as follows:

```
$ /opt/VRTS/bin/dbed_ckptremove -S PROD \
-c Checkpoint_971672042_wr001
```

## Cloning the Oracle instance using `dbed_clonedb`

You can use the `dbed_clonedb` command to clone an Oracle instance using a Storage Checkpoint.

Cloning an existing database using a Storage Checkpoint must be done on the same host.

You have the option to manually or automatically recover the database when using the `dbed_clonedb` command:

- Manual (interactive) recovery, which requires using the `-i` option, of the clone database allows the user to control the degree of recovery by specifying which archive log files are to be replayed.
- Automatic (non-interactive) recovery, which is the default usage of the command, recovers the entire database and replays all of the archive logs. You will not be prompted for any archive log names.

Before cloning the Oracle instance, the following conditions must be met:

- |                               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prerequisites                 | <ul style="list-style-type: none"> <li>■ You must first create a Storage Checkpoint. See <a href="#">“Creating Storage Checkpoints using dbed_ckptcreate”</a> on page 257.</li> <li>■ You must be logged in as the database administrator.</li> <li>■ Make sure you have enough space and system resources to create a clone database on your system.</li> <li>■ A clone database takes up as much memory and machine resources as the primary database.</li> </ul>                                          |
| Usage notes                   | <ul style="list-style-type: none"> <li>■ The <code>dbed_clonedb</code> command is used to create a copy of a database, cloning all existing database files to new locations.</li> <li>■ The <code>ORACLE_SID</code> and <code>ORACLE_HOME</code> environment variables must be set to the primary database.</li> <li>■ It is assumed that the user has a basic understanding of the database recovery process.</li> <li>■ See the <code>dbed_clonedb(1M)</code> manual page for more information.</li> </ul> |
| Limitations for SF Oracle RAC | <ul style="list-style-type: none"> <li>■ Note that the database cloning using Instant Checkpoint is not supported for SF Oracle RAC.</li> <li>■ When you clone the database by using Checkpoint, the node can be any node in the same SF Oracle RAC cluster but the archive log destination is required to be on CFS file system. Otherwise, you must manually copy the archive log files.</li> </ul>                                                                                                        |

[Table 15-3](#) lists the options for the `dbed_clonedb` command.

**Table 15-3** `dbed_clonedb` command options

Option	Description
<code>-S CLONE_SID</code>	Specifies the name of the new Oracle SID, which will be the name of the new database instance.
<code>-m MOUNT_POINT</code>	Indicates the new mount point of the Storage Checkpoint.
<code>-c CKPT_NAME</code>	Indicates the name of the Storage Checkpoint.
<code>-i</code>	Runs the command in interactive mode where you must respond to prompts by the system. The default mode is non-interactive. (Optional)
<code>-o unmount</code>	Shuts down the clone database and unmounts the Storage Checkpoint file system.

**Table 15-3** dbed\_clonedb command options

Option	Description
-o restartdb	Mounts the Storage Checkpoint file system and starts the clone database. The -o restartdb option will not attempt to recover the clone database.
-d	Used with the -o umount option. If the -d option is specified, the Storage Checkpoint used to create the clone database will be removed along with the clone database.

**To clone an Oracle instance with manual Oracle recovery**

- ◆ Use the dbed\_clonedb command as follows:

```
$ /opt/VRTS/bin/dbed_clonedb -S NEW9 -m /local/oracle10g/1 \
-c Checkpoint_988813047 -i
Primary Oracle SID is TEST10g
New Oracle SID is NEW9
Checkpoint_988813047 not mounted at /local/oracle10g/1
Mounting Checkpoint_988813047 at /local/oracle9/1
Using environment-specified parameter file
/local/oracle9/links/dbs/initTEST10g.ora
Default Oracle parameter file found:
/local/oracle9/links/dbs/initTEST10g.ora
Copying /local/oracle9/links/dbs/initTEST10g.ora to
/local/oracle9/1/testvol
Control file 'ora_control2' path not explicitly specified in
init file; assuming ORACLE_HOME/dbs

All redo-log files found
Copying initTEST10g.ora to initNEW9.ora
in /local/oracle9/1/testvol
Altering db_name in initNEW9.ora
Altering control file locations in initNEW9.ora
Creating new link for clone database init file
Creating archive log directory

About to start up new database and begin reconfiguration

Database NEW9 is being reconfigured
Altering clone database archive log directory
Updating log_archive_dest in clone database init file
Found archive log destination at /testvol

The latest archive log(s) must now be applied. To apply the
logs, open a new window and perform the following steps:

1. copy required archive log(s) from primary to clone:
primary archive logs in /testvol
clone archive logs expected in /local/oracle9/1/testvol
```

```
2. ORACLE_SID=NEW9; export ORACLE_SID # sh and ksh, OR
 setenv ORACLE_SID NEW9 #csh
3. /local/oracle9/links/bin/sqlplus /nolog
4. CONNECT / AS SYSDBA
5. RECOVER DATABASE UNTIL CANCEL USING BACKUP CONTROLFILE
6. enter the archive log(s) you wish to apply
7. EXIT
```

Press <Return> after you have completed the above steps.  
<Return>

Resetting logs on new database NEW9  
Database instance NEW9 is up and running

### To clone an Oracle instance with automatic Oracle recovery

- ◆ Use the `dbed_clonedb` command as follows:

```
$ /opt/VRTS/bin/dbed_clonedb -S NEW9 -m /local/oracle9/1 \
-c Checkpoint_988813047
Primary Oracle SID is TEST10g
New Oracle SID is NEW9
Checkpoint_988813047 not mounted at /local/oracle9/1
Mounting Checkpoint_988813047 at /local/oracle9/1
Using environment-specified parameter file
 /local/oracle9/links/dbs/initTEST10g.ora
Default Oracle parameter file found:
 /local/oracle9/links/dbs/initTEST10g.ora
Copying /local/oracle9/links/dbs/initTEST10g.ora
 to /local/oracle9/1/testvol
Control file 'ora_control2' path not explicitly specified in
init file; assuming ORACLE_HOME/dbs

All redo-log files found
Copying initTEST10g.ora to initNEW9.ora
 in /local/oracle10g/1/testvol
Altering db_name in initNEW10g.ora
Altering control file locations in initNEW9.ora
Creating new link for clone database init file
Creating archive log directory

About to start up new database and begin reconfiguration
Database NEW9 is being reconfigured
Starting automatic (full) database recovery
Shutting down clone database
Altering clone database archive log directory
Updating log_archive_dest in clone database init file
Found archive log destination at /testvol
Mounting clone database
Resetting logs on new database NEW10g
Database instance NEW9 is up and running
```

**To shut down the clone database and unmount the Storage Checkpoint**

- ◆ Use the `dbed_clonedb` command as follows:  
`$ /opt/VRTS/bin/dbed_clonedb -S NEW9 -o umount`

**To mount a Storage Checkpoint file system and start the clone database**

- ◆ Use the `dbed_clonedb` command as follows:  
`$ /opt/VRTS/bin/dbed_clonedb -S NEW9 -o restartdb`  
Database instance NEW9 is up and running.

**To delete a clone database and the Storage Checkpoint used to create it**

- ◆ Use the `dbed_clonedb` command as follows:  
`$ /opt/VRTS/bin/dbed_clonedb -S NEW9 -o umount -d`

# Using FlashSnap for backup and recovery

This chapter contains the following topics:

- [About Veritas Database FlashSnap](#)
- [Planning to use Database FlashSnap](#)
- [Preparing hosts and storage for Database FlashSnap](#)
- [Summary of database snapshot steps](#)
- [Creating a snapplan \(dbed\\_vmchecksnap\)](#)
- [Validating a snapplan \(dbed\\_vmchecksnap\)](#)
- [Displaying, copying, and removing a snapplan \(dbed\\_vmchecksnap\)](#)
- [Creating a snapshot \(dbed\\_vmsnap\)](#)
- [Backing up the database from snapshot volumes \(dbed\\_vmclonedb\)](#)
- [Cloning a database \(dbed\\_vmclonedb\)](#)
- [Resynchronizing the snapshot to your database](#)
- [Removing a snapshot volume](#)

## About Veritas Database FlashSnap

Veritas Database FlashSnap helps to create a point-in-time copy of a database for backup and off-host processing. Database FlashSnap allows you to make backup copies of your volumes online with minimal interruption to users.

Veritas Database FlashSnap is included with Storage Foundation for Oracle RAC.

Database FlashSnap lets you capture an online image of an actively changing database at a given instant, known as a snapshot. You can perform backups and off-host processing tasks on snapshots while providing continuous availability of your critical data. Database FlashSnap commands can be executed from the command-line.

Veritas Database FlashSnap offers you a flexible way to efficiently manage multiple point-in-time copies of your data, and reduce resource contention on your business-critical servers.

Database FlashSnap allows database administrators to create a consistent copy of a database without root privileges by creating a snapshot. A snapshot copy of the database is referred to as a *database snapshot*.

In Veritas Storage Foundation for Oracle RAC, you can use Database FlashSnap on the same cluster that the database resides on. Database FlashSnap does not support reverse resynchronization feature for SF Oracle RAC.

Database FlashSnap can significantly reduce the time it takes to backup your database, increase the availability of your production database, and still maintain your production database's performance.

---

**Note:** To use Database FlashSnap, you must have Storage Foundation for Oracle RAC on all systems on which you intend to use Database FlashSnap.

---

To use Database FlashSnap, you must first configure the volumes used by the database.

---

**Note:** For SF Oracle RAC database, when you issue the commands, replace `$ORACLE_SID` with `$ORACLE_SID=instance_name` and provide the instance name on which the instance is running.

---

See [“Preparing hosts and storage for Database FlashSnap”](#) on page 277.

## Solving typical database problems with Database FlashSnap

Database FlashSnap is designed to enable you to use database snapshots to overcome the following types of problems encountered in enterprise database environments:

- In many companies, there is a clear separation between the roles of system administrators and database administrators. Creating database snapshots typically requires superuser (root) privileges, privileges that database administrators do not usually have.

- In some companies, database administrators are granted root privileges, but managing storage is typically not central to their job function or their core competency.
- Creating database snapshots is a complex process, especially in large configurations where thousands of volumes are used for the database. One mistake can render the snapshots useless.

Because it does not require root privileges, Database FlashSnap overcomes these obstacles by enabling database administrators to create consistent snapshots of the database more easily. The snapshots can be utilized for repetitive use.

## About Database FlashSnap applications

The following are typical applications of Veritas Database FlashSnap:

- *Database Backup and Restore:* Enterprises require 24/7 online data availability. They cannot afford the downtime involved in backing up critical data offline. By creating a clone database or a duplicate volume snapshot of data, and then using it to back up your data, your business-critical applications can continue to run without extended down time or impacted performance. After a clone database or snapshot volume is created, it can be used as a source to back up the original database.
- *Decision-Support Analysis and Reporting:* Operations such as decision-support analysis and business reporting may not require access to real-time information. You can direct such operations to use a clone database that you have created from snapshots using Veritas Database FlashSnap, rather than allowing them to compete for access to the primary volume or database. When required, you can quickly resynchronize the clone database with the primary database to get up-to-date information.
- *Application Development and Testing:* Development or service groups can use a clone database created with Database FlashSnap as a test database for new applications. A clone database provides developers, system testers, and quality assurance groups with a realistic basis for testing the robustness, integrity, and performance of new applications.
- *Logical Error Recovery:* Logical errors caused by an administrator or an application program can compromise the integrity of a database. You can recover a database by restoring the database files from a volume snapshot or by recovering logical objects (such as tables, for example) from a clone database created from volume snapshots using Database FlashSnap. These solutions are faster than fully restoring database files from tape or other backup media.

## Using Database FlashSnap

The system administrator needs to configure storage according to the requirements specified in the snapplan.

See [“Preparing hosts and storage for Database FlashSnap”](#) on page 277.

Database FlashSnap allows you to check the storage setup against requirements set forth in the snapplan. Depending on the results, the database administrator may need to modify the snapplan or the system administrator may need to adjust the storage configuration. Properly configuring storage is the only aspect of using Database FlashSnap that requires the system administrator’s participation.

To use Database FlashSnap, a database administrator must first define their snapshot requirements. For example, they need to determine whether off-host processing is required and, if it is, which host should be used for it. In addition, it is also important to consider how much database downtime can be tolerated. Database snapshot requirements are defined in a file called a snapplan. The snapplan specifies snapshot options that will be used when creating a snapshot image (such as whether the snapshot mode will be online, offline, or instant).

See [“Creating a snapplan \(dbed\\_vmchecksnap\)”](#) on page 293.

After creating the snapplan, the database administrator must validate it to ensure that it is correct. During validation the snapplan is copied to the repository before using it to create a snapshot. Depending on the validation results, the database administrator may need to modify the snapplan or the system administrator may need to adjust the storage configuration.

After storage is configured as specified in the snapplan and the snapplan has been validated, the database administrator can create snapshots of the database and create database clones based on the snapshots on either the same host or a secondary one.

A database clone can be used on a secondary host for off-host processing, including decision-support analysis and reporting, application development and testing, database backup, and logical error recovery. After a user has finished using the clone on a secondary host, the database administrator can shut down the clone and move the snapshot database back to the primary host. Regardless of whether a snapshot is used on the primary or secondary host, it can be resynchronized with the primary database using Database FlashSnap. Database FlashSnap utilizes Veritas Volume Manager FastResync to quickly resynchronize the changed section between the primary and snapshot.

See *Veritas Volume Manager User’s Guide* for details about the Volume Manager FastResync.

Database FlashSnap can also be used to recover the primary copy of the database if it becomes corrupted by overwriting it with the snapshot.

## Using Database FlashSnap commands

The Database FlashSnap feature consists of three commands:

- `dbed_vmchecksnap` (used on the master node)  
Creates and validates the snapshot plan used to create a snapshot image of an Oracle database. You can also use `dbed_vmchecksnap` to copy, list, or remove a snapplan or make sure the storage is configured properly for the task. `dbed_vmchecksnap` is also used on the slave node to list the snapplan.
- `dbed_vmsnap` (used on the master node)  
Creates a snapshot image of an Oracle database by splitting the mirror volumes used by the database.

---

**Note:** Snapplan creation and validation is only supported on CVM master node.

---

- `dbed_vmclonedb` (used on the master or slave nodes)  
Mounts and starts a clone database using snapshot volumes. It can also shut down a clone database and deport its volumes, as well as restart a clone database that has been shut down. The snapshot image can be brought up on any of the cluster nodes.

---

**Note:** `dbed_vmclonedb` does not support instant snapshot for database cloning.

---

All of these commands can be executed by the Oracle database administrator and do not require superuser (root) privileges.

---

**Note:** Database FlashSnap operations can be executed from the command-line.

---

## Using Database FlashSnap options

Database FlashSnap offers three options for creating database snapshots. The option you choose is specified in the snapplan.

- *online*  
With this option, the tablespaces are put into online backup mode before the snapshot is created. This type of snapshot is also a valid database backup. Select this option if you are performing a point-in-time recovery from logical errors.
- *instant*

With this option, the database can be up and running, and the tablespaces do not need to be put into online backup mode before the snapshot is created. However, all the file systems used by the database (including those containing the online redo logs and control files) are temporarily frozen and the cache is flushed before the snapshot is created. By freezing the file systems, the snapshot will be a consistent point-in-time image of the database from which a database clone can be created.

An instant snapshot can be used to guard against data corruption or for off-host decision-support queries. However, it is not a valid database backup and cannot be used to perform a point-in-time recovery or off-host backup since tablespaces are not put into online backup mode before the snapshot is created. The instant option is much faster than the online option.

- *offline*

The offline option can be used to clone or back up a database. With this option, the database must be shut down when the snapshot is created and online redo logs are required. This type of snapshot is a valid database backup.

In this release of Veritas Storage Foundation for Oracle RAC, Database FlashSnap supports third mirror break-off snapshots only. Third mirror break-off snapshots are fully synchronized, full-sized snapshots.

See the *Veritas Volume Manager Administrator's Guide* for more information.

## Planning to use Database FlashSnap

Before using Database FlashSnap, you must first determine your intended application. You will then need to make the following decisions:

- Which snapshot mode is appropriate: online, offline, or instant?
- Will you need one or two hosts?

### Selecting the snapshot mode

If your purpose is to use the snapshot for backup or to recover the database after logical errors have occurred, choose the online option. In the event that your production database is offline, choose offline. If you intend to use the snapshot for decision-support analysis, reporting, development, or testing, choose instant. An instant snapshot is not suitable for recovery because it is not necessarily an exact copy of the primary database.

# Preparing hosts and storage for Database FlashSnap

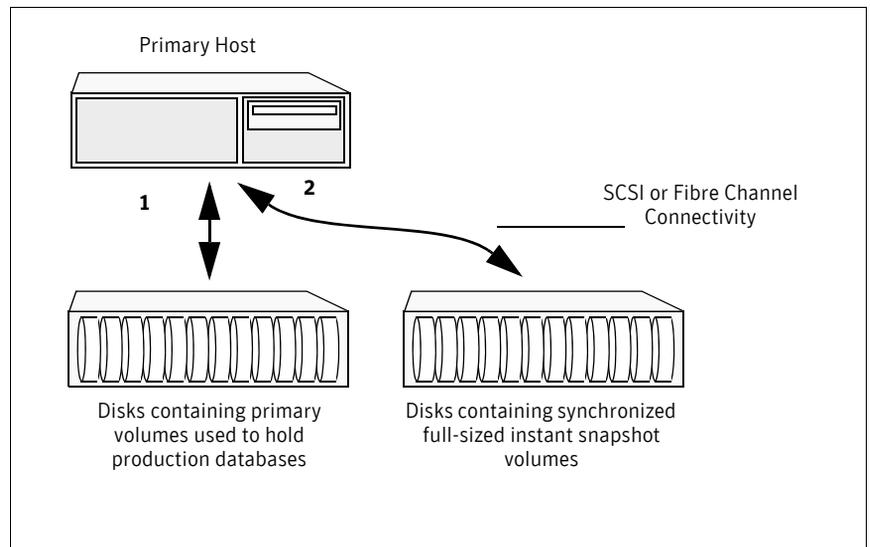
## Setting up hosts

Database FlashSnap requires sufficient Veritas Volume Manager disk space, and can be used on the same host that the database resides on (the primary host) or on a secondary host. Setting up a storage configuration for Database FlashSnap operations is a system administrator's responsibility and requires superuser (root) privileges. Database FlashSnap utilities do not address setting up an appropriate storage configuration.

### Single-host configuration

Figure 16-1 shows the suggested arrangement for implementing Database FlashSnap solutions on the primary host to avoid disk contention.

**Figure 16-1** Example of a Database Flashsnap solution on a primary host

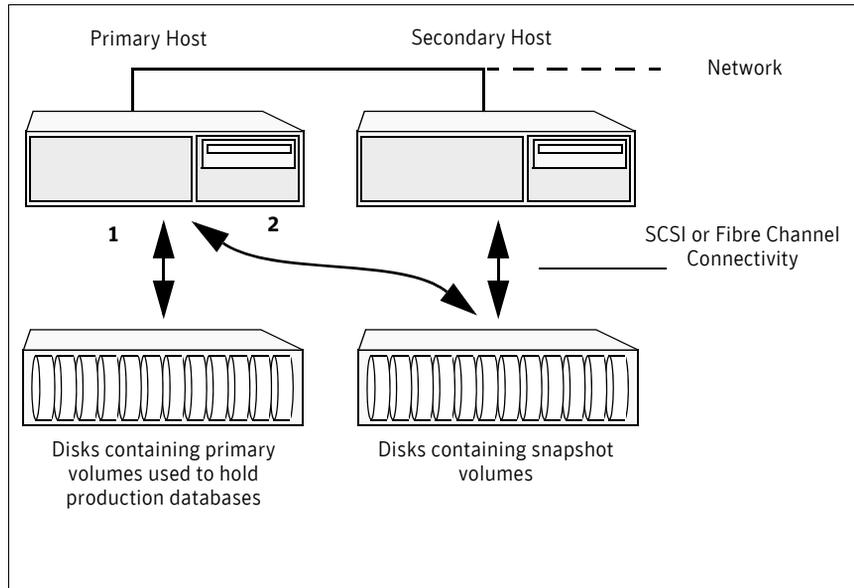


### Two-host configuration

Figure 16-2 illustrates an off-host Database FlashSnap solution. This figure shows a Database FlashSnap configuration with two hosts that allows CPU-intensive and I/O-intensive operations to be performed for online backup and decision support without degrading the performance of the primary host

running the production database. A two-host configuration also allows the snapshot database to avoid contending for I/O resources on the primary host. For off-host processing applications, both the primary and secondary hosts need to share the storage in which the snapshot database is created. Both the primary and secondary hosts must be able to access the disks containing the snapshot volumes.

**Figure 16-2** Example of an off-host Database FlashSnap solution



## Host and storage requirements

Before using Database FlashSnap, ensure that:

- All files are on VxFS file systems over VxVM volumes. Raw devices are not supported.
- Symbolic links to datafiles are not supported.
- ORACLE\_HOME is on a separate file system.
- Archive logs are on a separate VxFS file system and are on a different VxFS file system than Oracle datafiles or ORACLE\_HOME.
- The database does not contain BFILES and external tables.
- Oracle datafiles, archive logs, redo logs, and control files are in a single disk group.

In addition, before attempting to use Database FlashSnap with two hosts, ensure that:

- The versions of Veritas Storage Foundation for Oracle RAC on the primary and secondary hosts are the same.
- The same version of Oracle is installed on both hosts
- The Oracle binaries and datafiles are on different volumes and disks.
- The UNIX login for the database user and group must be the same on both hosts.
- You have a Veritas Storage Foundation for Oracle RAC license on both hosts.

## Creating a snapshot mirror of a volume or volume set used by the database

With Database FlashSnap, you can mirror the volumes used by the database to a separate set of disks, and those mirrors can be used to create a snapshot of the database. These snapshot volumes can be split and placed in a separate disk group. This snapshot disk group can be imported on a separate host, which shares the same storage with the primary host. The snapshot volumes can be resynchronized periodically with the primary volumes to get recent changes of the datafiles. If the primary datafiles become corrupted, you can quickly restore them from the snapshot volumes. Snapshot volumes can be used for a variety of purposes, including backup and recovery, and creating a clone database.

You must create snapshot mirrors for all of the volumes used by the database datafiles before you can create a snapshot of the database. This section describes the procedure used to create snapshot mirrors of volumes.

You can use the `vxsnap` CLI command to create a snapshot mirror.

Prerequisites

- You must be logged in as superuser (root).
- The disk group must be version 110 or later.  
For more information on disk group versions, see the `vxclg (1M)` online manual page.
- Be sure that a data change object (DCO) and a DCO log volume are associated with the volume for which you are creating the snapshot.
- Persistent FastResync must be enabled on the existing database volumes and disks must be assigned for the snapshot volumes.  
FastResync optimizes mirror resynchronization by tracking updates to stored data that have been missed by a mirror. When a snapshot mirror is reattached to its primary volumes, only the updates that were missed need to be re-applied to resynchronize it. FastResync increases the efficiency of the volume snapshot mechanism to better support operations such as backup and decision support. For detailed information about FastResync, see the *Veritas Volume Manager Administrator's Guide*.
- Snapshot mirrors and their associated DCO logs should be on different disks than the original mirror plexes, and should be configured correctly for creating snapshots by the system administrator.
- When creating a snapshot mirror, create the snapshot on a separate controller and separate disks from the primary volume.
- Allocate separate volumes for archive logs.
- Do not place any datafiles, including control files, in the `$ORACLE_HOME/dbs` directory.

## Usage Notes

- Create a separate disk group for Oracle database-related files.
- Do not share volumes between Oracle database files and other software.
- ORACLE\_HOME cannot be included in the snapshot mirror.
- Resynchronization speed varies based on the amount of data changed in both the primary and snapshot volumes during the break-off time.
- Do not share any disks between the original mirror and the snapshot mirror.
- Snapshot mirrors for datafiles and archive logs should be created so that they do not share any disks with the data of the original volumes. If they are not created in this way, the VxVM disk group cannot be split and, as a result, Database FlashSnap will not work.

**Note:** Database FlashSnap commands support third-mirror break-off snapshots only. The snapshot mirror must be in the SNAPDONE state.

The example procedure is for existing volumes without existing snapshot plexes or associated snapshot volumes. In the example procedure, *volume\_name* is the name of either a volume or a volume set.

**To create a snapshot mirror of a volume or volume set**

- 1 To prepare the volume for being snapshot, use the `vxsnap prepare` command:

```
vxsnap -g diskgroup prepare volume \
alloc="storage_attribute ..."
```

The `vxsnap prepare` command automatically creates a DCO and DCO volumes and associates them with the volume, and enables Persistent FastResync on the volume. Persistent FastResync is also set automatically on any snapshots that are generated from a volume on which this feature is enabled.

For enabling persistent FastResync on a volume in VxVM 4.1, either from the command line or from within a script, use the `vxsnap prepare` command as described above.

- 2 To verify that FastResync is enabled on the volume, use the `vxprint` command:

```
vxprint -g diskgroup -F%fastresync volume_name
```

This returns on if FastResync is on. Otherwise, it returns off.

- 3 To verify that a DCO and DCO log volume are attached to the volume, use the `vxprint` command:

```
vxprint -g diskgroup -F%hasdcolog volume_name
```

This returns on if a DCO and DCO log volume are attached to the volume. Otherwise, it returns off.

- 4 Create a mirror of a volume:

```
vxsnap -g diskgroup addmir volume_name alloc= diskname
```

There is no option for creating multiple mirrors at the same time. Only one mirror can be created at a time.

- 5 List the available mirrors:

```
vxprint -g diskgroup -F%name -e"pl_v_name in \"volume_name\""
```

The following two steps enable database FlashSnap to locate the correct mirror plexes when creating snapshots.

- 6 Set the `dbed_flashsnap` tag for the data plex you want to use for breaking off the mirror. You can choose any tag name you like, but it needs to match the tag name specified in the `snapplan`.

```
vxedit -g diskgroup set putil2=dbed_flashsnap plex_name
```

- 7 Verify that the `dbed_flashsnap` tag has been set to the desired data plex:

```
vxprint -g diskgroup -F%name -e"pl_v_name in \"volume_name\" \"
&& p2 in \"dbed_flashsnap\""
```

If you require a backup of the data in the snapshot, use an appropriate utility or operating system command to copy the contents of the snapshot to tape or to some other backup medium.

This example shows the steps involved in creating a snapshot mirror for the volume `data_vol` belonging to the disk group `PRODDg`.

- 1 Prepare the volume `data_vol` for mirroring:

```
vxsnap -g PRODDg prepare data_vol alloc=PRODDg01
```

- 2 Verify that `FastResync` is enabled:

```
vxprint -g PRODDg -F%fastresync data_vol
on
```

- 3 Verify that a DCO and a DCO log are attached to the volume:

```
vxprint -g PRODDg -F%hasdcolog data_vol
on
```

- 4 Create a snapshot mirror of `data_vol`:

```
vxsnap -g PRODDg addmir data_vol alloc=PRODDg02
```

- 5 List the data plexes:

```
vxprint -g PRODDg -F%name -e"pl_v_name in \"data_vol\""
data_vol-01
data_vol-02
```

- 6 Choose the plex that is in the SNAPDONE state. Use the `vxprint -g diskgroup` command to identify the plex that is in the SNAPDONE state.
- 7 Decide which data plex you want to use and set the `dbed_flashsnap` tag for it:

```
vxedit -g PRODDg set putil2=dbed_flashsnap data_vol-02
```

- 8 Verify that the `dbed_flashsnap` tag has been set to the desired data plex, `data_vol-02`:

```
vxprint -g PRODDg -F%name -e"pl_v_name in \"data_vol\" \"
&& p2 in \"dbed_flashsnap\""
data_vol-02
```

- 9 To verify that the snapshot volume was created successfully, use the `vxprint -g dg` command as follows:

```
vxprint -g PRODDg

v data_vol fsgen
ENABLED 4194304 - ACTIVE - -
pl data_vol-01 data_vol
ENABLED 4194304 - ACTIVE - -
sd PRODDg03-01 data_vol-01
ENABLED 4194304 0 - - -
pl data_vol-02 data_vol
ENABLED 4194304 - SNAPDONE - -
sd PRODDg02-01 data_vol-02
ENABLED 4194304 0 - - -
dc data_vol_dco data_vol
- - - - -
v data_vol_dcl gen
ENABLED 560 - ACTIVE - -
pl data_vol_dcl-01 data_vol_dcl ENABLED
560 - ACTIVE - -
sd PRODDg01-01 data_vol_dcl-01 ENABLED
560 0 - - -
pl data_vol_dcl-02 data_vol_dcl DISABLED
560 - DCOSNP - -
sd PRODDg02-02 data_vol_dcl-02 ENABLED
560 0 - - -
```

Identify that the specified plex is in the SNAPDONE state. In this example, it is `data_vol-02`.

The snapshot mirror is now ready to be used.

## Upgrading existing volumes to use Veritas Volume Manager 5.0

The procedure described in this section describes how to upgrade a volume created using a version older than VxVM 5.0 so that it can take advantage of database FlashSnap.

---

**Note:** The plexes of the DCO volume require persistent storage space on disk to be available. To make room for the DCO plexes, you may need to add extra disks to the disk group, or reconfigure existing volumes to free up space in the disk group. Another way to add disk space is to use the disk group move feature to bring in spare disks from a different disk group.

