Veritas Storage Foundation™ and High Availability Solutions 6.0.1 Disaster Recovery Implementation Guide - AIX

January 2014

Symantec™

Veritas Storage Foundation™ and High Availability Solutions Disaster Recovery Implementation Guide

Product version: 6.0.1

Document version: 6.0.1 Rev 2

Legal Notice

Symantec Corporation
350 Ellis Street
Mountain View, CA 94043

http://www.symantec.com

Technical Support

Symantec Technical Support maintains support centers globally. Technical Support's primary role is to respond to specific queries about product features and functionality. The Technical Support group also creates content for our online Knowledge Base. The Technical Support group works collaboratively with the other functional areas within Symantec to answer your questions in a timely fashion. For example, the Technical Support group works with Product Engineering and Symantec Security Response to provide alerting services and virus definition updates.

Symantec's support offerings include the following:

- A range of support options that give you the flexibility to select the right amount of service for any size organization
- Telephone and/or Web-based support that provides rapid response and up-to-the-minute information
- Upgrade assurance that delivers software upgrades
- Global support purchased on a regional business hours or 24 hours a day, 7 days a week basis
- Premium service offerings that include Account Management Services

For information about Symantec's support offerings, you can visit our website at the following URL:

www.symantec.com/business/support/index.jsp

All support services will be delivered in accordance with your support agreement and the then-current enterprise technical support policy.

Contacting Technical Support

Customers with a current support agreement may access Technical Support information at the following URL:

www.symantec.com/business/support/contact_techsupp_static.jsp

Before contacting Technical Support, make sure you have satisfied the system requirements that are listed in your product documentation. Also, you should be at the computer on which the problem occurred, in case it is necessary to replicate the problem.

When you contact Technical Support, please have the following information available:

- Product release level
- Hardware information

- Available memory, disk space, and NIC information
- Operating system
- Version and patch level
- Network topology
- Router, gateway, and IP address information
- Problem description:
  - Error messages and log files
  - Troubleshooting that was performed before contacting Symantec
  - Recent software configuration changes and network changes

Licensing and registration

If your Symantec product requires registration or a license key, access our technical support Web page at the following URL:

www.symantec.com/business/support/

Customer service

Customer service information is available at the following URL:

www.symantec.com/business/support/

Customer Service is available to assist with non-technical questions, such as the following types of issues:

- Questions regarding product licensing or serialization
- Product registration updates, such as address or name changes
- General product information (features, language availability, local dealers)
- Latest information about product updates and upgrades
- Information about upgrade assurance and support contracts
- Information about the Symantec Buying Programs
- Advice about Symantec's technical support options
- Nontechnical presales questions
- Issues that are related to CD-ROMs or manuals

Support agreement resources

If you want to contact Symantec regarding an existing support agreement, please contact the support agreement administration team for your region as follows:

Asia-Pacific and Japan	customercare_apac@symantec.com
Europe, Middle-East, and Africa	semea@symantec.com
North America and Latin America	supportsolutions@symantec.com

Documentation

Product guides are available on the media in PDF format. Make sure that you are using the current version of the documentation. The document version appears on page 2 of each guide. The latest product documentation is available on the Symantec website.

https://sort.symantec.com/documents

Your feedback on product documentation is important to us. Send suggestions for improvements and reports on errors or omissions. Include the title and document version (located on the second page), and chapter and section titles of the text on which you are reporting. Send feedback to:

doc_feedback@symantec.com

For information regarding the latest HOWTO articles, documentation updates, or to ask a question regarding product documentation, visit the Storage and Clustering Documentation forum on Symantec Connect.

https://www-secure.symantec.com/connect/storage-management/forums/storage-and-clustering-documentation

About Symantec Connect

Symantec Connect is the peer-to-peer technical community site for Symantec's enterprise customers. Participants can connect and share information with other product users, including creating forum posts, articles, videos, downloads, blogs and suggesting ideas, as well as interact with Symantec product teams and Technical Support. Content is rated by the community, and members receive reward points for their contributions.

http://www.symantec.com/connect/storage-management

Contents

Technical Support 4

Section 1 Introducing Veritas Storage Foundation and High Availability Solutions for disaster recovery 11

Chapter 1 About supported disaster recovery scenarios 12

About disaster recovery scenarios 12
About campus cluster configuration 14
VCS campus cluster requirements 14
How VCS campus clusters work 16
Typical VCS campus cluster setup 19
About replicated data clusters 21
How VCS replicated data clusters work 22
About global clusters 23
How VCS global clusters work 23
User privileges in global clusters 24
VCS global clusters: The building blocks 25
Disaster recovery feature support across Storage Foundation and High Availability Solutions 6.0.1 products 31
Replication agent support in virtual environments 34

Chapter 2 Planning for disaster recovery 36

Planning for cluster configurations 36
Planning a campus cluster setup 36
Planning a replicated data cluster setup 37
Planning a global cluster setup 38
Planning for data replication 38
Data replication options 38
Data replication considerations 39

Section 2 Implementing campus clusters 40

Chapter 3 Setting up campus clusters for VCS and SFHA 41

About setting up a campus cluster configuration 41
Preparing to set up a campus cluster configuration 41
Configuring I/O fencing to prevent data corruption 42
Configuring VxVM disk groups for campus cluster configuration 42
Configuring VCS service group for campus clusters 44
Fire drill in campus clusters 45
About the DiskGroupSnap agent 45
About running a fire drill in a campus cluster 45
Configuring the fire drill service group 46
Running a successful fire drill in a campus cluster 46

Chapter 4 Setting up campus clusters for SFCFS, SFRAC 48

About setting up a campus cluster for disaster recovery for SFCFS HA or SF Oracle RAC 48
Preparing to set up a campus cluster in a parallel cluster database environment 51
Configuring I/O fencing to prevent data corruption 52
Configuring VxVM disk groups for a campus cluster in a parallel cluster database environment 54
Configuring VCS service groups for a campus cluster for SFCFS HA and SF Oracle RAC 58
Tuning guidelines for parallel campus clusters 59
Best practices for a parallel campus cluster 59

Section 3 Implementing replicated data clusters 61

Chapter 5 Configuring a replicated data cluster using VVR 62

About setting up a replicated data cluster configuration 62
About typical replicated data cluster configuration 62
About setting up replication 63
Configuring the service groups 64
Configuring the service group dependencies 65
About migrating a service group 65
Fire drill in replicated data clusters 66

Chapter 6 Configuring a replicated data cluster using third-party replication 67

About setting up a replicated data cluster configuration using third-party replication 67
About typical replicated data cluster configuration using third-party replication 68
About setting up third-party replication 68
Configuring the service groups for third-party replication 69
Fire drill in replicated data clusters using third-party replication 69

Section 4 Implementing global clusters 70

Chapter 7 Configuring global clusters for VCS and SFHA 71

Installing and Configuring Veritas Cluster Server 71
Setting up VVR replication 71
About configuring VVR replication 72
Best practices for setting up replication 72
Creating a Replicated Data Set 74
Synchronizing the Secondary and starting replication 90
Starting replication when the data volumes are zero initialized 96
Setting up third-party replication 97
Fire drill in global clusters 98

Chapter 8 Configuring a global cluster with Storage Foundation Cluster File System or Storage Foundation for Oracle RAC 99

About global clusters 99
About replication for parallel global clusters using Storage Foundation and High Availability (SFHA) Solutions 100
About setting up a global cluster environment for parallel clusters 101
Configuring the primary site 102
Configuring the secondary site 105
Setting up replication between parallel global cluster sites 110
Testing a parallel global cluster configuration 117

Chapter 9 Configuring a global cluster with Veritas Volume Replicator and Storage Foundation Cluster File

System or Storage Foundation for Oracle RAC 119

About configuring a parallel global cluster using Veritas Volume Replicator (VVR) for replication 120
Setting up replication on the primary site using VVR 122
Creating the data and SRL volumes on the primary site 122
Setting up the Replicated Volume Group on the primary site 124
Setting up replication on the secondary site using VVR 125
Creating the data and SRL volumes on the secondary site 126
Editing the /etc/vx/vras/.rdg files 127
Setting up IP addresses for RLINKs on each cluster 127
Setting up the disk group on secondary site for replication 128
Starting replication of the primary site database volume to the secondary site using VVR 130
Configuring Veritas Cluster Server to replicate the database volume using VVR 132
Modifying the Veritas Cluster Server (VCS) configuration on the primary site 137
Modifying the VCS configuration on the secondary site 142
Replication use cases for global parallel clusters 147

Section 5 Reference 155

Appendix A Sample configuration files 156

Sample Storage Foundation for Oracle RAC configuration files 156
sfrac02_main.cf file 156
sfrac07_main.cf and sfrac08_main.cf files 157
sfrac09_main.cf and sfrac10_main.cf files 159
sfrac11_main.cf file 162
sfrac12_main.cf and sfrac13_main.cf files 163
Sample fire drill service group configuration 166
About sample main.cf files for Veritas Storage Foundation (SF) for Oracle RAC 168
Sample main.cf for Oracle 10g for CVM/VVR primary site 169
Sample main.cf for Oracle 10g for CVM/VVR secondary site 174

Section 1

Introducing Veritas Storage Foundation and High Availability Solutions for disaster recovery

- Chapter 1. About supported disaster recovery scenarios
- Chapter 2. Planning for disaster recovery

Chapter 1

About supported disaster recovery scenarios

This chapter includes the following topics:

- About disaster recovery scenarios
- About campus cluster configuration
- About replicated data clusters
- About global clusters
- Disaster recovery feature support across Storage Foundation and High Availability Solutions 6.0.1 products
- Replication agent support in virtual environments

About disaster recovery scenarios

Symantec Storage Foundation offers cost–effective, short-distance disaster recovery with active configurations and long distance replication solutions to effectively manage disaster recovery requirements.

This guide allows you to configure campus clusters, global clusters, and replicated clusters for disaster recovery failover using the following Storage Foundation and High Availability Solutions products:

- Storage Foundation Cluster File System High Availability (SFCFSHA)
- Storage Foundation™ for Oracle® RAC (SF Oracle RAC)
- Veritas Cluster Server (VCS)
- Veritas Volume Replicator (VVR)

See the *Veritas Storage Foundation Cluster File System Administrator's Guide* for more information on configuring SFCFSHA.

See the *Veritas Storage Foundation for Oracle RAC Administrator's Guide* for more information on configuring SF Oracle RAC.

See the *Veritas Cluster Server Administrator's Guide* for more information on configuring VCS.

See the *Veritas Storage Foundation and High Availability Solutions Replication Administrator's Guide* for more information on configuring VVR.

Table 1-1 lists key use cases for campus cluster, global cluster, and replicated data cluster disaster recovery configurations.

Table 1-1 Key use cases for disaster recovery configurations

Use case description	Recommended disaster recovery configuration
Disaster Recovery of business-critical applications from the production site to a geographically distributed Disaster Recovery (DR) site. ■ Distance between the two sites exceeds 80 KM or 50 miles ■ Application data is made available at the DR site through replication ■ Application is expected to be active at only one site at any point of time (Active/Passive)	Veritas Cluster Server HA/DR with Global Cluster Option (GCO) See " How VCS global clusters work" on page 23. SFRAC with Global cluster option (GCO)
Disaster Recovery of business-critical applications from the production site to a geographically distributed Disaster Recovery (DR) site. ■ Distance between the two sites is less than 80 KM or 50 miles ■ Application data is made available at the DR site through replication ■ Application is expected to be active at only one site at any point of time (Active/Passive)	Veritas Cluster Server HA/DR with Replicated Data Cluster (RDC) See " How VCS replicated data clusters work" on page 22.

Table 1-1 Key use cases for disaster recovery configurations *(continued)*

Use case description	Recommended disaster recovery configuration
Disaster Recovery of business-critical applications from the production site to a geographically distributed Disaster Recovery (DR) site. ■ Distance between the two sites is less than 80 KM or 50 miles ■ Application data is made available at the DR site through remote mirroring ■ Application is expected to be active at only one site at any point of time (Active/Passive) ■ Automatic application failover within a site, automated failover across sites	Veritas Cluster Server HA/DR with Campus Cluster See " How VCS campus clusters work" on page 16. Veritas Storage Foundation for remote mirroring
High Availability of business-critical applications across two geographically distributed sites. ■ Distance between the two sites is less than 80 KM or 50 miles ■ Application data is made available at the DR site through remote mirroring ■ Application is expected to be simultaneously active at both the sites (Active/Active)	Veritas Cluster Server with Campus Cluster See " How VCS campus clusters work" on page 16. Veritas Storage Foundation Cluster File System for remote mirroring and parallel cross-site access SFRAC with Campus Cluster for remote mirroring and parallel cross-site access

About campus cluster configuration

The campus cluster configuration provides local high availability and disaster recovery functionality in a single VCS cluster. This configuration uses data mirroring to duplicate data at different sites. There is no host or array replication involved.

VCS supports campus clusters that employ disk groups mirrored with Veritas Volume Manager.

VCS campus cluster requirements

Review the following requirements for VCS campus clusters:

- You must install VCS.

You must enable the HA/DR license if you want to manually control a service group failover across sites or system zones.

- You must have a single VCS cluster with at least one node in each of the two sites, where the sites are separated by a physical distance of no more than 80 kilometers.
- You must have redundant network connections between nodes. All paths to storage must also be redundant.

  Symantec recommends the following in a campus cluster setup:

  - A common cross-site physical infrastructure for storage and LLT private networks.
  - Technologies such as Dense Wavelength Division Multiplexing (DWDM) for network and I/O traffic across sites. Use redundant links to minimize the impact of network failure.
- Symantec recommends that you configure I/O fencing to prevent data corruption in the event of link failures.
- You must install Veritas Volume Manager with the FMR license and the Site Awareness license.
- You must configure storage to meet site-based allocation and site-consistency requirements for VxVM.
  - All the nodes in the site must be tagged with the appropriate VxVM site names.
  - All the disks must be tagged with the appropriate VxVM site names.
  - The VxVM site names of both the sites in the campus cluster must be added to the disk groups.
  - The allsites attribute for each volume in the disk group must be set to on. (By default, the value is set to on.)
  - The siteconsistent attribute for the disk groups must be set to on.
- Oracle requires that all of the nodes use IP addresses from the same subnet.
- Each host at a site must be connected to a storage switch. The switch must have access to storage arrays at all the sites..
- Symantec recommends a common cross-site physical infrastructure for storage and LLT private networks
- SF Oracle RAC campus clusters require mirrored volumes with storage allocated from both sites.

How VCS campus clusters work

This topic describes how VCS works with VxVM to provide high availability in a campus cluster environment.

In a campus cluster setup, VxVM automatically mirrors volumes across sites. To enhance read performance, VxVM reads from the plexes at the local site where the application is running. VxVM writes to plexes at both the sites.

In the event of a storage failure at a site, VxVM detaches all the disks at the failed site from the disk group to maintain data consistency. When the failed storage comes back online, VxVM automatically reattaches the site to the disk group and recovers the plexes.

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

When service group or system faults occur, VCS fails over the service groups or the nodes based on the values you set for the service group attributes SystemZones and AutoFailOver.

For campus cluster setup, you must define the SystemZones attribute in such a way that the nodes at each site are grouped together.

Depending on the value of the AutoFailOver attribute, VCS failover behavior is as follows:

0	VCS does not fail over the service group or the node.
1	VCS fails over the service group to another suitable node. VCS chooses to fail over the service group within the same site before choosing a node in the other site. By default, the AutoFailOver attribute value is set to 1.
2	VCS fails over the service group if another suitable node exists in the same site. Otherwise, VCS waits for administrator intervention to initiate the service group failover to a suitable node in the other site. This configuration requires the HA/DR license enabled. Symantec recommends that you set the value of AutoFailOver attribute to 2.

Sample definition for these service group attributes in the VCS main.cf is as follows:

```
group oragroup1 (
    SystemList = { node1=0, node2=1, node3=2, node4=3 }
    SystemZones = { node1=0, node2=0, node3=1, node4=1 }
    AutoFailOver = 2
```

. . .
)

Table 1-2 lists the possible failure scenarios and how VCS campus cluster recovers from these failures.