---

---

**Note:** Existing snapshot volumes created by the vxassist command are not supported. A combination of snapshot volumes created by vxassist and vxsnap are not supported.

---

**To upgrade an existing volume created with an earlier version of VxVM:**

- 1 Upgrade the disk group that contains the volume, to a version 120 or higher, before performing the remainder of the procedure described in this section. Use the following command to check the version of a disk group:

```
vxdg list diskgroup
```

To upgrade a disk group to the latest version, use the following command:

```
vxdg upgrade diskgroup
```

- 2 If the volume to be upgraded has a DRL plex or subdisk from an earlier version of VxVM, use the following command to remove this:

```
vxassist [-g diskgroup] remove log volume [nlog=n]
```

Use the optional attribute nlog=n to specify the number, n, of logs to be removed. By default, the vxassist command removes one log.

- 3 For a volume that has one or more associated snapshot volumes, use the following command to reattach and resynchronize each snapshot:

```
vxsnap [-g diskgroup] snapback snapvol
```

If persistent FastResync was enabled on the volume before the snapshot was taken, the data in the snapshot plexes is quickly resynchronized from the original volume. If persistent FastResync was not enabled, a full resynchronization is performed.

- 4 Use the following command to turn off persistent FastResync for the volume:

```
vxvol [-g diskgroup] set fastresync=off volume
```

- 5 Use the following command to dissociate a DCO object from an earlier version of VxVM, DCO volume and snap objects from the volume:

```
vxassist [-g diskgroup] remove log volume logtype=dco
```

- 6 Use the following command on the volume to upgrade it:

```
vxsnap [-g diskgroup] prepare volume
alloc="disk_name1,disk_name2"
```

Provide two disk names to avoid overlapping the storage of the snapshot DCO plex with any other non-moving data or DCO plexes.

The `vxsnap prepare` command automatically enables persistent FastResync on the volume and on any snapshots that are generated from it. It also associates a DCO and DCO log volume with the volume to be snapshot.

- 7 To view the existing DCO plexes and see whether there are enough for the existing data plexes, enter:
 

```
vxprint -g diskgroup
```

 There needs to be one DCO plex for each existing data plex.
- 8 If there are not enough DCO plexes for the existing data plexes, create more DCO plexes:
 

```
vxsnap [-g diskgroup] addmir dco_volume_name [alloc=disk_name]
```

 where `dco_volume_name` is the name of the DCO volume you are creating.
- 9 If the plex is in a SNAPDONE state, convert it to an ACTIVE state:
 

```
vxplex [-g diskgroup] convert state=ACTIVE data_plex
```
- 10 Convert the data plexes to a SNAPDONE state and associate a DCO plex with the data plex that will be used for snapshot operations:
 

```
vxplex [-g diskgroup] -o dco_plex_name convert \ state=SNAPDONE data_plex
```

 where `dco_plex_name` is the name of the DCO plex you are creating.

In this example, the volume, `data_vol`, is upgraded to make use of VxVM 5.0 features.

- 1 Upgrade the disk group, `PRODDg`.
 

```
vxdg upgrade PRODDg
```
- 2 Remove the DRL plexes or subdisks, belonging to an earlier version of VxVM, from the volume to be upgraded.
 

```
vxassist -g PRODDg remove log data_vol logtype=drl
```
- 3 Reattach any snapshot volume back to the primary volume to be upgraded.
 

```
vxsnap -g PRODDg snapback SNAP-data_vol
```
- 4 Turn off FastResync on the volume to be upgraded.
 

```
vxvol -g PRODDg set fastresync=off data_vol
```
- 5 Disassociate and remove any older DCO object and DCO volumes.
 

```
vxassist -g PRODDg remove log data_vol logtype=dco
```
- 6 Upgrade the volume by associating a new DCO object and DCO volume.
 

```
vxsnap -g PRODDg prepare data_vol alloc="PRODDg01 PRODDg02"
```
- 7 View the existing DCO plexes and plex state.

*Scenario 1*

In this scenario, there are enough DCO plexes for the data plexes. Also, no data plex is associated with a DCO plex.

```
vxprint -g PRODDg

v data_vol fsgen
ENABLED 4194304 - ACTIVE - -
pl data_vol-01 data_vol
ENABLED 4194304 - ACTIVE - -
sd PRODDg01-01 data_vol-01
ENABLED 4194304 0 - - -
pl data_vol-04 data_vol
ENABLED 4194304 - SNAPDONE - -
sd PRODDg02-03 data_vol-04
ENABLED 4194304 0 - - -
dc data_vol_dco data_vol
- - - - -

v data_vol_dcl gen
ENABLED 560 - ACTIVE - -
pl data_vol_dcl-01 data_vol_dcl
ENABLED 560 - ACTIVE - -
sd PRODDg01-02 data_vol_dcl-01
ENABLED 560 0 - - -
pl data_vol_dcl-02 data_vol_dcl
ENABLED 560 - ACTIVE - -
sd PRODDg02-02 data_vol_dcl-02
ENABLED 560 0 - - -

■ Convert the data plex state from SNAPDONE to ACTIVE.
vxplex -g PRODDg convert state=ACTIVE data_vol-04

■ Associate the data plex with a new DCO plex and convert it back to a
SNAPDONE state.
vxplex -g PRODDg -o dcoplex=data_vol_dcl-02 \
convert state=SNAPDONE data_vol-04

vxprint -g PRODDg

pl data_vol-03 -
DISABLED 4194304 - - - -
sd PRODDg02-01 data_vol-03
ENABLED 4194304 0 - - -

v data_vol fsgen
ENABLED 4194304 - ACTIVE - -
pl data_vol-01 data_vol
ENABLED 4194304 - ACTIVE - -
sd PRODDg01-01 data_vol-01
ENABLED 4194304 0 - - -
pl data_vol-04 data_vol
ENABLED 4194304 - SNAPDONE - -
sd PRODDg02-03 data_vol-04
ENABLED 4194304 0 - - -
```

```

dc data_vol_dco data_vol
- - - - -
v data_vol_dcl gen
ENABLED 560 - ACTIVE - -
pl data_vol_dcl-01 data_vol_dcl
ENABLED 560 - ACTIVE - -
sd PRODDg01-02 data_vol_dcl-01
ENABLED 560 0 - - -
pl data_vol_dcl-02 data_vol_dcl
DISABLED 560 - DCOSNP - -
sd PRODDg02-02 data_vol_dcl-02
ENABLED 560 0 - - -

```

### Scenario 2

In this scenario, there are fewer DCO plexes than data plexes.

```
vxprint -g PRODDg
```

```

pl data_vol-03 -
DISABLED 4194304 - - - -
sd PRODDg02-01 data_vol-03
ENABLED 4194304 0 - - -
v data_vol fsgen
ENABLED 4194304 - ACTIVE - -
pl data_vol-01 data_vol
ENABLED 4194304 - ACTIVE - -
sd PRODDg01-01 data_vol-01
ENABLED 4194304 0 - - -
pl data_vol-04 data_vol
ENABLED 4194304 - ACTIVE - -
sd PRODDg02-03 data_vol-04
ENABLED 4194304 0 - - -
dc data_vol_dco data_vol
- - - - -
v data_vol_dcl gen
ENABLED 560 - ACTIVE - -
pl data_vol_dcl-01 data_vol_dcl
ENABLED 560 - ACTIVE - -
sd PRODDg01-02 data_vol_dcl-01
ENABLED 560 0 - - -

```

- Add a DCO plex to the DCO volume using the vxassist mirror command.

```
vxsnap -g PRODDg addmir data_vol_dcl alloc=PRODDg02
```

- Associate the data plex with the new DCO plex and convert it to a SNAPDONE state.

```
vxplex -g PRODDg -o dcoplex=data_vol_dcl-02 \
convert state=SNAPDONE data_vol-04
```

```
vxprint -g PRODDg
```

```

pl data_vol-03 -
DISABLED 4194304 - - - -

```

**Summary of database snapshot steps**

v	data_vol	fsgen				
ENABLED	4194304		-	ACTIVE	-	-
pl	data_vol-01	data_vol				
ENABLED	4194304		-	ACTIVE	-	-
sd	PRODdg01-01	data_vol-01				
ENABLED	4194304		0	-	-	-
pl	data_vol-04	data_vol				
ENABLED	4194304		-	SNAPDONE	-	-
sd	PRODdg02-03	data_vol-04				
ENABLED	4194304		0	-	-	-
dc	data_vol_dco	data_vol				
-	-		-	-	-	-
v	data_vol_dcl	gen				
ENABLED	560		-	ACTIVE	-	-
pl	data_vol_dcl-01	data_vol_dcl				
ENABLED	560		-	ACTIVE	-	-
sd	PRODdg01-02	data_vol_dcl-01				
ENABLED	560		0	-	-	-
pl	data_vol_dcl-02	data_vol_dcl				
DISABLED	560		-	DCOSNP	-	-
sd	PRODdg02-02	data_vol_dcl-02				
ENABLED	560		0	-	-	-

## Summary of database snapshot steps

You can use Database FlashSnap commands to create a snapshot of your entire database on the same host or on a different one. Three types of snapshots can be created: online, offline, or instant.

If the `SNAPSHOT_MODE` specified in the `snappan` is set to online, `dbed_vmsnap` first puts the tablespaces to be snapshot into backup mode. After the snapshot is created, the tablespaces are taken out of backup mode, the log files are switched to ensure that the extra redo logs are archived, and a snapshot of the archive logs is created.

If the `SNAPSHOT_MODE` is set to offline, the database must be shut down before the snapshot is created. Online redo logs and control files are required and will be used to ensure a full database recovery.

If the `SNAPSHOT_MODE` is set to instant, tablespaces are not put into and out of backup mode. Online redo logs and control files are required and will be used to ensure a full database recovery.

Both online and offline snapshots provide a valid backup copy of the database. You can use the snapshot as a source for backing up the database or creating a clone database for decision-support purposes. Instant snapshots do not represent a valid backup copy for point-in-time recovery.

The sections that follow explain how to create snapshots of all volumes on a database using the `snappan`. Optionally, you can use the `VxVM` command

(`vxsnap`) to create volume snapshots. However, unlike the Database FlashSnap commands, the `vxsnap` command does not automate disk group content reorganization functions. For more information about the `vxsnap` command, see *Veritas Volume Manager Administrator's Guide*.

---

**Note:** Make sure the volumes used by the database are configured properly before attempting to take a snapshot. This requires superuser (root) privileges.

---

---

**Note:** Anytime you change the structure of the database (for example, by adding or deleting datafiles, converting PFILE to SPFILE, or converting SPFILE to PFILE), you must run `dbed_update`.

---

---

**Note:** Database FlashSnap commands must be run by the Oracle database administrator.

---

### To create a snapshot image of a database

- 1 Create a snapshot mirror of a volume or volume set.  
See [“To create a snapshot mirror of a volume or volume set”](#) on page 281.
- 2 Use the `dbed_vmchecksnap` command to create a snapplan template and check the volume configuration to ensure that it is valid for creating volume snapshots of the database.

The snapplan contains detailed database and volume configuration information that is needed for snapshot creation and resynchronization. You can modify the snapplan template with a text editor.

The `dbed_vmchecksnap` command can also be used to:

List all snapplans associated with a specific ORACLE\_SID

```
dbed_vmchecksnap -o list
```

Remove the snapplan from the SFDB repository

```
dbed_vmchecksnap -o remove -f SNAPPLAN
```

Copy a snapplan from the SFDB repository to your local directory

```
dbed_vmchecksnap -o copy -f SNAPPLAN
```

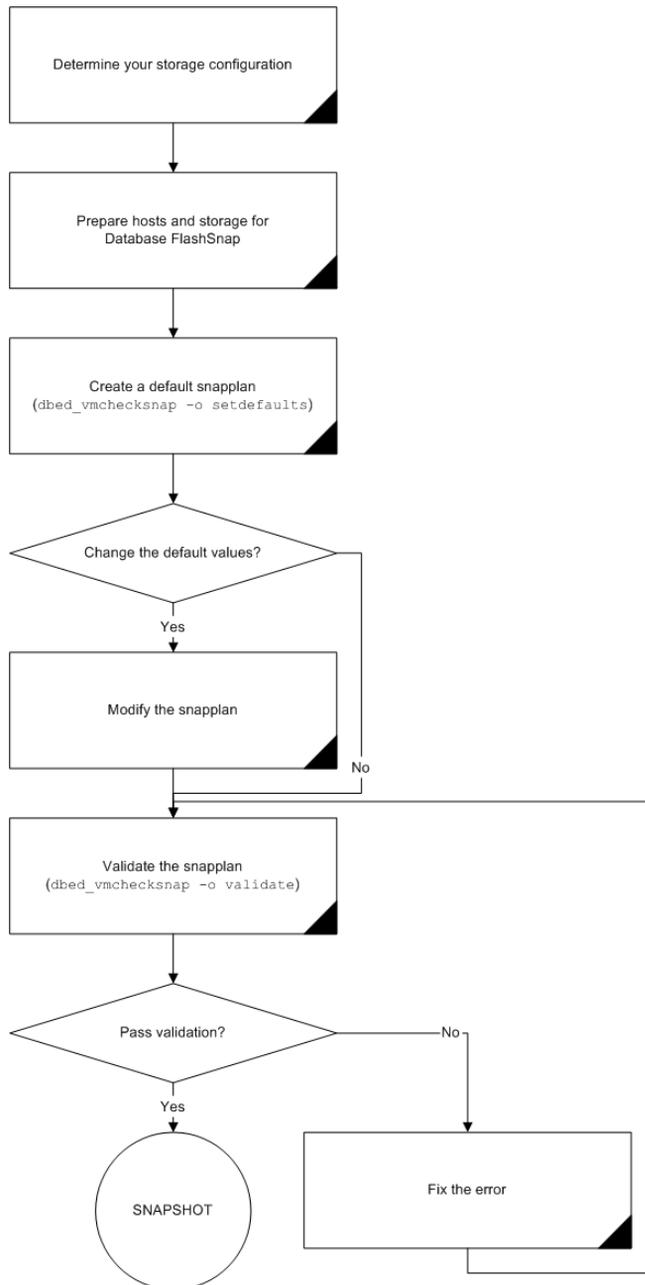
See [“Creating a snapplan \(dbed\\_vmchecksnap\)”](#) on page 293.

- 3 Use the `dbed_vmsnap` command to create snapshot volumes for the database.  
See [“Creating a snapshot \(dbed\\_vmsnap\)”](#) on page 305.

- 4 On the secondary host, use the `dbed_vmclonedb` command to create a clone database using the disk group deported from the primary host. See “Cloning a database (`dbed_vmclonedb`)” on page 313.  
If the primary and secondary hosts specified in the snapplan are different, the `dbed_vmclonedb` command imports the disk group that was deported from the primary host, recovers the snapshot volumes, mounts the file systems, recovers the database, and brings the database online with a different Oracle SID name than the primary host. If the secondary host is different, the database name can be same. You can use the `-o recoverdb` option to let `dbed_vmclonedb` perform an automatic database recovery, or you can use the `-o mountdb` option to perform your own point-in-time recovery and bring up the database manually. For a point-in-time recovery, the snapshot mode must be online.  
You can also create a clone on the primary host. Your snapplan settings specify whether a clone should be created on the primary or secondary host.
- 5 You can now use the clone database to perform database backup and other off-host processing work.
- 6 The clone database can be used to reverse resynchronize the original volume from the data in the snapshot, or can be discarded by rejoining the snapshot volumes with the original volumes (that is, by resynchronizing the snapshot volumes) for future use.

Figure 16-3 depicts the sequence of steps leading up to taking a snapshot using Database FlashSnap.

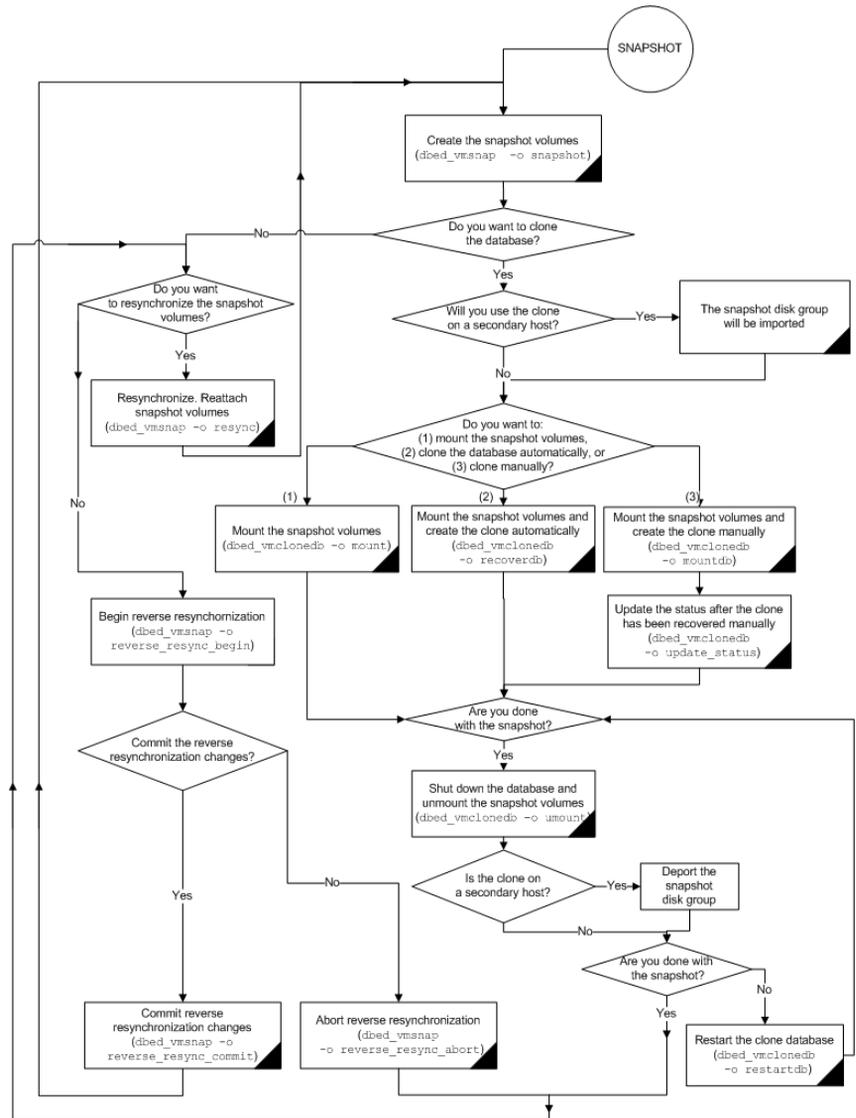
Figure 16-3 Prerequisites for creating a snapshot of your database



There are many actions you can take after creating a snapshot of your database using Database FlashSnap. You can create a clone of the database for backup and off-host processing purposes. You can resynchronize the snapshot volumes with the primary database.

[Figure 16-4](#) is a flow chart, which depicts the actions that you can perform after creating a snapshot of your database using Database FlashSnap.

Figure 16-4 Actions you can perform after creating a snapshot of your database



## Creating a snapplan (dbed\_vmchecksnap)

The `dbed_vmchecksnap` command creates a snapplan that `dbed_vmsnap` uses to create a snapshot of an Oracle database. The snapplan specifies snapshot scenarios (such as online, offline, or instant).

You can name a snapplan file whatever you choose. Each entry in the snapplan file is a line in *parameter=argument* format.

When using `dbed_vmchecksnap` to create or validate a snapplan, the following parameters are set

**Table 16-4** Parameter values for `dbed_vmchecksnap`

Parameter	Value
SNAPSHOT_VERSION	Specifies the snapshot version for this major release of SF Oracle RAC.
PRIMARY_HOST	The name of the host where the primary database resides.
SECONDARY_HOST	The name of the host where the database will be imported.
PRIMARY_DG	The name of the VxVM disk group used by the primary database.
SNAPSHOT_DG	The name of the disk group containing the snapshot volumes.  The snapshot volumes will be put into this disk group on the primary host and deported. The secondary host will import this disk group to start a clone database.
ORACLE_SID	The name of the Oracle database. By default, the name of the Oracle database is included in the snapplan.
ARCHIVELOG_DEST	The full path of the archive logs.  There are several archive log destinations that can be used for database recovery if you are multiplexing the archive logs. You must specify which archive log destination to use.  It is recommended that you have the archive log destination on a separate volume if SNAPSHOT_ARCHIVE_LOG is yes.
SNAPSHOT_ARCHIVE_LOG	yes or no  Specifies whether to create a snapshot of the archive log volumes. Specify yes to split the archive log volume mirrors and deport them to the secondary host. When using the Oracle remote archive log destination feature to send the archive logs to the secondary host, you can specify no to save some space.  Because the archive logs may not always be delivered to the secondary host reliably, it is recommended that you specify yes.

**Table 16-4** Parameter values for dbed\_vmchecksnap

Parameter	Value
SNAPSHOT_MODE	<p>online or offline or instant</p> <p>Specifies whether the database snapshot should be online, offline, or instant.</p> <ul style="list-style-type: none"> <li>■ If the snapshot is created while the database is online, the <code>dbed_vmsnap</code> command will put the tablespaces into backup mode. After <code>dbed_vmsnap</code> finishes creating the snapshot, it will take the tablespaces out of backup mode, switch the log files to ensure that the extra redo logs are archived, and create a snapshot of the archived logs.</li> <li>■ If the database is offline, it is not necessary to put the tablespaces into backup mode. The database must be shut down before creating an offline snapshot.</li> <li>■ If the database snapshot is instant, <code>dbed_vmsnap</code> will skip putting the tablespace into backup mode.</li> </ul> <p><b>Note:</b> If <code>SNAPSHOT_MODE</code> is set to offline or instant, a two-host configuration is required and the <code>-r relocate_path</code> option is not allowed.</p>
SNAPSHOT_PLAN_FOR	<p>The default value is database and cannot be changed.</p> <p>Specifies the database object for which you want to create a snapshot.</p>
SNAPSHOT_PLEX_TAG	<p>Specifies the snapshot plex tag. Use this variable to specify a tag for the plexes to be snapshot. The maximum length of the <code>plex_tag</code> is 15 characters. The default plex tag is <code>dbed_flashsnap</code>.</p>
SNAPSHOT_VOL_PREFIX	<p>Specifies the snapshot volume prefix. Use this variable to specify a prefix for the snapshot volumes split from the primary disk group. A volume name cannot be more than 32 characters. You should consider the length of the volume name when assigning the prefix.</p>
ALLOW_REVERSE_RESYNC	<p>By default, reverse resynchronization is off (set equal to no).</p> <p>SF Oracle RAC does not support reverse resynchronization. So, the value must always be set to no.</p>
SNAPSHOT_MIRROR	<p>Specifies the number of plexes to be snapshot. The default value is 1.</p>

When you first run `dbed_vmchecksnap`, use the `-o setdefaults` option to create a snapplan using default values for variables. You may then edit the file manually to set the variables for different snapshot scenarios.

---

**Note:** You cannot access Database FlashSnap commands (`dbed_vmchecksnap`, `dbed_vmsnap`, and `dbed_vmc1onedb`) with the SFDB menu utility.

---

Before creating a snapplan, make sure the following conditions have been met:

- |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prerequisites | <ul style="list-style-type: none"><li>■ You must have configured the storage for Database Flashsnap.<br/>See <a href="#">“Preparing hosts and storage for Database FlashSnap”</a> on page 277.</li><li>■ You must be the Oracle database administrator.</li><li>■ The disk group must be version 110 or later. For more information on disk group versions, see the <code>vxvdxg (1M)</code> manual page.</li><li>■ Be sure that a DCO and DCO volume are associated with the volume for which you are creating the snapshot.</li><li>■ Snapshot plexes and their associated DCO logs should be on different disks than the original plexes, and should be configured correctly for creating snapshots by the system administrator.</li><li>■ Persistent FastResync must be enabled on the existing database volumes and disks must be assigned for the snapshot volumes.</li><li>■ The database must be running in archive log mode. Archive log mode is set in the Oracle initialization parameter file (<code>init.ora</code>).</li><li>■ The Oracle database must have at least one mandatory archive destination.<br/>See <a href="#">“Establishing a mandatory archive destination”</a> on page 300.</li><li>■ <code>ORACLE_HOME</code> cannot reside on disk which will be used for snapshot.</li></ul> |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

## Usage Notes

- The snapplan must be created on the primary host.
- After creating the snapplan using the `dbed_vmchecksnap` command, you can use a text editor to review and update the file, if necessary.
- It is recommended that you create a local working directory to store your snapplans in.
- See the `dbed_vmchecksnap (1M)` online manual page for more information.
- If the `SNAPSHOT_MODE` for the database is set to online, the primary and secondary hosts can be the same. If the `SNAPSHOT_MODE` is set to offline or instant, the primary and secondary hosts must be different.

## To create a snapplan

- 1 Change directories to the working directory you want to store your snapplan in.

```
$ cd /working_directory
```

- 2 Create a snapplan with default values using the `dbed_vmchecksnap` command:

```
$ /opt/VRTS/bin/dbed_vmchecksnap -S ORACLE_SID \
-H ORACLE_HOME -f SNAPPLAN -o setdefaults -t host_name \
[-p PLEX_TAG]
```

- 3 Open the snapplan file in a text editor and modify it as needed.

In this example, a snapplan, `snap1`, is created for a snapshot image in a single-host configuration and default values are set. The host is named `host1` and the working directory is `/export/snap_dir`.

```
$ cd /export/snap_dir
$ /opt/VRTS/bin/dbed_vmchecksnap -S PROD \
-H /oracle/product/10g -f snap1 -o setdefaults -t host1
Snapplan snap1 for PROD.
=====
SNAPSHOT_VERSION=5.0
PRIMARY_HOST=host1
SECONDARY_HOST=host1
PRIMARY_DG=PRODDG
SNAPSHOT_DG=SNAP_PRODDG
ORACLE_SID=PROD
ARCHIVELOG_DEST=/prod_ar
SNAPSHOT_ARCHIVE_LOG=yes
SNAPSHOT_MODE=online
SNAPSHOT_PLAN_FOR=database
SNAPSHOT_PLEX_TAG=dbed_flashsnap
SNAPSHOT_VOL_PREFIX=SNAP_
ALLOW_REVERSE_RESYNC=no
SNAPSHOT_MIRROR=1
```

In this other example, a snapplan, snap2, is created for a snapshot image in a two-host configuration, and default values are set. The primary host is host1, the secondary host is host2, and the working directory is /export/snap\_dir.

```
$ cd /export/snap_dir
$ /opt/VRTS/bin/dbed_vmchecksnap -S PROD \
-H /oracle/product/10g -f snap2 -o setdefaults -t host2
Snapplan snap2 for PROD.
=====
SNAPSHOT_VERSION=5.0
PRIMARY_HOST=host1
SECONDARY_HOST=host2
PRIMARY_DG=PRODDg
SNAPSHOT_DG=SNAP_PRODDg
ORACLE_SID=PROD
ARCHIVELOG_DEST=/mytest/arch
SNAPSHOT_ARCHIVE_LOG=yes
SNAPSHOT_MODE=online
SNAPSHOT_PLAN_FOR=database
SNAPSHOT_PLEX_TAG=dbed_flashsnap
SNAPSHOT_VOL_PREFIX=SNAP_
ALLOW_REVERSE_RESYNC=no
SNAPSHOT_MIRROR=1
```

By default, a snapplan's SNAPSHOT\_PLEX\_TAG value is set as dbed\_flashsnap. You can use the -p option to assign a different tag name. Make use of the -p option when creating the snapplan with the setdefaults option.

In the following example, the -p option is used with setdefaults to assign my\_tag as the SNAPSHOT\_PLEX\_TAG value.

```
dbed_vmchecksnap -S $ORACLE_SID -H $ORACLE_HOME -O \
setdefaults -p my_tag -f snap1 -t PROD
Snapplan snap1 for PROD
=====
SNAPSHOT_VERSION=5.0
PRIMARY_HOST=host1
SECONDARY_HOST=host2
PRIMARY_DG=PRODDg
SNAPSHOT_DG=SNAP_PRODDg
ORACLE_SID=PROD
ARCHIVELOG_DEST=/arch_data
SNAPSHOT_ARCHIVE_LOG=yes
SNAPSHOT_MODE=online
SNAPSHOT_PLAN_FOR=database
SNAPSHOT_PLEX_TAG=my_tag
SNAPSHOT_VOL_PREFIX=SNAP_
ALLOW_REVERSE_RESYNC=no
SNAPSHOT_MIRROR=1
```

## Creating multi-mirror snapshots

To make Database Snapshots highly available, the snapped snapshot volume should contain more than one mirror. This makes the snapshot volumes available even if one of the mirrors gets disabled. Snapshot volumes can be mounted and the entire database snapshot is usable even if one of the mirror gets disabled. The multi-mirror snapshots are enabled by `SNAPSHOT_MIRROR=<n>` in the snapplan.

---

**Note:** There are no changes to the command-line usage or arguments for the Flashsnap tools.

---



---

**Note:** Before taking the snapshot, make sure all tagged snapshot mirrors are in SNAPDONE state.

---

The following sample explains the setup and the procedure for taking multi-mirror snapshots.

### To set up and take multi-mirror snapshots

- 1 Add the second mirror and DCO log. When allocating storage for the second mirror and DCO logs, make sure the snap volumes are splittable. If snap volumes are not splittable, `dbed_vmchecksnap` fails with appropriate errors.