Table 1-2 Failure scenarios in campus cluster

Failure	Description and recovery
Node failure	■ A node in a site fails. If the value of the AutoFailOver attribute is set to 1, VCS fails over the Oracle service group to another system within the same site that is defined in the SystemZones attribute. ■ All nodes in a site fail. If the value of the AutoFailOver attribute is set to 1, VCS fails over the Oracle service group to a system in the other site that is defined in the SystemZones attribute. If the value of the AutoFailOver attribute is set to 2, VCS requires administrator intervention to initiate the Oracle service group failover to a system in the other site. If the value of the AutoFailOver attribute is set to 0, VCS requires administrator intervention to initiate a fail over in both the cases of node failure.
Application failure	The behavior is similar to the node failure.
Storage failure - one or more disks at a site fails	VCS does not fail over the service group when such a storage failure occurs. VxVM detaches the site from the disk group if any volume in that disk group does not have at least one valid plex at the site where the disks failed. VxVM does not detach the site from the disk group in the following cases: ■ None of the plexes are configured on the failed disks. ■ Some of the plexes are configured on the failed disks, and at least one plex for a volume survives at each site. If only some of the disks that failed come online and if the vxrelocd daemon is running, VxVM relocates the remaining failed disks to any available disks. Then, VxVM automatically reattaches the site to the disk group and resynchronizes the plexes to recover the volumes. If all the disks that failed come online, VxVM automatically reattaches the site to the disk group and resynchronizes the plexes to recover the volumes.

Table 1-2 Failure scenarios in campus cluster *(continued)*

Failure	Description and recovery
Storage failure - all disks at both sites fail	VCS acts based on the DiskGroup agent's PanicSystemOnDGLoss attribute value. See the *Veritas Cluster Server Bundled Agents Reference Guide* for more information.
Site failure	All nodes and storage at a site fail. Depending on the value of the AutoFailOver attribute, VCS fails over the Oracle service group as follows: ■ If the value is set to 1, VCS fails over the Oracle service group to a system in the other site that is defined in the SystemZones attribute. ■ If the value is set to 2, VCS requires administrator intervention to initiate the Oracle service group failover to a system in the other site. Because the storage at the failed site is inaccessible, VCS imports the disk group in the application service group with all devices at the failed site marked as NODEVICE. When the storage at the failed site comes online, VxVM automatically reattaches the site to the disk group and resynchronizes the plexes to recover the volumes.
Network failure (LLT interconnect failure)	Nodes at each site lose connectivity to the nodes at the other site The failure of private interconnects between the nodes can result in split brain scenario and cause data corruption. Review the details on other possible causes of split brain and how I/O fencing protects shared data from corruption. Symantec recommends that you configure I/O fencing to prevent data corruption in campus clusters.

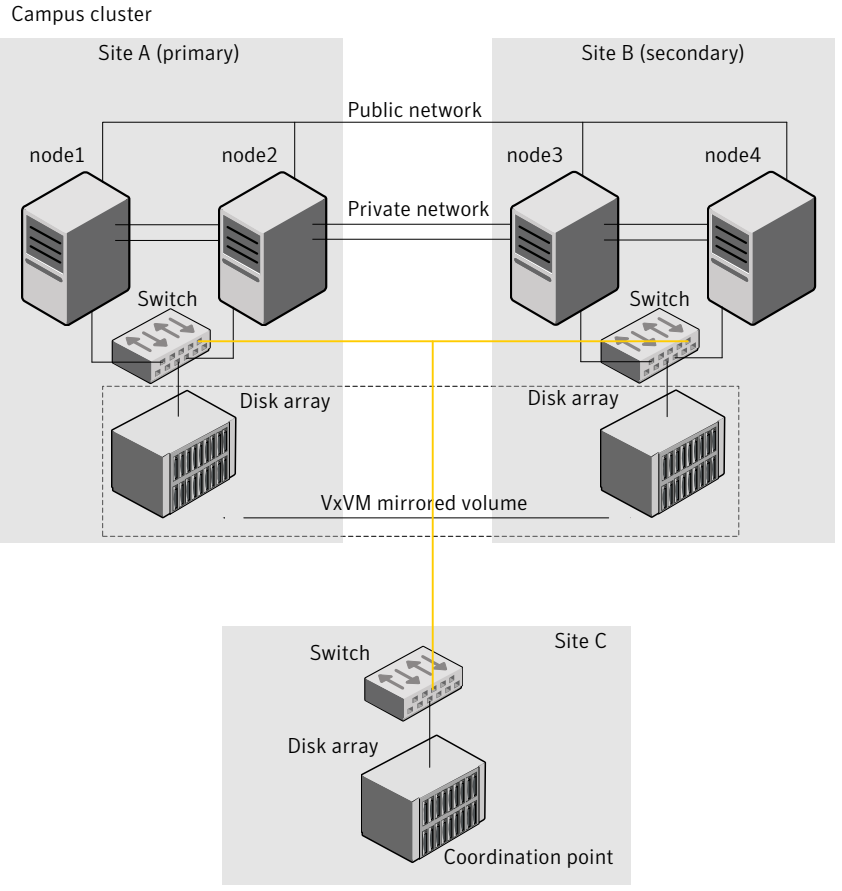
Table 1-2 Failure scenarios in campus cluster *(continued)*

Failure	Description and recovery
Network failure (LLT and storage interconnect failure)	Nodes at each site lose connectivity to the storage and the nodes at the other site Symantec recommends that you configure I/O fencing to prevent split brain and serial split brain conditions. ■ If I/O fencing is configured: The site that loses the race commits suicide. When you restore the network connectivity, VxVM detects the storage at the failed site, reattaches the site to the disk group, and resynchronizes the plexes to recover the volumes. ■ If I/O fencing is not configured: If the application service group was online at site A during such failure, the application service group remains online at the same site. Because the storage is inaccessible, VxVM detaches the disks at the failed site from the disk group. At site B where the application service group is offline, VCS brings the application service group online and imports the disk group with all devices at site A marked as NODEVICE. So, the application service group is online at both the sites and each site uses the local storage. This causes inconsistent data copies and leads to a site-wide split brain. When you restore the network connectivity between sites, a serial split brain may exist. See the *Veritas Storage Foundation Administrator's Guide* for details to recover from a serial split brain condition.

Typical VCS campus cluster setup

Figure 1-1 depicts a typical VCS campus cluster setup.

Figure 1-1 Typical VCS campus cluster setup



VCS campus cluster typically has the following characteristics:

- Single VCS cluster spans multiple sites.
  In the sample figure, VCS is configured on four nodes: node 1 and node 2 are located at site A and node 3 and node 4 at site B.
- I/O fencing is configured with one coordinator disk from each site of the campus cluster and another coordinator disk from a third site.
  Figure 1-1 illustrates a typical setup with disk-based I/O fencing. You can also configure server-based I/O fencing.
  Mix mode fencing with two coordinator disks from each site and a CP server on third site is also supported.

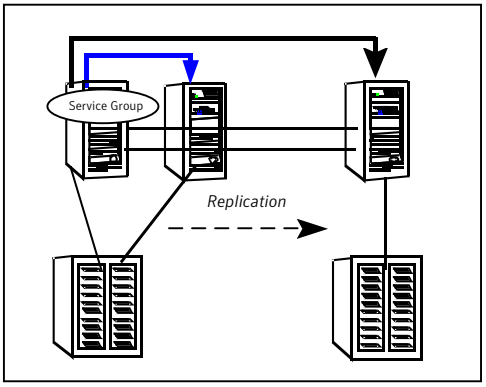
- The shared data is located on mirrored volumes on a disk group configured using Veritas Volume Manager.
- The volumes that are required for the application have mirrors on both the sites.
- All nodes in the cluster are tagged with the VxVM site name. All disks that belong to a site are tagged with the corresponding VxVM site name.
- The disk group is configured in VCS as a resource of type DiskGroup and is mounted using the Mount resource type.

About replicated data clusters

In a replicated data cluster no shared disks exist. Instead, a data replication product synchronizes copies of data between nodes or sites. Replication can take place at the application, host, and storage levels. Application-level replication products, such as Oracle DataGuard, maintain consistent copies of data between systems at the SQL or database levels. Host-based replication products, such as Veritas Volume Replicator, maintain consistent storage at the logical volume level. Storage-based or array-based replication maintains consistent copies of data at the disk or RAID LUN level.

Figure 1-2 shows a hybrid shared storage and replicated data cluster, in which different failover priorities are assigned to nodes according to particular service groups.

Figure 1-2 Shared storage replicated data cluster



You can also configure replicated data clusters without the ability to fail over locally, but this configuration is not recommended.

See “ How VCS replicated data clusters work” on page 22.

How VCS replicated data clusters work

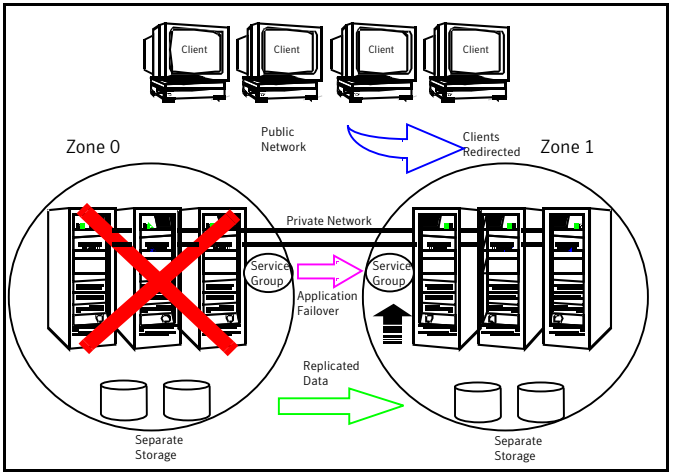
To understand how a replicated data cluster configuration works, let us take the example of an application configured in a VCS replicated data cluster. The configuration has two system zones:

- Primary zone (zone 0) comprising nodes located at the primary site and attached to the primary storage
- Secondary zone (zone 1) comprising nodes located at the secondary site and attached to the secondary storage

The application is installed and configured on all nodes in the cluster. Application data is located on shared disks within each RDC zone and is replicated across RDC zones to ensure data concurrency. The application service group is online on a system in the current primary zone and is configured to fail over in the cluster.

Figure 1-3 depicts an application configured on a VCS replicated data cluster.

Figure 1-3 A VCS replicated data cluster configuration



In the event of a system or application failure, VCS attempts to fail over the application service group to another system within the same RDC zone. However, in the event that VCS fails to find a failover target node within the primary RDC zone, VCS switches the service group to a node in the current secondary RDC zone (zone 1). VCS also redirects clients once the application is online on the new location.

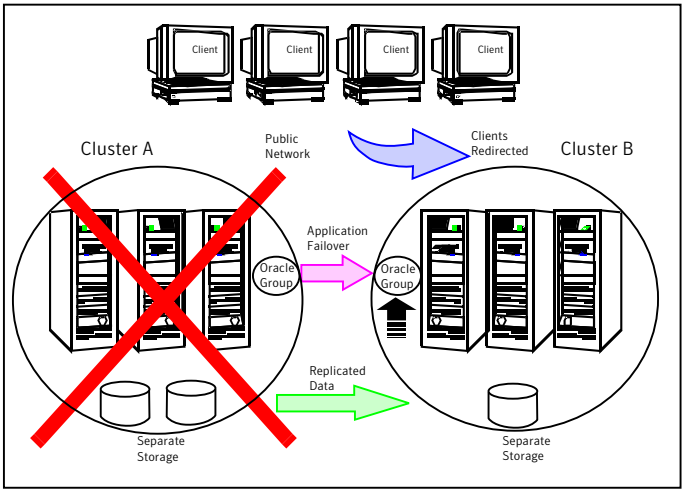
About global clusters

A global cluster links clusters at separate locations and enables wide-area failover and disaster recovery.

Local clustering provides local failover for each site or building. Campus and replicated cluster configurations offer protection against disasters that affect limited geographic regions. Large scale disasters such as major floods, hurricanes, and earthquakes can cause outages for an entire city or region. In such situations, you can ensure data availability by migrating applications to sites located considerable distances apart.

Figure 1-4 shows a global cluster configuration.

Figure 1-4 Global cluster



In a global cluster, if an application or a system fails, the application is migrated to another system within the same cluster. If the entire cluster fails, the application is migrated to a system in another cluster. Clustering on a global level also requires the replication of shared data to the remote site.

See " How VCS global clusters work" on page 23.

How VCS global clusters work

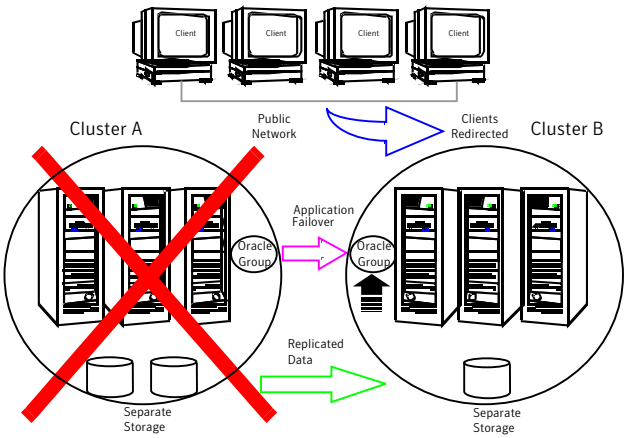
Local clustering provides local failover for each site or building. But, these configurations do not provide protection against large-scale disasters such as major floods, hurricanes, and earthquakes that cause outages for an entire city or region. The entire cluster could be affected by an outage.

In such situations, VCS global clusters ensure data availability by migrating applications to remote clusters located considerable distances apart.

Let us take the example of an Oracle database configured in a VCS global cluster. Oracle is installed and configured in both clusters. Oracle data is located on shared disks within each cluster and is replicated across clusters to ensure data concurrency. The Oracle service group is online on a system in cluster A and is configured to fail over globally, on clusters A and B.

Figure 1-5 shows a sample global cluster setup.

Figure 1-5 Sample global cluster setup



VCS continuously monitors and communicates events between clusters. Inter-cluster communication ensures that the global cluster is aware of the state of the service groups that are configured in the global cluster at all times.

In the event of a system or application failure, VCS fails over the Oracle service group to another system in the same cluster. If the entire cluster fails, VCS fails over the service group to the remote cluster, which is part of the global cluster. VCS also redirects clients once the application is online on the new location.

User privileges in global clusters

VCS permits a cross-cluster online or offline operation only if the user initiating the operation has one of the following privileges:

- Group administrator or group operator privileges for the group on the remote cluster
- Cluster administrator or cluster operator privileges on the remote cluster

VCS permits a cross-cluster switch operation only if the user initiating the operation has the following privileges:

- Group administrator or group operator privileges for the group on both clusters
- Cluster administrator or cluster operator privileges on both clusters

VCS global clusters: The building blocks

VCS extends clustering concepts to wide-area high availability and disaster recovery with the following:

- Remote cluster objects
  See " Visualization of remote cluster objects" on page 25.
- Global service groups
  See "About global service groups" on page 26.
- Global cluster management
  See "About global cluster management" on page 26.
- Serialization
  See "About serialization–The Authority attribute" on page 27.
- Resiliency and right of way
  See "About resiliency and "Right of way"" on page 28.
- VCS agents to manage wide-area failover
  See " VCS agents to manage wide-area failover" on page 28.
- Split-brain in two-cluster global clusters
  See "About the Steward process: Split-brain in two-cluster global clusters" on page 28.
- Secure communication
  See " Secure communication in global clusters" on page 30.

Visualization of remote cluster objects

VCS enables you to visualize remote cluster objects using any of the supported components that are used to administer VCS.

You can define remote clusters in your configuration file, `main.cf`. The Remote Cluster Configuration wizard provides an easy interface to do so. The wizard updates the `main.cf` files of all connected clusters with the required configuration changes.

About global service groups

A global service group is a regular VCS group with additional properties to enable wide-area failover. The global service group attribute ClusterList defines the list of clusters to which the group can fail over. The service group must be configured on all participating clusters and must have the same name on each cluster. The Global Group Configuration Wizard provides an easy interface to configure global groups.

About global cluster management

VCS enables you to perform operations (online, offline, switch) on global service groups from any system in any cluster. You must log on with adequate privileges for cluster operations.

See "User privileges in global clusters" on page 24.

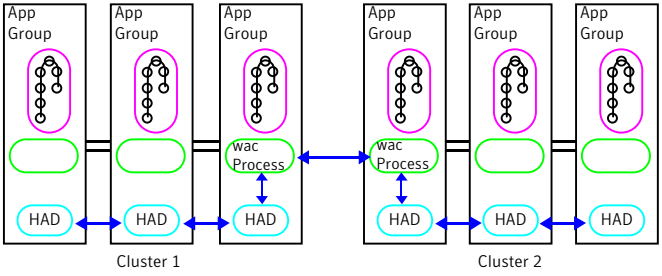
You can bring service groups online or switch them to any system in any cluster. If you do not specify a target system, VCS uses the FailOverPolicy to determine the system.

Management of remote cluster objects is aided by inter-cluster communication enabled by the wide-area connector (wac) process.

About the wide-area connector process

Figure 1-6 is an illustration of the wide-area connector process.

Figure 1-6 Wide-area connector (wac) process



The wac process runs on one system in each cluster and connects with peers in remote clusters. It receives and transmits information about the status of the cluster, service groups, and systems. This communication enables VCS to create a consolidated view of the status of all the clusters configured as part of the global cluster. The process also manages wide-area heartbeating to determine the health of remote clusters. The process also transmits commands between clusters and returns the result to the originating cluster.

VCS provides the option of securing the communication between the wide-area connectors.

See " Secure communication in global clusters" on page 30.

About the wide-area heartbeat agent

The wide-area heartbeat agent manages the inter-cluster heartbeat. Heartbeats are used to monitor the health of remote clusters. VCS wide-area hearbeat agents include Icmp and IcmpS. While other VCS resource agents report their status to VCS engine, heartbeat agents report their status directly to the WAC process. The heartbeat name must be the same as the heartbeat type name. You can add only one heartbeat of a specific heartbeat type.

You can create custom wide-area heartbeat agents. For example, the VCS replication agent for SRDF includes a custom heartbeat agent for Symmetrix arrays.

You can add heartbeats using the `hahb -add` *`heartbeatname`* command and change the default values of the heartbeat agents using the `hahb -modify` command.

About serialization–The Authority attribute

VCS ensures that multi-cluster service group operations are conducted serially to avoid timing problems and to ensure smooth performance. The Authority attribute prevents a service group from coming online in multiple clusters at the same time. Authority is a persistent service group attribute and it designates which cluster has the right to bring a global service group online. The attribute cannot be modified at runtime.

If two administrators simultaneously try to bring a service group online in a two-cluster global group, one command is honored, and the other is rejected based on the value of the Authority attribute.

The attribute prevents bringing a service group online in a cluster that does not have the authority to do so. If the cluster holding authority is down, you can enforce a takeover by using the command `hagrp -online -force` *`service_group`*. This command enables you to fail over an application to another cluster when a disaster occurs.

Note: A cluster assuming authority for a group does not guarantee the group will be brought online on the cluster. The attribute merely specifies the right to attempt bringing the service group online in the cluster. The presence of Authority does not override group settings like frozen, autodisabled, non-probed, and so on, that prevent service groups from going online.

You must seed authority if it is not held on any cluster.

Offline operations on global groups can originate from any cluster and do not require a change of authority to do so, because taking a group offline does not necessarily indicate an intention to perform a cross-cluster failover.

About the Authority and AutoStart attributes

The attributes Authority and AutoStart work together to avoid potential concurrency violations in multi-cluster configurations.

If the AutoStartList attribute is set, and if a group's Authority attribute is set to 1, the VCS engine waits for the wac process to connect to the peer. If the connection fails, it means the peer is down and the AutoStart process proceeds. If the connection succeeds, HAD waits for the remote snapshot. If the peer is holding the authority for the group and the remote group is online (because of takeover), the local cluster does not bring the group online and relinquishes authority.

If the Authority attribute is set to 0, AutoStart is not invoked.

About resiliency and "Right of way"

VCS global clusters maintain resiliency using the wide-area connector process and the ClusterService group. The wide-area connector process runs as long as there is at least one surviving node in a cluster.

The wide-area connector, its alias, and notifier are components of the ClusterService group.

VCS agents to manage wide-area failover

VCS agents now manage external objects that are part of wide-area failover. These objects include replication, DNS updates, and so on. These agents provide a robust framework for specifying attributes and restarts, and can be brought online upon fail over.

About the Steward process: Split-brain in two-cluster global clusters

Failure of all heartbeats between any two clusters in a global cluster indicates one of the following:

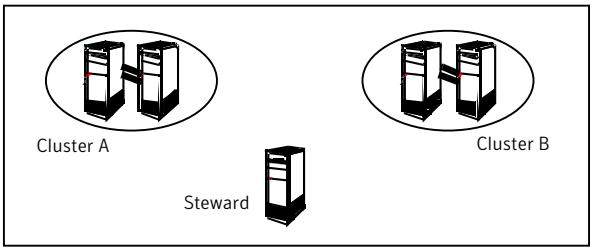
- The remote cluster is faulted.
- All communication links between the two clusters are broken.

In global clusters with more than three clusters, VCS queries the connected clusters to confirm that the remote cluster is truly down. This mechanism is called inquiry.

In a two-cluster setup, VCS uses the Steward process to minimize chances of a wide-area split-brain. The process runs as a standalone binary on a system outside of the global cluster configuration.

Figure 1-7 depicts the Steward process to minimize chances of a split brain within a two-cluster setup.

Figure 1-7 Steward process: Split-brain in two-cluster global clusters



When all communication links between any two clusters are lost, each cluster contacts the Steward with an inquiry message. The Steward sends an ICMP ping to the cluster in question and responds with a negative inquiry if the cluster is running or with positive inquiry if the cluster is down. The Steward can also be used in configurations with more than two clusters. VCS provides the option of securing communication between the Steward process and the wide-area connectors.

See " Secure communication in global clusters" on page 30.

In non-secure configurations, you can configure the steward process on a platform that is different to that of the global cluster nodes. Secure configurations have not been tested for running the steward process on a different platform.

For example, you can run the steward process on a Windows system for a global cluster running on AIX systems. However, the VCS release for AIX contains the steward binary for AIX only. You must copy the steward binary for Windows from the VCS installation directory on a Windows cluster, typically `C:\Program Files\VERITAS\Cluster Server`.

A Steward is effective only if there are independent paths from each cluster to the host that runs the Steward. If there is only one path between the two clusters, you must prevent split-brain by confirming manually via telephone or some messaging system with administrators at the remote site if a failure has occurred. By default, VCS global clusters fail over an application across cluster boundaries with administrator confirmation. You can configure automatic failover by setting the ClusterFailOverPolicy attribute to Auto.

For more information on configuring the Steward process, see the *Veritas Cluster Server Administrator's Guide*.

The default port for the steward is 14156.

Secure communication in global clusters

In global clusters, VCS provides the option of making the following types of communication secure:

- Communication between the wide-area connectors.
- Communication between the wide-area connectors and the Steward process.

For secure authentication, the wide-area connector process gets a security context as an account in the local authentication broker on each cluster node.

The WAC account belongs to the same domain as HAD and Command Server and is specified as:

```
name = WAC
domain = VCS_SERVICES@cluster_uuid
```

You must configure the wide-area connector process in all clusters to run in secure mode. If the wide-area connector process runs in secure mode, you must run the Steward in secure mode.

Migrating from non-secure to secure setup for CP server and VCS cluster communication

The following procedure describes how to migrate from a non-secure to secure set up for the coordination point server (CP server) and VCS cluster.

To migrate from non-secure to secure setup for CP server and VCS cluster

1. Stop VCS on all cluster nodes that use the CP servers.

   ```
   # hastop -all
   ```

2. Stop fencing on all the VCS cluster nodes of all the clusters.

   ```
   # /etc/init.d/vxfen.rc stop
   ```

3. Stop all the CP servers using the following command on each CP server:

   ```
   # hagrp -offline CPSSG -any
   ```

4. Ensure that security is configured for communication on CP Servers as well as all the clients.

   See the *Veritas Cluster Server Installation Guide* for more information.

5 ■ If CP server is hosted on an SFHA cluster, perform this step on each CP server.
Bring the mount resource in the CPSSG service group online.

```
# hares -online cpsmount -sys local_system_name
```

Complete the remaining steps.

■ If CP server is hosted on a single-node VCS cluster, skip to step 8 and complete the remaining steps.

6 After the mount resource comes online, move the `credentials` directory from the default location to shared storage.

```
# mv /var/VRTSvcs/vcsauth/data/CPSERVER /etc/VRTSvcs/db/
```

7 Create softlinks on all the nodes of the CP servers.

```
# ln -s /etc/VRTScps/db/CPSERVER \
/var/VRTSvcs/vcsauth/data/CPSERVER
```

8 Edit `/etc/vxcps.conf` on each CP server to set security=1.

9 Start CP servers by using the following command:

```
# hagrp -online CPSSG -any
```

10 Edit `/etc/VRTSvcs/conf/config/main.cf` on the first node of the cluster and remove the UseFence=SCSI3 attribute.

Start VCS on the first node and then on all other nodes of the cluster.

11 Reconfigure fencing on each cluster by using the installer.

```
# /opt/VRTS/install/installvcs<version> -fencing
```

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

Disaster recovery feature support across Storage Foundation and High Availability Solutions 6.0.1 products

Disaster recovery solutions and use cases are based on the shared availability and disaster recovery features of Veritas Storage Foundation and High Availability (SFHA) Solutions products. Clustering and disaster recovery features are available

separately through Veritas Cluster Server (VCS) as well as through the SFHA Solutions products which include VCS as a component.

Table 1-3 lists high availability and disaster recovery features available in SFHA Solutions products.

Table 1-3 High availability and disaster recovery feature support in SFHA Solutions products

High availability and disaster recovery features	VCS	VCS HA/DR	SF Std. HA	SF Ent. HA	SFCFS HA	SFRAC	SF Sybase CE
Clustering for high availability (HA)	Y	Y	Y	Y	Y	Y	Y
Database and application/ISV agents	Y	Y	Y	Y	Y	Y	Y
Advanced failover logic	Y	Y	Y	Y	Y	Y	Y
Data integrity protection with I/O fencing	Y	Y	Y	Y	Y	Y	Y
Advanced virtual machines support	Y	Y	Y	Y	Y	Y	Y
Virtual Business Services	Y	Y	Y	Y	Y	Y	N
Replication agents for VVR	N	Y	O	O	O	O	O
Replication agents for third-party array-based replication	N	Y	O	O	O	O	N

Table 1-3 High availability and disaster recovery feature support in SFHA Solutions products *(continued)*

High availability and disaster recovery features	VCS	VCS HA/DR	SF Std. HA	SF Ent. HA	SFCFS HA	SFRAC	SF Sybase CE
Replicated Data Cluster	N	Y	O	O	O	N	N
Campus or stretch cluster	N	Y	O	O	O	O	N
Global clustering (GCO) using VVR	N	Y	O	O	O	O	O
Global clustering (GCO) using third-party array-based replication	N	Y	O	O	O	O	N
Fire Drill	N	Y	O	O	O	O	O

- Y=Feature is included in your license.
- O=Feature is not included in your license but may be licensed separately.
- N=Feature is not supported with your license.

The following SFHA Solutions products support multiple third-party replication options:

- Veritas Cluster Server (VCS)
- Veritas Storage Foundation High Availability (SFHA)
- Veritas Storage Foundation Cluster File System High Availability (SFCFSHA)
- Veritas Storage Foundation for Oracle RAC (SF Oracle RAC)

Veritas Storage Foundation for Sybase CE supports VVR replication only at this time.

For current information on third-party replication support:

See: https://sort.symantec.com/agents and select **Replication Agents** under **Agent type**.

Table 1-4 Replication support for databases across SFHA Solutions 6.0.1 products

Database replication support	VCS	VCS HA/DR	SF Std. HA	SF Ent. HA	SFCFS HA	SFRAC	SF Syb CE
DB2	Y	Y	Y	Y	Y	N	N
Single instance Oracle	Y	Y	Y	Y	Y	Y	N
Oracle RAC	N	N	N	N	N	Y	N
Sybase	Y	Y	Y	Y	Y	N	N
Syabase ASE CE	Y	Y	Y	Y	Y	N	Y

Single instance Oracle and Oracle RAC replication support includes Storage Foundation for Databases (SFDB) tools replication support.

Replication agent support in virtual environments

All VCS-supported replication agents listed on SORT are supported inside LPARs (virtual) with virtual devices. The LPAR support matrix is the same as the VCS support matrix. If VCS is supported inside a vitual machine on a virtualization platform, then the replication agent is also supported.

Pre-requisite for replication agent support in virtual environments:

Make the following disks visible to the LPAR as pass thru devices (NPIV) for the following replication agents:

- SRDF: Gatekeeper
- HTC agents: Command Device

Exception:

Firedrill functionality is not supported in virtual environments for the following replication agents:

- EMC MirrorView
- HP-UX EVA CA

Only Firedrill functionality is affected: these replication agents can still be used to manage replication inside LPARs.

Chapter 2

Planning for disaster recovery

This chapter includes the following topics:

- Planning for cluster configurations
- Planning for data replication

Planning for cluster configurations

Storage Foundation and High Availability Solutions provides various disaster recovery configurations, such as campus clusters, global clusters for multi-site clusters. In multi-site clusters, the nodes can be placed in different parts of a building, in separate buildings, or in separate cities. The distance between The nodes depends on The type of disaster from which protection is needed and on The technology used to replicate data. Storage Foundation and High Availability supports various replication technologies for data replication.

To protect clusters against outages caused by disasters, the cluster components must be geographically separated.

Planning a campus cluster setup

A campus cluster is also known as a stretch cluster or remote mirror configuration. In a campus cluster, the hosts and storage of a cluster span multiple sites separated by a few miles.

Keep in mind the following best practices when you configure a Storage Foundation campus cluster:

- Campus cluster sites are typically connected using a redundant high-capacity network that provides access to storage and private network communication

between the cluster nodes. A single DWDM link can be used for both storage and private network communication.

- Tag the disks or enclosures that belong to a site with the corresponding VxVM site name. VxVM allocates storage from the correct site when creating or resizing a volume and when changing a volume's layout if the disks in the VxVM disk group that contain the volume are tagged with the site name.
- Tag each host with the corresponding VxVM site name. Make sure the read policy of the volumes is set to `SITEREAD`. This setting ensures that the reads on the volumes are satisfied from the local site's plex.
- Turn on the `allsites` attribute for all volumes that have data required by the application, to make sure they are evenly mirrored. Each site must have at least one mirror of all volumes hosting application data, including the FlashSnap log volume.
- Turn on the `siteconsistent` attribute for the disk groups and the volumes to enable site-aware plex detaches. Snapshot volumes need not be site-consistent.
- In the case of a two-site campus cluster, place the third coordinator disk on the third site. You may use iSCSI disk on the third site as an alternative to Dark Fiber connected FC-SAN or a Coordination Point Server (CPS), as a third coordination point.
- Make sure that a DCO log version 20 or higher is attached to the volumes to enable Fast Resync operations.
- Set the CVM disk detach policy as `global` or `local` for all disk groups containing data volumes.
  For OCR and voting disk, it is recommended to have the disk group policy as `local` detach policy.

Planning a replicated data cluster setup

The VCS replicated data cluster (RDC) configuration allows you to provide a robust and easy-to manage disaster recovery protection for your applications. For example you can convert a single instance database configured for local high availability in a VCS cluster to a disaster-protected RDC infrastructure using Veritas Volume Replicator or a supported third-party replication technology to replicate changed data.

Keep in mind the following best practicies when you configure an RDC:

- Make sure the sites and systems at each site are identified correctly for use when defining system zones in an RDC.
- Make sure there are dual dedicated LLT links between the replicated nodes.

- Since the sites used in the RDC configuration are within metro limits, synchronous replication is typically used. Make sure the replication technology that you plan to use supports synchronous replication mode.

The RDC can also be configured using supported third-party replication technologies.

See "Planning for data replication" on page 38.

Planning a global cluster setup

Global clusters provide the ability to fail over applications between geographically distributed clusters when a disaster occurs.