Tag the newly added mirror with the same tag as that of the first mirror.

Assume that the volume has `fastresync = on`, has `dcolog = on`, and already has one SNAPDONE mirror and is tagged with `dbed_flashsnap`.

```
vxsnap -g dg_a addmir dg_a_vo11 alloc=dg_a03
vxedit -g dg_a set puti12=dbed_flashsnap dg_a_vo11-03
```

- 2 Add `SNAPSHOT_MIRROR` keyword to the snapplan. Here is a sample snapplan.

```
SNAPSHOT_VERSION=5.0
PRIMARY_HOST=host1
SECONDARY_HOST=host1
PRIMARY_DG=PRODdg
SNAPSHOT_DG=SNAP_PRODdg
ORACLE_SID=PROD
ARCHIVELOG_DEST=/prod_ar
SNAPSHOT_ARCHIVE_LOG=yes
SNAPSHOT_MODE=online
SNAPSHOT_PLAN_FOR=database
SNAPSHOT_PLEX_TAG=dbed_flashsnap
SNAPSHOT_VOL_PREFIX=SNAP_
ALLOW_REVERSE_RESYNC=no
SNAPSHOT_MIRROR=2
```

## Establishing a mandatory archive destination

When cloning a database using Database FlashSnap, the Oracle database must have at least one mandatory archive destination.

See “Cloning a database (dbed\_vmclonedb)” on page 313.

If no mandatory archive destination is set, dbed\_vmchecksnap results in this error message:

```
SFORA dbed_vmchecksnap ERROR V-81-5677 Could not find a mandatory,
primary and valid archive destination for database PROD.
```

Please review the LOG\_ARCHIVE\_DEST\_n parameters and check v\$archive\_dest.

This example shows how to establish a mandatory archive destination using SQL\*Plus:

```
alter system set log_archive_dest_1 = 'LOCATION=/ora_mnt/oracle/
oradata/PROD/archivelogs MANDATORY [REOPEN]' [scope=both];
```

For more information about Oracle parameters for archiving redo logs, see your Oracle documentation.

## Validating a snapplan (dbed\_vmchecksnap)

After creating a snapplan, the next steps are to validate the snapplan parameters and check whether the snapshot volumes have been configured correctly for creating snapshots. If validation is successful, the snapplan is copied to the repository. The snapplan is validated using the dbed\_vmchecksnap command with the -o validate option.

Consider the following prerequisites and notes before validating a snapplan:

- Prerequisites
- The database must be up and running while executing the dbed\_vmchecksnap command.

## Usage Notes

- The `dbed_vmchecksnap` command must be run as the Oracle database administrator.
- After validating the snapplan, you have the option of modifying the snapplan file to meet your storage configuration requirements.
- When using `dbed_vmchecksnap` to validate the snapplan and storage, you can save the validation output. The system administrator can use this information to adjust the storage setup if the validation fails.
- If a snapplan is updated or modified, you must revalidate it. It is recommended that snapplans are revalidated when changes are made in the database disk group.
- The `dbed_vmchecksnap` command can be used on the primary or secondary host.
- See the `dbed_vmchecksnap (1M)` manual page for more information.

## To validate a snapplan

- 1 Change directories to the working directory your snapplan is stored in:

```
$ cd /working_directory
```

- 2 Validate the snapplan using the `dbed_vmchecksnap` command:

```
$ /opt/VRTS/bin/dbed_vmchecksnap -S-D ORACLE_SID \
-H ORACLE_HOME -f SNAPPLAN -o validate
```

---

**Note:** In HA environment, you must modify the default snapplan, use the virtual host name defined for the resource group for the PRIMARY\_HOST and/or SECONDARY\_HOST, and run validation.

---

In the following example, a snapplan, `snap1`, is validated for a snapshot image in a single-host configuration. The primary host is `host1` and the working directory is `/export/snap_dir`.

```
$ cd /export/snap_dir
$ /opt/VRTS/bin/dbed_vmchecksnap -S PROD -H \
/oracle/product/10g \
-f snap1 -o validate
```

```
PRIMARY_HOST is host1
```

```
SECONDARY_HOST is host1
```

```
The version of PRIMARY_DG-PRODDg is 110.
```

```
The primary diskgroup PRODDg is a shared disk group
SNAPSHOT_DG is SNAP_PRODDg
```

```
SNAPSHOT_MODE is online
```

```
The database is running in archivelog mode.
```

```
ARCHIVELOG_DEST is /prod_ar
```

```
SNAPSHOT_PLAN_FOR is database
```

```
SNAPSHOT_ARCHIVE_LOG is yes
```

```
ARCHIVELOG_DEST=/prod_ar is mount on /dev/vx/dsk/PRODDg/prod_ar.
```

```
Examining Oracle volume and disk layout for snapshot
```

```
Volume prod_db on PRODDg is ready for snapshot.
Original plex and DCO log for prod_db is on PRODDg01.
Snapshot plex and DCO log for prod_db is on PRODDg02.
```

```
SNAP_PRODDg for snapshot will include: PRODDg02
```

```
ALLOW_REVERSE_RESYNC is no
```

```
The snapplan snap1 has been created.
```

In the following example, a snapplan, snap2, is validated for a snapshot image in a two-host configuration. The primary host is host1, the secondary host is host2, and the working directory is /export/snap\_dir.

```
$ cd /export/snap_dir
$ /opt/VRTS/bin/dbed_vmchecksnap -S PROD -H \
/oracle/product/10g -f snap2 -o validate
PRIMARY_HOST is host1
```

```
SECONDARY_HOST is host2
```

```
The version of PRIMARY_DG-PRODDg is 110.
```

```
The primary diskgroup PRODDg is a shared disk group
SNAPSHOT_DG is SNAP_PRODDg
```

```
SNAPSHOT_MODE is online
```

```
The database is running in archivelog mode.
```

```
ARCHIVELOG_DEST is /mytest/arch
```

```
SNAPSHOT_PLAN_FOR is database
```

```
SNAPSHOT_ARCHIVE_LOG is yes
```

```
ARCHIVELOG_DEST=/mytest/arch is mount on /dev/vx/dsk/PRODDg/
arch.
```

```
Examining Oracle volume and disk layout for snapshot.
```

```
Volume arch on PRODDg is ready for snapshot.
Original plex and DCO log for arch is on PRODDg01.
Snapshot plex and DCO log for arch is on PRODDg02.
```

```
Volume prod_db on PRODDg is ready for snapshot.
Original plex and DCO log for prod_db is on PRODDg01.
Snapshot plex and DCO log for prod_db is on PRODDg04.
```

```
SNAP_PRODDg for snapshot will include: PRODDg02
```

```
ALLOW_REVERSE_RESYNC is no
```

```
The snapplan snap2 has been created.
```

## Displaying, copying, and removing a snapplan (dbed\_vmchecksnap)

Consider these notes before listing all snapplans for a specific Oracle database, displaying a snapplan file, or copying and removing snapplans.

### Usage Notes

- If the local snapplan is updated or modified, you must revalidate it.
- If the database schema or disk group is modified, you must revalidate it after running `dbed_update`.

## Displaying a snapplan

You can use the `dbed_vmchecksnap` command to list all available snapplans, and then use the `dbed_vmchecksnap` command to display detailed information for a particular snapplan.

### To list all available snapplans for a specific Oracle database

- ◆ Use the `dbed_vmchecksnap` command as follows:

```
$ /opt/VRTS/bin/dbed_vmchecksnap -S ORACLE_SID -o list
```

In the following example, all available snapplans are listed for the database PROD.

```
$ /opt/VRTS/bin/dbed_vmchecksnap -S PROD -o list
```

```
The following snapplan(s) are available for PROD:
```

SNAP_PLAN	SNAP_STATUS
DB_STATUS	SNAP_READY
snap1	init_full
init	yes

**Displaying, copying, and removing a snapplan (dbed\_vmchecksnap)**

```

snap2 init_full
init yes

```

---

**Note:** The command output displays all available snapplans, their snapshot status (SNAP\_STATUS), database status (DB\_STATUS), and whether a snapshot may be taken (SNAP\_READY).

For Database FlashSnap status information, see the *Veritas Storage Foundation for Oracle Administrator's Guide*.

---

**To display detailed information for a snapplan**

- ◆ Use the `dbed_vmchecksnap` command as follows:

```

$ /opt/VRTS/bin/dbed_vmchecksnap -S \
 ORACLE_SID -f SNAPPLAN -o list

```

In the following example, the snapplan `snap1` is displayed.

```

$ /opt/VRTS/bin/dbed_vmchecksnap -S PROD -f snap1 -o list
SNAPSHOT_VERSION=5.0
PRIMARY_HOST=host1
SECONDARY_HOST=host1
PRIMARY_DG=PRODdg
SNAPSHOT_DG=SNAP_PRODdg
ORACLE_SID=PROD
ARCHIVELOG_DEST=/prod_ar
SNAPSHOT_ARCHIVE_LOG=yes
SNAPSHOT_MODE=online
SNAPSHOT_PLAN_FOR=database
SNAPSHOT_PLEX_TAG=dbed_flashsnap
SNAPSHOT_VOL_PREFIX=SNAP_
ALLOW_REVERSE_RESYNC=no
SNAPSHOT_MIRROR=1
STORAGE_INFO
PRODdg02
SNAP_PLEX=prod_ar-02

STATUS_INFO
SNAP_STATUS=init_full
DB_STATUS=init

```

## Copying a snapplan

If you want to create a snapplan similar to an existing snapplan, you can simply create a copy of the existing snapplan and modify it. To copy a snapplan from the SFDB repository to your current directory, the snapplan must not already be present in the current directory.

**To copy a snapplan from the SFDB repository to your current directory**

- ◆ Use the `dbed_vmchecksnap` command as follows:

```
$ /opt/VRTS/bin/dbed_vmchecksnap -S ORACLE_SID \
-f SNAPPLAN -o copy
```

In the following example, the snapplan, snap1, is copied from the SFDB repository to the current directory.

```
$ /opt/VRTS/bin/dbed_vmchecksnap -S PROD \
-f snap1 -o copy
Copying 'snap1' to '/export/snap_dir'
```

## Removing a snapplan

A snapplan can be removed from a local directory or repository if the snapplan is no longer needed.

### To remove a snapplan from the SFDB repository

- ◆ Use the `dbed_vmchecksnap` command as follows:

```
$ /opt/VRTS/bin/dbed_vmchecksnap -S ORACLE_SID -f\
SNAPPLAN -o remove
```

In the following example, the snapplan, snap1, is removed from the SFDB repository.

```
$ /opt/VRTS/bin/dbed_vmchecksnap -S PROD -f snap1 -o remove
The snapplan snap1 has been removed.
```

## Creating a snapshot (dbed\_vmsnap)

The `dbed_vmsnap` command creates a snapshot of an Oracle database by splitting the mirror volumes used by the database into a snapshot database. You can use the snapshot image on either the same host as the database or on a secondary host provided storage is shared by the two hosts.

The snapshot image created by `dbed_vmsnap` is a frozen image of an Oracle database's datafiles. The `dbed_vmsnap` command ensures that a backup control file is created when the snapshot database is created, which allows for complete data recovery, if needed.

For Database FlashSnap status information, see the *Veritas Storage Foundation for Oracle Administrator's Guide*.

---

**Note:** You cannot access Database FlashSnap commands (dbed\_vmchecksnap, dbed\_vmsnap, and dbed\_vmc1onedb) with the SFDB menu utility.

---

- |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prerequisites | <ul style="list-style-type: none"><li>■ You must be logged in as the Oracle database administrator.</li><li>■ You must create and validate a snapplan using dbed_vmchecksnap before you can create a snapshot image with dbed_vmsnap.</li></ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Usage Notes   | <ul style="list-style-type: none"><li>■ The dbed_vmsnap command can only be used on the primary host.</li><li>■ Do not share volumes between Oracle database files and other software.</li><li>■ When creating a snapshot volume, create the snapshot on a separate controller and on separate disks from the primary volume.</li><li>■ Make sure your archive log destination is separate from your Oracle database volumes.</li><li>■ Do not place any datafiles, including control files, in the \$ORACLE_HOME/dbs directory.</li><li>■ Resynchronization speed varies based on the amount of data changed in both the primary and secondary volumes when the mirror is broken off.</li><li>■ See the dbed_vmsnap (1M) manual page for more information.</li></ul> |

### To create a snapshot

- 1 Change directories to the working directory in which your snapplan is stored:  

```
$ cd /working_directory
```
- 2 If SNAPSHOT\_MODE is set to offline in the snapplan, shut down the database.
- 3 Create the snapshot image using the dbed\_vmsnap command:  

```
$ /opt/VRTS/bin/dbed_vmsnap -S ORACLE_SID -f SNAPPLAN \
-o snapshot [-F]
```

---

**Note:** To force snapshot creation, use the `-F` option. The `-F` option can be used after a snapshot operation has failed and the problem was fixed without using Veritas Storage Foundation for Oracle commands. (That is, the volumes were synchronized without using Veritas Storage Foundation for Oracle commands.) In this situation, the status of the snapplan will appear as unavailable for creating a snapshot. The `-F` option ignores the unavailable status, checks for the availability of volumes, and creates the snapshot after the volumes pass the availability check.

---

---

**Note:** After the snapshot is created, `dbed_vmsnap` returns values you will need to run `dbed_vmclonedb`. These values include the snapshot disk group, the snapplan name, and the SFDB repository volume for a two-host configuration. Make a note of these values so you have them when running `dbed_vmclonedb`.

---

---

**Note:** You can also use the command `dbed_vmchecksnap -f snapplan -o list` to access the information regarding the snapshot disk group, the snapplan name, and the SFDB repository.

---

The snapshot volumes now represent a consistent backup copy of the database. You can backup the database by copying the snapshot volumes to tape or other backup media.

See “[Backing up the database from snapshot volumes \(dbed\\_vmclonedb\)](#)” on page 308.

You can also create another Oracle database for decision-support purposes.

See “[Cloning a database \(dbed\\_vmclonedb\)](#)” on page 313.

In this example, a snapshot image of the database, PROD, is created for a single-host configuration. In this case, the `SECONDARY_HOST` parameter is set the same as the `PRIMARY_HOST` parameter in the snapplan.

```
$ /opt/VRTS/bin/dbed_vmsnap -S PROD -f snap1 -o snapshot
```

```
dbed_vmsnap started at 2006-03-02 14:15:27
SFDB repository is up to date.
The database is running in archivelog mode.
A snapshot of ORACLE_SID PROD is in DG SNAP_PRODDg.
Snapplan snap1 is used for the snapshot.
```

If `-r <relocate_path>` is used in `dbed_vmclonedb`, make sure `<relocate_path>` is created and owned by Oracle DBA. Otherwise, the following mount points need to be created and owned by Oracle DBA:

```
/prod_db.
/prod_ar.
```

```
dbed_vmsnap ended at 2005-03-02 14:16:11
```

In this example, a snapshot image of the primary database, PROD, is created for a two-host configuration. In this case, the `SECONDARY_HOST` parameter specifies a different host name than the `PRIMARY_HOST` parameter in the snapplan.

```
$ /opt/VRTS/bin/dbed_vmsnap -S PROD -f snap2 -o snapshot
```

```
dbed_vmsnap started at 2005-03-02 23:01:10
SFDB repository is up to date.
The database is running in archivelog mode.
A snapshot of ORACLE_SID PROD is in DG SNAP_PRODDG.
Snapplan snap2 is used for the snapshot.
SFDB repository volume is SNAP_arch.
```

If `-r <relocate_path>` is used in `dbed_vmclonedb`, make sure `<relocate_path>` is created and owned by Oracle DBA. Otherwise, the following mount points need to be created and owned by Oracle DBA:

```
/prod_db.
/prod_ar.
```

```
dbed_vmsnap ended at 2005-03-02 23:02:58
```

## Backing up the database from snapshot volumes (dbed\_vmclonedb)

Snapshots are most commonly used as a source for backing up a database. The advantage of using snapshot volumes is that the backup will not contest the I/O bandwidth of the physical devices. Making the snapshot volumes available on a secondary host will eliminate the extra loads put on processors and I/O adapters by the backup process on the primary host.

A clone database can also serve as a valid backup of the primary database. You can back up the primary database to tape using snapshot volumes.

Figure 16-5 shows a typical configuration when snapshot volumes are located on the primary host.

**Figure 16-5** Example system configuration for database backup on the primary host

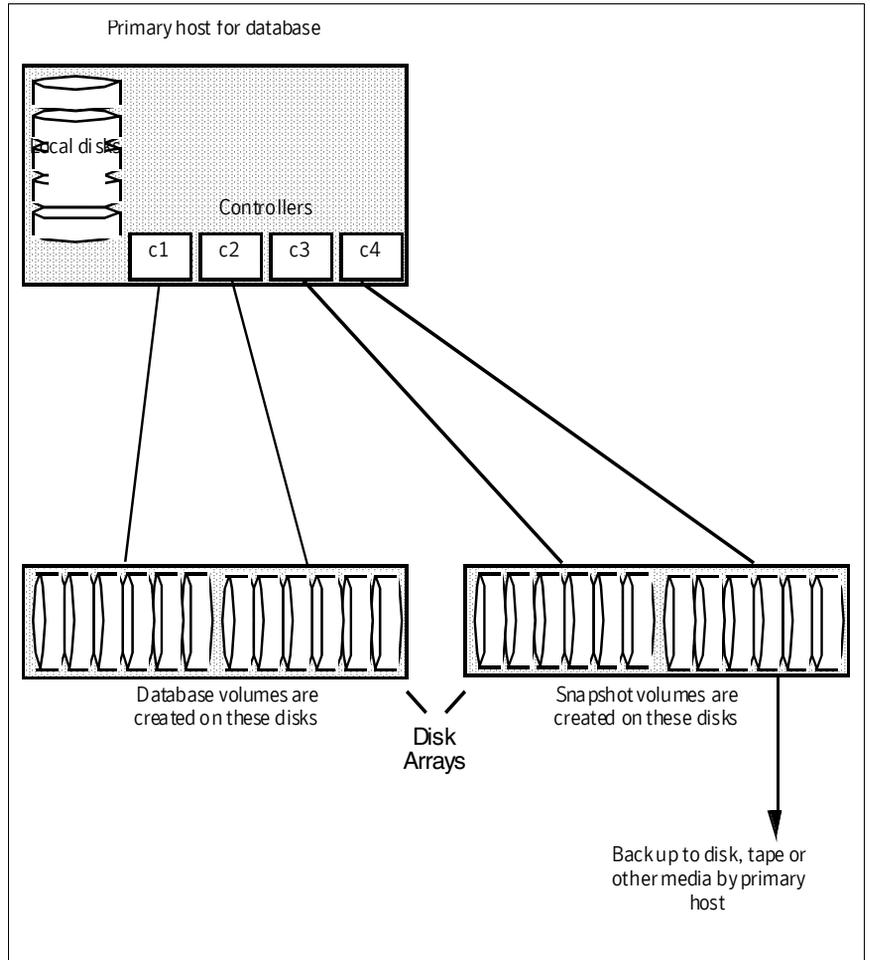
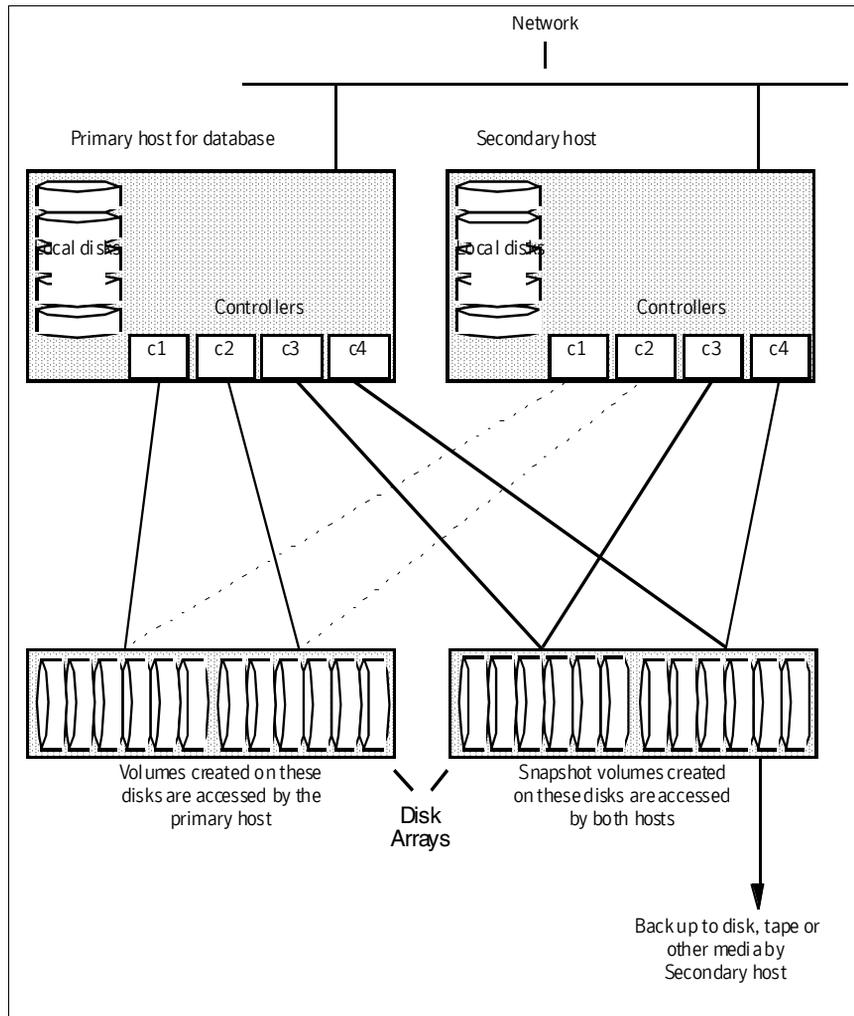


Figure 16-6 shows a typical configuration when snapshot volumes are used on a secondary host.

**Figure 16-6** Example system configuration for database backup on a secondary host



- Prerequisites
- You must be logged in as the Oracle database administrator to use `dbed_vmclonedb` command.
  - Before you can use the `dbed_vmclonedb` command, you must validate a snapplan and create a snapshot. See “[Summary of database snapshot steps](#)” on page 288. See “[Validating a snapplan \(dbed\\_vmchecksnap\)](#)” on page 300. See “[Creating a snapshot \(dbed\\_vmsnap\)](#)” on page 305.
  - The volume snapshot must contain the entire database.
  - Before you can use the `dbed_vmclonedb` command with the `-r relocate_path` option (which specifies the initial mount point for the snapshot image), the system administrator must create the mount point and then change the owner to the Oracle database administrator.
- Usage Notes
- The `dbed_vmclonedb` command can be used on the secondary host.
  - In a single-host configuration, the primary and secondary hosts are the same.
  - In a single-host configuration, `-r relocate_path` is required.
  - In a node in the cluster configuration, the `sfdbvol=vol_name` option is required.
  - If `SNAPSHOT_MODE` is set to `offline` or `instant`, a two-host configuration is required and `-r relocate_path` is not allowed.
  - See the `dbed_vmclonedb (1M)` manual page for more information.

---

**Note:** You cannot access Database FlashSnap commands (`dbed_vmchecksnap`, `dbed_vmsnap`, and `dbed_vmclonedb`) with the SFDB menu utility.

---

## Mounting the snapshot volumes and backing up

Before using the snapshot volumes to do a backup, you must first mount them.

### To mount the snapshot volumes

- ◆ Use the `dbed_vmclonedb` command as follows:

```
$ /opt/VRTS/bin/dbed_vmclonedb -S ORACLE_SID -g snap_dg \
-o mount,new_sid=new_sid,server_name=server_name \
-f SNAPPLAN [-H ORACLE_HOME] [-r relocate_path]
```

You can now back up an individual file or a group of files under a directory onto the backup media.

In this example, snapshot volumes are mounted.

```
$ /opt/VRTS/bin/dbed_vmclonedb -S PROD -g SNAP_PRODDg \
-o mount,new_SID=NEWPROD,server_name=host1 \
-f snap1 -r /clone/single
dbed_vmclonedb started at 2006-03-02 15:35:41
Mounting /clone/single/prod_db on /dev/vx/dsk/SNAP_PRODDg/
SNAP_prod_db.
Mounting /clone/single/prod_ar on /dev/vx/dsk/SNAP_PRODDg/
SNAP_prod_ar.
dbed_vmclonedb ended at 2006-03-02 15:35:50
```

### To mount a Storage Checkpoint carried over from the snapshot volumes to a secondary host

- 1 On the secondary host, list the Storage Checkpoints carried over from the primary database using:

```
$ /opt/VRTS/bin/dbed_ckptdisplay -S ORACLE_SID -n
```

- 2 You can mount one of the listed Storage Checkpoints using:

```
$ /opt/VRTS/bin/dbed_ckptmount -S ORACLE_SID -c CKPT_NAME \
-m MOUNT_POINT
```

Note the following limitations:

- Any mounted Storage Checkpoints must be unmounted before running the following commands:  

```
$ /opt/VRTS/bin/dbed_vmclonedb -o umount,new_sid=new_sid \
-f SNAPPLAN
```
- It is only possible to mount a Storage Checkpoint carried over with the snapshot volumes in a two-host configuration if the snapshot volumes were mounted with the `dbed_vmclonedb` command with the `-o mount` option without the use of `-r relocate_path`.
- Storage Checkpoints carried over with the snapshot volumes can be mounted before a clone database is created using `dbed_vmclonedb` with the `-o mount` option. After a clone database is created using `dbed_vmclonedb` with the `-o recoverdb` option, however, Storage Checkpoints are no longer present.

### To back up the database using the snapshot

- ◆ Copy the snapshot volumes to tape or other appropriate backup media.

---

**Note:** If you use the Oracle online backup method, you must also back up all the archived log files in order to do a complete restore and recovery of the database.

---

## Cloning a database (dbed\_vmclonedb)

Veritas Storage Foundation lets you create a clone database using snapshot volumes. You can use snapshots of a primary database to create a clone of the database at a given point in time. You can then implement decision-support analysis and report generation operations that take their data from the database clone rather than from the primary database to avoid introducing additional burdens on the production database.

A clone database can also serve as a valid backup of the primary database.

See [“Backing up the database from snapshot volumes \(dbed\\_vmclonedb\)”](#) on page 308.

You can also back up the primary database to tape using snapshot volumes.

The resynchronization functionality of Database FlashSnap allows you to quickly refresh the clone database with up-to-date information from the primary database. Reducing the time taken to update decision-support data also lets you generate analysis reports more frequently.

## Using Database FlashSnap to Clone a Database

In a single-host configuration, the `dbed_vmclonedb` command creates a clone database on the same host. The command can also be used to shut down the clone database and unmount its file systems. When creating or unmounting the clone database in a single-host configuration, `-r relocate_path` is required so that the clone database’s file systems use different mount points than those used by the primary database.

When used in a two-host configuration, the `dbed_vmclonedb` command imports the snapshot disk group `SNAP_dg`, mounts the file systems on the snapshot volumes, and starts a clone database. It can also reverse the process by shutting down the clone database, unmounting the file systems, and deporting the snapshot disk group. When creating the clone off host, `-o sfdvvol=vol_name` is required.

---

**Warning:** When creating a clone database, all Storage Checkpoints in the original database are discarded.

---

- |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prerequisites | <ul style="list-style-type: none"><li>■ You must be logged in as the Oracle database administrator.</li><li>■ Before you can use the <code>dbed_vmclonedb</code> command, you must validate a snapplan and create a snapshot. See <a href="#">“Summary of database snapshot steps”</a> on page 288. See <a href="#">“Validating a snapplan (dbed_vmchecksnap)”</a> on page 300. See <a href="#">“Creating a snapshot (dbed_vmsnap)”</a> on page 305.</li><li>■ The volume snapshot must contain the entire database.</li><li>■ The system administrator must provide the database administrator with access to the necessary volumes and mount points.</li><li>■ Before you can use the <code>dbed_vmclonedb</code> command with the <code>-r relocate_path</code> option (which specifies the initial mount point for the snapshot image), the system administrator must create the mount point and then change the owner to the Oracle database administrator.</li><li>■ If <code>SNAPSHOT_MODE</code> is set to offline or instant, a two-host configuration is required and <code>-r relocate_path</code> is not allowed.</li><li>■ The Oracle database must have at least one mandatory archive destination. See <a href="#">“Establishing a mandatory archive destination”</a> on page 300.</li></ul> |
| Usage Notes   | <ul style="list-style-type: none"><li>■ The <code>dbed_vmclonedb</code> command can be used on the secondary host.</li><li>■ In a single-host configuration, <code>-r relocate_path</code> is required. This command is also needed if the name of the clone database is different than the primary database.</li><li>■ The initialization parameters for the clone database are copied from the primary database. This means that the clone database takes up the same memory and machine resources as the primary database. If you want to reduce the memory requirements for the clone database, shut down the clone database and then start it up again using a different <code>init.ora</code> file that has reduced memory requirements. If the host where <code>dbed_vmclonedb</code> is run has little available memory, you may not be able to start up the clone database and the cloning operation may fail.</li><li>■ See the <code>dbed_vmclonedb (1M)</code> manual page for more information.</li></ul>                                                                                                                                                                                                                                                                                      |

### To mount a database and recover it manually

- 1 Start and mount the clone database to allow manual database recovery:
 

```
$ /opt/VRTS/bin/dbed_vmclonedb -S ORACLE_SID -g snap_dg \
-o mountdb,new_sid=new_sid,server_name=server_name\
[,sfdbvol=vol_name] -f SNAPPLAN \
[-H ORACLE_HOME] [-r relocate_path]
```
- 2 Recover the database manually.
- 3 Update the snapshot status information for the clone database in the SFDB repository:
 

```
$ /opt/VRTS/bin/dbed_vmclonedb -o update_status,\
new_sid=new_sid,server_name=server_name -f SNAPPLAN \
[-r relocate_path]
```

### Example: Mounting the file systems without bringing up the clone database

In this example, file systems are mounted without bringing up the clone database. The clone database must be manually created and recovered before it can be used. This example is for a clone created on the same host as the primary database.

```
$ /opt/VRTS/bin/dbed_vmclonedb -S PROD -g SNAP_PRODDg \
-o mountdb,new_sid=NEWPROD,server_name=host1 \
-f snap1 -r /clone
dbed_vmclonedb started at 2006-03-02 15:34:41
Mounting /clone/prod_db on /dev/vx/dsk/SNAP_PRODDg/SNAP_prod_db.
Mounting /clone/prod_ar on /dev/vx/dsk/SNAP_PRODDg/SNAP_prod_ar.
All redo-log files found.
Altering instance_name paramter in initpune.ora.
Altering instance_number paramter in initpune.ora.
Altering thread paramter in initpune.ora.
Starting automatic database recovery.
Database NEWPROD (SID=NEWPROD) is in recovery mode.
If the database NEWPROD is recovered manually, you must run
dbed_vmclonedb -o update_status to change the snapshot status.
dbed_vmclonedb ended at 2006-03-02 15:34:59
```

The database is recovered manually using `dbinitdb` command.