Global clustering involves two steps:

1. Replication of data between the sites
2. Configuring VCS clusters at the geographically distant sites and establishing a global cluster connection between them

The following aspects need to be considered when you design a disaster recovery solution:

- The amount of data lost in the event of a disaster (Recovery Point Objective)
- The acceptable recovery time after the disaster (Recovery Time Objective)

Planning for data replication

When planning for data replication, it is important to review the various hardware and software replication technologies and to review important considerations including the required level of data throughput.

Data replication options

Disaster recovery solutions support various hardware and software replication technologies.

Examples of hardware replication options	■ Hitachi True Copy ■ IBM Metro Mirror ■ IBM SVC ■ EMC Mirror View
Examples of software replication options	■ Veritas Volume Replicator (VVR) ■ Oracle Data Guard

A complete list of supported replication technologies is listed on the Symantec Web site:

https://sort.symantec.com/agents

Data replication considerations

When you choose a replication solution, one of the important factors that you need to consider is the required level of data throughput. Data throughput is the rate at which the application is expected to write data. The impact of write operations on replication are of more significance than that of the read operations.

In addition to the business needs discussed earlier, the following factors need to be considered while choosing the replication options:

- Mode of replication
- Network bandwidth
- Network latency between the two sites
- Ability of the remote site to keep up with the data changes at the first site

Section 2

Implementing campus clusters

- Chapter 3. Setting up campus clusters for VCS and SFHA
- Chapter 4. Setting up campus clusters for SFCFS, SFRAC

Chapter 3

Setting up campus clusters for VCS and SFHA

This chapter includes the following topics:

- About setting up a campus cluster configuration
- Fire drill in campus clusters
- About the DiskGroupSnap agent
- About running a fire drill in a campus cluster

About setting up a campus cluster configuration

You must perform the following tasks to set up a campus cluster:

- Preparing to set up a campus cluster configuration
- Configuring I/O fencing to prevent data corruption
- Configuring VxVM disk groups for campus cluster configuration
- Configuring VCS service group for campus clusters

Preparing to set up a campus cluster configuration

Before you set up the configuration, review the VCS campus cluster requirements.

See " VCS campus cluster requirements" on page 14.

To prepare to set up a campus cluster configuration

1 Set up the physical infrastructure.

- Set up access to the local storage arrays and to remote storage arrays on each node.
- Set up private heartbeat network.

See " Typical VCS campus cluster setup" on page 19.

2 Install VCS on each node to form a cluster with at least one node in each of the two sites.

See the *Veritas Cluster Server Installation Guide* for instructions.

3 Install VxVM on each node with the required licenses.

See the *Veritas Storage Foundation and High Availability Installation Guide* for instructions.

Configuring I/O fencing to prevent data corruption

Perform the following tasks to configure I/O fencing to prevent data corruption in the event of a communication failure.

See the *Veritas Cluster Server Installation Guide* for more details.

To configure I/O fencing to prevent data corruption

1 Set up the storage at a third site.

You can extend the DWDM to the third site to have FC SAN connectivity to the storage at the third site. You can also use iSCSI targets as the coordinator disks at the third site.

2 Set up I/O fencing.

Configuring VxVM disk groups for campus cluster configuration

Follow the procedure to configure VxVM disk groups for remote mirroring.

See the *Veritas Storage Foundation Cluster File System High Availability Administrator's Guide* for more information on the VxVM commands.

To configure VxVM disk groups for campus cluster configuration

1 Set the site name for each host:

```
# vxdctl set site=sitename
```

The site name is stored in the /etc/vx/volboot file. Use the following command to display the site names:

```
# vxdctl list | grep siteid
```

2 Set the site name for all the disks in an enclosure:

```
# vxdisk settag site=sitename encl:enclosure
```

To tag specific disks, use the following command:

```
# vxdisk settag site=sitename disk
```

3 Verify that the disks are registered to a site.

```
# vxdisk listtag
```

4 Create a disk group with disks from both the sites.

```
# vxdg -s init diskgroup siteA_disk1 siteB_disk2 layout=layout
```

5 Configure site-based allocation on the disk group that you created for each site that is registered to the disk group.

```
# vxdg -g diskgroup addsite sitename
```

6 Configure site consistency on the disk group.

```
# vxdg -g diskgroup set siteconsistent=on
```

7 Create one or more mirrored volumes in the disk group.

```
# vxassist -g diskgroup make volume size layout=layout
```

With the Site Awareness license installed on all hosts, the volume that you create has the following characteristics by default:

- The `allsites` attribute is set to `on`; the volumes have at least one plex at each site.
- The volumes are automatically mirrored across sites.
- The read policy `rdpol` is set to `siteread`.
- The volumes inherit the site consistency value that is set on the disk group.

Configuring VCS service group for campus clusters

Follow the procedure to configure the disk groups under VCS control and set up the VCS attributes to define failover in campus clusters.

To configure VCS service groups for campus clusters

1. Create a VCS service group (app_sg) for the application that runs in the campus cluster.

   ```
   hagrp -add app_sg
   hagrp -modify app_sg SystemList node1 0 node2 1 node3 2 node4 3
   ```

2. Set up the system zones. Configure the SystemZones attribute for the service group.

   ```
   hagrp -modify app_sg SystemZones node1 0 node2 0 node3 1 node4 1
   ```

3. Set up the group fail over policy. Set the value of the AutoFailOver attribute for the service group.

   ```
   hagrp -modify app_sg AutoFailOver 2
   ```

4. For the disk group you created for campus clusters, add a DiskGroup resource to the VCS service group app_sg.

   ```
   hares -add dg_res1 DiskGroup app_sg
   hares -modify dg_res1 DiskGroup diskgroup_name
   hares -modify dg_res1 Enabled 1
   ```

5. Configure the application and other related resources to the app_sg service group.
6. Bring the service group online.

Fire drill in campus clusters

Fire drill tests the disaster-readiness of a configuration by mimicking a failover without stopping the application and disrupting user access.

The process involves creating a fire drill service group, which is similar to the original application service group. Bringing the fire drill service group online on the remote node demonstrates the ability of the application service group to fail over and come online at the site, should the need arise.

Fire drill service groups do not interact with outside clients or with other instances of resources, so they can safely come online even when the application service group is online. Conduct a fire drill only at the remote site; do not bring the fire drill service group online on the node hosting the original application.

About the DiskGroupSnap agent

The DiskGroupSnap agent verifies the VxVM disk groups and volumes for site awareness and disaster readiness in a campus cluster environment. To perform a fire drill in campus clusters, you must configure a resource of type DiskGroupSnap in the fire drill service group.

Note: To perform fire drill, the application service group must be online at the primary site.

During fire drill, the DiskGroupSnap agent does the following:

- For each node in a site, the agent correlates the value of the SystemZones attribute for the application service group to the VxVM site names for that node.
- For the disk group in the application service group, the agent verifies that the VxVM site tags are defined for the disk group.
- For the disk group in the application service group, the agent verifies that the disks at the secondary site are not tagged with the same VxVM site name as the disks at the primary site.
- The agent verifies that all volumes in the disk group have a plex at each site.

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

About running a fire drill in a campus cluster

This topic provides information on how to run a fire drill in campus clusters.

Do the following tasks to perform fire drill:

- Configuring the fire drill service group
- Running a successful fire drill in a campus cluster

Configuring the fire drill service group

This topic provides information on how to configure the fire drill service group.

To configure the fire drill service group

1. Configure a fire drill service group similar to the application service group with the following exceptions:
   - The AutoFailOver attribute must be set to 0.
   - Network-related resources must not be configured.
   - The disk group names for the DiskGroup and the Mount resources in the fire drill service group must be appended with "_fd".
     For example, if the value of the DiskGroup attribute in the application service group is ccdg, then the corresponding value in the fire drill service group must be ccdg_fd.
     If the value of the BlockDevice attribute for the Mount resource in the application service group is /dev/vx/dsk/ccdg/ccvol, then the corresponding value in the fire drill service group must be /dev/vx/dsk/ccdg_fd/ccvol.
2. Add a resource of type DiskGroupSnap. Define the TargetResName and the FDSiteName attributes for the DiskGroupSnap resource.

   See the *Veritas Cluster Server Bundled Agent Reference Guide* for attribute descriptions.
3. Create a dependency such that the DiskGroup resource depends on the DiskGroupSnap resource.
4. Create a group dependency such that the fire drill service group has an offline local dependency on the application service group.

Running a successful fire drill in a campus cluster

Bring the fire drill service group online on a node within the system zone that does not have the application running. Verify that the fire drill service group comes online. This action validates that your solution is configured correctly and the production service group will fail over to the remote site in the event of an actual failure (disaster) at the local site.

You must take the fire drill service group offline before you shut down the node or stop VCS locally on the node where the fire drill service group is online or where

the disk group is online. Otherwise, after the node restarts you must manually reattach the fire drill site to the disk group that is imported at the primary site.

Note: For the applications for which you want to perform fire drill, you must set the value of the FireDrill attribute for those application resource types to 1. After you complete fire drill, reset the value to 0.

To run a successful fire drill

1. Set the FireDrill attribute for the application resource type to 1 to prevent the agent from reporting a concurrency violation when the application service group and the fire drill service group are online at the same time.
2. Bring the fire drill service group online.

   If the fire drill service group does not come online, review the VCS engine log to troubleshoot the issues so that corrective action can be taken as necessary in the production service group.

   Warning: You must take the fire drill service group offline after you complete the fire drill so that the failover behavior of the application service group is not impacted. Otherwise, when a disaster strikes at the primary site, the application service group cannot fail over to the secondary site due to resource conflicts.

   ---

3. After you complete the fire drill, take the fire drill service group offline.
4. Reset the FireDrill attribute for the application resource type to 0.

Chapter 4

Setting up campus clusters for SFCFS, SFRAC

This chapter includes the following topics:

- About setting up a campus cluster for disaster recovery for SFCFS HA or SF Oracle RAC
- Preparing to set up a campus cluster in a parallel cluster database environment
- Configuring I/O fencing to prevent data corruption
- Configuring VxVM disk groups for a campus cluster in a parallel cluster database environment
- Configuring VCS service groups for a campus cluster for SFCFS HA and SF Oracle RAC
- Tuning guidelines for parallel campus clusters
- Best practices for a parallel campus cluster

About setting up a campus cluster for disaster recovery for SFCFS HA or SF Oracle RAC

Campus clusters:

- Are connected using a high speed cable that guarantees network access between the nodes
- Provide local high availability and disaster recovery functionality in a single cluster

- Employ shared disk groups mirrored across sites with Veritas Volume Manager (VxVM)
- Are supported for Storage Foundation and High Availability (SFHA) Solutions products including Storage Foundation Cluster File System(SFCFS HA) and Storage Foundation (SF) for Oracle RAC

Note: Campus clusters are not supported for Storage Foundation for Sybase CE at this time.

The following high-level tasks illustrate the setup steps for a campus cluster in a parallel cluster database environment. The example values are given for SF for Oracle RAC and should be adapted for an SFCFS HA cluster using another database application.

The following high-level tasks illustrate the setup steps for a parallel campus cluster in an SF for Oracle RAC environment.

Table 4-1 Tasks for setting up a parallel campus cluster for disaster recovery

Task	Description
Prepare to set up campus cluster configuration	See "Preparing to set up a campus cluster in a parallel cluster database environment" on page 51.
Configure I/O fencing to prevent data corruption	See "Configuring I/O fencing to prevent data corruption" on page 42.
Prepare to install Oracle RAC Clusterware and database binaries	See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.
Configure VxVM disk groups for campus cluster	See "Configuring VxVM disk groups for a campus cluster in a parallel cluster database environment" on page 54.
Install Oracle RAC Clusterware and database binaries	For Oracle RAC, see the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*. For SFCFS HA, see your database documentation.
Configure VCS service groups	See "Configuring VCS service groups for a campus cluster for SFCFS HA and SF Oracle RAC" on page 58.

The sample SF Oracle RAC configuration illustrates the configuration procedures with a four-node campus cluster with two nodes at each site. Each node is running SF Oracle RAC 6.0.1.

Figure 4-1 Sample SF Oracle RAC configuration

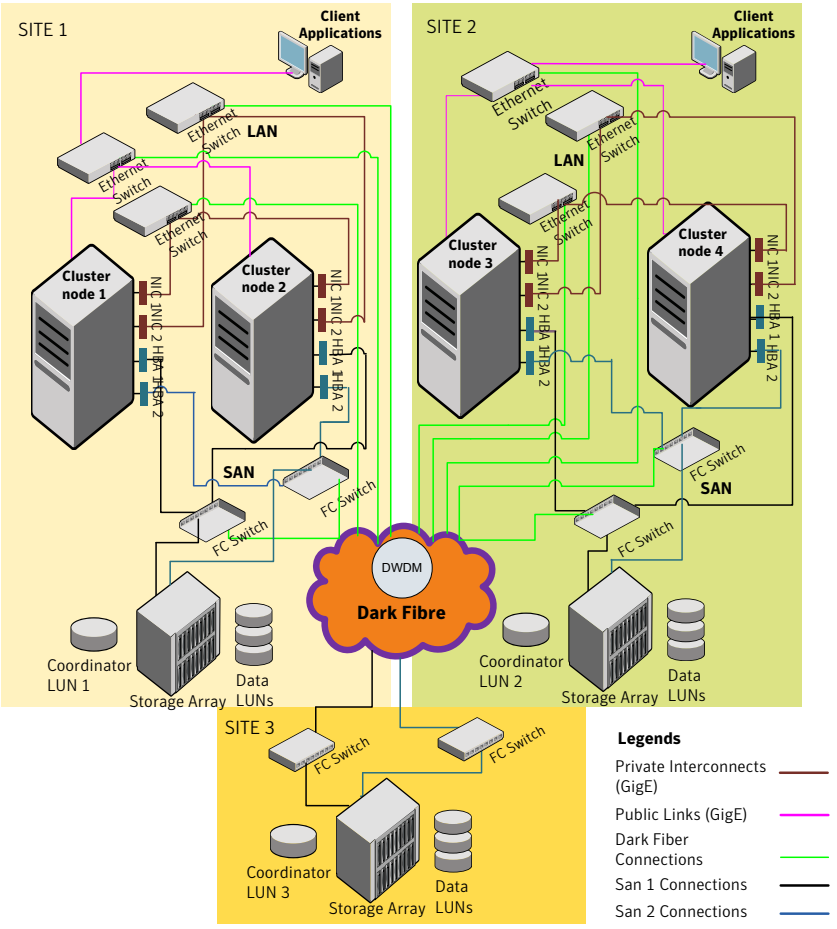


Table 4-2 Sample setup for an SF Oracle RAC campus cluster

Site	Hardware description
Site 1: site1 Cluster name: clus1	Servers: sys1 and sys2 Shared LUNs: disk01 disk02 disk03 disk04 (used as coordinator disk) disk05
Site 2: Site name: site2 Cluster name: clus1	Servers: sys3 and sys4 Shared LUNs: disk06 disk07 disk08 disk09 (used as coordinator disk)
Site 3: Site name: site3 Cluster name: clus1	Shared LUN disk10 (used as coordinator disk)

Although a Coordination Point (CP) server is not used in the current example, it can also be used instead of a third site for a coordinator disk.

Preparing to set up a campus cluster in a parallel cluster database environment

To verify your configuration is supported, review the product requirements and licensing information:

See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*.

See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

You will need to verify the following to setup a campus cluster:

- Hardware requirements for Veritas Storage Foundation Custer File System High Availability (SFCFS HA) or Veritas Storage Foundation for Oracle RAC (SF Oracle RAC)
- License requirements: in addition to your product with HA/DR, you will need:
  - FlashSnap license
  - Site awareness license

With keyless licensing, your enterprise product keys enable all of the above features.

The following preparation must be completed before configuring the campus cluster.

To prepare to set up a campus cluster

1 Configure the physical infrastructure for campus cluster:
   - Set up access to the local storage arrays and to remote storage arrays on each node. The storage link will extend to the third site as well.
   - Set up the private heartbeat network
     See " Typical VCS campus cluster setup" on page 19.

2 Install the operating system on all the nodes of the cluster.

   See your operating system documenation.

3 Install and configure either SFCFS HA or SF Oracle RAC on all nodes on both the sites.

   See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*.

   See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

   For a setup example, steps are provided to install and configure SF Oracle RAC 6.0.1 on all four nodes. Your installation and configuration steps will necessarily differ to reflect your configuration details.

Configuring I/O fencing to prevent data corruption

Perform the following tasks to configure I/O fencing to prevent data corruption in the event of a communication failure.

To configure I/O fencing to prevent data corruption

1 After installing and configuring SFCFS HA or SF Oracle RAC, configure I/O fencing for data integrity.

   See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*.

See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

2 Set up the storage at a third site.

You can extend the DWDM to the third site to have FC SAN connectivity to the storage at the third site. You can also use iSCSI targets as the coordinator disks at the third site.

For example:

Enable I/O fencing by using the coordinator disks from all the three sites.

```
# vxdisksetup -i disk04 format=cdsdisk


# vxdisksetup -i disk09 format=cdsdisk
# vxdisksetup -i disk10 format=cdsdisk
# hastop -all
# vxdg init fencedg disk10 disk04 disk09
# vxdg -g fencedg set coordinator=on
# vxdg deport fencedg
# vxdg -t import fencedg
# vxdg deport fencedg
```

Edit the main.cf to add "UseFence = SCSI3"

```
# vi /etc/VRTSvcs/conf/config/main.cf
# more /etc/vxfendg
fencedg
# more /etc/vxfentab
/dev/vx/rdmp/disk10
/dev/vx/rdmp/disk04
/dev/vx/rdmp/disk09
# cp /etc/vxfen.d/vxfenmode_scsi3_dmp /etc/vxfemode
# /sbin/init.d/vxfen stop
# /sbin/init.d/vxfen start
Starting vxfen..
Checking for /etc/vxfendg
Starting vxfen.. Done
```

On all nodes, start VCS:

```
# hastart
```

Set the site name for each host.

```
# site=site1
# vxdctl set site=site2
# vxdctl set site=site2
```

3 Start I/O fencing by using the disks from all three sites.

Configuring VxVM disk groups for a campus cluster in a parallel cluster database environment

After configuring I/O fencing for data integrity, you must configure the VxVM disk groups for a campus cluster before installing your database by configuring VxVM disk groups for remote mirroring.

For the example configuration, the database is Oracle RAC.

To configure VxVM disk groups for Oracle RAC on an SF for Oracle RAC campus cluster

1 Initialize the disks as CDS disks

```
# vxdisksetup -i disk01 format=cdsdisk
# vxdisksetup -i disk02 format=cdsdisk
# vxdisksetup -i disk03 format=cdsdisk
# vxdisksetup -i disk05 format=cdsdisk
# vxdisksetup -i disk06 format=cdsdisk
# vxdisksetup -i disk07 format=cdsdisk
# vxdisksetup -i disk08 format=cdsdisk
```

2 Set the site name for each host:

```
# vxdctl set site=sitename
```

The site name is stored in the /etc/vx/volboot file. To display the site names:

```
# vxdctl list | grep siteid
```

For example, for a four node cluster with two nodes at each site, mark the sites as follows:

On the nodes at first site:

```
# vxdctl set site=site1
```

On the nodes at second site:

```
# vxdctl set site=site2
```

3 Obtain the enclosure name using the following command:

```
# vxdmpadm listenclosure
ENCLR_NAME       ENCLR_TYPE     ENCLR_SNO  STATUS    ARRAY_TYPE  LUN_COUNT
============================================================================
ams_wms0         AMS_WMS        75040638   CONNECTED   A/A-A     35
hds9500-alua0    HDS9500-ALUA   D600145E   CONNECTED   A/A-A      9
hds9500-alua1    HDS9500-ALUA   D6001FD3   CONNECTED   A/A-A      6
disk             Disk           DISKS      CONNECTED   Disk       2
```

4 Set the site name for all the disks in an enclosure.

```
# vxdisk settag site=sitename encl:ENCLR_NAME
```

5 Run the following command if you want to tag only the specific disks:

```
# vxdisk settag site=sitename disk
```

For example:

```
# vxdisk settag site=site1 disk01
# vxdisk settag site=site1 disk02
# vxdisk settag site=site1 disk03
# vxdisk settag site=site2 disk06
# vxdisk settag site=site2 disk08
```

6 Verify that the disks are registered to a site.

```
# vxdisk listtag
```

For example:

```
# vxdisk listtag
DEVICE          NAME                          VALUE
disk01           site                          site1
disk02           site                          site1
disk03           site                          site1
disk04           site                          site1
disk05           site                          site1
disk06           site                          site2
disk07           site                          site2
disk08           site                          site2
disk09           site                          site2
```

7 Create a disk group for OCR and Vote Disks and another for Oracle data, with disks picked from both the sites. While the example below shows a single disk group, you can create as many as you need.

```
# vxdg -s init ocrvotedg disk05 disk07

# vxdg -s init oradatadg disk01 disk06
```

8 Enable site-based allocation on the disk groups for each site.

```
# vxdg -g ocrvotedg addsite site1

# vxdg -g ocrvotedg addsite site2

# vxdg -g oradatadg addsite site1

# vxdg -g oradatadg addsite site2
```

9 If you are using an enclosure, set the tag on the enclosure for both sites.

```
# vxdg -o retain -g ocrvotedg settag encl:3pardata0 site=site1

# vxdg -o retain -g ocrvotedg settag encl:3pardata1 site=site2

# vxdg -o retain -g oradatadg settag encl:3pardata0 site=site1

# vxdg -o retain -g oradatadg settag encl:3pardata1 site=site2
```

10 Configure site consistency for the disk groups.

```
# vxdg -g ocrvotedg set siteconsistent=on

# vxdg -g oradatadg set siteconsistent=on
```

11 Create one or more mirrored volumes in the disk group.

```
# vxassist -g ocrvotedg make ocrvotevol 2048m nmirror=2

# vxassist -g oradatadg make oradatavol 10200m nmirror=2

# vxassist -g ocrvotedg make ocrvotevol 2048m nmirror=2 \
    allsites=on siteconsistent=on

# vxassist -g oradatadg make oradatavol 10200m nmirror=2 \
    allsites=on siteconsistent=on
```

12 To verify the site awareness license, use the `vxlicrep` command. The Veritas VolumeManager product section should indicate: Site Awareness = Enabled

With the Site Awareness license installed on all hosts, the volume created has the following characteristics by default.

- The all sites attribute is set to ON; the volumes have at least one mirror at each site.
- The volumes are automatically mirrored across sites.
- The read policy (rdpol) is set to siteread.
  The read policy can be displayed using the `vxprint -ht` command.
- The volumes inherit the site consistency value that is set on the disk group.

13 From the CVM master, start the volumes for all the disk groups.

```
# vxvol -g ocrvotedg startall
# vxvol -g oradatadg startall
```

14 Create a file system on each volume and mount the same.

```
# mkfs -V vxfs /dev/vx/rdsk/ocrvotedg/ocrvotevol

# mkfs -V vxfs /dev/vx/rdsk/oradatadg/oradatavol

# mount -V vxfs -o cluster /dev/vx/dsk/ocrvotedg/ocrvotevol /ocrvote

# mount -V vxfs -o cluster /dev/vx/dsk/oradatadg/oradatavol /oradata
```

15 Create seperate directories for OCR and Vote file as follows:

```
# mkdir -p /ocrvote/ocr

# mkdir -p /ocrvote/vote
```

16 After creating directories, change the permissions of these directories to Oracle or Grid user:

```
# chown -R user:group /ocrvote
```

Also change the ownership of /oradata to Oracle user:

```
# chown user:group /oradata
```

Note: One Vote Disk is sufficient since it is already mirrored by VxVM.

17 Install you database software.

For Oralce RAC:

- Insall Oracle Clusterware/GRID
- Install Oracle RAC binaries
- Perform library linking of Oracle binaries
- Create the database on /oradata. For detailed steps,
  See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

Configuring VCS service groups for a campus cluster for SFCFS HA and SF Oracle RAC

Follow the procedure below to configure the disk groups under Storage Foundation (SF) for Oracle RAC control and set up the Veritas Cluster Server (VCS) attributes to define parallel applications in campus clusters. The Cluster Volume Manager (CVM) and Oracle service groups can be configured by editing the VCS configuration file, the main.cf, to define the service groups.

To configure the VCS service groups

1 Configure the disk groups under SFCFS HA or SF Oracle RAC control and set up the VCS attributes to define parallel applications in campus clusters. The CVM and Oracle service groups can be configured by editing the VCS configuration file, main.cf, to define the service groups.

2 Configure the SystemZones attribute in the service group definition as explained previously.

See "Configuring VCS service group for campus clusters" on page 44.

3 Group the hosts at each physical site into a single logical SystemZone. This will enable the failover applications to try to come up on local nodes before they try to come up on a remote site.

   Not all SFCFS HA or SF Oracle RAC service groups are parallel. In the sample configuration file, hosts sys1 and sys2 should be configured in zone 0 and hosts sys3 and sys4 in zone 1. In the event of a failure, this setting instructs VCS to failover the group first within the same site and then across the sites.

4 After configuring your service groups and before putting your configuration into production, you can verify your configuration resilience by means of testing various failure scenarios.

See "sfrac11_main.cf file" on page 162.

Tuning guidelines for parallel campus clusters

An important consideration while tuning a campus cluster in a Storage Foundation Cluster File System High Availability (SFCFS HA) or Storage Foundation (SF) for Oracle RAC environment is setting the LLT peerinact time. Follow the guidelines below to determine the optimum value of peerinact time:

- Calculate the roundtrip time using lltping (1M).
- Evaluate LLT heartbeat time as half of the round trip time.
- Set the LLT peer trouble time as 2-4 times the heartbeat time.
- LLT peerinact time should be set to be more than 4 times the heart beat time.

Best practices for a parallel campus cluster

The following best practices ensure a robust Storage Foundation Cluster File System High Availability (SFCFS HA) or Storage Foundation (SF) for Oracle RAC campus cluster:

- Tag all the mirrored volumes in the campus cluster with appropriate site names. VxVM allocates storage from the correct site when creating or resizing a volume and when changing a volume's layout if the volume is tagged with site name.
- All volumes that have data required by the application must be evenly mirrored. Each site must have at least one mirror of all volumes hosting application data, including the FlashSnap log volume.
- Do not enable site consistency on VxVM snapshot volumes.
- Use redundant links for storage and private interconnects. DWDM can be used for storage and heartbeat together. Another redundant DWDM link can be used

to prevent single point of failure. Separate switches and multiplexer / de-multiplexer devices should be used.

- Use Coordination Point Server as the third coordination point.
- Use the procedure for online replacement of coordination points, to replace disk based or Coordination Point Server based coordination points.

Section 3

Implementing replicated data clusters

- Chapter 5. Configuring a replicated data cluster using VVR
- Chapter 6. Configuring a replicated data cluster using third-party replication

Chapter 5

Configuring a replicated data cluster using VVR

This chapter includes the following topics:

- About setting up a replicated data cluster configuration
- About migrating a service group
- Fire drill in replicated data clusters

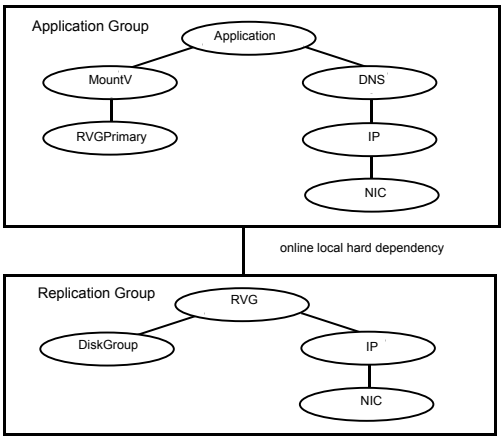
About setting up a replicated data cluster configuration

This topic describes the steps for planning, configuring, testing, and using the VCS RDC configuration to provide a robust and easy-to-manage disaster recovery protection for your applications. It describes an example of converting a single instance Oracle database configured for local high availability in a VCS cluster to a disaster-protected RDC infrastructure. The solution uses Veritas Volume Replicator to replicate changed data.

About typical replicated data cluster configuration

Figure 5-1 depicts a dependency chart of a typical RDC configuration.

Figure 5-1 Dependency chart of a typical RDC configuration



In this example, a single-instance application is configured as a VCS service group (DataGroup) on a four-node cluster, with two nodes in the primary RDC system zone and two in the secondary RDC system zone. In the event of a failure on the primary node, VCS fails over the application to the second node in the primary zone.

The process involves the following steps:

- Setting Up Replication
- Configuring the Service Groups
- Configuring the Service Group Dependencies

About setting up replication

Veritas Volume Replicator (VVR) technology is a license-enabled feature of Veritas Volume Manager (VxVM), so you can convert VxVM-managed volumes into replicated volumes managed using VVR. In this example, the process involves grouping the Oracle data volumes into a Replicated Volume Group (RVG), and creating the VVR Secondary on hosts in another VCS cluster, located in your DR site.

When setting up VVR, it is a best practice to use the same DiskGroup and RVG name on both sites. If the volume names are the same on both zones, the Mount resources will mount the same block devices, and the same Oracle instance will start on the secondary in case of a failover.

Configuring the service groups

This topic describes how to configure service groups.

To configure the replication group

1. Create a hybrid service group (oragrp_rep) for replication. You can use the VvrRvgGroup template to create the service group.
2. Copy the DiskGroup resource from the application to the new group. Configure the resource to point to the disk group that contains the RVG.
3. Configure new resources of type IP and NIC.
4. Configure a new resource of type RVG in the service group.
5. Set resource dependencies as per the following information:
   - RVG resource depends on the IP resource
   - RVG resource depends on the DiskGroup resource
     IP resource depends on the NIC resource
6. Set the SystemZones attribute of the child group, oragrp_rep, such that all nodes in the primary RDC zone are in system zone 0 and all nodes in the secondary RDC zone are in system zone 1.

To configure the application service group

1. In the original Oracle service group (oragroup), delete the DiskGroup resource.
2. Add an RVGPrimary resource and configure its attributes.

   Set the value of the RvgResourceName attribute to the name of the RVG type resource that will be promoted and demoted by the RVGPrimary agent.

   Set the AutoTakeover and AutoResync attributes from their defaults as desired.
3. Set resource dependencies such that all Mount resources depend on the RVGPrimary resource. If there are a lot of Mount resources, you can set the TypeDependencies attribute for the group to denote that the Mount resource type depends on the RVGPRimary resource type.

4 Set the SystemZones attribute of the Oracle service group such that all nodes in the primary RDC zone are in system zone 0 and all nodes in the secondary RDC zone are in zone 1. The SystemZones attribute of both the parent and the child group must be identical.

5 If your setup uses BIND DNS, add a resource of type DNS to the oragroup service group. Set the Hostname attribute to the canonical name of the host or virtual IP address that the application uses on that cluster. This ensures DNS updates to the site when the group is brought online. A DNS resource would be necessary only if the nodes in the primary and the secondary RDC zones are in different IP subnets.

Configuring the service group dependencies

Set an online local hard group dependency from application service group to the replication service group to ensure that the service groups fail over and switch together.

1 In the Cluster Explorer configuration tree, select the cluster name.

2 In the view panel, click the **Service Groups** tab. This opens the service group dependency graph.

3 Click **Link**.

4 Click the parent group oragroup and move the mouse toward the child group, oragroup_rep.

5 Click the child group oragroup_rep.

6 On the Link Service Groups dialog box, click the online local relationship and the hard dependency type and click **OK**.

About migrating a service group

In the RDC set up for the Oracle database, consider a case where the primary RDC zone suffers a total failure of the shared storage. In this situation, none of the nodes in the primary zone see any device.

The Oracle service group cannot fail over locally within the primary RDC zone, because the shared volumes cannot be mounted on any node. So, the service group must fail over, to a node in the current secondary RDC zone.

The RVGPrimary agent ensures that VVR volumes are made writable and the DNS agent ensures that name services are resolved to the DR site. The application can be started at the DR site and run there until the problem with the local storage is corrected.

If the storage problem is corrected, you can switch the application to the primary site using VCS.

Fire drill in replicated data clusters

You can use fire drills to test the configuration's fault readiness by mimicking a failover without stopping the application in the primary data center. To set up a disaster recovery fire drill, you have the option to create and configure the fire drill service group manually, or using the Fire Drill Setup wizard.