The database status (database\_recovered) needs to be updated for a clone database on the primary host after manual recovery has been completed.

```
$ /opt/VRTS/bin/dbed_vmclonedb -o update_status,\
new_sid=NEWPROD,server_name=host1 \
-f snap1 -r /clone
dbed_vmclonedb started at 2006-03-02 15:35:16
The snapshot status has been updated.
dbed_vmclonedb ended at 2006-03-02 15:35:42
```

## Example: Mounting the file systems without recovering the clone database

In this example, file systems are mounted without recovering the clone database. The clone database must be manually recovered before it can be used. This example is for a clone created on a secondary host.

```
$ /opt/VRTS/bin/dbed_vmclonedb -S -g SNAP_PRODDg \
-o mountdb,new_sid=NEWPROD,vxdbavol=SNAP_arch -f snap2
dbed_vmclonedb started at 2006-03-09 23:26:50
Mounting /clone/arch on /dev/vx/dsk/SNAP_PRODDg/SNAP_arch.
Mounting /clone/prod_db on /dev/vx/dsk/SNAP_PRODDg/SNAP_prod_db.
All redo-log files found.
Altering instance_name paramter in initpune.ora.
Altering instance_number paramter in initpune.ora.
Altering thread paramter in initpune.ora.
Starting automatic database recovery.
Database NEWPROD (SID=NEWPROD) is in recovery mode.
```

If the database NEWPROD is recovered manually, you must run `dbed_vmclonedb -o update_status` to change the snapshot status.

```
dbed_vmclonedb ended at 2006-03-09 23:27:17
```

The database is recovered manually.

The snapshot status (database\_recovered) is updated for a clone database on a secondary host after manual recovery has been completed.

```
$ /opt/VRTS/bin/dbed_vmclonedb -o update_status,new_sid=NEWPROD \
-f snap2
dbed_vmclonedb started at 2006-03-09 23:34:01
The snapshot status has been updated.
dbed_vmclonedb ended at 2006-03-09 23:34:35
```

### To clone the database automatically

- ◆ Use the `dbed_vmclonedb` command as follows:

```
$ /opt/VRTS/bin/dbed_vmclonedb -S ORACLE_SID -g snap_dg \
-o recoverdb,new_sid=new_sid,server_name=server_name\
[,sfdbv1=vol_name] -f SNAPPLAN \
[-H ORACLE_HOME] [-r relocate_path]
```

Where:

<i>ORACLE_SID</i>	Represents the name of the Oracle database used to create the snapshot.
<i>snap_dg</i>	Represents the name of the diskgroup that contains all the snapshot volumes.
<i>new_sid</i>	Specifies the ORACLE_SID for the clone database.
<i>server_name</i>	Specifies the server name
<i>SNAPPLAN</i>	Represents the name of the snapplan file.

<i>ORACLE_HOME</i>	Represents the <i>ORACLE_HOME</i> setting for the <i>ORACLE_SID</i> database.
<i>relocate_path</i>	Represents the name of the initial mount point for the snapshot image.

---

**Note:** When cloning a database on a secondary host, ensure that *PRIMARY\_HOST* and *SECONDARY\_HOST* parameters in the snapplan file are different.

---

When the `-o recoverdb` option is used with `dbed_vmclonedb`, the clone database is recovered automatically using all available archive logs. If the `-o recoverdb` option is not used, you can perform point-in-time recovery manually.

In the following example, a clone of the primary database is automatically created on the same host as the primary database.

```
$ /opt/VRTS/bin/dbed_vmclonedb -S PROD -g SNAP_PRODDg \
-o recoverdb,new_sid=NEWPROD,server_name=host1 -f snap1 -r /clone
dbed_vmclonedb started at 2006-03-02 14:42:10
Mounting /clone/prod_db on /dev/vx/dsk/SNAP_PRODDg/SNAP_prod_db.
Mounting /clone/prod_ar on /dev/vx/dsk/SNAP_PRODDg/SNAP_prod_ar.
All redo-log files found.
Altering instance_name paramter in initpune.ora.
Altering instance_number paramter in initpune.ora.
Altering thread paramter in initpune.ora.
Starting automatic database recovery.
Database NEWPROD (SID=NEWPROD) is running.
dbed_vmclonedb ended at 2006-03-02 14:43:05
```

## Shutting down the clone database and unmounting file systems

When you are done using the clone database, you can shut it down and unmount all snapshot file systems with the `dbed_vmclonedb -o umount` command. If the clone database is used on a secondary host that has shared disks with the primary host, the `-o umount` option also deports the snapshot disk group.

---

**Note:** Any Storage Checkpoints mounted need to be unmounted before running `dbed_vmclonedb -o umount` command.

---

### To shut down the clone database and unmount all snapshot file systems

- ◆ Use the `dbed_vmclonedb` command as follows:

```
$ /opt/VRTS/bin/dbed_vmclonedb -o umount,new_sid=new_sid \
```

```
-f SNAPPLAN [-r relocate_path]
```

In this example, the clone database is shut down and file systems are unmounted for a clone on the same host as the primary database (a single-host configuration).

```
$ /opt/VRTS/bin/dbed_vmclonedb -o umount,new_sid=NEWPROD \
-f snap1 -r /clone
dbed_vmclonedb started at 2006-03-02 15:11:22
dbed_vmclonedb ended at 2006-03-02 15:11:47
```

In this example, the clone database is shut down, file systems are unmounted, and the snapshot disk group is deported for a clone on a secondary host (a two-host configuration).

```
$ /opt/VRTS/bin/dbed_vmclonedb -o umount,new_sid=NEWPROD \
-f snap2
dbed_vmclonedb started at 2006-03-09 23:09:21
dbed_vmclonedb ended at 2006-03-09 23:09:50
```

## Restarting a Clone Database

If the clone database is down as a result of using `dbed_vmclonedb -o umount` or rebooting the system, you can restart it with the `-o restartdb` option.

---

**Note:** This option can only be used when a clone database is created successfully. If the clone database is recovered manually, `-o update_status` must be run to update the status before `-o restartdb` will work.

---

### To start the clone database

- ◆ Use the `dbed_vmclonedb` command as follows:

```
$ /opt/VRTS/bin/dbed_vmclonedb -S ORACLE_SID -g snap_dg \
-o restartdb,new_sid=new_sid,server_name=server_name\
-f SNAPPLAN [-H ORACLE_HOME] \
[-r relocate_path]
```

In this example, the clone database is re-started.

```
$ /opt/VRTS/bin/dbed_vmclonedb -S PROD -g SNAP_PRODDg \
-o restartdb,new_sid=NEWPROD,server_name=host1\
-f snap1 -r /clone
dbed_vmclonedb started at 2006-03-02 15:14:49
Mounting /clone/prod_db on /dev/vx/dsk/SNAP_PRODDg/SNAP_prod_db.
Mounting /clone/prod_ar on /dev/vx/dsk/SNAP_PRODDg/SNAP_prod_ar.
Oracle instance NEWPROD successfully started.
dbed_vmclonedb ended at 2006-03-02 15:15:19
```

## Recreating Oracle tempfiles

After a clone database is created and opened, the tempfiles are added if they were residing on the snapshot volumes. If the tempfiles were not residing on the same file systems as the datafiles, `dbed_vmsnap` does not include the underlying volumes in the snapshot. In this situation, `dbed_vmclonedb` issues a warning message and you can then recreate any needed tempfiles on the clone database as described in the following procedure.

### To recreate the Oracle tempfiles

- 1 If the tempfiles were not residing on the same file systems as the datafiles, `dbed_vmclonedb` will display the WARNING and INFO messages similar to the following:

```
WARNING: Not all tempfiles were included in snapshot for
$ORACLE_SID, there is no snapshot volume for /clone_path/
temp02.dbf.
WARNING: Could not recreate tempfiles for $ORACLE_SID due to
lack of free space.
INFO: The sql script for adding tempfiles to $ORACLE_SID is at
/tmp/add_tf.$ORACLE_SID.sql.
```

where `$ORACLE_SID` is the name of the clone database.

- 2 A script named `add_tf.$ORACLE_SID.sql` is provided in the `/tmp` directory for the purpose of recreating Oracle tempfiles. This script contains the SQL\*Plus commands to recreate the missing tempfiles.
- 3 Make a copy of the `/tmp/add_tf.$ORACLE_SID.sql` script and open it to view the list of missing tempfiles.

An example of the `add_tf.$ORACLE_SID.sql` script is shown below:

```
$ cat /tmp/add_tf.$ORACLE_SID.sql
-- Commands to add tempfiles to temporary tablespaces.
-- Online tempfiles have complete space information.
-- Other tempfiles may require adjustment.
ALTER TABLESPACE TEMP ADD TEMPFILE
'/clone_path/temp01.dbf'
SIZE 4194304 REUSE AUTOEXTEND ON NEXT 1048576 MAXSIZE 33554432 ;
ALTER TABLESPACE TEMP ADD TEMPFILE
'/clone_path/temp02.dbf' REUSE;
ALTER DATABASE TEMPFILE '/clone_path2/temp02.dbf'
OFFLINE;
```

- 4 Evaluate whether you need to recreate any temp files. If you want to recreate tempfiles, proceed to the next step.
- 5 In the `add_tf.$ORACLE_SID.sql` file, edit the sizes and default path names of the tempfiles as needed to reside on cloned volumes configured for database storage.

---

**Warning:** Do not run the script without first editing it because path names may not exist and the specified mount points may not contain sufficient space.

---

- 6 After you have modified the `add_tf.$ORACLE_SID.sql` script, execute it against your clone database.
- 7 After you have successfully run the script, you may delete it.

## Resynchronizing the snapshot to your database

When you have finished using a clone database or want to refresh it, you can resynchronize it with the original database. This is also known as refreshing the snapshot volume or merging the split snapshot image back to the current database image. After resynchronizing, the snapshot can be retaken for backup or decision-support purposes.

You can resynchronize the snapshot from the original volume.

### Prerequisites

- You must be logged in as the Oracle database administrator.
- Before you can resynchronize the snapshot image, you must validate a snapplan and create a snapshot.  
See [“Summary of database snapshot steps”](#) on page 288.  
See [“Validating a snapplan \(dbed\\_vmchecksnap\)”](#) on page 300.  
See [“Creating a snapshot \(dbed\\_vmsnap\)”](#) on page 305.
- If a clone database has been created, shut it down and unmount the file systems using the `dbed_vmclonedb -o umount` command. This command also deports the disk group if the primary and secondary hosts are different.  
See [“Shutting down the clone database and unmounting file systems”](#) on page 317.
- The Oracle database must have at least one mandatory archive destination.  
See [“Establishing a mandatory archive destination”](#) on page 300.

## Usage Notes

- The `dbed_vmsnap` command can only be executed on the primary host.
- In a two-host configuration, the `dbed_vmsnap` command imports the disk group that was deported from the secondary host and joins the disk group back to the original disk group. The snapshot volumes again become plexes of the original volumes. The snapshot is then resynchronized.
- See the `dbed_vmsnap (1M)` manual page for more information.

---

**Note:** You cannot access Database FlashSnap commands (`dbed_vmchecksnap`, `dbed_vmsnap`, and `dbed_vmcloneadb`) with the SFDB menu utility.

---

**To resynchronize the snapshot image**

- ◆ Use the `dbed_vmsnap` command as follows:

```
$ /opt/VRTS/bin/dbed_vmsnap -S ORACLE_SID -f SNAPPLAN -o resync
```

In this example, the snapshot image is resynchronized with the primary database.

```
$ /opt/VRTS/bin/dbed_vmsnap -S PROD -f snap1 -o resync
dbed_vmsnap started at 2006-03-02 16:19:05
The option resync has been completed.
dbed_vmsnap ended at 2006-03-02 16:19:26
```

Now, you can again start creating snapshots.

## Removing a snapshot volume

If a snapshot volume is no longer needed, you can remove it and free up the disk space for other uses by using the `vxedit rm` command.

## Prerequisites

- You must be logged in as superuser.
- If the volume is on a mounted file system, you must unmount it before removing the volume.

**To remove a snapplan and snapshot volume**

- 1 To remove the snapshot and free up the storage used by it:
  - If the snapshot has been taken, remove the snapshot as follows:
 

```
vxsnap -g diskgroup dis snapshot_volume
vxvol -g diskgroup stop snapshot_volume
vxedit -g diskgroup -rf rm snapshot_volume
```

- If the snapshot has not been taken and the snapshot plex (mirror) exists, remove the snapshot as follows:

```
vxsnap -g diskgroup rmmir volume
```

- 2 Remove the DCO and DCO volume:

```
vxsnap -g diskgroup unprepare volume
```

- 3 Remove the snapplan.

```
/opt/VRTS/bin/dbed_vmchecksnap -S PROD -f snapplan -o remove
```

For example, the following commands will remove a snapshot volume from disk group PRODDg:

```
vxsnap -g PRODDg dis snap_v1
```

```
vxvol -g PRODDg stop snap_v1
```

```
vxedit -g PRODDg -rf rm snap_v1
```

# Troubleshooting SF Oracle RAC and optimizing I/O performance

- [Chapter 17, “Investigating I/O performance for SF Oracle RAC: Storage Mapping” on page 325](#)
- [Chapter 18, “Troubleshooting SF Oracle RAC” on page 339](#)



# Investigating I/O performance for SF Oracle RAC: Storage Mapping

This chapter contains the following topics:

- [About Storage Mapping in SF Oracle RAC](#)
- [Understanding Storage Mapping](#)
- [Verifying Veritas Storage Mapping setup](#)
- [Using vxstorage\\_stats](#)
- [Using dbed\\_analyzer](#)
- [Oracle File Mapping \(ORAMAP\)](#)
- [About arrays for Storage Mapping and statistics](#)

## About Storage Mapping in SF Oracle RAC

The storage mapping feature is available with SF Oracle RAC and enables you to map datafiles to physical devices. Storage mapping enables you to map datafiles to physical devices. You may obtain and view detailed storage topology information using the `vxstorage_stats` and `dbed_analyzer` commands. You may also use the Oracle Enterprise Manager to access storage mapping information.

## Understanding Storage Mapping

Access to mapping information is important since it allows for a detailed understanding of the storage hierarchy in which files reside, information that is critical for effectively evaluating I/O performance.

Mapping files to their underlying device is straightforward when datafiles are created directly on a raw device. With the introduction of host-based volume managers and sophisticated storage subsystems that provide RAID features, however, mapping files to physical devices has become more difficult.

With the SF Oracle RAC Storage Mapping option, you can map datafiles to physical devices. Veritas Storage Mapping relies on Veritas Mapping Service (VxMS), a library that assists in the development of distributed SAN applications that must share information about the physical location of files and volumes on a disk.

The Veritas Storage Mapping option supports Oracle's set of storage APIs called Oracle Mapping ("ORAMAP" for short) that lets Oracle determine the mapping information for files and devices.

Oracle provides a set of dynamic performance views (v\$ views) that shows the complete mapping of a file to intermediate layers of logical volumes and physical devices. These views enable you to locate the exact disk on which any specific block of a file resides. You can use these mappings, along with device statistics, to evaluate I/O performance.

The Veritas Storage Mapping option supports a wide range of storage devices and allows for "deep mapping" into EMC, Hitachi, and IBM Enterprise Storage Server ("Shark") arrays. Deep mapping information identifies the physical disks that comprise each LUN and the hardware RAID information for the LUNs.

You can view storage mapping topology information and I/O statistics using the following:

- |                                      |                                                                                                                                                                                                               |
|--------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>vxstorage_stats</code> command | This command displays the complete I/O topology mapping of specific datafiles through intermediate layers like logical volumes down to actual physical devices.                                               |
| <code>dbed_analyzer</code> command   | This command retrieves tablespace-to-physical disk mapping information for all the datafiles in a specified database. It also provides information about the amount of disk space being used by a tablespace. |

---

**Note:** For SF Oracle RAC database, when you issue the commands, replace `$ORACLE_SID` with `$ORACLE_SID=instance_name` and provide the instance name on which the instance is running.

---

In addition, you can also use the Oracle Enterprise Manager GUI to display storage mapping information after file mapping has occurred. Oracle Enterprise Manager does not display I/O statistics information.

## Verifying Veritas Storage Mapping setup

Before using the Veritas Storage Mapping option, verify that the features are set up correctly:

### To verify that your system is using the Veritas Storage Mapping option

- 1 Verify that you have a license key for the storage mapping option.

```
/opt/VRTS/bin/vxlictest -n "VERITAS Mapping Services" -f \
"Found_Edi_map"
Found_Edi_map feature is licensed
```

- 2 Verify that the VRTSvxmsa depot is installed.

```
swlist VRTSvxmsa
VRTSvxmsa 4.4-REVbuild010-2006.03.07 VxMS
Application Deployment Package
VRTSvxmsa.ADMIN 4.4-REVbuild010-2006.03.07 VERITAS
Federated Mapping Service
VRTSvxmsa.LIBRARIES 4.4-REVbuild010-2006.03.07 libraries
VRTSvxmsa.LOGGING 4.4-REVbuild010-2006.03.07 logging
VRTSvxmsa.PLUGINS 4.4-REVbuild010-2006.03.07 plugins
```

## Using vxstorage\_stats

The `vxstorage_stats` command displays detailed storage mapping information and I/O statistics about one or more VxFS files. The mapping information and

I/O statistics are recorded only for VxFS files and VxVM volumes.

In `vxstorage_stats` command output, I/O topology information appears first followed by summary statistics for each object.

The command syntax is as follows:

```
/opt/VRTSdbed/bin/vxstorage_stats -m -s [-i interval -c count] -f
filename
```

### Prerequisites

- You must log in as the database administrator (typically, the user ID oracle) or superuser.

Usage Notes

- The `-s` option displays the file statistics for the specified file.
- The `-c count` option specifies the number of times to display statistics within the interval specified by `-i interval`.
- The `-i interval` option specifies the interval frequency for displaying updated I/O statistics.
- The `-f filename` option specifies the file to display I/O mapping and statistics for.
- The `-m` option displays the I/O topology for the specified file.
- For more information, see the `vxstorage_stats(1m)` online manual page.

## Displaying Storage Mapping information

To display storage mapping information

- ◆ Use the `vxstorage_stats` command with the `-m` option to display storage mapping information:

```
$ /opt/VRTSdbed/bin/vxstorage_stats -m -f file_name
```

For example:

```
$ /opt/VRTSdbed/bin/vxstorage_stats -m -f /oradata/system01.dbf
```

Output similar to the following is displayed:

TY	NAME	NSUB	DESCRIPTION	SIZE(sectors)	OFFSET(sectors)
PROPERTIES					
fi	/oradata/system01.dbf	1	FILE	2621442048(B)	4718592(B)
Extents:3 Sparse Extents:0					
v	myindex	1	MIRROR	16777216	0
pl	vxvm:mydb/myindex-01	3	STRIPE	16779264	0
Stripe_size:2048					
rd	/dev/vx/rdmp/c3t1d3s3	1	PARTITION	5593088	0
sd	/dev/rdisk/c3t1d3s3	1	PARTITION	17674560	960
sd	c3t1d3	2	MIRROR	17677440	0
da	EMC000184502242:02:0c:02	0	DISK	143113019	0
da	EMC000184502242:31:0c:02	0	DISK	143113019	0
rd	/dev/vx/rdmp/c3t1d15s4	1	PARTITION	5593088	0
sd	/dev/rdisk/c3t1d15s4	1	PARTITION	17669760	5760
sd	c3t1d15	2	MIRROR	17677440	0
da	EMC000184502242:01:0c:02	0	DISK	143113019	0

---

**Note:** For file type (fi), the SIZE column is number of bytes; for volume (v), plex (pl), sub-disk (sd), and physical disk (da), the SIZE column is in 512-byte blocks. Stripe sizes are given in sectors.

---

## Displaying I/O statistics information

### To display I/O statistics information

- ◆ Use the `vxstorage_stats` command with the `-s` option to display I/O statistics information:

```
$ /opt/VRTSdbed/bin/vxstorage_stats -s -f file_name
```

For example:

```
$ /opt/VRTSdbed/bin/vxstorage_stats -s -f \
/data/system01.dbf
```

Output similar to the following is displayed:

OBJECT	I/O OPERATIONS		I/O BLOCKS(512 byte)		AVG TIME(ms)	
	READ	WRITE	B_READ	B_WRITE	AVG_RD	AVG_WR
/data/system01.dbf	2	2479	8	5068810	0.00	53.28
/dev/vx/rdsk/mydb/myindex	101	2497	1592	5069056	12.18	52.78
vxvm:mydb/myindex-01	101	2497	1592	5069056	12.18	52.76
/dev/rdsk/c3t1d3s3	131	1656	2096	1689696	14.43	39.09
c3t1d3	131	1656	2096	1689696	14.43	39.09
EMC000184502242:02:0c:02	8480	231019	275952	23296162	-	-
EMC000184502242:31:0c:02	3244	232131	54808	23451325	-	-
/dev/rdsk/c3t1d15s4	0	1652	0	1689606	0.00	46.47
c3t1d15	0	1652	0	1689606	0.00	46.47
EMC000184502242:01:0c:02	23824	1188997	1038336	32407727	-	-
EMC000184502242:32:0c:02	5085	852384	135672	29956179	-	-
/dev/rdsk/c3t1d2s4	14	1668	200	1689834	18.57	34.19
c3t1d2	14	1668	200	1689834	18.57	34.19
EMC000184502242:16:0c:02	4406	271155	121368	23463948	-	-
EMC000184502242:17:0c:02	3290	269281	55432	23304619	-	-

### To display storage mapping and I/O statistics information at repeated intervals

- ◆ Use the `vxstorage_stats` command with the `-i interval` and `-c count` options. The `-i interval` option specifies the interval frequency for displaying updated I/O statistics and the `-c count` option specifies the number of times to display statistics:

```
$ /opt/VRTSdbed/bin/vxstorage_stats [-m] [-s] \
[-i interval -c count] -f file_name
```

For example, type the following command to display statistics two times with a time interval of two seconds:

```
$ /opt/VRTSdbed/bin/vxstorage_stats -s -i2 -c2 \
-f /data/system01.dbf
```

Output similar to the following is displayed:

OBJECT	I/O OPERATIONS		I/O BLOCKS(512 byte)		AVG TIME(ms)	
	READ	WRITE	B_READ	B_WRITE	AVG_RD	AVG_WR
/oradata/system01.dbf	11917	13102	95336	594402	2.84	278.81
/dev/vx/rdsk/ramdg/oradata	11959	14729	105839	704811	3.11	295.20
vxvm:ramdg/oradata-01	11947	14729	107783	704799	3.14	295.23

/dev/rhdisk13	5974	8366	53663	352207	2.95	246.65
EMC000184502242:02:0c:01	5	733	320	49897	-	-
EMC000184502242:31:0c:01	1	1230	8	174752	-	-
/dev/rhdisk41	4705	8249	73526	339740	6.01	250.18
EMC000184502242:02:0c:01	5	733	320	49897	-	-
EMC000184502242:31:0c:01	1	1230	8	174752	-	-
/dev/rhdisk14	3895	6264	91672	295491	11.49	279.42
EMC000184502242:01:0d:01	0	0	0	0	-	-
EMC000184502242:32:0d:01	1	1336	8	190456	-	-
/dev/rhdisk42	4407	4840	119576	116696	13.12	287.46
EMC000184502242:01:0d:01	0	0	0	0	-	-
EMC000184502242:32:0d:01	1	1336	8	190456	-	-

## Using dbed\_analyzer

Effectively performing a parallel backup requires an understanding of which tablespaces reside on which disks. If two tablespaces reside on the same disk, for example, backing them up in parallel will not reduce their downtime.

The `dbed_analyzer` command provides tablespace-to-physical disk mapping information for all the datafiles in a specified tablespace, list of tablespaces, or an entire database. (In contrast, the `vxstorage_stats` command provides this information on a per-file basis only.) In addition, `dbed_analyzer` provides information about the amount of disk space they are using.

### Prerequisites

- You must log in as the database administrator (typically, the user ID oracle).

### Usage Notes

- The `-o sort=tbs` option provides the layout of the specified tablespaces on the physical disk as well as the amount of disk space they are using.
- The `-o sort=disk` option provides the name of the disks containing the specified tablespaces as well as the amount of disk space the tablespaces are using.
- The `-f filename` option specifies the name of a file containing a list of the tablespaces for which to obtain mapping information.
- The `-t tablespace` option specifies the name of a tablespace for which to obtain mapping information.
- If `-f filename` or `-t tablespace` is not specified then all the tablespaces in the database will be analyzed.
- For more information, see the `dbed_analyzer(1M)` online manual page.

## Obtaining Storage Mapping information for a list of tablespaces

### To obtain storage mapping information sorted by tablespace

- ◆ Use the `dbed_analyzer` command with the `-f filename` and `-o sort=tbs` options:

```
$ /opt/VRTSdbed/bin/dbed_analyzer -S $ORACLE_SID \
-H $ORACLE_HOME -o sort=tbs -f filename
```

For example,

```
$ /opt/VRTSdbed/bin/dbed_analyzer -S PROD -H /usr1/oracle \
-o sort=tbs -f /tmp/tbsfile
```

Output similar to the following is displayed in the file `tbsfile`:

TBSNAME	DATAFILE	DEVICE	SIZE(sectors)
SYSTEM	/usr1/oracle/rw/DATA/PROD.dbf	c3t21000020379DBD5Fd0	819216
TEMP	/usr1/oracle/rw/DATA/temp_20000	c3t21000020379DBD5Fd0	1021968
TEMP	/usr1/oracle/rw/DATA/temp_20001	c3t21000020379DBD5Fd0	2048016
SYSAUX	/usr1/oracle/rw/DATA/sysaux.dbf	c3t21000020379DBD5Fd0	819216
ITEM	/usr1/oracle/rw/DATA/item_1000	c3t21000020379DBD5Fd0	1021968
ITM_IDX	/usr1/oracle/rw/DATA/itm_idx_2000	c3t21000020379DBD5Fd0	1021968
PRODID_IDX	/usr1/oracle/rw/DATA/prodid_idx_3000	c3t21000020379DBD5Fd0	1021968
QTY_IDX	/usr1/oracle/rw/DATA/qty_idx_7000	c3t21000020379DBD5Fd0	1021968
ROLL_1	/usr1/oracle/rw/DATA/roll_1_5000	c3t21000020379DBD5Fd0	1021968
ROLL_2	/usr1/oracle/rw/DATA/roll_2_6000	c3t21000020379DBD5Fd0	1021968
ORDERS	/usr1/oracle/rw/DATA/orders_4000	c3t21000020379DBD5Fd0	1021968
ORD_IDX	/usr1/oracle/rw/DATA/ord_idx_10000	c3t21000020379DBD5Fd0	1021968
QTY_IDX	/usr1/oracle/rw/DATA/qty_idx_7001	c3t21000020379DBD5Fd0	1024016
ITM_IDX	/usr1/oracle/rw/DATA/itm_idx_2001	c3t21000020379DBD5Fd0	1024016
ROLL_1	/usr1/oracle/rw/DATA/roll_1_5001	c3t21000020379DBD5Fd0	1024016
QTY_IDX	/usr1/oracle/rw/DATA/qty_idx_7002	c3t21000020379DBD5Fd0	1024016
ROLL_2	/usr1/oracle/rw/DATA/roll_2_6001	c3t21000020379DBD5Fd0	1024016
ITEM	/usr1/oracle/rw/DATA/item_1001	c3t21000020379DBD5Fd0	4096016

### To obtain storage mapping information sorted by disk

- ◆ Use the `dbed_analyzer` command with the `-f filename` and `-o sort=disk` options:

```
$ /opt/VRTSdbed/bin/dbed_analyzer -S $ORACLE_SID \
-H $ORACLE_HOME -o sort=disk -f filename
```

For example,

```
$ /opt/VRTSdbed/bin/dbed_analyzer -S PROD -H /usr1/oracle \
-o sort=disk -f /tmp/tbsfile
```

Output similar to the following is displayed in the file `tbsfile`:

DEVICE	TBSNAME	DATAFILE	SIZE(sectors)
c3t21000020379DBD5Fd0	SYSTEM	/usr1/oracle/rw/DATA/PROD.dbf	819216
c3t21000020379DBD5Fd0	TEMP	/usr1/oracle/rw/DATA/temp_20000	1021968
c3t21000020379DBD5Fd0	TEMP	/usr1/oracle/rw/DATA/temp_20001	2048016
c3t21000020379DBD5Fd0	SYSAUX	/usr1/oracle/rw/DATA/sysaux.dbf	819216
c3t21000020379DBD5Fd0	ITEM	/usr1/oracle/rw/DATA/item_1000	1021968

c3t21000020379DBD5Fd0	ITM_IDX	/usr1/oracle/rw/DATA/itm_idx_2000	1021968
c3t21000020379DBD5Fd0	PRODID_IDX	/usr1/oracle/rw/DATA/prodid_idx_3000	1021968
c3t21000020379DBD5Fd0	QTY_IDX	/usr1/oracle/rw/DATA/qty_idx_7000	1021968
c3t21000020379DBD5Fd0	ROLL_1	/usr1/oracle/rw/DATA/roll_1_5000	1021968
c3t21000020379DBD5Fd0	ROLL_2	/usr1/oracle/rw/DATA/roll_2_6000	1021968
c3t21000020379DBD5Fd0	ORDERS	/usr1/oracle/rw/DATA/orders_4000	1021968
c3t21000020379DBD5Fd0	ORD_IDX	/usr1/oracle/rw/DATA/ord_idx_10000	1021968
c3t21000020379DBD5Fd0	QTY_IDX	/usr1/oracle/rw/DATA/qty_idx_7001	1024016
c3t21000020379DBD5Fd0	ITM_IDX	/usr1/oracle/rw/DATA/itm_idx_2001	1024016
c3t21000020379DBD5Fd0	ROLL_1	/usr1/oracle/rw/DATA/roll_1_5001	1024016
c3t21000020379DBD5Fd0	QTY_IDX	/usr1/oracle/rw/DATA/qty_idx_7002	1024016
c3t21000020379DBD5Fd0	ROLL_2	/usr1/oracle/rw/DATA/roll_2_6001	1024016
c3t21000020379DBD5Fd0	ITEM	/usr1/oracle/rw/DATA/item_1001	4096016

## Oracle File Mapping (ORAMAP)

Veritas has defined and implemented two libraries:

- For HP-UX IA: `libvxoramap_64.so`
- For HP-UX PA: `libvxoramap_64.sl`

These two libraries provide a mapping interface to Oracle 10g release 2. These two libraries serve as a bridge between the Oracle's set of storage APIs (ORAMAP) and Veritas Federated Mapping Service (VxMS), a library that assists in the development of distributed SAN applications that must share information about the physical location of files and volumes on a disk.