See the *Veritas Cluster Server Administrator's Guide*.

Chapter 6

Configuring a replicated data cluster using third-party replication

This chapter includes the following topics:

- About setting up a replicated data cluster configuration using third-party replication
- About typical replicated data cluster configuration using third-party replication
- About setting up third-party replication
- Configuring the service groups for third-party replication
- Fire drill in replicated data clusters using third-party replication

About setting up a replicated data cluster configuration using third-party replication

The VCS replicated data cluster (RDC) configuration provides robust and easy-to manage disaster recovery protection for your applications. You can convert an application configured for local high availability in a VCS cluster to a disaster-protected RDC infrastructure. When configuring an RDC you can use a supported third-party replication technology to replicate application data.

Review the best practices and planning considerations prior to setting up an RDC using third-party replication.

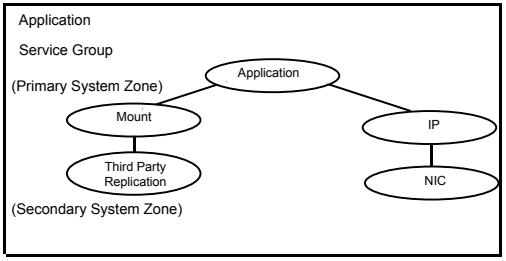
See “Planning for data replication” on page 38.

See “Planning a replicated data cluster setup” on page 37.

About typical replicated data cluster configuration using third-party replication

Figure 6-1 depicts a dependency chart of a typical RDC configuration using third-party replication.

Figure 6-1 Dependency chart of a typical RDC configuration using third-party replication



In this example, an application is configured as a VCS service group on a four-node cluster, with two nodes in the primary RDC system zone and two in the Secondary RDC system zone. In the event of a failure on the Primary node, VCS fails over the application to the second node in the Primary zone. When there is no system available for failover in the Primary system zone, VCS will failover the application to a system in the Secondary system zone.

Note: Some third-party replication software such as Oracle Dataguard require special configuration in RDC environment. Please refer to respective agent's Installation and Configuration Guide for more details.

The process involves the following steps:

- Setting up replication
- Configuring the service groups

About setting up third-party replication

A typical replicated data cluster configuration process involves setting up replication and configuring the service groups. You can set up replication for the RDC using various third-party replication technologies.

Review the best practices for setting up third-party replication.

See "Setting up third-party replication" on page 97.

Once the replication is configured, verify that the replicated storage is visible to all the nodes in the cluster according to the system zones.

Configuring the service groups for third-party replication

The original application service group will have the required application, storage, and network resources configured. You must set the SystemZones attribute of the application service group so that all the nodes in the Primary RDC zone are in system zone 0 and all nodes in the Secondary RDC zone are in zone 1.

To configure the service groups for third-party replication

1. Add the replication resource to the application service group
2. Configure the replication agent resources using guidelines from the replication agent's Installation and Configuration Guide
3. Localize the required resource attributes as per system-zone requirements.
4. Set the dependency between the storage (mount) resource and the replication resource.

Fire drill in replicated data clusters using third-party replication

A fire drill procedure verifies the fault-readiness of a disaster recovery configuration. This procedure is performed without stopping the application at the Primary site and disrupting user access.

A fire drill is performed at the Secondary site using a special service group for fire drills. The fire drill service group is identical to the application service group, but uses a fire drill resource in place of the replication agent resource. The fire drill service group uses a copy of the data that is used by the application service group.

Almost all third-party replication agents have a corresponding fire drill agent bundled with them. These fire drill agents are responsible for taking a snapshot inside the storage array and importing the snapshot disks on the hosts. The fire drill agents are named similar to the replication agents with the suffix “Snap” appended to the agent name. For example, the fire drill agent for EMC SRDF is SRDFSnap and the fire drill agent for Hitachi TrueCopy (HTC) agent is HTCSnap.

For more information about configuring fire drills, please refer the Installation and Configuration Guide of the appropriate third-party replication agent.

Section 4

Implementing global clusters

- Chapter 7. Configuring global clusters for VCS and SFHA
- Chapter 8. Configuring a global cluster with Storage Foundation Cluster File System or Storage Foundation for Oracle RAC
- Chapter 9. Configuring a global cluster with Veritas Volume Replicator and Storage Foundation Cluster File System or Storage Foundation for Oracle RAC

Chapter 7

Configuring global clusters for VCS and SFHA

This chapter includes the following topics:

- Installing and Configuring Veritas Cluster Server
- Setting up VVR replication
- Setting up third-party replication
- Fire drill in global clusters

Installing and Configuring Veritas Cluster Server

To create a global cluster environment, you must first install and configure Veritas Cluster Server (VCS) at the primary site, and then set up the application for high availability on both the primary and secondary sites. You can install and configure VCS using the script-based installer, the Web-based installer, response files, or manually.

For more information on installing and configuring VCS, see the *Veritas Cluster Server Installation Guide*.

Setting up VVR replication

After you have configured a VCS cluster at the primary site, and have set up the application for high availability on both the primary and secondary sites, you must set up data replication.

About configuring VVR replication

You can configure and administer Veritas Volume Replicator (VVR) using one of the following interfaces:

Command line interface (CLI)	You can use the command line interface of VVR to configure, administer, and monitor VVR in a distributed environment. The *Veritas Storage Foundation™ and High Availability Solutions Disaster Recovery Implementation Guide* (this guide) gives instructions on configuring, administering, and monitoring VVR using the Veritas product installer.

This topic explains how to set up a Replicated Data Set (RDS) using the command-line interface. VVR enables you to set up replication either when the data volumes are zero initialized or contain valid data. Make sure you follow the best practices or recommendations described to ensure successful configuration of VVR.

Detailed examples are also available on how to configure and set up a simple VVR configuration. Read this information before you start setting up replication.

Before setting up a Replicated Data Set, decide how you plan to lay out your VVR configuration.

To configure and set up replication, perform the following tasks in the order presented below.

See "Creating a Replicated Data Set" on page 74.

See "Synchronizing the Secondary and starting replication" on page 90.

Note: The procedure to set up replication is the same either when the application is running or stopped, unless noted otherwise.

Best practices for setting up replication

Set up replication according to the following best practices:

- Create one RVG for each application, rather than for each server. For example, if a server is running three separate databases that are being replicated, create three separate RVGs for each database. Creating three separate RVGs helps to avoid write-order dependency between the applications and provides three separate SRLs for maximum performance per application.
- Create one RVG per disk group. Creating one RVG per disk group enables you to efficiently implement application clustering for high availability, where only

one RVG needs to be failed over by the service group. If the disk group contains more than one RVG, the applications using the other RVGs would have to be stopped to facilitate the failover. You can use the Disk Group Split feature to migrate application volumes to their own disk groups before associating the volumes to the RVG.

- Plan the size and layout of the data volumes based on the requirement of your application.
- Plan the size of the network between the Primary and each Secondary host.
- Lay out the SRL appropriately to support the performance characteristics needed by the application. Because all writes to the data volumes in an RVG are first written to the SRL, the total write performance of an RVG is bound by the total write performance of the SRL. For example, dedicate separate disks to SRLs and if possible dedicate separate controllers to the SRL.
- Size the SRL appropriately to avoid overflow.
  The Veritas Volume Replicator Advisor (VRAdvisor), a tool to collect and analyze samples of data, can help you determine the optimal size of the SRL.
- Include all the data volumes used by the application in the same RVG. This is mandatory.
- Provide dedicated bandwidth for VVR over a separate network. The RLINK replicates data critical to the survival of the business. Compromising the RLINK compromises the business recovery plan.
- Use the same names for the data volumes on the Primary and Secondary nodes. If the data volumes on the Primary and Secondary have different names, you must map the name of the Secondary data volume to the appropriate Primary data volume.
- Use the same name and size for the SRLs on the Primary and Secondary nodes because the Secondary SRL becomes the Primary SRL when the Primary role is transferred.
- Mirror all data volumes and SRLs. This is optional if you use hardware-based mirroring.
- The `vradmin` utility creates corresponding RVGs on the Secondary of the same name as the Primary. If you choose to use the `vxmake` command to create RVGs, use the same names for corresponding RVGs on the Primary and Secondary nodes.
- Associate a DCM to each data volume on the Primary and the Secondary if the DCMs had been removed for some reason. By default, the `vradmin createpri` and `vradmin addsec` commands add DCMs if they do not exist.

- If you are setting up replication in a shared environment, before you do so, determine the node that is performing the most writes by running the `vxstat` command on each node for a suitable period of time, and then after you set up replication, specify that node as the logowner. Note that the logowner is not supported as Secondary.
- In a shared disk group environment, the cluster master server node will be selected as the logowner by default.
- The on-board write cache should not be used with VVR. The application must also store data to disk rather than maintaining it in memory. The takeover system, which could be a peer primary node in case of clustered configurations or the secondary site, must be capable of accessing all required information. This requirement precludes the use of anything inside a single system inaccessible by the peer. NVRAM accelerator boards and other disk caching mechanisms for performance are acceptable, but must be done on the external array and not on the local host.

Creating a Replicated Data Set

To create a Replicated Data Set (RDS), perform the following tasks in the order presented below:

- Create a Primary Replicated Volume Group (RVG) of the RDS
  You can also associate volume-set component volumes to an RDS.
- Add a Secondary to the RDS
- Change the Replication Settings for the Secondary

In a shared disk group environment, the `vradmin`, `vrstat`, and `vrnotify` commands can be issued on any node in the cluster. However, the `vradmin createpri`, `vxibc`, `vxrlink` (other than informational commands), and `vxrvg` commands must be issued from the CVM master node.

In a SAN disk group environment, the `vradmin` commands can be issued on the volume server or the volume client, provided the data volumes are attached to the host on which the commands are issued.

Creating a Primary RVG of an RDS

The first step in creating an RDS is creating its Primary RVG. VVR enables you to create a Primary RVG of an RDS using the `vradmin createpri` command.

The `vradmin createpri` command enables you to associate existing data volumes and the Storage Replicator Log (SRL) to the Primary RVG.

The `vradmin createpri` command performs the following operations:

- Creates the Primary RVG on the host on which the command is issued.
- Enables or starts the Primary RVG.
- Associates DCMs to the data volumes in the RVG.
- Associates the specified data volumes and SRL to the RVG.
- Associates the specified volume sets (if any) to the RVG.

Note: Specify the volume set name in the command, not the names of each component volume. Specifying the component volume name causes the command to fail.

VVR does not support RAID-5 volumes, that is, volumes with usage type `raid5` are not supported. Data volumes must be of usage type `gen` or `fsgen`. However, data volumes can be configured on hardware-based RAID-5 disks.

Dirty Region Logs (DRLs) are not needed with VVR because VVR uses the SRL to recover volumes, not the DRLs. If any of the data volumes or the SRL has a DRL, the `vradmin createpri` command removes the DRL before the data volume is associated to the RVG.

By default, the `vradmin createpri` command adds DCMs to the data volumes, if they have not already been added. The `vradmin createpri` command creates the DCM of an appropriate default size based on the size of the volume and mirrors the DCM by default. To create and add a DCM of a size that is different from the default, associate the DCM of the required size to the data volumes before running the `vradmin createpri` command.

Note: The `vradmin createpri` command will fail if there are not enough drives to mirror the DCM. You may need to associate the DCM to an appropriately sized data volume.

The `-nodcm` option when used with the `vradmin createpri` command associates data volumes to the RVG but does not add DCMs to the data volumes.

If you want to associate additional volumes to the RVG after creating the RVG, use the `vradmin addvol` command.

Prerequisites for creating a Primary RVG of an RDS

Before creating a Primary RVG of an RDS, the following prerequisites must be met:

- The data volumes and SRL must exist on the Primary. If the data volumes and SRL do not exist on the Primary, create them. To associate a volume set to the RVG, the volume set must exist on the Primary.
- The SRL cannot be a volume set or a component volume of a volume set.
- The data volumes and SRL must be started. If the data volumes and SRL are not started, start them. When a data volume is started, its state is active.
- The data volumes used by the application must exist in the same RVG. Include the data volumes used by the application in the same RVG.
- In a SAN disk group environment, if the application resides on the volume client, all the Primary data volumes must be attached to the volume client or unattached from the volume client.
- Make sure you include the appropriate loopback address(es) in the `/etc/hosts file`.
  - If your environment only uses IPv4, you must include an IPv4 loopback address in the `/etc/hosts` file. The following is a sample entry:

    ```
    127.0.0.1       localhost       loopback
    ```

  - If your environment only uses IPv6, you must include an IPv6 loopback address in the `/etc/hosts` file.

    ```
    ::1             localhost       loopback
    ```

  - If your environment uses both IPv4 and IPv6, the `/etc/hosts` file must include both loopback addresses.

    ```
    127.0.0.1       localhost       loopback
    ::1             localhost       loopback
    ```

To create a Primary RVG of an RDS

Issue the following command on the host on which you want to create the Primary RVG:

```
# vradmin -g diskgroup createpri rvgname \
      dv01_name,dv02_name... srl_name
```

The argument *rvgname* is the name of the RVG to be created.

The argument *dv01_name,dv02_name,...* is a comma-separated list of the names of the data volumes to be associated to the RVG. Each item can be an independent data volume name, or the name of a volume set. To associate a volume set to the

RVG, specify the name of the volume set, not the names of the individual component volumes.

Note: In previous releases, component volumes could be associated directly to an RVG. Beginning in Release 5.0, the volume set itself is associated to the RVG, enabling VVR to verify consistency between the volume sets on the Primary and the Secondary RVGs. The `vradmin createpri` command fails if a component volume of the volume set and the volume set itself are each specified for an RVG.

The argument *`srl_name`* is the name of the SRL to be associated to the RVG.

Use `-nodcm` option if you do not want DCMs to be added to the data volumes. By default, DCMs are added automatically.

Example - Creating a Primary RVG containing a data volume

This example shows how to create a Primary RVG `hr_rvg` in the disk group `hrdg`, which contains the data volumes `hr_dv01` and `hr_dv02`, and the volume `hr_srl` that is to be used as the SRL. This example automatically adds DCMs to the data volumes.

```
# vradmin -g hrdg createpri hr_rvg hr_dv01,hr_dv02 hr_srl
```

Example - Creating a Primary RVG containing a volume set

This example shows how to create a Primary RVG `hr_rvg` in the disk group `hrdg`, which contains the volume set `hr_vset`, the data volumes `hr_dv01` and `hr_dv02`, and the volume `hr_srl` that is to be used as the SRL.

```
# vradmin -g hrdg createpri hr_rvg hr_dv01,hr_dv02,hr_vset \
hr_srl
```

If the volume set includes the component volumes `hr_vsetdv01` and `hr_vsetdv02`, these volumes are associated to the RVG `hr_rvg`. This example automatically adds DCMs to the data volumes, including the component volumes `hr_vsetdv01` and `hr_vsetdv02`.

Adding a Secondary to an RDS

After creating the Primary RVG of the RDS, go on to adding a Secondary. Use the `vradmin addsec` command to add a Secondary RVG to an RDS. This command can also be used to add additional Secondary RVGs. The `vradmin addsec` command can be issued from any host that is already in the RDS.

Note: Run the `vradmin addsec` from the Primary node. If you run this command from the node being added as the Secondary, the command fails.

The `vradmin addsec` command performs the following operations by default:

- Creates and adds a Secondary RVG of the same name as the Primary RVG to the specified RDS on the Secondary host. By default, the Secondary RVG is added to the disk group with the same name as the Primary disk group. Use the option `-sdg` with the `vradmin addsec` command to specify a different disk group on the Secondary.
- If any of the data volumes or the SRL on the Secondary has a DRL, the DRL is removed before the data volume is associated to the RVG. DRLs are not needed with VVR because VVR uses the SRL to recover volumes, not the DRLs.
- Automatically adds DCMs to the Primary and Secondary data volumes if they do not have DCMs. Use the `-nodcm` option to specify that DCMs are not to be added to the data volumes.