With Veritas Storage Mapping option, you can view the complete I/O topology mapping of datafiles through intermediate layers like logical volumes down to actual physical devices. You can use this information to determine the exact location of an Oracle data block on a physical device and to help identify hot spots.

## Mapping components

Review the information on the mapping components in the System Global Area (SGA) and Oracle's representation of these components. You will need an understanding of these components to interpret the mapping information in Oracle's dynamic performance views.

The mapping information in Oracle's dynamic performance views consists of:

- File components  
A mapping file component is a mapping structure describing a file. It provides a set of attributes for a file, including the file's size, number of extents, and type. File components are exported to the user through `V$MAP_FILE`.

- **File extent components**  
A mapping file extent component describes a contiguous group of blocks residing on one element. The description specifies the device offset, the extent size, the file offset, the extent type (Data or Parity), and the name of the element where the extent resides.
- **Element components**  
A mapping element component is a mapping structure that describes a storage component within the I/O stack. Elements can be mirrors, stripes, partitions, RAID5, concatenated elements, and disks.  
This component contains information about the element's mapping structure, such as the element's size, type, number of subelements, and a brief description. Element components are exported to the user through V\$MAP\_ELEMENT.
- **Subelement components**  
A mapping subelement component describes the link between an element and the next element in the I/O stack. The subelement component contains the subelement number, size, the element name for the subelement, and the element offset. Subelement components are exported to the user through V\$MAP\_SUBELEMENT.

These four types of mapping components completely describe the mapping information for an Oracle instance.

## Storage Mapping views

The mapping information that is captured is presented in Oracle's dynamic performance views. [Table 17-5](#) provides brief descriptions of these views. For more detailed information, refer to your Oracle documentation.

**Table 17-5** Storage Mapping views

View	Description
V\$MAP_LIBRARY	Contains a list of all the mapping libraries that have been dynamically loaded by the external process.
V\$MAP_FILE	Contains a list of all the file mapping structures in the shared memory of the instance.
V\$MAP_FILE_EXTENT	Contains a list of all the file extent mapping structures in the shared memory of the instance.
V\$MAP_ELEMENT	Contains a list of all the element mapping structures in the SGA of the instance.

**Table 17-5** Storage Mapping views

View	Description
V\$MAP_EXT_ELEMENT	Contains supplementary information for all element mapping structures.
V\$MAP_SUBELEMENT	Contains a list of all subelement mapping structures in the shared memory of the instance.
V\$MAP_COMP_LIST	Describes the component list associated with the element name.
V\$MAP_FILE_IO_STACK	Contains the hierarchical arrangement of storage containers for the file. This information is displayed as a series of rows. Each row represents a level in the hierarchy.

## Verifying Oracle file mapping setup

To verify that \$ORACLE\_HOME is set up for Oracle file mapping (ORAMAP)

- 1 Verify \$ORACLE\_HOME is ready for Oracle file mapping (ORAMAP):
 

```
cd $ORACLE_HOME/rdbms/filemap/bin
ls -l
-r-xr-x--- 1 root system 900616 Apr 08 19:16 fmpu1
-r-sr-xr-x 1 root system 14614 Apr 08 19:16 fmpu1hp
```
- 2 Confirm the following items and make the appropriate corrections:
  - root owns fmpu1hp and the setud bit is set.
  - The permissions for fmpu1hp are -r-sr-xr-x.
  - The permissions for fmpu1 are -r-xr-x---
- 3 If any of these items is not set as specified, make the appropriate corrections.

## Enabling Oracle file mapping

To enable Oracle file mapping with the Veritas Storage Mapping option

- 1 Ensure that the file filemap.ora exists and contains a valid entry for the Veritas mapping library for Oracle storage mapping.
 

```
cd $ORACLE_HOME/rdbms/filemap/etc
cat filemap.ora
```

 For 64-bit Oracle, the filemap.ora file should contain the following setting:

```
For PA lib=VERITAS:/opt/VRTSdbed/lib/libvxoramap_64.sl
```

For IA `lib=VERITAS:/opt/VRTSdbed/lib/libvxoramap_64.so`

- 2 After verifying that the system is using the Veritas library for Oracle storage mapping, set the `file_mapping` initialization parameter to true.

```
SQL> alter system set file_mapping=true;
```

The `file_mapping` initialization parameter is set to false by default. You do not need to shut down the instance to set this parameter. Setting `file_mapping=true` starts the FMON background process.

If you want storage mapping to be enabled whenever you start up an instance, set the `file_mapping` initialization parameter to true in the `init.ora` file.

## Accessing dynamic performance views

### To access dynamic performance views

- 1 Confirm that the Veritas mapping library for Oracle file mapping has been enabled:

```
SQL> select lib_idx idx, lib_name name, vendor_name vname, \
path_name path from v$map_library;
```

- 2 After storage mapping has been enabled, Oracle datafiles can be mapped using the `DBMS_STORAGE_MAP` package.

For more information about various features and capabilities of the `DBMS_STORAGE_MAP` package, see your Oracle documentation.

- 3 Use SQL commands to display the mapping information that is captured in Oracle's dynamic performance views.

- To display the contents of `v$map_file` for a Quick I/O file:

```
SQL> select file_name name, file_map_idx idx, \
file_status status, file_type type, file_structure str, \
file_size fsize, file_nexts nexts from v$map_file;
```

- To display the contents of `v$map_file_extnt`:

```
SQL> select * from v$map_file_extnt;
```

- To display the contents of `v$map_element`:

```
SQL> select elem_idx idx, elem_name, elem_type type, \
elem_size, elem_nsubelem nsub, elem_descr, stripe_size \
from v$map_element;
```

- To display the contents of `v$map_subelement`:

```
SQL> select * from v$map_subelement;
```

- To display all the elements within the I/O stack for a specific file:

```
SQL> with fv as
2 (select file_map_idx, file_name from v$map_file
3 where file_name = '/ora92/dbs/qio10m.dbf')
```

```
4 select
5 fv.file_name, lpad(' ', 4 * (level - 1)) || \
 el.elem_name elem_name, el.elem_size, el.elem_type, \
 el.elem_descr
6 from
7 v$map_subelement sb, v$map_element el, fv,
8 (select unique elem_idx from v$map_file_io_stack io, fv
9 where io.file_map_idx = fv.file_map_idx) fs
10 where el.elem_idx = sb.child_idx
11 and fs.elem_idx = el.elem_idx
12 start with sb.parent_idx in
13 (select distinct elem_idx
14 from v$map_file_extent fe, fv
15 where fv.file_map_idx = fe.file_map_idx)
16 connect by prior sb.child_idx = sb.parent_idx;
```

## Using Oracle Enterprise Manager

Oracle Enterprise Manager is a web-based GUI for managing Oracle databases. You can use this GUI to perform a variety of administrative tasks such as creating tablespaces, tables, and indexes; managing user security; and backing up and recovering your database. You can also use Oracle Enterprise Manager to view performance and status information about your database instance.

From Oracle Enterprise Manager, you can view storage mapping information and a graphical display of the storage layout. Storage mapping information cannot be viewed with the Oracle 10g version of the Oracle Enterprise Manager client. However, the Oracle9*i* version of Oracle Enterprise Manager can be used with Oracle 10g to view storage mapping information.

For more information about Oracle Enterprise Manager, refer to your Oracle documentation.

### To view storage information

- 1 To view storage information, start Oracle Enterprise Manager and select a database from the left navigational pane (the object tree) of the Oracle Enterprise Manager Console.
- 2 Expand the Databases icon and select the desired database. The Database Connect Information window appears.
- 3 Enter a user name and password to log in to the database and click **OK**.
- 4 In the object tree, expand the **Storage** icon.
- 5 Under the **Storage** icon, expand the **Datfiles** icon.
- 6 Select the datafile for which you want to view storage layout information.
- 7 In the right pane, click the **Storage Layout** tab.

- 8 Expand the objects to display their storage layout.  
Within the Oracle Enterprise Manager Console, you can point to an object on the screen and a description of the object is displayed in a pop-up field. If an object name or path appears truncated, point to it and the pop-up field will display the full object name and path.
- 9 By default, storage layout information appears in tabular format. That is, the **Tabular Display** icon is selected. To view a graphical display of the storage layout, click the **Graphical Display** icon.
- 10 Expand the objects to display their storage layout information graphically.
- 11 To exit, choose **Exit** from the **File** menu.

## About arrays for Storage Mapping and statistics

SF Oracle RAC provides “deep” mapping information and performance statistics for supported storage arrays. Deep mapping information consists of identifying the physical disks that comprise each LUN and the hardware RAID information for the LUNs.

---

**Note:** To use deep mapping, you must have Oracle 9.2.0.3. or later installed.

---

Veritas Array Integration Layer (VAIL) software interfaces third-party hardware storage arrays with Veritas storage software. VAIL providers are software modules that enable Veritas applications to discover, query, and manage third-party storage arrays. The vx\_emc\_symmetrix provider manages EMC Symmetrix arrays.

For the most up-to-date array support information, see the appropriate hardware compatibility list (HCL) on the Veritas Technical Support Web page at: <http://entsupport.symantec.com>

If you want to use storage array information accessible through the VAIL providers, install VAIL and perform any required configuration for the storage arrays and VAIL providers. To use deep mapping services and performance statistics for supported storage arrays, you must install both VAIL and Veritas Mapping Services (VxMS).

You will need to install required third-party array CLIs and APIs on the host where you are going to install VAIL. If you install any required CLI or API after you install VAIL, rescan the arrays so that SF Oracle RAC can discover them.

For details on supported array models, see the *Veritas Array Integration Layer Array Configuration Guide*.



# Troubleshooting SF Oracle RAC

This chapter contains the following topics:

- [About troubleshooting SF Oracle RAC](#)
- [Troubleshooting tips](#)
- [Troubleshooting Oracle](#)
- [Troubleshooting fencing](#)
- [Troubleshooting ODM](#)
- [Troubleshooting VCSIPC](#)
- [Troubleshooting CVM](#)
- [Troubleshooting interconnects](#)
- [Troubleshooting SF Oracle RAC checkpoint](#)

## About troubleshooting SF Oracle RAC

Review the troubleshooting options, known problems, and their solutions.

### Running scripts for engineering support analysis

You can use a set of three scripts that gather information about the configuration and status of your cluster and its various modules. The scripts also identify package information, debugging messages, console messages, and information about disk groups and volumes. You can forward the output of each of these scripts to Veritas customer support who can analyze the information

and assist you in solving any problems. [Table 18-6](#) lists the scripts that you can use to gather the required information.

**Table 18-6** Scripts for engineering support analysis

Script	Description
getdbac	This script gathers information about the SF Oracle RAC modules. The file <code>/tmp/vcsopslog.time_stamp.sys_name.tar.Z</code> contains the script's output. On each node, enter: <pre># /opt/VRTSvcs/bin/getdbac -local</pre>
getcomms	This script gathers information about the GAB and LLT modules. The file <code>/tmp/commslog.time_stamp.tar</code> contains the script's output. On each node, enter: <pre># /opt/VRTSgab/getcomms -local</pre>
hagetcf	This script gathers information about the VCS cluster and the status of resources. The script's output is placed in a tar file, <code>/tmp/vcsconf.sys_name.tar</code> , on each node. On each node, enter: <pre># /opt/VRTSvcs/bin/hagetcf</pre>

## Troubleshooting tips

Review the information on the error logs that you must access.

- To check the Oracle installation error log, you must access:  
`$ORACLE_BASE/oraInventory/logs/installActionsdate_time.log`  
 This file contains errors that occurred during installation. It clarifies the nature of the error and at exactly which point it occurred during the installation. If there are any installation problems, you must send this file to Tech Support for debugging the issue.
- To check the Veritas log file, you must access:  
`/var/VRTSvcs/log/engine_A.log`  
 This file contains all actions performed by HAD. Verify if there are any CVM or PrivNIC errors logged in this file, because the errors may prove to be critical.

# Troubleshooting Oracle

For help resolving issues with Oracle components, check the following information:

- [Oracle log files](#)
- [Oracle notes](#)
- [Oracle troubleshooting topics](#)

## Oracle log files

Review the information on the files that you must access.

- To check the Oracle log file: `$ORA_CRS_HOME/log/<hostname>/crsd/`  
 This file contains the logs pertaining to the CRS resources such as the virtual IP, Listener, and database instances. It indicates some configuration errors or Oracle problems, because CRS does not directly interact with any of the Veritas components.
- To check for core dumps: `$ORA_CRS_HOME/log/<hostname>/crsd/`  
 Core dumps for the crsd.bin daemon are written here. Use this file for further debugging.
- To check the Oracle css log file: `$ORA_CRS_HOME/log/<hostname>/cssd/`  
 The css logs indicate actions such as reconfigurations, missed checkins, connects, and disconnects from the client CSS listener. If there are membership issues, they will show up here. If there are communication issues over the private networks, they are logged here. The ocspd process interacts with vcsmm for cluster membership.
- To check for ocspd core dumps: `$ORA_CRS_HOME/log/<hostname>/cssd/`  
 Core dumps from the ocspd and the pid for the css daemon whose death is treated as fatal are located here. If there are abnormal restarts for css the core files are found here.

## Oracle notes

Review the following Oracle notes:

259301.1	CRS and 10g Real Application Clusters
280589.1	How to install Oracle 10g CRS on a cluster where one or more nodes are not to be configured to run CRS.
265769.1	10g RAC: Troubleshooting CRS Reboots

279793.1	How to Restore a Lost Vote Disk in 10g
239998.1	10g RAC: How to Clean Up After a Failed CRS Install Two items missing in this Oracle note are: <ul style="list-style-type: none"><li>■ Remove the <code>/etc/oracle/ocr.loc</code> file. This file contains the location for the Cluster registry. If this file is not removed then during the next installation the installer will not query for the OCR location and will pick it from this file.</li><li>■ If there was a previous 9i Oracle installation, then remove the following file: <code>/var/opt/oracle/srvConfig.loc</code>. If this file is present the installer will pick up the Vote disk location from this file and may create the error “the Vote disk should be placed on a shared file system” even before specifying the Vote disk location.</li></ul>
272332.1	CRS 10g Diagnostic Collection Guide

## Oracle troubleshooting topics

Headings indicate likely symptoms or procedures required for a solution.

### Oracle user must be able to read `/etc/llttab` file

Check the permissions of the `/etc/llttab` file; the oracle user must be allowed to read it.

### Error when starting an Oracle instance

If the VCSMM driver (the membership module) is not configured, an error is displayed on starting the Oracle instance that resembles:

```
ORA-29702: error occurred in Cluster Group Operation
```

To start the driver, enter the following command:

```
/sbin/vcsmmconfig -c
```

The command included in the `/etc/vcsmmtab` file enables the VCSMM driver to be started at system boot.

### Instance numbers must be unique (error code 205)

If you encounter error code 205 when the `skgxnreg` function fails (look in the Oracle trace files to find the error returned), make sure there is a unique instance number specified in the `$ORACLE_HOME/dbs/init${ORACLE_SID}.ora` file on each node.

## ORACLE\_SID must be unique (error code 304)

If you encounter error code 304 when the `skgxnreg` function fails (look in the Oracle trace file to find the error returned), make sure that the `ORACLE_SID` environment variable specified during Oracle startup is unique on each node in your cluster. Also, make sure that the `SID` attribute for the Oracle resource in the `main.cf` is specified locally and is unique.

## Oracle log files show shutdown called even when not shutdown manually

The Oracle enterprise agent calls shutdown if monitoring of the Oracle/Netlsnr resources fails for some reason. On all cluster nodes, look at the following VCS and Oracle agent log files for any errors or status:

```
/var/VRTSvcs/log/engine_A.log
/var/VRTSvcs/log/Oracle_A.log
```

## Set MLOCK privilege for DBA user

If `ASYNCH_IO` errors occur during select and update queries on the Oracle database, the workaround involves setting the `MLOCK` privilege for the `dba` user.

### To set MLOCK privilege for DBA user

- 1 Give the `MLOCK` privilege to the `dba` group:
 

```
setprivgrp dba MLOCK
```
- 2 Create the `/etc/privgroup` file and add the line:
 

```
dba MLOCK
```
- 3 Verify the availability of `MLOCK` privilege for the `dba` group:
 

```
/usr/bin/getprivgrp dba
```

# Troubleshooting fencing

Headings indicate likely symptoms or procedures required for a solution.

## Node is unable to join cluster while another node is being ejected

A cluster that is currently fencing out (ejecting) a node from the cluster prevents a new node from joining the cluster until the fencing operation is completed. The following are example messages that appear on the console for the new node:

```
...VCS FEN ERROR V-11-1-25 ... Unable to join running cluster
..VCS FEN ERROR V-11-1-25 ... since cluster is currently
fencing
```

```
...VCS FEN ERROR V-11-1-25 ... a node out of the cluster.
```

```
...VCS GAB.. Port b closed
```

If you see these messages when the new node is booting, the `vxfen` startup script on the node makes up to five attempts to join the cluster. If this is not sufficient to allow the node to join the cluster, restart the new node or attempt to restart `vxfen` driver with the command:

```
/sbin/init.d/vxfen start
```

## vxfersthdw fails when SCSI TEST UNIT READY command fails

If you see a message resembling:

```
Issuing SCSI TEST UNIT READY to disk reserved by other node
FAILED.
Contact the storage provider to have the hardware configuration
fixed.
```

The disk array does not support returning success for a SCSI TEST UNIT READY command when another host has the disk reserved using SCSI-3 persistent reservations. This happens with Hitachi Data Systems 99XX arrays if bit 186 of the system mode option is not enabled.

## Removing existing keys from disks

Review the procedure to remove the registration and reservation keys created by another node from a disk.

### To remove the registration and reservation keys from disk

- 1 Create a file to contain the access names of the disks:

```
vi /tmp/disklist
```

For example:

```
/dev/vx/rdmp/c1t12d0
```

- 2 Read the existing keys:

```
vxferadm -g all -f /tmp/disklist
```

The output from this command displays the key:

```
Device Name: /dev/vx/rdmp/c1t12d0
```

```
Total Number Of Keys: 1
```

```
key[0]:
```

```
Key Value [Numeric Format]: 65,49,45,45,45,45,45,45
```

```
Key Value [Character Format]: A1-----
```

- 3 If you know on which node the key was created, log in to that node and enter the following command:

```
vxferadm -x -kA1 -f /tmp/disklist
```

The key is removed.

- 4 If you do not know on which node the key was created, follow [step 5](#) through [step 7](#) to remove the key.
- 5 Register a second key "A2" temporarily with the disk:
 

```
vxfenadm -m -kA2 -f /tmp/disklist
Registration completed for disk path /dev/vx/rdmp/c1t12d0
```
- 6 Remove the first key from the disk by preempting it with the second key:
 

```
vxfenadm -p -kA2 -f /tmp/disklist -vA1
key: A2----- preempted the key: A1----- on disk
/dev/vx/rdmp/c1t0d11s2
```
- 7 Remove the temporary key assigned in [step 5](#).
 

```
vxfenadm -x -kA2 -f /tmp/disklist
Deleted the key : [A2-----] from device /dev/vx/rdmp/c1t12d0
No registration keys exist for the disk.
```

## System panics to prevent potential data corruption

When a node experiences a split brain condition and is ejected from the cluster, it panics and displays the following console message:

```
VXFFEN:vxfen_plat_panic: Local cluster node ejected from cluster
to prevent potential data corruption.
```

### How vxfen driver checks for pre-existing split brain condition

The vxfen driver functions to prevent an ejected node from rejoining the cluster after the failure of the private network links and before the private network links are repaired.

For example, suppose the cluster of system 1 and system 2 is functioning normally when the private network links are broken. Also suppose system 1 is the ejected system. When system 1 restarts before the private network links are restored, its membership configuration does not show system 2; however, when it attempts to register with the coordinator disks, it discovers system 2 is registered with them. Given this conflicting information about system 2, system 1 does not join the cluster and returns an error from vxfenconfig that resembles:

```
vxfenconfig: ERROR: There exists the potential for a preexisting
split-brain. The coordinator disks list no nodes which are in
the current membership. However, they also list nodes which are
not in the current membership.
```

```
I/O Fencing Disabled!
```

---

**Note:** During the system boot, because the HP-UX rc sequencer redirects the stderr of all rc scripts to the file /etc/rc.log, the error messages will not be printed on the console. It will be logged in the /etc/rc.log file.

---

Also, the following information is displayed on the console:

```
<date> <system name> vxfen: WARNING: Potentially a preexisting
<date> <system name> split-brain.
<date> <system name> Dropping out of cluster.
<date> <system name> Refer to user documentation for steps
<date> <system name> required to clear preexisting split-brain.
<date> <system name>
<date> <system name> I/O Fencing DISABLED!
<date> <system name>
<date> <system name> gab: GAB:20032: Port b closed
```

---

**Note:** If `syslogd` is configured with the `-D` option, then the informational message will not be printed on the console. The messages will be logged in the system buffer. The system buffer can be read with the `dmesg` command.

---

However, the same error can occur when the private network links are working and both systems go down, system 1 restarts, and system 2 fails to come back up. From the view of the cluster from system 1, system 2 may still have the registrations on the coordinator disks.

### Case 1: system 2 up, system 1 ejected (actual potential split brain)

Determine if system1 is up or not. If it is up and running, shut it down and repair the private network links to remove the split brain condition. restart system 1.

### Case 2: system 2 down, system 1 ejected (apparent potential split brain)

- 1 Physically verify that system 2 is down.
- 2 Verify the systems currently registered with the coordinator disks. Use the following command:

```
vxfenadm -g all -f /etc/vxfentab
```

The output of this command identifies the keys registered with the coordinator disks.

- 3 Clear the keys on the coordinator disks as well as the data disks using the command `/opt/VRTSvcs/rac/bin/vxfenclearpre`.  
See [“Clearing keys after split brain using vxfenclearpre command”](#) on page 347.

- 4 Make any necessary repairs to system 2 and restart.

## Clearing keys after split brain using vxfenclearpre command

When you have encountered a split brain condition, use the `vxfenclearpre` command to remove SCSI-3 registrations and reservations on the coordinator disks as well as on the data disks in all shared disk groups.

### To clear keys after split brain

- 1 Shut down all other nodes in the cluster that have access to the shared storage. This prevents data corruption.
- 2 Start the script:
 

```
cd /opt/VRTSvcs/vxfen/bin
./vxfenclearpre
```
- 3 Read the script's introduction and warning. Then, you can choose to let the script run.

Do you still want to continue: [y/n] (default : n) **y**

Informational messages resembling the following may appear on the console of one of the nodes in the cluster when a node is ejected from a disk/LUN:

```
<date> <system name> scsi: WARNING: /sbus@3,0/lpfs@0,0/
sd@0,1(sd91):
<date> <system name> Error for Command: <undecoded cmd 0x5f>
Error Level: Informational
<date> <system name> scsi: Requested Block: 0 Error Block 0
<date> <system name> scsi: Vendor: <vendor> Serial Number:
0400759B006E
<date> <system name> scsi: Sense Key: Unit Attention
<date> <system name> scsi: ASC: 0x2a (<vendor unique code
0x2a>), ASCQ: 0x4, FRU: 0x0
```

**These informational messages may be ignored.**

Cleaning up the coordinator disks...

Cleaning up the data disks for all shared disk groups...

Successfully removed SCSI-3 persistent registration and reservations from the coordinator disks as well as the shared data disks.

```
Reboot the server to proceed with normal cluster startup...
#
```

- 4 Restart all nodes in the cluster.

## Adding or removing coordinator disks

Review the following information to:

- Replace coordinator disk in the coordinator disk group
- Destroy a coordinator disk group

---

**Note:** Adding or removing coordinator disks requires all services be shut down.

---

Note the following about the procedure:

- A coordinator disk group requires three disks/LUNs.
- When adding a disk, add the disk to the disk group `vx fenceoordg` and retest the group for support of SCSI-3 persistent reservations.
- You can destroy the coordinator disk group such that no registration keys remain on the disks. The disks can then be used elsewhere.

### To remove and replace a disk in the coordinator disk group

- 1 Log in as superuser on one of the cluster nodes.
- 2 If VCS is running, shut it down:  

```
hastop -all
```
- 3 Stop the VCSMM driver on each node:  

```
/sbin/init.d/vcsmm stop
```
- 4 Stop I/O fencing on all nodes:  

```
/sbin/init.d/vxfen stop
```

This removes any registration keys on the disks.
- 5 Import the coordinator disk group. The file `/etc/vxfendg` includes the name of the disk group (typically, `vx fenceoordg`) that contains the coordinator disks, so use the command:  

```
vxdg -tfc import `cat /etc/vxfendg`
```

where:

  - t specifies that the disk group is imported only until the node restarts.
  - f specifies that the import is to be done forcibly, which is necessary if one or more disks is not accessible.
  - C specifies that any import blocks are removed.
- 6 To remove disks from the disk group, use the VxVM disk administrator utility, `vx diskadm`.  
You may also destroy the existing coordinator disk group. For example:  

```
vxdg destroy vx fenceoordg
```

- 7 Add the new disk to the node, initialize it as a VxVM disk, and add it to the `vxfencoorddg` disk group.  
 See “[Creating the coordinator disk group](#)” on page 108.
- 8 Test the recreated disk group for SCSI-3 persistent reservations compliance.  
 See “[Testing the coordinator disk group using vxfcntlhdw -c](#)” on page 390.
- 9 After replacing disks in a coordinator disk group, deport the disk group:  

```
vxpdg deport `cat /etc/vxfendg`
```
- 10 On each node, start the I/O fencing driver:  

```
/sbin/init.d/vxfen start
```
- 11 On each node, start the VCSMM driver:  

```
/sbin/init.d/vcsmm start
```
- 12 If necessary, restart VCS on each node:  

```
hastart
```

## Troubleshooting ODM

Headings indicate likely symptoms or procedures required for a solution.

### File system configured incorrectly for ODM shuts down Oracle

Linking Oracle with the Veritas ODM libraries provides the best file system performance. Review the instructions on creating the link and confirming that Oracle uses the libraries. Shared file systems in RAC clusters without ODM libraries linked to Oracle may exhibit slow performance and are *not* supported.

If ODM cannot find the resources it needs to provide support for cluster file systems, it does not allow Oracle to identify cluster files and causes Oracle to fail at startup. Run the following command:

```
cat /dev/odm/cluster
cluster status: enabled
```

If the status is “enabled,” ODM is supporting cluster files. Any other cluster status indicates that ODM is not supporting cluster files. Other possible values include:

- |          |                                                                                                                                                                                                                                                                                                                                                                                                              |
|----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| pending  | ODM cannot yet communicate with its peers, but anticipates being able to eventually.                                                                                                                                                                                                                                                                                                                         |
| failed   | ODM cluster support has failed to initialize properly. Check console logs.                                                                                                                                                                                                                                                                                                                                   |
| disabled | ODM is not supporting cluster files. If you think it should, check: <ul style="list-style-type: none"> <li>■ <code>/dev/odm</code> mount options in <code>/etc/vfstab</code>. If the “<code>nocluster</code>” option is being used, it can force the “<code>disabled</code>” cluster support state.</li> <li>■ Make sure the <code>VRTSgms</code> (group messaging service) package is installed.</li> </ul> |

If `/dev/odm` is not mounted, no status can be reported.

## Troubleshooting VCSIPC

Headings indicate likely symptoms or procedures required for a solution.

### VCSIPC wait warning messages in Oracle trace/log files

When Gigabit Ethernet interconnections are used, a high load can cause LMX/LLT to flow-control VCSIPC, resulting in warning messages to be reported in the Oracle trace file. The default location for the trace file is `$ORACLE_HOME/rdbms/log`; it may have changed if the parameters `background_dump_dest` or `user_dump_dest` have been changed. The messages resemble:

```
.
Unix process pid; 9560, image: oracle@MCB4800 (LMS0)
*** 2003-03-22 10:18:46.370
*** SESSION ID:(5.1) 2003-03-22 10:18:44.387
VCSIPC wait: WARNING: excessive poll done, 1001 times
VCSIPC wait: WARNING: excessive poll done, 1001 times
VCSIPC wait: WARNING: excessive poll done, 1002 times
VCSIPC wait: WARNING: excessive poll done, 1003 times
VCSIPC wait: WARNING: excessive poll done, 1004 times
VCSIPC wait: WARNING: excessive poll done, 1005 times
.
```

As a workaround, you can change the LLT lowwater mark, highwater mark, and window values for flow control. Please contact Veritas support for more information about changing these values.

### VCSIPC errors in Oracle trace/log files

If you see any VCSIPC errors in the Oracle trace/log files, check `/var/adm/` messages for any LMX error messages. If you see messages that contain any of the following:

```
. . . out of buffers
. . . out of ports
. . . no minors available
```

See [“About LMX tunable parameters”](#) on page 408.

If you see any VCSIPC warning messages in Oracle trace/log files that resemble:  
connection invalid

or,

```
Reporting communication error with node
```

check whether the Oracle Real Application Cluster instance on the other node is still running or has been restarted. The warning message indicates that the VCSIPC/LMX connection is no longer valid.

# Troubleshooting CVM

Headings indicate likely symptoms or procedures required for a solution.

## Shared disk group cannot be imported

If you see a message resembling:

```
vxvm:vxconfigd:ERROR:vold_pgr_register(/dev/vx/rdmp/disk_name) :
local_node_id<0
Please make sure that CVM and vxfen are configured and operating
correctly
```

This message is displayed when CVM cannot retrieve the node ID of the local system from the vxfen driver. This usually happens when port b is not configured. Verify that the vxfen driver is configured by checking the GAB ports with the command:

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

Port b must exist on the local system.

## CVMVolDg does not go online even though CVMCluster is online

When the CVMCluster resource goes online, the shared disk groups are automatically imported. If the disk group import fails for some reason, the CVMVolDg resources fault. Clearing and offlining the CVMVolDg type resources does not fix the problem.

### To workaround the issue

- 1 Fix the problem causing the import of the shared disk group to fail.
- 2 Offline the service group containing the resource of type CVMVolDg as well as the service group containing the CVMCluster resource type.
- 3 Bring the service group containing the CVMCluster resource online.
- 4 Bring the service group containing the CVMVolDg resource online.

# Troubleshooting interconnects

Headings indicate likely symptoms or procedures required for a solution.

## Restoring communication between host and disks after cable disconnection

If a fiber cable is inadvertently disconnected between the host and a disk, you can restore communication between the host and the disk without restarting.

### To restore communication between host and disks

- 1 Reconnect the cable.
- 2 Use the `format` command to verify that the host sees the disks. It may take a few minutes before the host is capable of seeing the disk.
- 3 Issue the following `vxctl` command to force the VxVM configuration daemon `vxconfigd` to rescan the disks:

```
vxctl enable
```

## Troubleshooting SF Oracle RAC checkpoint

Headings indicate likely symptoms or procedures required for a solution.

### SF Oracle RAC checkpoint feature requires DBA group to have MLOCK privileges

If MLOCK privileges are not set for the dba group, the checkpoint creation may fail with the following error:

```
Could not find ascii control file under
oracle home/admin/dbname/udump
```

To fix this issue, you must set the MLOCK privilege for the dba user.

See [“Set MLOCK privilege for DBA user”](#) on page 343.

# SF Oracle RAC Reference

- [Appendix A, “Sample VCS configuration files for SF Oracle RAC” on page 355](#)
- [Appendix B, “SF Oracle RAC agents” on page 369](#)
- [Appendix C, “Creating a starter database” on page 381](#)
- [Appendix D, “I/O fencing testing and scenarios” on page 387](#)
- [Appendix E, “Configuring the Symantec License Inventory Agent” on page 401](#)
- [Appendix F, “Tunable kernel driver parameters” on page 407](#)
- [Appendix G, “Error messages” on page 411](#)



# Sample VCS configuration files for SF Oracle RAC

This appendix contains the following topics:

- [About sample main.cf files](#)
- [Sample main.cf for Oracle 10g without Oracle agent](#)
- [Sample main.cf for Oracle 10g with Oracle agent](#)
- [Sample main.cf for Oracle 10g for CVM/VVR primary site](#)
- [Sample main.cf for Oracle 10g for CVM/VVR secondary site](#)

## About sample main.cf files

You can examine the VCS configuration file, `main.cf`, to verify SF Oracle RAC installation and configuration. The `main.cf` file is located in the folder `/etc/VRTSvcs/conf/config`.

- All sample configuration assume that Oracle binaries are installed on local disks and that they are managed by the operating system. These file systems must be specified in the file `/etc/fstab`.
- For Oracle 10g, the sample configurations assume that CRS binaries are installed on local disks and that they are managed by the operating system. These file systems must be specified in the file `/etc/fstab`.
- The “cluster” definition in all of the configurations must specify `UseFence=SCSI3`.

Review the following sample configuration files:

## Sample main.cf for Oracle 10g without Oracle agent

Configuration details:

- Configuration file name: 10g\_simple\_main.cf
- Use for single Oracle 10g database only
- Has only one parallel service group: cvm
- cvm group includes PrivNIC and Application resource for CSSD

```
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"
include "OracleTypes.cf"
include "PrivNIC.cf"

cluster rac_cluster101 (
 UserNames = { admin = bopHo }
 Administrators = { admin }
 UseFence = SCSI3
)

system galaxy (
)

system nebula (
)

group cvm (
 SystemList = { galaxy = 0, nebula = 1 }
 AutoFailOver = 0
 Parallel = 1
 AutoStartList = { galaxy, nebula }
)

Application cssd (
 Critical = 0
 StartProgram = "/opt/VRTSvcs/rac/bin/cssd-online"
 StopProgram = "/opt/VRTSvcs/rac/bin/cssd-offline"
 CleanProgram = "/opt/VRTSvcs/rac/bin/cssd-clean"
 MonitorProgram = "/opt/VRTSvcs/rac/bin/cssd-monitor"
 OnlineRetryLimit = 20
)

CVMVolDg ocrvote_voldg (
 Critical = 0
 CVMDiskGroup = ocrvotedg
 CVMVolume = { ocrvol, votevol }
 CVMActivation = sw
)
```

```
CFSMount oradata_mnt (
 Critical = 0
 MountPoint = "/oradata"
 BlockDevice = "/dev/vx/dsk/oradatadg/oradatavol"
)

CVMVoldg oradata_voldg (
 Critical = 0
 CVMDiskGroup = oradatadg
 CVMVolume = { oradatavol }
 CVMActivation = sw
)

CFSfsckd vxfsckd (
)

CVMCluster cvm_clus (
 CVMClustName = rac_cluster101
 CVMNodeId = { galaxy = 1, nebula = 2 }
 CVMTransport = gab
 CVMTimeout = 200
)

CVMVxconfigd cvm_vxconfigd (
 Critical = 0
 CVMVxconfigdArgs = { syslog }
)

PrivNIC ora_priv (
 Critical = 0
 Device = { lan1 = 0, lan2 = 2 }
 Address@galaxy = "192.168.12.1"
 Address@nebula = "192.168.12.2"
 NetMask = "255.255.240.0"
)

cssd requires ocrvote_voldg
cssd requires oradata_mnt
cssd requires ora_priv

oradata_mnt requires oradata_voldg

oradata_voldg requires cvm_clus
ocrvote_voldg requires cvm_clus
oradata_mnt requires vxfsckd

cvm_clus requires cvm_vxconfigd
```

## Sample main.cf for Oracle 10g with Oracle agent

Configuration details:

- Configuration file name: 10g\_main.cf
- For multiple Oracle 10g databases
- Has two parallel service groups: cvm and oradb1\_grp
- oradb1\_grp depends on cvm
- oradb1\_grp has Oracle and oradata mount resource

```
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"
include "OracleTypes.cf"
include "PrivNIC.cf"

cluster rac_cluster101 (
 UserNames = { admin = bopHo }
 Administrators = { admin }
 UseFence = SCSI3
)

system galaxy (
)

system nebula (
)

group oradb1_grp (
 SystemList = { galaxy = 0, nebula = 1 }
 AutoFailOver = 0
 Parallel = 1
 AutoStartList = { galaxy, nebula }
)

Oracle ora1 (
 Critical = 0
 Sid @galaxy = vrts1
 Sid @nebula = vrts2
 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 (
 CVMDiskGroup = oradatadg
 CVMVolume = { oradatavol }
```

```
 CVMActivation = sw
)

requires group cvm online local firm
oral requires oradata_mnt
oradata_mnt requires oradata_voldg

group cvm (
 SystemList = { galaxy = 0, nebula = 1 }
 AutoFailOver = 0
 Parallel = 1
 AutoStartList = { galaxy, nebula }
)

Application cssd (
 Critical = 0
 StartProgram = "/opt/VRTSvcs/rac/bin/cssd-online"
 StopProgram = "/opt/VRTSvcs/rac/bin/cssd-offline"
 CleanProgram = "/opt/VRTSvcs/rac/bin/cssd-clean"
 MonitorProgram = "/opt/VRTSvcs/rac/bin/cssd-monitor"
 OnlineRetryLimit = 20
)

CVMVoldg ocrvote_voldg (
 Critical = 0
 CVMDiskGroup = ocrvotedg
 CVMVolume = { ocrvol, votevol }
 CVMActivation = sw
)

CFSfsckd vxfsckd (
)

CVMCluster cvm_clus (
 CVMClustName = rac_cluster101
 CVMNodeId = { galaxy = 1, nebula = 2 }
 CVMTransport = gab
 CVMTimeout = 200
)

CVMVxconfigd cvm_vxconfigd (
 Critical = 0
 CVMVxconfigdArgs = { syslog }
)

PrivNIC ora_priv (
 Critical = 0
 Device = { lan1 = 0, lan2 = 1 }
 Address@galaxy = "192.168.12.1"
 Address@nebula = "192.168.12.2"
 NetMask = "255.255.240.0"
```

```
cssd requires ocrvote_voldg
cssd requires ora_priv
ocrvote_voldg requires cvm_clus
cvm_clus requires cvm_vxconfigd
```

## Sample main.cf for Oracle 10g for CVM/VVR primary site

Configuration details:

- Configuration file name: cvmvvr\_primary\_main.cf
- More general purpose, can have multiple Oracle databases

```
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"
include "OracleTypes.cf"
include "PrivNIC.cf"
include "/etc/VRTSvcs/conf/config/VVRTypes.cf"

cluster rac_cluster101 (
 UserNames = { admin = bopHo }
 ClusterAddress = "10.10.10.101"
 Administrators = { admin }
 UseFence = SCSI3
)

remoteclass rac_cluster102 (
 ClusterAddress = "10.11.10.102"
)

heartbeat Icmp (
 ClusterList = { rac_cluster102 }
 Arguments @rac_cluster102 = { "10.11.10.102" }
)

system galaxy (
)

system nebula (
)

group ClusterService (
 SystemList = { galaxy = 0, nebula = 1 }
 AutoStartList = { galaxy, nebula }
 OnlineRetryLimit = 3
 OnlineRetryInterval = 120
)

Application wac (
```

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

IP gcoip (
 Device = lan0
 Address = "10.10.10.101"
 NetMask = "255.255.240.0"
)

NIC csgnic (
 Device = lan0
 NetworkHosts = { "10.10.12.2", "10.10.12.3" }
)

gcoip requires csgnic
wac requires gcoip

group RVGgroup (
 SystemList = { galaxy = 0, nebula = 1 }
 Parallel = 1
 AutoStartList = { galaxy, nebula }
)

CVMVoldg racdata_voldg (
 CVMDiskGroup = oradatadg
 CVMActivation = sw
)

RVGShared racdata_rvg (
 RVG = rac1_rvg
 DiskGroup = oradatadg
)

requires group cvm online local firm
racdata_rvg requires racdata_voldg

group cvm (
 SystemList = { galaxy = 0, nebula = 1 }
 AutoFailOver = 0
 Parallel = 1
 AutoStartList = { galaxy, nebula }
)

Application cssd (
 Critical = 0
 StartProgram = "/opt/VRTSvcs/ops/bin/cssd-online"
 StopProgram = "/opt/VRTSvcs/ops/bin/cssd-offline"
 CleanProgram = "/opt/VRTSvcs/ops/bin/cssd-clean"
 MonitorProgram = "/opt/VRTSvcs/ops/bin/cssd-monitor"
```

```
 OnlineRetryLimit = 20
)

CFSfsckd vxfsckd (
)

CVMCluster cvm_clus (
 CVMClustName = rac_cluster101
 CVMNodeId = { galaxy = 1, nebula = 2 }
 CVMTransport = gab
 CVMTimeout = 200
)

CVMVoldg ocrvote_voldg (
 Critical = 0
 CVMDiskGroup = ocrvotedg
 CVMVolume = { ocrvol, votevol }
 CVMActivation = sw
)

CVMVxconfigd cvm_vxconfigd (
 Critical = 0
 CVMVxconfigdArgs = { syslog }
)

PrivNIC ora_priv (
 Critical = 0
 Device = { lan1 = 0, lan2 = 1}
 Address@galaxy = "192.168.12.1"
 Address@nebula = "192.168.12.2"
 NetMask = "255.255.240.0"

 cssd requires ocrvote_voldg
 cssd requires ora_priv
 ocrvote_voldg requires cvm_clus
 vxfsckd requires cvm_clus
 cvm_clus requires cvm_vxconfigd
)

group oradb1_grp (
 SystemList = { galaxy = 0, nebula = 1 }
 Parallel = 1
 ClusterList = { rac_cluster101 = 0, rac_cluster102 = 1 }
 OnlineRetryInterval = 300
 ClusterFailOverPolicy = Manual
 AutoStartList = { galaxy, nebula }
)

CFSMount oradata_mnt (
 MountPoint = "/oradata"
 BlockDevice = "/dev/vx/dsk/oradatadg/oradatavol"
)

```

```
Oracle ora1 (
 Critical = 0
 Sid @galaxy = vrts1
 Sid @nebula = vrts2
 Owner = oracle
 Home = "/app/oracle/orahome"
 StartUpOpt = SRVCTLSTART
 ShutDownOpt = SRVCTLSTOP
)

RVGSharedPri ora_vvr_sharedpri (
 RvgResourceName = racdata_rvg
 OnlineRetryLimit = 0
)

requires group RVGgroup online local firm
ora1 requires oradata_mnt
oradata_mnt requires ora_vvr_sharedpri

group rlogowner (
 SystemList = { galaxy = 0, nebula = 1 }
 AutoStartList = { galaxy, nebula }
 OnlineRetryLimit = 2
)

IP logowner_ip (
 Device = lan0
 Address = "10.10.9.101"
 NetMask = "255.255.240.0"
)

NIC nic (
 Device = lan0
)

RVGLogowner logowner (
 RVG = rac1_rvg
 DiskGroup = oradatadg
)

requires group RVGgroup online local firm
logowner requires logowner_ip
logowner_ip requires nic
```

## Sample main.cf for Oracle 10g for CVM/VVR secondary site

Configuration details:

- Configuration file name: cvmvvr\_secondary\_main.cf

■ More general purpose, can have multiple Oracle databases

```
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"
include "OracleTypes.cf"
include "PrivNIC.cf"
include "/etc/VRTSvcs/conf/config/VVRTypes.cf"

cluster rac_cluster102 (
 UserNames = { admin = bopHo }
 ClusterAddress = "10.11.10.102"
 Administrators = { admin }
 UseFence = SCSI3
)

remotecluster rac_cluster101 (
 ClusterAddress = "10.10.10.101"
)

heartbeat Icmp (
 ClusterList = { rac_cluster101 }
 Arguments @rac_cluster101 = { "10.10.10.101" }
)

system mercury (
)

system jupiter (
)

group ClusterService (
 SystemList = { mercury = 0, jupiter = 1 }
 AutoStartList = { mercury, jupiter }
 OnlineRetryLimit = 3
 OnlineRetryInterval = 120
)

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

IP gcoip (
 Device = lan0
 Address = "10.11.10.102"
 NetMask = "255.255.240.0"
)

NIC csgnic (
 Device = lan0
```

```

 NetworkHosts = { "10.10.12.2", "10.10.12.3" }
)

 gcoip requires csgnic
 wac requires gcoip

group RVGgroup (
 SystemList = { mercury = 0, jupiter = 1 }
 Parallel = 1
 AutoStartList = { mercury, jupiter }
)

CVMVolDg racdata_voldg (
 CVMDiskGroup = oradatadg
 CVMActivation = sw
)

RVGShared racdata_rvg (
 RVG = rac1_rvg
 DiskGroup = oradatadg
)

requires group cvm online local firm
racdata_rvg requires racdata_voldg

group cvm (
 SystemList = { mercury = 0, jupiter = 1 }
 AutoFailOver = 0
 Parallel = 1
 AutoStartList = { mercury, jupiter }
)

Application cssd (
 Critical = 0
 StartProgram = "/opt/VRTSvcs/ops/bin/cssd-online"
 StopProgram = "/opt/VRTSvcs/ops/bin/cssd-offline"
 CleanProgram = "/opt/VRTSvcs/ops/bin/cssd-clean"
 MonitorProgram = "/opt/VRTSvcs/ops/bin/cssd-monitor"
 OnlineRetryLimit = 20
)

CFSfsckd vxfsckd (
)

CVMCluster cvm_clus (
 CVMClustName = rac_cluster102
 CVMNodeId = { mercury = 1, jupiter = 2 }
 CVMTransport = gab
 CVMTimeout = 200
)

CVMVolDg ocrvote_voldg (

```

```
 Critical = 0
 CVMDiskGroup = ocrvotedg
 CVMVolume = { ocrvol, votevol }
 CVMActivation = sw
)

CVMVxconfigd cvm_vxconfigd (
 Critical = 0
 CVMVxconfigdArgs = { syslog }
)

PrivNIC ora_priv (
 Critical = 0
 Device = { lan1 = 0, lan2 = 1 }
 Address@mercury = "192.168.12.3"
 Address@jupiter = "192.168.12.4"
 NetMask = "255.255.240.0"
)

cssd requires ocrvote_voldg
cssd requires ora_priv
ocrvote_voldg requires cvm_clus
vxfsckd requires cvm_clus
cvm_clus requires cvm_vxconfigd

group oradbl_grp (
 SystemList = { mercury = 0, jupiter = 1 }
 Parallel = 1
 ClusterList = { rac_cluster101 = 0, rac_cluster102 = 1 }
 OnlineRetryInterval = 300
 ClusterFailOverPolicy = Manual
 AutoStartList = { mercury, jupiter }
)

CFSMount oradata_mnt (
 MountPoint = "/oradata"
 BlockDevice = "/dev/vx/dsk/oradatadg/oradatavol"
)

Oracle ora1 (
 Critical = 0
 Sid @mercury = vrts1
 Sid @jupiter = vrts2
 Owner = oracle
 Home = "/app/oracle/orahome"
 StartUpOpt = SRVCTLSTART
 ShutDownOpt = SRVCTLSTOP
)

RVGSharedPri ora_vvr_sharedpri (
 RvgResourceName = racdata_rvg
 OnlineRetryLimit = 0
)
```

```
)

 requires group RVGgroup online local firm
 oral requires oradata_mnt
 oradata_mnt requires ora_vvr_sharedpri

group rlogowner (
 SystemList = { mercury = 0, jupiter = 1 }
 AutoStartList = { mercury, jupiter }
 OnlineRetryLimit = 2
)

IP logowner_ip (
 Device = lan0
 Address = "10.11.9.102"
 NetMask = "255.255.240.0"
)

NIC nic (
 Device = lan0
)

RVGLogowner logowner (
 RVG = rac1_rvg
 DiskGroup = oradatadg
)

requires group RVGgroup online local firm
logowner requires logowner_ip
logowner_ip requires nic
```



# SF Oracle RAC agents

This appendix contains the following topics:

- [About configuring the VCS agents in SF Oracle RAC](#)
- [About CVMCluster agent](#)
- [About CVMVxconfigd agent](#)
- [About the CVMVolDg agent](#)
- [About the CFSSMount agent](#)
- [About PrivNIC agent](#)
- [About CSSD agent](#)

## About configuring the VCS agents in SF Oracle RAC

VCS agents in SF Oracle RAC include:

- CVMCluster agent
- CVMVxconfigd agent
- CVMVolDG agent
- CFSSMount agent

The SF Oracle RAC installation automatically configures the CVMCluster resource and the CVMVxconfigd resource.

You must configure the CVMVolDg agent for each disk group that is used by an Oracle service group. Configure a disk group for only a single Oracle service group. If the database uses cluster file systems, configure the CFSSMount agent for each volume in the disk group.

You can use the information on the entry points and attributes of the CVMCluster, CVMVxconfigd, CVMVolDg, and CFSSMount agents to make necessary changes to the configuration.

Refer to the *Veritas Cluster Server User's Guide* for information on how to modify the VCS configuration.

## About CVMCluster agent

The CVMCluster agent controls system membership on the cluster port associated with VxVM.

### Entry points for CVMCluster agent

The entry points for the CVMCluster agent include:

- Online        Joins a system to the CVM cluster port. Enables Volume Manager cluster functionality by autoimporting shared disk groups.
- Offline       Removes a system from the CVM cluster port.
- Monitor       Monitors the state of CVM cluster membership for a system.

### Attribute definition for CVMCluster agent

[Table B-7](#) describes modifiable attributes of the CVMCluster resource type.

**Table B-7**        CVMCluster agent attributes

Attribute	Dimension	Description
CVMClustName	string-scalar	Name of the cluster.
CVMNodeId{}	string-association	Associative list. The first part names the system; the second part contains the LLT ID number for the system.
CVMTransport	string-scalar	Cluster messaging mechanism. Default = gab
CVMTimeout	integer-scalar	Timeout in seconds for CVM cluster reconfiguration. Default = 200

### Resource type definition for CVMCluster agent

Review the type definition that the CVMTypes.cf file includes. The CVMNodeAddr, PortConfigd, and PortKmsgd attributes are not applicable in an SF Oracle RAC environment because GAB, the required mechanism for cluster communication, does not use these attributes.

```
type CVMCluster (
 static int NumThreads = 1
 static int OnlineRetryLimit = 2
 static int OnlineTimeout = 400
 static str ArgList[] = { CVMTransport, CVMClustName,
 CVMNodeAddr, CVMNodeId, PortConfigd, PortKmsgd,
 CVMTimeout }
 NameRule = ""
 str CVMClustName
 str CVMNodeAddr{}
 str CVMNodeId{}
 str CVMTransport
 int PortConfigd
 int PortKmsgd
 int CVMTimeout
)
```

## Sample configuration for CVMCluster agent

You can configure a resource of type CVMCluster in the CVM service group as shown in the sample definition. Also, review the extensive main.cf that includes the CVMCluster resource.

See “[Sample VCS configuration files for SF Oracle RAC](#)” on page 355.

```
CVMCluster cvm_clus (
 Critical = 0
 CVMClustName = RACcluster1
 CVMNodeId = { galaxy = 0, nebula = 1 }
 CVMTransport = gab
 CVMTimeout = 200
)
```

## About 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 configuration information 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.

Symantec 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.

## Entry points for CVMVxconfigd agent

The entry points for the CVMVxconfigd agent include:

Online	Starts the vxconfigd daemon.
Offline	N/A.
Monitor	Monitors whether vxconfigd daemon is running.

## Attribute definition for CVMVxconfigd agent

[Table B-8](#) describes modifiable attributes of the CVMVxconfigd resource type.

**Table B-8** CVMVxconfigd agent attribute

Attribute	Dimension	Description
CVMVxconfigdArgs	keylist	List of arguments sent to the <code>online</code> entry point. Symantec recommends always specifying the <code>syslog</code> option.

## Resource type definition for CVMVxconfigd agent

The `CVMTypes.cf` file includes the following type definition.

```
type CVMVxconfigd (
 static int FaultOnMonitorTimeouts = 2
 static int RestartLimit = 5
 static str ArgList[] { CVMVxconfigdArgs }
 static str Operations = OnOnly
 keylist CVMVxconfigdArgs
)
```

## Sample configuration for CVMVxconfigd agent

You can configure a resource of type `CVMVxconfigd` in the CVM service group as shown in the sample definition. Also, review the extensive `main.cf` that includes the `CVMVxconfigd` resource.

See [Appendix A](#), “”, “[Sample VCS configuration files for SF Oracle RAC](#)” on page 355.

```
CVMVxconfigd cvm_vxconfigd (
 Critical = 0
 CVMVxconfigdArgs = { syslog }
)
.
.
cvm_clus requires cvm_vxconfigd
```

```
// resource dependency tree
//
// group cvm
// {
// CVMCluster cvm_clus
// {
// CVMVxconfigd cvm_vxconfigd
// }
// }
```

## About the CVMVolDg agent

The CVMVolDg agent represents and controls CVM disk groups and CVM volumes within the disk groups. The global nature of CVM disk groups and volumes requires importing them only once on the CVM master node.

Configure the CVMVolDg agent for each disk group used by an Oracle service group. Configure a disk group for only a single Oracle service group. If the database uses cluster file systems, configure the CFMount agent for each volume in the disk group.

## Entry points for CVMVolDg agent

The entry points for the CVMVolDg agent include:

Online	Starts all volumes in the shared disk group specified by the CVMVolume attribute. Sets the disk group activation mode to shared-write if the value of the CVMActivation attribute is sw. You can set the activation mode on both slave and master systems.
Offline	Cleans up the temporary files created by the online entry point.
Monitor	Monitors specified critical volumes in the disk group. The CVMVolume attribute specifies these volumes. SF Oracle RAC requires specifying at least one volume in a disk group.
Clean	Cleans up the temporary files created by the online entry point.

## Attribute definition for CVMVolDg agent

[Table B-9](#) describes modifiable attributes of the CVMVolDg resource type.

**Table B-9** CVMVolDg agent attributes

Attribute	Dimension	Description
CVMDiskGroup	string-scalar	Names the disk group.

**Table B-9** CVMVolDg agent attributes

Attribute	Dimension	Description
CVMVolume	string-keylist	Lists critical volumes in the disk group. SF Oracle RAC requires specifying at least one volume in the disk group.
CVMActivation	string-scalar	Sets the activation mode for the disk group. Default = sw

## Resource type definition for CVMVolDg agent

The CVMTypes.cf file includes the CVMVolDg type definition.

```

type CVMVolDg (
 static keylist RegList = { CVMActivation }
 static str ArgList[] = { CVMDiskGroup, CVMVolume,
 CVMActivation }
 str CVMDiskGroup
 keylist CVMVolume[]
 str CVMActivation
 temp int voldg_stat
)

```

## Sample configuration for CVMVolDg agent

Each Oracle service group requires a defined CVMVolDg resource type.

```

CVMVolDg ora_voldg (
 CVMDiskGroup = oradata1
 CVMVolume = { oradata1, oradata2 }
 CVMActivation = sw
)

```

## About the CFSSMount agent

The CFSSMount agent brings online, takes offline, and monitors a cluster file system mount point. The agent executable is /opt/VRTSvcs/bin/CFSSMount/CFSSMountAgent. The CFSSMount type definition is in the /etc/VRTSvcs/conf/config/CFSSTypes.cf file.

## Entry points for CFSSMount agent

The entry points for the CFSSMount agent include:

Online           Mounts a block device in cluster mode.

Offline	Unmounts the file system, forcing the process if necessary, and sets primary to secondary if necessary.
Monitor	Determines whether the file system is mounted. Checks mount status using the <code>fsclustadm</code> command.
Clean	Generates a null operation for a cluster file system mount.

## Attribute definition for CFSMount agent

[Table B-10](#) lists modifiable attributes of the CFSMount Agent resource type.

**Table B-10** CFSMount agent attributes

Attribute	Dimension	Description
MountPoint	string-scalar	Directory for the mount point.
BlockDevice	string-scalar	Block device for the mount point.
NodeList (optional)	string-keylist	List of nodes on which to mount. If NodeList is NULL, the agent uses the service group system list.
MountOpt (optional)	string-scalar	Options for the mount command. To create a valid MountOpt attribute string: <ul style="list-style-type: none"> <li>■ Use the VxFS type-specific options only.</li> <li>■ Do not use the <code>-o</code> flag to specify the VxFS-specific options.</li> <li>■ Do not use the <code>-F</code> VxFS file system type option.</li> <li>■ Be aware the cluster option is not required.</li> <li>■ Specify options in comma-separated list.</li> </ul> For example: <pre>ro ro,cluster blkclear,mincache=closesync</pre>
Policy (optional)	string-scalar	List of nodes to assume primaryship of the cluster file system if the primary node fails. If set to NULL or if none of the hosts specified in the list is active when the primary node fails, a node is randomly selected from the set of active nodes to assume primaryship.
Primary (Not set by user)	string-scalar	Information only. Stores the primary node name for a VxFS file system. The value is automatically modified in the configuration file when an unmounted file system is mounted or another node becomes the primary node. The user does not set this attribute and user programs do not rely on this attribute.

## Resource type Definition for CFSSMount agent

The CFSTypes.cf file includes the CFSSMount agent type definition.