  The `vradmin addsec` command creates the DCM of an appropriate default size based on the size of the volume and mirrors the DCM by default. To create and add a DCM of a size that is different from the default, associate the DCM of the required size to the data volumes before running the `vradmin addsec` command.
- Associates to the Secondary RVG, existing data volumes of the same names and sizes as the Primary data volumes; it also associates an existing volume with the same name as the Primary SRL, as the Secondary SRL.
- If the Primary RVG includes a volume set, the `vradmin addsec` command associates the corresponding volume set to the Secondary, if the volume set exists on the Secondary. The volume set on the Secondary must include volumes of the same name, lengths and indices as the component volumes on the Primary. If the volume set exists on the Secondary and the volume set configuration is correct except that it does not include all of the component volumes corresponding to those in the volume set on the Primary, the `vradmin addsec` command attempts to add the remaining component volumes to the volume set on the Secondary and then associate the volume set to the Secondary RVG. This command succeeds if all of the remaining component volumes exist on the Secondary with the same names, lengths, and indices as the component volumes on the Primary. However, if any of the component volumes do not exist on the Secondary or have a mismatched name, length, or index, the `vradmin addsec` command fails with the appropriate error message.

  If the volume set does not exist on the Secondary, but the component volumes exist with the same names, lengths, and indices, the `vradmin addsec` command creates the volume set on the Secondary and then associates it to the Secondary RVG.

- Creates and associates to the Primary and Secondary RVGs respectively, the Primary and Secondary RLINKs with default RLINK names *rlk_remotehost_rvgname*. If you choose to use names other than the default, use the `prlink` and `srlink` attributes of the `vradmin addsec` command to specify the Primary and Secondary RLINK names.
  See "Example - Creating a Primary RVG containing a volume set" on page 77.

Note: For replication in asynchronous mode with secondary logging, the SRL size on both the Primary and Secondary must be the same.

Best practices for adding a Secondary to an RDS

When you add a Secondary to an RDS, we recommend the following best practices:

- Determine the network and IP addresses to use. Add all participating system names and IP addresses to the `/etc/hosts` files on each system or to the name server database of your name service. Make sure the IP addresses are available (that is, plumbed and up) on the appropriate hosts for your configuration.
- Plan ahead for application clustering by configuring the IP addresses used for replication as virtual IP addresses. For each replicated data set, the Primary and the Secondary cluster should each have one unique virtual IP address to use as the address for the RLINK. If you do this, you can place VVR under cluster control without having to modify the IP address of the RLINK later. Changing the IP address of an RLINK requires pausing replication.
- Plan the bandwidth of the network based on your requirement. You can choose to use either the UDP protocol or TCP protocol for network communication between the Primary and Secondary. Also, plan to operate in a firewall environment.
- We recommend that you use the following naming conventions for RLINKs. By default, VVR follows the following naming conventions for RLINKs:
  Primary RLINK: *rlk_remotehost_rvgname*. For example:
  `rlk_london_hr_rvg`
  Secondary RLINK: *rlk_remotehost_rvgname*. For example:
  `rlk_seattle_hr_rvg`
- If you have multiple secondaries in your RDS setup, VVR automatically creates RLINKs between ever pair of secondaries. By doing this, the additional secondaries will be automatically added to the RDS after the migrate operation has completed successfully.
- Associate a DCM to each data volume on the Primary and the Secondary to use the SRL Protection and Failback Logging features.

Prerequisites for adding a Secondary to an RDS

On the Secondary to be added, do the following:

- Create a disk group with the same name as the Primary disk group.
- Create data volumes of the same names and lengths as the Primary data volumes.
- Create an SRL of the same name as the Primary SRL. Note that the SRL cannot be a volume set or a component volume of a volume set.
- If the Primary RVG includes a volume set, make sure that the component volumes on the Secondary to be added have identical names, lengths, and indices as the component volumes on the Primary.
- Make sure the `/etc/vx/vras/.rdg` file on the Secondary host to be added to the RDS contains the Primary disk group ID. Ensure that each disk group ID entry in the `.rdg` file is on a separate line.

  Refer to the `.rdg` file for the sample format for the disk group ID entry.

  The `vradmin addsec` command checks whether the Primary RVG is authorized to create a corresponding Secondary RVG on the specified Secondary host. A Primary is determined as authorized if the `/etc/vx/vras/.rdg` file on the specified Secondary host contains the Primary disk group ID. If the Primary contains multiple RVGs in the same disk group, only one entry is required. A plus (+) sign in the `/etc/vx/vras/.rdg` file on the Secondary host indicates that all Primary RVGs on all hosts are authorized to create a Secondary RVG on the specified Secondary host.

  The `/etc/vx/vras/.rdg` file on the Secondary host is only used for authorization checking when a Secondary is added, or when remote data volumes are synchronized or verified. To perform these operations after a Secondary takes over from the Primary, the original Primary host should also have an `/etc/vx/vras/.rdg` file containing the disk group ID for the new Primary host.

  To display the Primary disk group ID, issue the following command on the Primary host:

  ```
  # vxprint -l diskgroup
  ```

  For example, to enable host `seattle` to create an RVG on Secondary host `london` the `.rdg` file on the host `london` must have the following entries, each on a new line.

  ```
  1083007373.10.seattle
  ```

- In a SAN disk group environment, if the application resides on the volume client, the Secondary data volumes must be attached to the corresponding volume client on the Secondary.

- In a SAN disk group environment, if the application resides on the volume client, the Primary volume server must have network connection to the Secondary volume server and Secondary volume client.
- In a SAN disk group environment, if the application resides on the volume client, the hostname or IP address for the Secondary must be available on the volume client on the Secondary.
- In a SAN disk group environment, if the application resides on the volume server, the hostname or IP address for the Secondary must be available on the volume server on the Secondary.

To add a Secondary to an RDS

```
# vradmin -g local_diskgroup addsec local_rvgname pri_hostname \
      sec_hostname
```

The argument *local_diskgroup* is the name of the disk group on the local host.

The argument *local_rvgname* is the name of the RVG on the local host.

The arguments *pri_hostname* and *sec_hostname* are either resolvable hostnames or IP addresses for the Primary and the Secondary hosts. These names are used as `local_host` and `remote_host` attributes while creating RLINKs. The `local_host` and `remote_host` specify the network connection to use for the Primary and Secondary RLINKs.

Use the `-nodcm` option if you do not want to add DCMs to the data volumes. By default, DCMs are automatically added unless the `-nodcm` option is specified.

Note: By default, SRL protection on the new Primary and Secondary RLINKs is set to `autodcm`. If you specify the `-nodcm` option, the `vradmin addsec` command disables SRL protection.

Note that the Secondary RVG is added to the disk group with the same name as the Primary disk group, unless specified otherwise using the `-sdg` option.

Example 1:

This example shows how to add a Secondary host `london_priv` to the RDS, which contains the RVG `hr_rvg`. For replication, this example uses a private network with the Primary hostname `seattle_priv`, Secondary hostname `london_priv`. On the Secondary, the RVG is added to the same disk group as the Primary, that is, `hrdg`. This example automatically adds DCMs to the data volumes.

```
# vradmin -g hrdg addsec hr_rvg seattle_priv london_priv
```

Example 2:

This example shows how to add the Secondary host `london_priv` to the RDS, which contains the RVG `hr_rvg`. It creates the Secondary with the specific Primary and Secondary RLINK names `to_london` and `to_seattle`. The RLINK connects the Primary host `seattle_priv` and the Secondary host `london_priv`. On the Secondary, the RVG is added to the same disk group as the Primary, that is, `hrdg`.

```
# vradmin -g hrdg addsec hr_rvg seattle_priv london_priv \
prlink=to_london srlink=to_seattle
```

Example 3:

This example shows how to add a Secondary host `london-v6_priv` to the RDS, which contains the RVG `hr_rvg`. For replication, this example uses a private IPv6 network with the Primary hostname `seattle-v6_priv`, Secondary hostname `london-v6_priv`. Both hostnames `london-v6_priv` and `seattle-v6_priv` resolve to IPv6 addresses belonging to the private IPv6 network. On the Secondary, the RVG is added to the same disk group as the Primary, that is, `hrdg`. This example automatically adds DCMs to the data volumes.

```
# vradmin -g hrdg addsec hr_rvg seattle-v6_priv london-v6_priv
```

Example 4:

This example shows how to add a Secondary host `london-v6` to the RDS, which contains the RVG `hr_rvg`. It creates the Secondary with the specific Primary and Secondary RLINK names `to_london-v6` and `to_seattle-v6`. The RLINK connects the Primary host `seattle-v6` and the Secondary host `london-v6`, which resolve to IPv6 addresses `aaaa:bbbb:cccc:dddd:eeee:ffff:gggg:hhhh and pppp:qqqq:rrrr:ssss:wwww:xxxx:yyyy:zzzz` respectively. On the Secondary, the RVG is added to the same disk group as the Primary, that is, `hrdg`. This example also automatically adds DCMs to the data volumes.

```
# vradmin -g hrdg addsec hr_rvg aaaa:bbbb:cccc:dddd:eeee:ffff:gggg:hhhh \
pppp:qqqq:rrrr:ssss:wwww:xxxx:yyyy:zzzz prlink=to_london-v6 \
srlink=to_seattle-v6
```

Changing the replication settings for a Secondary

When you add a Secondary to an RDS, the default replication attributes of the Secondary are set to `synchronous=off`, `latencyprot=off`, `srlprot=autodcm`, `packet_size=8400` and `bandwidth_limit=none`.

To display the default replication settings for the Secondary, use the following form of the `vxprint` command:

```
vxprint -g diskgroup -Pl
```

If you are using the UDP protocol, this form of the `vxprint` command also shows the default packet size.

You can set up the replication mode, latency protection, SRL protection, transport protocol, packet size, and the bandwidth used by VVR using the replication attributes, such as `synchronous`, `latencyprot`, and `srlprot`. These attributes are of the form `attribute=value`. Each attribute setting could affect replication and must be set up with care.

The `vradmin set` command enables you to change the replication settings between the Primary and a Secondary. This command can be issued from any host in the RDS. It enables you to perform the following tasks:

- See "Setting the mode of replication for a Secondary" on page 83.
- See "Setting the latency protection for a Secondary" on page 84.
- See "Setting the SRL overflow protection for a Secondary" on page 86.
- See "Setting the network transport protocol for a Secondary" on page 87.
- See "Setting the packet size for a Secondary" on page 87.
- See "Setting the bandwidth limit for a Secondary" on page 88.

The `vradmin set` command changes the corresponding attributes on both the Primary and Secondary RLINK. The attributes `synchronous`, `latencyprot`, and `srlprot` are only active on the Primary RLINK; however, the Secondary attributes are already set up and ready for use if the Primary role is transferred to the Secondary.

Setting the mode of replication for a Secondary

You can set up VVR to replicate to a Secondary in synchronous or asynchronous mode by setting the `synchronous` attribute of the RLINK to `override`, or `off` respectively.

Setting the `synchronous` attribute to `override` puts the RLINK in synchronous mode. During normal operation, VVR replicates in synchronous mode, but if the RLINK becomes inactive due to a disconnection or administrative action, VVR switches temporarily to asynchronous mode and continues to receive updates from the application and store them in the SRL. After the connection is restored and the SRL is completely drained, the RLINK automatically switches back to synchronous mode. Most system administrators set the `synchronous` attribute to `override`.

The `vradmin` command does not allow you to set the `synchronous` attribute to `fail`. Use the `vxedit` command to set the attribute `synchronous=fail`. For more information on using the `vxedit` command, refer to the `vxedit` manual page.

Caution: if you use the `synchronous=fail` mode.

To enable asynchronous mode of replication

To set the replication to asynchronous mode, set the `synchronous` attribute to `off`.

```
# vradmin -g diskgroup set local_rvgname sec_hostname
synchronous=off
```

The argument *local_rvgname* is the name of the RVG on the local host and represents its RDS.

The argument *sec_hostname* is the name of the Secondary host displayed in the output of the *vradmin printrvg* command. If the RDS contains only one Secondary, the argument *sec_hostname* is optional.

Example - Setting the mode of replication to asynchronous for an RDS

To set the mode of replication to asynchronous for the RDS `hr_rvg` between the Primary `seattle` and the Secondary `london`, issue the following command on any host in the RDS:

```
# vradmin -g hrdg set hr_rvg london synchronous=off
```

To enable synchronous mode of replication

To set the synchronous attribute of the RLINK to `override`, use the following command:

```
# vradmin -g diskgroup set local_rvgname sec_hostname \
        synchronous=override
```

Example - Setting the mode of replication to synchronous for an RDS

To set the mode of replication to synchronous for the RDS `hr_rvg` between the Primary `seattle` and the Secondary `london`, issue the following command on any host in the RDS:

```
# vradmin -g hrdg set hr_rvg london synchronous=override
```

Setting the latency protection for a Secondary

The `vradmin set` command enables you to set the `latencyprot` attribute to `override`, `fail`, or `off`; it also enables you to specify a `latency_high_mark` and a `latency_low_mark`, which indicate when the protection becomes active or inactive.

Set the `latencyprot` attribute to enable latency protection between a Primary and a Secondary.

Note: Before enabling latency protection, be sure you understand how latency protection works when the Primary and Secondary are connected or disconnected.

To enable latency protection

1 Set the `latencyprot` attribute of the corresponding RLINKs on the Primary and Secondary.

To set the `latencyprot` attribute to `override`:

```
# vradmin -g diskgroup set local_rvgname sec_hostname \
        latencyprot=override
```

To set the `latencyprot` attribute to `fail`:

```
#    vradmin -g diskgroup set local_rvgname sec_hostname \
       latencyprot=fail
```

2 Set the `latency_high_mark` and the `latency_low_mark` attributes:

```
# vradmin -g diskgroup set local_rvgname sec_hostname \
        latency_high_mark=high_mark

# vradmin -g diskgroup set local_rvgname sec_hostname \
        latency_low_mark=low_mark
```

The argument *`local_rvgname`* is the name of the RVG on the local host and represents the RDS.

The argument *`sec_hostname`* is the name of the Secondary host as displayed in the output of the `vradmin printrvg` command.

Note that the value of `latency_high_mark` must be greater than the value of `latency_low_mark`. We recommend that the difference between the value of `latency_high_mark` and the value of `latency_low_mark` be a small number, for example, 50.

To disable latency protection

Setting the `latencyprot` attribute to `off` disables latency protection. This does not limit the number of waiting updates in the SRL.

To set the `latencyprot` attribute to `off`:

```
# vradmin -g diskgroup set local_rvgname sec_hostname
    latencyprot=off
```

The argument *local_rvgname* is the name of the RVG on the local host and represents the RDS.

The argument *sec_hostname* is the name of the Secondary host as displayed in the output of the `vradmin printrvg` command.

Setting the SRL overflow protection for a Secondary

VVR provides the following modes of SRL overflow protection: `autodcm`, `dcm`, `override`, `fail`, and `off`.

To enable SRL overflow protection

- Set the `srlprot` attribute of the corresponding RLINK to either `autodcm`, `dcm`, `override`, or `fail`.
  - To set the `srlprot` attribute to `autodcm`, use the following command:

    ```
    # vradmin -g diskgroup set local_rvgname sec_hostname \
            srlprot=autodcm
    ```

  - To set the `srlprot` attribute to `dcm`, use the following command:

    ```
    # vradmin -g diskgroup set local_rvgname sec_hostname \
            srlprot=dcm
    ```

  - To set the `srlprot` attribute to `override`, use the following command:

    ```
    # vradmin -g diskgroup set local_rvgname sec_hostname \
            srlprot=override
    ```

  - To set the `srlprot` attribute to `fail`, use the following command:

    ```
    # vradmin -g diskgroup set local_rvgname sec_hostname \
            srlprot=fail
    ```

  - To set the `srlprot` attribute to `off`, use the following command:

    ```
    # vradmin -g diskgroup set local_rvgname sec_hostname \
            srlprot=off
    ```

  The argument *local_rvgname* is the name of the RVG on the local host and represents the RDS.

  The argument *sec_hostname* is the name of the Secondary host as displayed in the output of the `vradmin printrvg` command.

Setting the network transport protocol for a Secondary

The value specified for the protocol attribute determines the protocol that will be used to communicate between the hosts. You can specify one of the following values for the protocol attribute.