```
type CFSSMount (
 static keylist RegList = { MountOpt, Policy, NodeList }
 static int FaultOnMonitorTimeouts = 1
 static int OnlineWaitLimit = 0
 static str ArgList[] = { MountPoint, BlockDevice,
 MountOpt }
 NameRule = resource.MountPoint
 str MountPoint
 str MountType
 str BlockDevice
 str MountOpt
 str Primary
 keylist NodeList
 keylist Policy
)
```

## Sample configuration for CFSSMount agent

Each Oracle service group requires a defined CFSSMount resource type.

```
CFSSMount ora_mount (
 MountPoint = "/oradata"
 BlockDevice = "/dev/vx/dsk/oradatadg/oradatavol1"
 Primary = nebula
)
```

## About PrivNIC agent

The PrivNIC resource can be used to maintain a “private IP address” that is locally highly available on LLT Ethernet interfaces. Such private IP addresses are required by the CRS daemons in Oracle10g to provide communication.

The PrivNIC agent relies on LLT to monitor the interfaces. It queries LLT to count the number of visible nodes on each of the LLT interfaces.

## Entry point for PrivNIC agent

The following table describes the `monitor` entry point used by the PrivNIC agent.

Entry Point	Description
Monitor	Queries LLT to make a list of nodes visible on every LLT network interface. It applies various filters to the list to arrive at a most desired failover decision and calculates a “winner” device on which to configure the IP address. The “winner” is compared to the currently active device where the IP is currently configured; if the active and winner device are different, the agent fails over the device.

## Attribute definition for PrivNIC agent

The following table describes the user-modifiable attributes of the PrivNIC resource type:

### Required Attributes

Attribute	Dimension	Description
Device	string - association	Specifies the network interface device as shown by the “ifconfig” command and the “network-id” associated with the interface; for example: <code>lan0=0</code> . Network-ids of the interfaces connected to the same physical network must match. The interface with the lower network-id has the higher preference for failover. Interfaces specified in the PrivNIC configuration must be exactly the same in name and total number as those which have been used for LLT configuration. Example: <pre>Device@galaxy = {lan0=0, lan1=1, lan2=2} Device@nebula = {lan0=0, lan1=1, lan2=2}</pre>
Address	string-scalar	The numerical private IP address. For example: <pre>Address = "192.168.12.1"</pre>
NetMask	string - association	The numerical netmask for the private IP address. For example: <pre>Address = "255.255.255.0"</pre>

## Optional Attributes

Attribute	Dimension	Description
DeviceTag	string - association	<p>Associates an LLT device “tag” with device via the network-id. If an LLT device tag (as specified in the <code>/etc/llttab</code> file) differs from the name of the network interface as shown in “ifconfig,” then DeviceTag must be specified for that interface.</p> <p>For example: in the common case, <code>/etc/llttab</code> contains:</p> <pre>link lan0 /dev/lan:0 - ether - - link lan1 /dev/lan:1 - ether - - link-lowpri lan0 /dev/lan:0 - ether - -</pre> <p>In the above case, DeviceTag does not need to be specified. However, if <code>/etc/llttab</code> contains:</p> <pre>link link1 /dev/lan:0 - ether - - link link2 /dev/lan:1 - ether - - link-lowpri spare /dev/lan:0 - ether - -</pre> <p>And,</p> <pre>Device@galaxy = { lan=0, lan1=1, lan2=2 }</pre> <p>DeviceTag needs to be specified as:</p> <pre>DeviceTag@galaxy = { spare=2 }</pre>
GabPort	string-scalar	<p>A single lower-case letter specifying the name of the GAB port to be used for filtering. “o” is the default. NULL disables GAB port filtering.</p> <p>Example: <code>GabPort = "b"</code></p>
UseVirtualIP	integer-scalar	<p>The default is 0, which specifies that the agent use the physical interface for configuring the private IP address when possible.</p> <p>The value 1 specifies that the agent always use the virtual interface for configuring the private IP address.</p> <p>The value 2 (which includes the functionality of the value 1) specifies the agent should complain if the private IP address already exists on a physical interface.</p>
UseSystemList	integer-scalar	<p>The value 1 specifies that the agent use the SystemList of the service group to filter the node list. The default is 0.</p>

Attribute	Dimension	Description
ExcludeNode	integer-vector	List of nodes to be permanently excluded from calculation.

## Resource type definition for PrivNIC agent

The following shows the content of the `PrivNIC.cf` file:

```
type PrivNIC (
 static str ArgList[] = { Device, DeviceTag, Address,
 NetMask, UseVirtualIP, GabPort,
 UseSystemList,
 ExcludeNode }
 static int OfflineMonitorInterval = 60
 static int MonitorTimeout = 300
 static str Operations = None

 str Device{}
 str DeviceTag{}
 str Address = ""
 str NetMask = ""
 int UseVirtualIP = 0
 str GabPort = "o"
 int UseSystemList = 0
 int ExcludeNode[]
)
```

## Sample configuration for PrivNIC agent

The following is a sample configuration using the PrivNIC agent.

```
group cvm (
 SystemList = { galaxy = 0, nebula = 1 }
 AutoFailOver = 0
 Parallel = 1
 AutoStartList = { galaxy, nebula }
)

PrivNIC ora_priv (
 Device = { lan0 = 0, lan1 = 1, lan2 = 5 }
 Address@galaxy = "192.168.12.1"
 Address@nebula = "192.168.12.2"
 NetMask = "255.255.255.0"
)
```

## About CSSD agent

The `cssd` resource is optional. It monitors the Oracle 10g `cssd` process. The purpose of the `cssd` resource is to ensure that the dependency of `cssd` on the OCR and VOTE resources and the PrivNIC (optional) resource are satisfied. If the `cssd` resource is online and any of its dependencies are brought offline, the machine will reboot. This agent allows this behavior to be avoided since the dependencies will be enforced by VCS.

The `cssd` resource should use the Application agent. The name of the resource is up to the user. The following are required attributes of the `cssd` resource:

Attribute Name	Required Value
<code>Critical</code>	<code>0</code>
<code>OnlineRetryLimit</code>	<code>20</code>
<code>StartProgram</code>	<code>/opt/VRTSvcs/rac/bin/cssd-online</code>
<code>StopProgram</code>	<code>/opt/VRTSvcs/rac/bin/cssd-offline</code>
<code>CleanProgram</code>	<code>/opt/VRTSvcs/rac/bin/cssd-clean</code>
<code>MonitorProgram</code>	<code>/opt/VRTSvcs/rac/bin/cssd-monitor</code>

An example `main.cf` entry is as follows:

```
Application cssd-resource (
 Critical = 0
 StartProgram = "/opt/VRTSvcs/rac/bin/cssd-online"
 StopProgram = "/opt/VRTSvcs/rac/bin/cssd-offline"
 CleanProgram = "/opt/VRTSvcs/rac/bin/cssd-clean"
 MonitorProgram = "/opt/VRTSvcs/rac/bin/cssd-monitor"
 OnlineRetryLimit = 20
)
```

# Creating a starter database

This appendix contains the following topics:

- [About creating a starter database](#)
- [Creating a starter database for Oracle 10g](#)

## About creating a starter database

The optional procedures to create a starter database describe methods for creating a starter database on shared storage.

Use your own tools or refer to the guidelines below on using the Oracle dbca (Database Creation Assistant) tool to create a database on shared raw VxVM volumes or shared VxFS file systems.

If you installed the Oracle binary on shared storage (CFS), Symantec recommends using a different file system for the Oracle datafiles. The Storage Foundation for Oracle RAC configuration wizard does not support the use of a single file system for the binary and the datafiles.

## Creating a starter database for Oracle 10g

Create the database in a shared raw volume or a cluster file system. Before you begin, review these prerequisites:

- CRS daemons must be running.  
To verify the status of CRS, type:  

```
$CRS_HOME/bin/crs_stat
```
- 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.

Depending on where you want to create the starter database, review one of the procedures:

- [Creating Oracle 10g database on raw volumes](#)
- [Creating Oracle 10g database on CFS](#)

Refer to the *Oracle Real Application Clusters Installation and Configuration Guide* for instructions on how to install the Oracle 10g database. You can create the database in a shared raw volume or in a cluster file system. The following paragraphs describe creating volumes for the database tablespaces or creating a file system for the database.

## Creating Oracle 10g database on raw volumes

Creating a starter database on raw VxVM volumes involves the following tasks:

- [Creating shared raw volume for database tablespaces](#)
- [Running the dbca utility for raw VxVM volumes](#)

### Creating shared raw volume for database tablespaces

You can create the database tablespaces on shared raw volume.

#### To create shared raw volume for database tablespaces

- 1 Log in as superuser.
- 2 On the master node, create a shared disk group:  

```
vxdg -s init oradatadg c4t1d2
```
- 3 Create a volume in the shared group for *each* of the required tablespaces. (Refer to the Oracle documentation to determine the tablespace requirements.) For example:  

```
vxassist -g oradatadg make oradata_system1 1000M
vxassist -g oradatadg make oradata_spfile1 10M
.
```
- 4 Deport and import the shared disk group to enable write access to it and to enable I/O fencing. On the CVM master node, enter:  

```
vxdg deport oradatadg
vxdg -s import oradatadg
vxvol -g oradatadg startall
vxdg -g oradatadg set activation=sw
```
- 5 On each other node, enter:  

```
vxdg -g oradatadg set activation=sw
```

- 6 Define the access mode and permissions for the volumes storing the Oracle data. For *each* volume listed in \$ORACLE\_HOME/raw\_config, use the vxedit (1M) command:

```
vxedit -g disk_group set group=group user=user mode=660 volume
```

For example:

```
vxedit -g oradatadg set group=dba user=oracle mode=660
oradata_system1
```

In this example, oradata\_system1 is the name of one of the volumes. Repeat the command to define access mode and permissions for each volume in the oradatadg.

- 7 You can now create the database using Oracle documentation.

## Running the dbca utility for raw VxVM volumes

Use the dbca utility on the master node to create a general-purpose database on raw VxVM volumes. The utility is a graphical user interface and requires setting the DISPLAY environment variable.

### To create a database on raw VxVM volumes

- 1 Make sure an oracle account is created on all nodes.
- 2 Verify that remsh works among all nodes under the oracle account.
- 3 From one node, log in as oracle user.
- 4 Create and start the Oracle listener using the NETCA utility. Refer to the *Oracle10g Installation Guide* for more information.  

```
$ netca
```
- 5 Run the dbca utility to create the database. Refer to Oracle documentation for details on the dbca utility.

## Creating Oracle 10g database on CFS

Creating a starter database on CFS involves the following tasks:

- [Creating a Cluster File System for the Oracle 10g database](#)
- [Running the dbca utility for cluster file system](#)

### Creating a Cluster File System for the Oracle 10g database

If you plan to use a cluster file system to store the Oracle database, you can use the following procedure to create the file system.

#### To create Cluster File System for database tablespaces

- 1 Create a disk group (for example: rac\_dg) :

- ```
# vxdg -s init oradatadg c4t2d3
```
- 2 Create a single shared volume (for example: oradatavol), large enough to contain a file system for all the tablespaces (see Oracle documentation the tablespace sizes). Assuming 6.8 GB are required for the tablespaces:


```
# vxassist -g oradatadg make oradatavol 6800M
```
 - 3 Deport and import the group in the shared mode to enable I/O fencing:


```
# vxdg deport oradatadg
# vxdg -s import oradatadg
```
 - 4 Set the activation mode (sw) to allow shared access to the disk group:


```
# vxdg -g oradatadg set activation=sw
```
 - 5 Start the volume in the disk group:


```
# vxvol -g oradatadg startall
```
 - 6 Create a VxFS file system in this volume. From one node, enter:


```
# mkfs -F vxfs -o largefiles /dev/vx/rdisk/oradatadg/oradatavol
```
 - 7 Create a mount point for the shared file system:


```
# mkdir /oradata
```
 - 8 From the same system, mount the file system:


```
# mount -F vxfs -o cluster /dev/vx/dsk/oradatadg/oradatavol \
/oradata
```
 - 9 Set “oracle” to be the owner of the file system, and set “755” as the permissions:


```
# chown -R oracle:oinstall /oradata
# chmod 755 /oradata
```
 - 10 On the other system(s), do [step 4](#) and [step 7](#) through [step 9](#).
You can now create the database; refer to Oracle documentation.

Running the dbca utility for cluster file system

As oracle user, use the dbca utility on the master node to create a general purpose database on a cluster file system. This utility is a graphical user interface and requires setting the DISPLAY environment variable.

To create a database on cluster file system

- 1 Make sure an oracle account is created on all nodes.
- 2 Verify that remsh works among all the nodes under the oracle account.
- 3 From one node, log in as oracle user.
- 4 Create and start the Oracle listener using NETCA utility.
Refer to the *Oracle10g Installation Guide* for more information.
\$ netca

- 5 Run the dbca utility. When starting the utility with a cluster file system, use the `-datafileDestination` option to specify the mount point.

For example:

```
$ dbca -datafileDestination /rac_ts
```

Refer to the Oracle documentation for details on the dbca utility.

I/O fencing testing and scenarios

This appendix contains the following topics:

- [I/O fencing of shared storage](#)
- [Verifying data storage arrays using the vxfsentsthdw utility](#)
- [How I/O fencing works in different event scenarios](#)
- [About vxfsenadm utility](#)

I/O fencing of shared storage

When multiple systems have access to the data on shared storage, the integrity of the data depends on the systems communicating with each other so that each is aware when the other is writing data. Usually this communication occurs in the form of heartbeats through the private networks between the systems. If the private links are lost, or even if one of the systems is hung or too busy to send or receive heartbeats, each system could be unaware of the other's activities with respect to writing data. This is a *split brain* condition and can lead to data corruption.

The I/O fencing capability of the Storage Foundation for Oracle RAC, which is managed by Veritas Volume Manager, prevents data corruption in the event of a split brain condition by using SCSI-3 persistent reservations for disks. This allows a set of systems to have registrations with the disk and a write-exclusive registrants-only reservation with the disk containing the data. This means that only these systems can read and write to the disk, while any other system can only read the disk. The I/O fencing feature fences out a system that no longer sends heartbeats to the other system by preventing it from writing data to the disk.

VxVM manages all shared storage subject to I/O fencing. It assigns the keys that systems use for registrations and reservations for the disks—including all paths—in the specified disk groups. The `vxfen` driver is aware of which systems have registrations and reservations with specific disks. To protect the data on shared disks, each system in the cluster *must* be configured to use I/O fencing.

Verifying data storage arrays using the `vxfcntlsthdw` utility

You can use the `vxfcntlsthdw` utility to verify that shared storage arrays to be used for data support SCSI-3 persistent reservations and I/O fencing. During the I/O fencing configuration, the testing utility is used to test a single disk. The utility has other options that may be more suitable for testing storage devices in other configurations. You also need to test coordinator disk groups.

See “[Setting up I/O fencing for SF Oracle RAC](#)” on page 104.

The utility, which you can run from one system in the cluster, tests the storage used for data by setting and verifying SCSI-3 registrations on the disk or disks you specify, setting and verifying persistent reservations on the disks, writing data to the disks and reading it, and removing the registrations from the disks. Refer also to the `vxfcntlsthdw(1M)` manual page.

General guidelines for using `vxfcntlsthdw`

- The utility requires two systems connected to the shared storage.

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

- The two nodes must have `ssh` (default) or `remsh` communication. If you use `remsh`, launch the `vxfcntlsthdw` utility with the `-n` option. After completing the testing process, remove permissions for communication and restore public network connections. See “[Removing permissions for communication](#)” on page 112.
- To ensure both systems are connected to the same disk during the testing, you can use the `vxfenadm -i diskpath` command to verify a disk’s serial number. See “[Verifying the nodes see the same disk](#)” on page 102.
- For disk arrays with many disks, use the `-m` option to sample a few disks before creating a disk group and using the `-g` option to test them all.

- When testing many disks with the `-f` or `-g` option, you can review results by redirecting the command output to a file.
- The utility indicates a disk can be used for I/O fencing with a message resembling:

```
The disk /dev/vx/rdmp/c4t8d0s2 is ready to be configured for I/O
Fencing on node nebula
```

 If the utility does not show a message stating a disk is ready, verification has failed.
- If the disk you intend to test has existing SCSI-3 registration keys, the test issues a warning before proceeding.

vxfentsthaw options

Table D-11 describes three methods the utility provides to test storage devices.

Table D-11 vxfentsthaw options

| vxfentsthaw option | Description | When to use |
|--------------------|--|--|
| -n | Utility uses <code>remsh</code> for communication. | Use when <code>remsh</code> is used for communication. |
| -r | Non-destructive testing. Testing of the disks for SCSI-3 persistent reservations occurs in a non-destructive way; that is, there is only testing for reads, not writes. May be used with <code>-m</code> , <code>-f</code> , or <code>-g</code> options. | Use during non-destructive testing. |
| -t | Testing of the return value of SCSI TEST UNIT (TUR) command under SCSI-3 reservations. A warning is printed on failure of TUR testing. | When you want to perform TUR testing. |
| -d | Use DMP devices.
May be used with <code>-c</code> or <code>-g</code> options. | By default, the script picks up the DMP paths for disks in the diskgroup.
If you want the script to use the raw paths for disks in the diskgroup, use the <code>-w</code> option. |

Table D-11 vxfcntlshdw options

| vxfcntlshdw option | Description | When to use |
|----------------------|--|--|
| -w | Use raw devices.
May be used with -c or -g options. | With the -w option, the script picks the raw paths for disks in the diskgroup.
By default, the script uses the -d option. |
| -c | Utility tests the coordinator disk group prompting for systems and devices, and reporting success or failure. | For testing disks in coordinator disk group. |
| -m | Utility runs manually, in interactive mode, prompting for systems and devices, and reporting success or failure.
May be used with -r and -t options.
-m is the default option. | For testing a few disks or for sampling disks in larger arrays. |
| -f <i>filename</i> | Utility tests system/device combinations listed in a text file.
May be used with -r and -t options. | For testing several disks. |
| -g <i>disk_group</i> | Utility tests all disk devices in a specified disk group.
May be used with -r and -t options. | For testing many disks and arrays of disks. Disk groups may be temporarily created for testing purposes and destroyed (ungrouped) after testing. |

Testing the coordinator disk group using vxfcntlshdw -c

Use the vxfcntlshdw utility to verify disks are configured to support I/O fencing. In this procedure, the vxfcntlshdw utility tests the three disks one disk at a time from each node.

From the node galaxy, the disks are /dev/vx/rdmp/c1t1d0, /dev/vx/rdmp/c2t1d0, and /dev/vx/rdmp/c3t1d0.

From the node nebula, the same disks are seen as /dev/vx/rdmp/c4t1d0, /dev/vx/rdmp/c5t1d0, and /dev/vx/rdmp/c6t1d0.

Note: To test the coordinator disk group using the vxfcntl utility, the utility requires that the coordinator disk group, vxfcntl, be accessible from two nodes.

To test the coordinator disk group using vxfcntl -c

- 1 Use the vxfcntl command with the -c option. For example:

```
# /opt/VRTSvcs/vxfen/bin/vxfcntl -c vxfcntl
```
- 2 Enter the nodes you are using to test the coordinator disks:
Enter the first node of the cluster:
galaxy
Enter the second node of the cluster:
nebula
- 3 Review the output of the testing process for both nodes for all disks in the coordinator disk group. Each disk should display output that resembles:

```
ALL tests on the disk /dev/vx/rmp/c1t1d0 have PASSED.  
The disk is now ready to be configured for I/O Fencing on node  
galaxy as a COORDINATOR DISK.  
  
ALL tests on the disk /dev/vx/rmp/c4t1d0 have PASSED.  
The disk is now ready to be configured for I/O Fencing on node  
nebula as a COORDINATOR DISK.
```
- 4 After you test all disks in the disk group, the vxfcntl disk group is ready for use.

Removing and replacing a failed disk

If a disk in the coordinator disk group fails verification, remove the failed disk or LUN from the vxfcntl disk group, replace it with another, and retest the disk group.

If you need to replace a disk in an active coordinator disk group, refer to the troubleshooting procedure.

See [“Adding or removing coordinator disks”](#) on page 348.

To remove and replace a failed disk

- 1 Use the vxdiskadm utility to remove the failed disk from the disk group. Refer to the *Veritas Volume Manager Administrator's Guide*.
- 2 Add a new disk to the node, initialize it, and add it to the coordinator disk group.
See [“Initializing disks”](#) on page 106.
See [“Setting up coordinator disk groups”](#) on page 107.
- 3 Retest the disk group.

Using the -r option for non-destructive testing

To test disk devices containing data you want to preserve, you can use the -r option with the -m, -f, or -g options, which are described in the following sections. For example, to use the -m option and the -r option, you can run the utility by entering:

```
# /opt/VRTSvcs/vxfen/bin/vxfentsthdw -rm
```

When invoked with the -r option, the utility does not use tests that write to the disks. Therefore, it does not test the disks for all of the usual conditions of use.

Using the -m option

Review the procedure to test the shared disks. The utility uses the -m option.

See [“Testing the shared disks for SCSI-3”](#) on page 102.

Using the -f option

Use the -f option to test disks that are listed in a text file. For example, you can create a file to test two disks shared by systems galaxy and nebula that might resemble:

```
galaxy /dev/vx/rdmp/c2t2d1 nebula /dev/vx/rdmp/c3t2d1  
galaxy /dev/vx/rdmp/c2t2d1 nebula /dev/vx/rdmp/c3t2d1
```

where the first disk is listed in the first line and is seen by galaxy as /dev/vx/rdmp/c2t2d1 and by nebula as /dev/vx/rdmp/c3t2d1. The other disk, in the second line, is seen as /dev/vx/rdmp/c2t2d2 from galaxy and /dev/vx/rdmp/c3t2d2 from nebula. Typically, the list of disks could be extensive.

Suppose you created the file named disks_blue. To test the disks, you would enter:

```
# /opt/VRTSvcs/vxfen/bin/vxfentsthdw -f disks_blue
```

The utility reports the test results one disk at a time, just as for the -m option.

You can redirect the test results to a text file. Precede the command with “yes” to acknowledge that the testing destroys any data on the disks to be tested.

Caution: Be advised that by redirecting the command’s output to a file, a warning that the testing destroys data on the disks cannot be seen until the testing is done.

For example:

```
# yes | /opt/VRTSvcs/vxfen/bin/vxfentsthdw -f disks_blue >  
blue_test.txt
```

Using the -g option

Use the -g option to test all disks within a disk group. For example, you create a temporary disk group consisting of all disks in a disk array and test the group.

Note: Do not import the test disk group as shared; that is, do not use the -s option.

The utility reports the test results one disk at a time. You can redirect the test results to a text file for review.

```
# /opt/VRTSvcs/vxfen/bin/vxfentsthdw -g red_disks_dg >
redtest.txt
```

After testing, destroy the disk group and put the disks into disk groups as you need.

Testing a disk with existing keys

If the utility detects that a coordinator disk has existing keys, you see a message that resembles:

```
There are Veritas I/O Fencing keys on the disk. Please make sure
that I/O Fencing is shut down on all nodes of the cluster before
continuing.
```

```
***** WARNING!!!!!!!!!! *****
```

```
THIS SCRIPT CAN ONLY BE USED IF THERE ARE NO OTHER ACTIVE NODES
IN THE CLUSTER! VERIFY ALL OTHER NODES ARE POWERED OFF OR
INCAPABLE OF ACCESSING SHARED STORAGE.
```

```
If this is not the case, data corruption will result.
```

```
Do you still want to continue : [y/n] (default: n) y
```

The utility prompts you with a warning before proceeding. You may continue as long as I/O fencing is not yet configured.

How I/O fencing works in different event scenarios

Table D-12 describes how I/O fencing works to prevent data corruption in different failure event scenarios. For each event, corrective operator actions are indicated.

Table D-12 I/O fencing scenarios

| Event | Node A: What happens? | Node B: What happens? | Operator action |
|---|---|--|--|
| Both private networks fail. | Node A races for majority of coordinator disks.

If Node A wins race for coordinator disks, Node A ejects Node B from the shared disks and continues. | Node B races for majority of coordinator disks.

If Node B loses the race for the coordinator disks, Node B removes itself from the cluster. | When Node B is ejected from cluster, repair the private networks before attempting to bring Node B back. |
| Both private networks function again after event above. | Node A continues to work. | Node B has crashed. It cannot start the database since it is unable to write to the data disks. | Restart Node B after private networks are restored. |
| One private network fails. | Node A prints message about an IOFENCE on the console but continues. | Node B prints message about an IOFENCE on the console but continues. | Repair private network. After network is repaired, both nodes automatically use it. |

Table D-12 I/O fencing scenarios

| Event | Node A: What happens? | Node B: What happens? | Operator action |
|---------------|--|--|---|
| Node A hangs. | <p>Node A is extremely busy for some reason or is in the kernel debugger.</p> <p>When Node A is no longer hung or in the kernel debugger, any queued writes to the data disks fail because Node A is ejected. When Node A receives message from GAB about being ejected, it removes itself from the cluster.</p> | <p>Node B loses heartbeats with Node A, and races for a majority of coordinator disks.</p> <p>Node B wins race for coordinator disks and ejects Node A from shared data disks.</p> | <p>Verify private networks function and restart Node A.</p> |

Table D-12 I/O fencing scenarios

| Event | Node A: What happens? | Node B: What happens? | Operator action |
|---|--|--|---|
| <p>Nodes A and B and private networks lose power. Coordinator and data disks retain power. Power returns to nodes and they restart, but private networks still have no power.</p> | <p>Node A restarts and I/O fencing driver (vxfen) detects Node B is registered with coordinator disks. The driver does not see Node B listed as member of cluster because private networks are down. This causes the I/O fencing device driver to prevent Node A from joining the cluster. Node A console displays:
Potentially a preexisting split brain. Dropping out of the cluster. Refer to the user documentation for steps required to clear preexisting split brain.</p> | <p>Node B restarts and I/O fencing driver (vxfen) detects Node A is registered with coordinator disks. The driver does not see Node A listed as member of cluster because private networks are down. This causes the I/O fencing device driver to prevent Node B from joining the cluster. Node B console displays:
Potentially a preexisting split brain. Dropping out of the cluster. Refer to the user documentation for steps required to clear preexisting split brain.</p> | <p>Refer to section in Troubleshooting chapter for instructions on resolving preexisting split brain condition.</p> |

Table D-12 I/O fencing scenarios

| Event | Node A: What happens? | Node B: What happens? | Operator action |
|---|---------------------------|--|---|
| <p>Node A crashes while Node B is down. Node B comes up and Node A is still down.</p> | <p>Node A is crashed.</p> | <p>Node B restarts and detects Node A is registered with the coordinator disks. The driver does not see Node A listed as member of the cluster. The I/O fencing device driver prints message on console:</p> <pre>Potentially a preexisting split brain. Dropping out of the cluster. Refer to the user documentation for steps required to clear preexisting split brain.</pre> | <p>Refer to section in Troubleshooting chapter for instructions on resolving preexisting split brain condition.</p> |

Table D-12 I/O fencing scenarios

| Event | Node A: What happens? | Node B: What happens? | Operator action |
|--|--|--|--|
| The disk array containing two of the three coordinator disks is powered off. | Node A continues to operate as long as no nodes leave the cluster. | Node B continues to operate as long as no nodes leave the cluster. | |
| Node B leaves the cluster and the disk array is still powered off. | Node A races for a majority of coordinator disks. Node A fails because only one of three coordinator disks is available. Node A removes itself from the cluster. | Node B leaves the cluster. | Power on failed disk array and restart I/O fencing driver to enable Node A to register with all coordinator disks. |

About vxfenadm utility

Administrators can use the `vxfenadm` command to troubleshoot and test fencing configurations. The command's options for use by administrators are:

- g read and display keys
- i read SCSI inquiry information from device
- m register with disks
- n make a reservation with disks
- p remove registrations made by other systems
- r read reservations
- x remove registrations

Registration key formatting

The key defined by VxVM associated with a disk group consists of seven bytes maximum. This key becomes unique among the systems when the VxVM prefixes it with the ID of the system. The key used for I/O fencing, therefore, consists of eight bytes.

| | | | | | | | |
|---------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 0 | | | | | | | 7 |
| Node ID | VxVM Defined |

The keys currently assigned to disks can be displayed by using the `vxfenadm` command.

For example, from the system with node ID 1, display the key for the disk `/dev/vx/rdmp/c1t12d0` by entering:

```
# vxfenadm -g /dev/vx/rdmp/c2t1d0s2
Reading SCSI Registration Keys...
Device Name: /dev/vx/rdmp/c1t12d0
Total Number of Keys: 1
key[0]:
    Key Value [Numeric Format]: 65,45,45,45,45,45,45,45
    Key Value [Character Format]: A-----
```

The `-g` option of `vxfenadm` displays all eight bytes of a key value in two formats. In the numeric format, the first byte, representing the Node ID, contains the system ID plus 65. The remaining bytes contain the ASCII values of the letters of the key, in this case, “-----.” In the next line, the node ID 0 is expressed as “A;” node ID 1 would be “B.”

Configuring the Symantec License Inventory Agent

This appendix includes the following topics:

- [About the Symantec License Inventory Manager](#)
- [When the Symantec License Inventory Agent is installed](#)
- [When the server and access points are installed](#)
- [What you can do with the agent after it is installed](#)
- [How to remove the agent](#)
- [How to order the Symantec License Inventory Manager license and media kit](#)

The Symantec License Inventory Manager installation disc is available separately. For information on how to order the full product, see “[How to order the Symantec License Inventory Manager license and media kit](#)” on page 405. The installation media provides online documentation with details on all topics discussed in this appendix.

Read the following Technical Support TechNote for the latest information on updates, patches, and software issues regarding this product:

<http://entsupport.symantec.com/docs/282183>

You can also download the *Symantec License Inventory Agent 4.1 Release Notes*, from this website.

About the Symantec License Inventory Manager

The Symantec License Inventory Manager (license inventory manager) is an enterprise asset management tracking tool that inventories Symantec Information Availability products in your network and consolidates critical information on the deployment of these products to facilitate license management and compliance tracking. Using the information provided by the license inventory manager, you can:

- Determine all the Symantec software products and licenses being used in your enterprise
- Achieve easier license self-compliance management
- Know your Enterprise License Agreement deployment status
- Reduce administrative overhead for managing license compliance
- Renew support and maintenance based on the licenses you have deployed
- Gain more control over your Symantec software usage
- Manage department chargebacks based on actual software usage
- Use more flexible licensing and pricing models
- Exploit detailed deployment data to perform return on investment analyses for purchased software

The license inventory manager is a three-tiered system that consists of a server tier, access point tier, and an agent tier. The server tier is the Symantec License Inventory Server, which consolidates and stores information that it gathers from the agents and access points.

The optional access point tier includes Symantec License Inventory Access Points and serves as a consolidation layer between the agents and server.

The agent tier includes Symantec License Inventory Agents, which are deployed on individual hosts in a network. Each agent gathers product information on the supported Symantec products that are installed on the agent's host, then sends the information to an access point or the server.

When the Symantec License Inventory Agent is installed

The Symantec product installer installs or upgrades the agent on the host with the Symantec product. The agent is installed in the following directory:

`/opt/SYMC1ma`

The agent is installed with a default configuration that minimizes its impact on a running system. The minimum configuration prevents remote communication with the agent to keep its data and interfaces secure.

When the server and access points are installed

The server and access points are not installed automatically. If you want to use the Symantec License Inventory Manager, you must manually install the server and, optionally, the access points. After you install the server and access points, the agents can gather information and you can create inventory reports.

You can install the server and access points from the Symantec License Inventory Manager installation disc.

What you can do with the agent after it is installed

If you are already participating in a Symantec sales program that requires the use of the agent, or if you want to order and deploy the Symantec License Inventory Manager, you can use the agent to track Symantec products on the systems on which it was installed. To use the agent, however, you must manually configure it to enable remote communication between the agent and its server or access point.

Complete instructions for reconfiguring the agent are provided in the *Symantec License Inventory Manager 4.1 Release Notes*. You can download this document from the following website:

<http://entsupport.symantec.com/docs/282183>

How to remove the agent

If you do not want to use the Symantec License Inventory Manager, you can remove the agent using the operating system package removal commands to remove the agent packages, which include SYMClma and VRTSsmf.

The server and access point also use the VRTSsmf package. If the server or access point is installed on this host with the agent, you can remove the SYMClma package, but not the VRTSsmf package. If neither the server nor the access point is installed on this host, you can remove both the SYMClma and VRTSsmf packages.

If you remove both packages, remove the SYMClma package first.

[Table E-13](#) lists the commands required to remove these packages on the supported platforms.

Table E-13 Package removal commands required to remove the agent

| Platform | Package removal command |
|----------|--|
| AIX | <code>installp -u VRTSlma</code>
<code>installp -u VRTSsmf</code> |
| HP-UX | <code>swremove SYMClma</code>
<code>swremove VRTSsmf</code> |
| Linux | <code>rpm evv SYMClma</code>
<code>rpm evv VRTSsmf</code> |
| Solaris | <code>pkgrm VRTSlma</code>
<code>pkgrm VRTSsmf</code> |

Later, you can reinstall the agent with the Symantec License Inventory Manager installation disc. This disc is available in the Symantec License Inventory Manager kit.

How to order the Symantec License Inventory Manager license and media kit

To order a Symantec License Inventory Manager license and media kit, contact your Symantec sales representative.

The installation media provides online documentation for the Symantec License Inventory Manager. You can contact your sales representative to order printed copies of the documentation. The documents you can order include:

- *Symantec License Inventory Manager Installation and Configuration Guide*
- *Symantec License Inventory Manager Administrator's Guide*
- *Symantec License Inventory Manager User's Guide*

Tunable kernel driver parameters

This appendix contains the following topics:

- [About tunable parameters](#)
- [About LMX tunable parameters](#)
- [About VXFEN tunable parameters](#)
- [Configuring the module parameters](#)

About tunable parameters

Symantec recommends not to change the tunable kernel parameters without assistance from Veritas support personnel. Several of the parameters preallocate memory for critical data structures; a change in their values could increase memory use or degrade performance.

Warning: Do not use the tunable parameters that are described to enhance performance without assistance from Veritas support personnel.

About LMX tunable parameters

[Table F-14](#) describes tunable parameters for the LMX driver.

Table F-14 LMX tunable parameters

| LMX Parameter | Description | Default Value | Maximum Value |
|----------------|---|---------------|---------------|
| lmx_minor_max | Specifies the maximum number of contexts system-wide. Each Oracle process typically has two LMX contexts.

“Contexts” and “minors” are used interchangeably in the documentation; “context” is an Oracle-specific term to specify the value in the lmx.conf file. | 8192 | 65535 |
| lmx_port_max | Specifies the number of communication endpoints for transferring messages from the sender to receiver in a uni-directional manner. | 4096 | 65535 |
| lmx_buffer_max | Specifies the number of addressable regions in memory to copy LMX data. | 4096 | 65535 |

If you see the message “no minors available” on one node, add a configuration parameter that increases the value for the maximum number of contexts. (While the “minors” term appears in the error message, use the “contexts” term in the configuration file.) Increasing the number of contexts on a node has some impact on the resources of that node.

See [“Configuring the module parameters”](#) on page 409.

About VXFEN tunable parameters

[Table F-15](#) describes tunable parameters for the VXFEN driver.

Table F-15 VXFEN tunable parameters

| vxfen Parameter | Description and Values: Default, Minimum, and Maximum |
|-----------------|---|
| vxfen_debug_sz | Size of debug log in bytes
<ul style="list-style-type: none"> ■ Values Default: 65536 Minimum: 65536 Maximum: 256K |

Table F-15 VXFEN tunable parameters

| vxfen Parameter | Description and Values: Default, Minimum, and Maximum |
|--|--|
| vxfen_max_delay and vxfen_min_delay (See below.) | In the event of a network partition, the smaller cluster delays before racing for the coordinator disks. The time delayed allows a larger sub-cluster to win the race for the coordinator disks. The vxfen_max_delay and vxfen_min_delay parameters define the delay in seconds. |
| vxfen_max_delay | Specifies the maximum number of seconds that the smaller sub-cluster waits before racing with larger clusters for control of the coordinator disks.

This value must be greater than the vxfen_min_delay value.
<ul style="list-style-type: none"> ■ Values Default: 60 Minimum: 0 Maximum: 600 |
| vxfen_min_delay | Specifies the minimum number of seconds that the smaller sub-cluster waits before racing with larger clusters for control of the coordinator disks. This value must be smaller than the vxfen_max_delay value.
<ul style="list-style-type: none"> ■ Values Default: 1 Minimum: 0 Maximum: 600 |

See [“Configuring the module parameters”](#) on page 409.

Configuring the module parameters

For the parameter changes to take effect, you must reconfigure the LMX or VXFEN modules.

To reconfigure the LMX or VXFEN modules

- 1 Configure the tunable parameter.

```
# /usr/sbin/kctune tunable=value
```

For example:

```
LMX          # /usr/sbin/kctune lmx_minor_max=16384
```

```
VXFEN       # /usr/sbin/kctune vxfen_min_delay=100
```

- 2 If you use Oracle 10g, stop CRS (if CRS is not under VCS control) and verify that CRS is stopped.
 - 3 Unmount CFS mounts (if mounts are not under VCS control).
 - 4 Stop VCS.
`# /opt/VRTSvcs/bin/hastop -local`
 - 5 Check that this node is registered at gab ports a, b, d, and o only. Ports f, h, v, and w should not be seen on this node.
`# gabconfig -a`
GAB Port Memberships
=====
- ```
Port a gen ada401 membership 0123
Port b gen ada40d membership 0123
Port d gen ada409 membership 0123
Port o gen ada406 membership 0123
```
- 6 Reboot the node.  
`galaxy> # /usr/sbin/shutdown -r`

# Error messages

This appendix contains the following topics:

- [About error messages](#)
- [LMX error messages](#)
- [VxVM error messages](#)
- [VXFEN driver error messages](#)

## About error messages

The error messages listed in this appendix are grouped by the software module in which the error occurs.

## LMX error messages

LMX error messages are of two types: critical and non-critical.

Gather information about systems and configurations for Veritas support personnel.

See [“Running scripts for engineering support analysis”](#) on page 339.

## LMX critical error messages

[Table G-16](#) lists critical error messages of LMX kernel module. These messages report critical errors when the system runs out of memory, LMX is unable to communicate with LLT, or you are unable to load or unload LMX.

**Table G-16** LMX critical error messages

Message ID	LMX Message
00001	lmxload packet header size incorrect ( <i>number</i> )
00002	lmxload invalid lmx_llt_port <i>number</i>
00003	lmxload context memory alloc failed
00004	lmxload port memory alloc failed
00005	lmxload buffer memory alloc failed
00006	lmxload node memory alloc failed
00007	lmxload msgbuf memory alloc failed
00008	lmxload tmp msgbuf memory alloc failed
00009	lmxunload node <i>number</i> conngrp not NULL
00010	lmxopen return, minor non-zero
00011	lmxopen return, no minors available
00012	lmxconnect lmxlltopen(1) err= <i>number</i>
00013	lmxconnect new connection memory alloc failed
00014	lmxconnect kernel request memory alloc failed
00015	lmxconnect mblk memory alloc failed
00016	lmxconnect conn group memory alloc failed
00017	lmxlltunregister: LLT unregister port <i>number</i> failed err= <i>number</i>
00018	lmxload contexts <i>number</i> > <i>number</i> , max contexts = system limit = <i>number</i>
00019	lmxload ports <i>number</i> > <i>number</i> , max ports = system limit = <i>number</i>
00020	lmxload buffers <i>number</i> > <i>number</i> , max buffers = system limit = <i>number</i>
00021	lmxload msgbuf <i>number</i> > <i>number</i> , max msgbuf size = system limit = <i>number</i>

## LMX non-critical error messages

Table G-17 contains LMX error messages that may appear during run-time.

If the display of these messages creates errors while running an Oracle application, use the `lmxconfig` command to turn off the display. For example:

```
/sbin/lmxconfig -e 0
```

To re-enable the display of the messages, type:

```
/sbin/lmxconfig -e 1
```

**Table G-17** LMX non-critical error messages

Message ID	LMX Message
06001	lmxreqlink duplicate kreq= <i>0xaddress</i> , req= <i>0xaddress</i>
06002	lmxreqlink duplicate ureq= <i>0xaddress</i> kr1= <i>0xaddress</i> , kr2= <i>0xaddress</i> req type = <i>number</i>
06003	lmxrequnlink not found kreq= <i>0xaddress</i> from= <i>number</i>
06004	lmxrequnlink_l not found kreq= <i>0xaddress</i> from= <i>number</i>
06101	lmxpollreq not in doneq CONN kreq= <i>0xaddress</i>
06201	lmxnewcontext lltinit fail err= <i>number</i>
06202	lmxnewcontext lltregister fail err= <i>number</i>
06301	lmxrecvport port not found unode= <i>number</i> node= <i>number</i> ctx= <i>number</i>
06302	lmxrecvport port not found (no port) ctx= <i>number</i>
06303	lmxrecvport port not found ugen= <i>number</i> gen= <i>number</i> ctx= <i>number</i>
06304	lmxrecvport dup request detected
06401	lmxinitport out of ports
06501	lmxsendport lltsend node= <i>number</i> err= <i>number</i>
06601	lmxinitbuf out of buffers
06602	lmxinitbuf fail ctx= <i>number</i> ret= <i>number</i>
06701	lmxsendbuf lltsend node= <i>number</i> err= <i>number</i>
06801	lmxconfig insufficient privilege, uid= <i>number</i>
06901	lmxlltnodestat: LLT getnodeinfo failed err= <i>number</i>

## VxVM error messages

[Table G-18](#) contains VxVM error messages that are related to I/O fencing.

**Table G-18** VxVM error messages for I/O fencing

Message	Explanation
vold_pgr_register( <i>disk_path</i> ): failed to open the vxfen device. Please make sure that the vxfen driver is installed and configured.	The vxfen driver is not configured. Follow the instructions in <a href="#">“Setting up coordinator disk groups”</a> on page 107 to set up these disks and start I/O fencing. You can then clear the faulted resources and bring the service groups online.
vold_pgr_register( <i>disk_path</i> ): Probably incompatible vxfen driver.	Incompatible versions of VxVM and the vxfen driver are installed on the system. Install the proper version of SF Oracle RAC.

## VXFEN driver error messages

[Table G-19](#) contains VXFEN driver error messages. VXFEN also displays some informational messages.

See [“VXFEN driver informational message”](#) on page 415.

See [“Node ejection informational messages”](#) on page 415.

**Table G-19** VXFEN driver error messages

Message	Explanation
VXFEN: Unable to register with coordinator disk with serial number: <i>xxxx</i>	This message appears when the vxfen driver is unable to register with one of the coordinator disks. The serial number of the coordinator disk that failed is printed.
VXFEN: Unable to register with a majority of the coordinator disks. Dropping out of cluster.	This message appears when the vxfen driver is unable to register with a majority of the coordinator disks. The problems with the coordinator disks must be cleared before fencing can be enabled.  This message is preceded with the message “VXFEN: Unable to register with coordinator disk with serial number <i>xxxx</i> .”

## VXFEN driver informational message

The following messages show the time required to fence data disks for nodes that left the cluster.

```
date and time VXFEN:00021:Starting to eject leaving nodes(s)
from data disks.
```

```
date and time VXFEN:00022:Completed ejection of leaving node(s)
from data disks.
```

## Node ejection informational messages

Informational messages may appear on the console of one of the cluster nodes when a node is ejected from a disk or LUN.

For example:

```
<date> <system name> scsi: WARNING: /sbus@3,0/lpfs@0,0/
sd@0,1(sd91):
<date> <system name> Error for Command: <undecoded cmd 0x5f>
Error Level: Informational
<date> <system name> scsi: Requested Block: 0 Error Block 0
<date> <system name> scsi: Vendor: <vendor> Serial Number:
0400759B006E
<date> <system name> scsi: Sense Key: Unit Attention
<date> <system name> scsi: ASC: 0x2a (<vendor unique code
0x2a>), ASCQ: 0x4, FRU: 0x0
```

You can ignore these messages.



# Glossary

## **Agent**

A process that starts, stops, and monitors all configured resources of a type, and reports their status to VCS.

## **Active/Active Configuration**

A failover configuration where each system runs a service group. If either fails, the other one takes over and runs both service groups. Also known as a symmetric configuration.

## **Active/Passive Configuration**

A failover configuration consisting of one service group on a primary system, and one dedicated backup system. Also known as an asymmetric configuration.

## **Cluster**

A cluster is one or more computers that are linked together for the purpose of multiprocessing and high availability. The term is used synonymously with VCS cluster, meaning one or more computers that are part of the same GAB membership.

## **Cluster Manager (Java Console)**

A Java-based graphical user interface to manage VCS clusters. It provides complete administration capabilities for a cluster, and can run on any system inside or outside the cluster, on any operating system that supports Java.

## **Cluster Manager (Web Console)**

A Web-based graphical user interface for monitoring and administering the cluster.

## **Disaster Recovery**

Administrators with clusters in physically disparate areas can set the policy for migrating applications from one location to another if clusters in one geographic area become unavailable due to an unforeseen event. Disaster recovery requires heartbeating and replication.

## **Disk Heartbeats (GABDISK)**

A way to improve cluster resiliency, GABDISK enables a heartbeat to be placed on a physical disk shared by all systems in the cluster.

## **Failover**

A failover occurs when a service group faults and is migrated to another system.

## **GAB**

*Group Atomic Broadcast* (GAB) is a communication mechanism of the VCS engine that manages cluster membership, monitors heartbeat communication, and distributes information throughout the cluster.

## **Global Service Group**

A VCS service group which spans across two or more clusters. The `ClusterList` attribute for this group contains the list of clusters over which the group spans.

**hashadow Process**

A process that monitors and, when required, restarts HAD.

**High Availability Daemon (HAD)**

The core VCS process that runs on each system. The HAD process maintains and communicates information about the resources running on the local system and receives information about resources running on other systems in the cluster.

**Jeopardy**

A node is in *jeopardy* when it is missing one of the two required heartbeat connections. When a node is running with one heartbeat only (in jeopardy), VCS does *not* restart the applications on a new node. This action of disabling failover is a safety mechanism that prevents data corruption.

**LLT**

*Low Latency Transport* (LLT) is a communication mechanism of the VCS engine that provides kernel-to-kernel communications and monitors network communications.

**main.cf**

The file in which the cluster configuration is stored.

**Monitor Program**

The Monitor Program informs the application agent whether the application process is online or offline, and properly returning service requests.

**Network Partition**

If all network connections between any two groups of systems fail simultaneously, a *network partition* occurs. When this happens, systems on both sides of the partition can restart applications from the other side resulting in duplicate services, or “split-brain.” A split brain occurs when two independent systems configured in a cluster assume they have exclusive access to a given resource (usually a file system or volume). The most serious problem caused by a network partition is that it affects the data on shared disks. See “[Jeopardy](#)” and “[Seeding](#)”.

**Node**

The physical host or system on which applications and service groups reside. When systems are linked by VCS, they become nodes in a cluster.

**N-to-1**

An N-to-1 configuration is based on the concept that multiple, simultaneous server failures are unlikely; therefore, a single backup server can protect multiple active servers. When a server fails, its applications move to the backup server. For example, in a 4-to-1 configuration, one server can protect four servers, which reduces redundancy cost at the server level from 100 percent to 25 percent.

**N-to-N**

N-to-N refers to multiple service groups running on multiple servers, with each service group capable of being failed over to different servers in the cluster. For example, consider a four-node cluster with each node supporting three critical database instances. If any

node fails, each instance is started on a different node, ensuring no single node becomes overloaded.

**N-to-M**

N-to-M (or Any-to-Any) refers to multiple service groups running on multiple servers, with each service group capable of being failed over to different servers in the same cluster, and also to different servers in a linked cluster. For example, consider a four-node cluster with each node supporting three critical database instances and a linked two-node back-up cluster. If all nodes in the four-node cluster fail, each instance is started on a node in the linked back-up cluster.

**Replication**

Replication is the synchronization of data between systems where shared storage is not feasible. The systems that are copied may be in local backup clusters or remote failover sites. The major advantage of replication, when compared to traditional backup methods, is that current data is continuously available.

**Resources**

Individual components that work together to provide application services to the public network. A resource may be a physical component such as a disk or network interface card, a software component such as Oracle8i or a Web server, or a configuration component such as an IP address or mounted file system.

**Resource Dependency**

A dependency between resources is indicated by the keyword “requires” between two resource names. This indicates the second resource (the child) must be online before the first resource (the parent) can be brought online. Conversely, the parent must be offline before the child can be taken offline. Also, faults of the children are propagated to the parent.

**Resource Types**

Each resource in a cluster is identified by a unique name and classified according to its type. VCS includes a set of pre-defined resource types for storage, networking, and application services.

**Seeding**

Seeding is used to protect a cluster from a preexisting network partition. By default, when a system comes up, it is not seeded. Systems can be seeded automatically or manually. Only systems that have been seeded can run VCS. Systems are seeded automatically only when: an unseeded system communicates with a seeded system or all systems in the cluster are unseeded and able to communicate with each other. See “[Network Partition](#)”.

**Service Group**

A service group is a collection of resources working together to provide application services to clients. It typically includes multiple resources, hardware- and software-based, working together to provide a single service.

**Service Group Dependency**

A service group dependency provides a mechanism by which two service groups can be linked by a dependency rule, similar to the way resources are linked.

**Shared Storage**

Storage devices that are connected to and used by two or more systems.

**SNMP Notification**

Simple Network Management Protocol (SNMP) developed to manage nodes on an IP network.

**State**

The current activity status of a resource, group or system. Resource states are given relative to both systems.

**System**

The physical system on which applications and service groups reside. When a system is linked by VCS, it becomes a node in a cluster. See “[Node](#).”

**types.cf**

The types.cf file describes standard resource types to the VCS engine; specifically, the data required to control a specific resource.

**Virtual IP Address**

A unique IP address associated with the cluster. It may be brought up on any system in the cluster, along with the other resources of the service group. This address, also known as the IP alias, should not be confused with the base IP address, which is the IP address that corresponds to the host name of a system.

# Index

## Symbols

.rhosts, editing to remove remsh permissions 159

## A

adding a node

    procedure for Oracle 10g 175

agents

    CFSMount 374  
    CVMCluster 370  
    CVMVolDg 373  
    CVMVxconfigd 371

attributes

    CFSMount agent 375  
    CVMCluster agent 370  
    CVMVolDg agent 373  
    CVMVxconfigd 372  
    UseFence 110

## C

centralized cluster management 90

CFSMount agent

    description 374

clone databases, creating 266

Cluster File System (CFS)

    overview 28

cluster management 91

cluster nodes

    adding 175  
    removing 175

Cluster Volume Manager (CVM)

    overview 27

commands

    dbed\_analyzer 330  
    dbed\_ckptcreate 252, 257  
    dbed\_ckptdisplay 258  
    dbed\_ckptremove 266  
    dbed\_ckptrollback 264  
    dbed\_ckptumount 264  
    dbed\_clonedb 266  
    dbed\_update 256

    format 352

    gcoconfig 209

    vcsmmconfig 342

    vradmin 219

    vxassist 214, 215, 384

    vxdctl enable 352

    vxdg list 56

    vxedit (set shared volume mode) 383

    vxfen start 109

    vxfenadm 398

    vxfenclearpre 347

    vxprint 220

    vxstorage\_stats 327

    vxvol 140, 214, 384

communications

    GAB 26

configurations

    editing for removed nodes 184

configuring CVM

    configuring CVM and Oracle service groups  
    manually 164

configuring Oracle

    modifying the VCS configuration 165

configuring Oracle 10g 161

    checking the configuration 164

configuring VCS

    Cluster Connector 90

    Cluster Management Console 90, 91

coordinator disks

    defined 417

    for I/O fencing 34

    setting up 107

cron 262

    scheduling Storage Checkpoints 262

crontab file 262

CSSD agent 380

CVMCluster agent

    description 370

    type definition 379

CVMTypes.cf file 379

CVMVolDg agent

    description 373

CVMVxconfigd agent  
description 371

## D

data disks  
for I/O fencing 34

database  
creating 159

databases  
creating 381  
upgrading 156

dbca  
description 381

dbed\_analyzer command 330

dbed\_ckptcreate command 252, 257

dbed\_ckptdisplay command 258

dbed\_ckptremove command 266

dbed\_ckptrollback command 264

dbed\_ckptumount command 264

dbed\_clonedb command 266

dbed\_update command 256

disk groups  
overview 56

disk groups and volumes 56

disk space  
required for SF Oracle RAC 44

disks  
adding and initializing 106  
coordinator 107  
testing with vxfcntl 102  
verifying node access 102

drivers  
tunable parameters 407

## E

environment variables  
MANPATH 68  
PATH 67

error messages  
LMX 412  
node ejection 415  
VXFEN 415  
VxVM errors related to I/O fencing 414

## F

file  
Oracle agent log file location 174

format command 352

## G

GAB  
overview 25  
port memberships 105

gcoconfig command 209

getcomms  
troubleshooting SF Oracle RAC 340

getdbac  
troubleshooting SF Oracle RAC 340

Global Cluster Option (GCO)  
overview 200

global clustering  
adding VVR types to VCS configuration 208  
configuring GCO 208  
configuring VCS to replicate database  
volumes 220  
illustration of dependencies 220, 221  
in SF Oracle RAC environment 199  
migration and takeover 231  
setting up replication 213

## H

hagetcf  
troubleshooting SF Oracle RAC 340

## I

I/O fencing  
event scenarios 394  
operations 35  
overview 33  
setting up 104  
starting 109  
testing and scenarios 394

Installation 2

installation  
of SF Oracle RAC 80  
utilities 80

installing  
Root Broker 50

installing Oracle 10g 147

installing SF Oracle RAC  
removing temporary remsh access  
permissions 159

IP address  
configuring public virtual IP addresses 145

**L**

## licenses

- obtaining 66
- removing keys 69

## Listener

- description 22, 383

## LMX

- error messages 412
- tunable parameters 408

**M**

## main.cf

- after SF Oracle RAC installation 113

## managing clusters, centrally 90

## MANPATH environment variable 68

## migrating Oracle 158

**O**

## OCR

- creating directories on CFS 139
- creating volumes on raw volumes 139

## operating systems

- supported 45

## Oracle

- adding a node for Oracle 10g 175
- adding or removing a node for Oracle 10g 175
- agent log file location 174
- applying patchsets 158
- configuring virtual IP addresses 145
- creating volumes for OCR and Vote-disk for Oracle 10g 139
- migrating 157
- removing a node for Oracle 10g 183
- supported versions 45
- upgrading 157

## Oracle 10g

- editing the CVM group 143
- upgrading databases 156

## Oracle Disk Manager (ODM)

- overview 29

## Oracle Enterprise Manager 336

## Oracle instance

- definition 20

**P**

## PATH environment variable 67

## preparing to install Oracle 10g 129

## PrivNIC agent 376

**R**

## registrations

- for I/O fencing 387
- key formatting 399

## removing a node

- from a cluster 183

## reservations

- description 34

## Root Broker

- installing 50

**S**

## scheduling Storage Checkpoints 262

## SCSI-3 persistent reservations

- verifying 104

## SF Oracle RAC

- configuring 83
- configuring GCO 208
- coordinator disks 107
- error messages 411
- high-level view 47
- I/O fencing 101, 387
- information required during installation 70
- overview of components 22
- overview of installation methods 80
- phases of installation and configuration 58
- sample configuration files 355
- shared storage 101
- Storage Checkpoints 243
- Storage Mapping 325
- Storage Rollback 243
- troubleshooting 339
- tunable parameters of kernel drivers 407
- using Storage Checkpoints 243
- using uninstallsfrac 193

## SFRAC

- adding VVR types to VCS configuration 208
- configuring VCS to replicate database volumes 220
- illustration of dependencies 220, 221
- migration and takeover 231
- setting up replication 213

## split brain

- description 33

## storage

- for I/O fencing 387

**Storage Checkpoints**

- backing up and recovering databases 244, 248
- creating 257
- description 244
- determining space requirements 245
- displaying 258
- performance 247
- removing 266
- scheduling 262
- unmounting 264
- using the CLI 253
- verifying 248

**Storage Mapping**

- configuring arrays 337
- dbed\_analyzer command 330
- description 326
- displaying information for a list of tablespaces 331
- enabling Oracle file mapping 334
- Oracle Enterprise Manager 336
- ORAMAP 332
- using the vxstorage\_stats command 327
- verifying feature setup 327
- verifying Oracle file mapping setup 334
- views 333, 335

**Storage Rollback 264**

- description 245
- guidelines for recovery 251

**Symantec Product Authentication Service 50****T**

troubleshooting 339

**U**

- uninstalling SF Oracle RAC 183
- uninstallsfrac
  - procedures 193

**V****VCS (Veritas Cluster Server)**

- agent log file location 174

**VCSIPC**

- errors in trace/log files 350
- overview 25, 32
- warnings in trace files 350

vcsmmconfig command 342

**Veritas Volume Replicator (VVR)**

overview 201

**volumes**

overview 56

**Vote-disk**

- creating directories on CFS 139
- creating volumes on raw volumes 139

vradmin command 219

vxassist command 214, 215, 384

vxdctl command 352

vxdg list command 56

**VXFEN**

informational messages 415

tunable parameters 408

vxfen command 109

vxfenadm command 398

vxfenclearpre command 347

vxfentsthdw utility 388

vxprint command 220

vxstorage\_stats command 327

**VxVM**

error messages related to I/O fencing 414

vxvol command 140, 214, 384