- UDP—The hosts communicate using the UDP/IP protocol.
  VVR automatically calculates the checksum for each data packet it replicates.
- TCP—The hosts communicate using the TCP/IP protocol, which is the default. If a protocol is not specified, then TCP is used as the protocol of communication between hosts.
  If you specify TCP, the VVR checksum is automatically disabled. VVR relies on the TCP checksum mechanism instead. Also, if a node in a replicated data set is using a version of VVR earlier than 5.1 SP1, VVR calculates the checksum regardless of the network protocol.
- STORAGE—Used for bunker replication. The Primary host and the bunker SRL communicate using STORAGE protocol. If the storage is directly accessible by the Primary, for example, DAS or NAS, set the protocol to STORAGE. If the bunker is replicating over IP, the protocol can be set to UDP or TCP.

Note: UDP, TCP, and STORAGE are case sensitive.

To set the network protocol

- To set the protocol for RDSs in disk group of version 110 or above, the following `vradmin` command can be used:

  ```
  # vradmin -g diskgroup set local_rvgname sec_hostname \
          protocol=protocol_name
  ```

  The argument *`protocol_name`* is the name of the protocol that the Primary will use to replicate to the Secondary. The protocol can be set to either TCP or UDP.

Setting the packet size for a Secondary

The packet size determines the number of bytes in a packet that are sent to the Secondary host. The packet size can be changed using the `packet_size` attribute for UDP mode only. If the protocol is set to TCP, the data is sent using the TCP stream.

for more information on the `packet_size` attribute.

To set the packet_size

```
# vradmin -g diskgroup set local_rvgname sec_hostname \
        packet_size=n
```

The argument *local_rvgname* is the name of the RVG on the local host and represents the RDS.

The argument *sec_hostname* is the name of the Secondary host as displayed in the output of the `vradmin printrvg` command.

The argument *n* represents the packet size in bytes.

The minimum value for the `packet_size` is 1300 bytes.

The maximum value of the `packet_size` is 65464 bytes.

Example - Setting the packet size between the Primary and Secondary

To set the packet size between the Primary host `seattle` and the Secondary host `london` to 1400 bytes, issue the following command on any host in the RDS:

```
# vradmin -g hrdg set hr_rvg london packet_size=1400
```

Setting the bandwidth limit for a Secondary

Use the `bandwidth_limit` attribute of the `vradmin set` command to set the limit on the network bandwidth used to replicate from the Primary to the Secondary. If `bandwidth_limit` is set to `none`, then VVR uses the available network bandwidth. The default value is `none`. To limit the network bandwidth used by VVR when synchronizing volumes that are not part of an RDS, use the `bandwidth_limit` attribute of the `vradmin syncvol` command.

To control the network bandwidth used for replication

To limit the bandwidth used for replication between the Primary and a Secondary in an RDS, issue the following command on any host in the RDS. In the command, you can either use the units of bandwidth `kbps`, `mbps`, or `gbps`, or abbreviate the units of bandwidth to `k`, `m`, `g`, respectively. The default unit of bandwidth is bits per second (bps).

```
# vradmin -g diskgroup set local_rvgname sec_hostname \
        bandwidth_limit=value
```

The argument *local_rvgname* is the name of the RVG on the local host and represents the RDS.

The argument *sec_hostname* is the name of the Secondary host as displayed in the output of the `vradmin printrvg` command.

Example: Limiting network bandwidth and disabling Bandwidth Throttling for the Secondary

To limit the bandwidth to 30 mbps for the RDS `hr_rvg` between the Primary `seattle` and the Secondary `london`, issue the following command on any host in the RDS:

```
# vradmin -g hrdg set hr_rvg london bandwidth_limit=30mbps
```

To disable Bandwidth Throttling for a Secondary

To disable Bandwidth Throttling for a Secondary in an RDS, issue the following command on any host in the RDS:

```
# vradmin -g diskgroup set local_rvgname sec_hostname \
     bandwidth_limit=none
```

The argument *local_rvgname* is the name of the RVG on the local host and represents the RDS.

The argument *sec_hostname* is the name of the Secondary host as displayed in the output of the `vradmin printrvg` command.

Example: Disabling Bandwidth Throttling between the Primary the Secondary and controlling the network bandwidth

To disable Bandwidth Throttling for replication between the Primary `seattle` and the Secondary `london` of RDS `hr_rvg`, issue the following command on any host in the RDS:

```
# vradmin -g hrdg set hr_rvg london bandwidth_limit=none
```

To control the network bandwidth used to synchronize volumes

To limit the network bandwidth used by VVR when synchronizing volumes that are not part of an RDS, issue the following command:

```
# vradmin -g diskgroup syncvol local_vols_list \
        remote_hostname.... bandwidth_limit=value
```

The argument *local_vols_list* is a comma-separated list of volumes on the local host. The names of the volumes on the local and remote hosts are assumed to be the same.

The argument *remote_hostname* is a space-separated list of names of the remote hosts on which the volumes to be resynchronized reside. It must be possible for IP to resolve the remote host names.

Example: Limiting network bandwidth used by VVR when using full synchronization

This example shows how to limit the network bandwidth used by VVR when using full synchronization to synchronize the remote volumes on host `london` with the local volumes `hr_dv01`, `hr_dv02`, `hr_dv03` in the disk group `hrdg` on the local host `seattle`. The names of the disk group and the volumes on the remote host are the same as the names of the disk group and volumes on the local host.

```
# vradmin -g hrdg -full syncvol hr_dv01,hr_dv02,hr_dv03 london  \
        bandwidth_limit=10mbps
```

Synchronizing the Secondary and starting replication

This section explains how to synchronize the Secondary and start replication.

Methods to synchronize the Secondary

You can synchronize the Secondary using the network, using block-level tape backup or by physically moving disks to the Secondary. Automatic synchronization is the recommended method to synchronize the Secondary. Use one of the following methods to synchronize the Secondary depending on your environment:

- Using the network
  - Automatic synchronization
  - Full synchronization with Storage Checkpoint
  - Difference-based synchronization with Storage Checkpoint
- Using block-level tape backup
  - Block-level tape backup and checkpointing
- Moving disks physically
  - Disk Group Split and Join

The following tables explain when and how to use the different synchronization methods:

Using the network to synchronize the Secondary

You can synchronize the Secondary over the network either when the application is active or inactive. SmartMove is supported with each of these methods.

Table 7-1 Synchronizing the Secondary using the network

To Synchronize the Secondary:	Perform:	Using This Command:
completely	automatic synchronization and start replication See "Using the automatic synchronization feature" on page 92.	`vradmin -a startrep`
completely	full synchronization with Storage Checkpoint	`vradmin -full -c checkpoint syncrvg`
when there is little difference between the data on the Primary and Secondary data volumes of the RDS	difference-based synchronization with Storage Checkpoint	`vradmin -c checkpoint syncrvg`

Using block-level tape backup to synchronize the Secondary

Table 7-2 shows how to synchronize the Secondary using block-level tape backup.

Table 7-2 Synchronizing the Secondary using block-level tape backup

To Synchronize the Secondary:	Do the following:	Using This Command:
completely and when a large amount of data must be moved from the Primary to the Secondary	1. Start a Primary Storage Checkpoint.	`vxrvg -c checkpoint checkstart` *`rvg_name`*
	2. Perform a block-level backup of the Primary.	
	3. End the Primary Storage Checkpoint.	`vxrvg -c checkpoint checkstart` *`rvg_name`*
	4. Restore the tapes on the Secondary and start replication to the Secondary using the Storage Checkpoint.	`vradmin -c checkpoint startrep`

Moving disks physically to synchronize the Secondary

Table 7-3 shows how to synchronize the Secondary by moving disks physically.

Table 7-3 Synchronizing the Secondary by moving disks physically

To Synchronize the Secondary:	Use This Feature	Using This Command:
completely by physically moving disks from the location of the Primary host to the location of Secondary host	Disk Group Split and Join	

Using the automatic synchronization feature

The Automatic Synchronization feature enables you to transfer the data on the Primary to the Secondary over the network. You can synchronize the Secondary using automatic synchronization either when the application is active or inactive.

The Automatic Synchronization procedure transfers data in the Primary data volumes to the Secondary by reading the Primary data volumes from start to finish and sending the data to the Secondary.

Note: Automatic Synchronization does not maintain the order of writes; therefore, the Secondary is inconsistent until the process is complete.

The Secondary becomes consistent after the automatic synchronization completes. To use Automatic Synchronization successfully, the network must be sized appropriately. Note that the synchronization will complete only if the Primary receives writes at a lesser rate than they can be sent to the Secondary. If the Primary receives writes at a faster rate than they can be sent to the Secondary, the synchronization might never complete, especially if the writes are dispersed widely in the volume.

This feature enables you to synchronize multiple Secondary hosts at the same time. When performing automatic synchronization to multiple Secondary hosts, synchronization proceeds at the rate of the slowest network.

VVR pauses synchronization if the Secondary fails or the network disconnects. If the Primary fails while synchronization is in progress, the synchronization continues from the point at which it had stopped when the Primary recovers.

Prerequisite for using Automatic Synchronization

- Each data volume in the Primary RVG must have a DCM associated to it. If data volumes do not have DCMs, an attempt to automatically synchronize a Secondary fails.

The `vradmin startrep` command when used with the option `-a` enables you to start replication and automatically synchronize the Secondary data volumes with

the Primary data volumes in an RDS; it brings the Secondary data volumes up-to-date with the Primary data volumes. You can use this command to synchronize the Secondary when the data volumes contain data and when the application is active or inactive. Replication to another Secondary can be started only after this automatic synchronization completes.

The `vradmin startrep` command can be issued from any host in the RDS. To check the status and progress of the automatic synchronization, use the `vxrlink status` command on the Primary RLINK.

To synchronize the Secondary and start replication using automatic synchronization, issue the following command:

```
# vradmin -g diskgroup -a startrep local_rvgname sec_hostname
```

The argument *`local_rvgname`* is the name of the RVG on the local host and represents its RDS.

The argument *`sec_hostname`* is the name of the Secondary host displayed in the output of the `vradmin printrvg` command. If the RDS contains only one Secondary, the *`sec_hostname`* is optional.

Example—Using the Automatic Synchronization Feature

In this example, the data volumes in the Primary RVG `hr_rvg` on host `seattle` contain valid data and the application is active. To start replication and synchronize the Secondary RVG `hr_rvg` on host `london`, issue the following command:

```
# vradmin -g hrdg -a startrep hr_rvg london
```

Notes on using automatic synchronization

Observe the following notes about using automatic synchronization:

- If you associate a new volume to an RDS while automatic synchronization is in progress, VVR does not automatically synchronize the newly associated data volume.
- In an RDS containing multiple Secondaries that have SRL overflow protection set to dcm, more than one Secondary may require the use of the DCM. If one Secondary is undergoing automatic synchronization and the RLINK of another Secondary is about to overflow, the Automatic Synchronization is abandoned and the DCM becomes active for the overflowing RLINK.
- If you try to automatically synchronize a new RLINK while an existing RLINK is using the DCM mechanism, the automatic synchronization fails.
- To remove a Secondary from a DCM resynchronization process, detach the corresponding Primary RLINK.

- If you try to dissociate a DCM from a data volume while the DCM is in use, the operation fails.
- If the DCM is detached because of I/O errors while the DCM is in use, the resynchronization is abandoned and the RLINKs that are being synchronized are detached.
- You can use automatic synchronization with SmartMove. This combination lets deploy your disaster recovery site must faster because you are replicating far less data.
  See "About SmartMove for VVR" on page 94.

Example for setting up replication using automatic synchronization

This example assumes that the RDS has been created using the example procedure.

You can synchronize the Secondary using automatic synchronization when the application is active or inactive.

To setup replication using automatic synchronization

- Start Secondary synchronization and replication using automatic synchronization by issuing the following command from any host in the RDS:

```
# vradmin -g hrdg -a startrep hr_rvg london
```

About SmartMove for VVR

The SmartMove for VVR feature enables VVR to leverage information from VxFS knowledge of the file system blocks in use to optimize the time and network bandwidth required for initial synchronization of replicated volumes. This feature is available when the volume being synchronized uses either full synchronization or difference based synchronization and it requires a VxFS file system to be mounted on top of it.

If you use SmartMove with automatic synchronization, you can deploy the disaster recovery site faster because you are replicating far less data than the storage provisioned on the system. To use automatic synchronization with SmartMove in a cluster volume replication (CVR) environment, the file system must be mounted on the logowner.

The default behavior is to use the SmartMove for VVR feature for initial synchronization. The commands that use SmartMove during initial synchronization are `vradmin syncrvg/syncvol/startrep` and `vxrlink -a att`.

To turn off SmartMove

- Enter:

```
# vxtune usefssmartmove none
```

The `vradmin verifydata` command has also been enhanced to leverage VxFS knowledge of file system blocks in use for verification.

About thin storage reclamation and VVR

Thin storage helps you optimize your array capacity by allocating storage to applications only when it is needed. When files are created and written to in the file system, storage is allocated from a free storage pool on the array.

However, when you delete files on the host, the storage is not automatically returned to the pool. The result is large amounts of allocated storage on the array that is now unused. You must reclaim this storage manually.

See "Determining if a thin reclamation array needs reclamation" on page 96.

In a VVR environment, you can reclaim storage on volumes configured under a replication volume group (RVG). You can reclaim storage at the disk, disk group, or file system level.

Thin storage reclamation is only supported for LUNs that have the `thinrclm` attribute. VxVM automatically discovers LUNs that support thin reclamation from thin-capable storage arrays. On the host, you can list devices that have the `thinonly` or `thinrclm` attributes.

Thin storage reclamation is not supported on volumes in an RVG that has full instant or space-optimized snapshots that are associated to it. The reclaim command may complete without an error, but the storage space is not reclaimed. Thin storage reclamation is not supported as reclamation on the volume triggers a data transfer from the primary volume to the snapshot volume. Moving this data to the snapshot volume triggers storage allocation at the backend. If there are multiple snapshots, copies of the same data are maintained, which requires more storage than reclaimed. In the case of space-optimized snapshots, the cache object size increases as it copies all the reclaimable data to the space-optimized snapshot.

If you have used Storage Foundation thin storage reclamation in another context, the commands are identical when you use it in a VVR environment.

When you use thin reclamation with VVR, keep in mind the following:

- The VxFS file system must be mounted on the Primary site before you can perform thin reclamation on the volume.

- When you reclaim storage on the Primary site, it is automatically reclaimed on the Secondary site - unless the Primary site is in data change map (DCM) mode or when autosync is running. The Primary site goes into DCM mode when its Storage Replicator Log (SRL) overflows.
- You can reclaim storage on the Primary and Secondary sites even if the sites use different types of arrays. The arrays can have different fixed size physical storage allocation units.
- You can reclaim storage during a rolling upgrade with no impact on your systems.

For detailed information on thin storage, as well procedures for reclaiming thin storage, see *Veritas Storage Foundation and High Availability Solutions Solutions Guide*.

Determining if a thin reclamation array needs reclamation

You can only perform thin storage reclamation on LUNs that have the `thinrclm` attribute. An array may be a good candidate for reclamation if:

- Your array-specific commands indicate that storage is nearing the maximum; for example, 90% is used.
- You receive an alert from the array that a certain storage level has been reached.

Even if storage capacity is reaching the maximum, that does not necessarily mean there is any data to reclaim; the data could still be needed. You should investigate further to determine if there are large blocks of deleted or unused data. If there are, run thin reclamation.

For more information reclaiming storage, see *Veritas Storage Foundation and High Availability Solutions Solutions Guide*.

Starting replication when the data volumes are zero initialized

Use the option `-f` with the `vradmin startrep` command to start replication when the Primary and Secondary data volumes are zero initialized. The `vradmin startrep` command can be issued from any host in an RDS.

To start replication to a Secondary in an RDS when the data volumes are zero initialized:

```
# vradmin -g diskgroup -f startrep local_rvgname sec_hostname
```

The argument *`local_rvgname`* is the name of the RVG on the local host and represents its RDS.

The argument *sec_hostname* is the name of the Secondary host as displayed in the output of the `vradmin printrvg` command. If the RDS contains only one Secondary, the sec_hostname is optional.

Example: Starting replication when the data volumes are zero initialized

To start replication from the Primary RVG `hr_rvg` on `seattle` to the Secondary RVG on host `london` when the data volumes are zero initialized, issue the following command from any host in the RDS:

```
# vradmin -g hrdg -f startrep hr_rvg london
```

Note: The `-f` option can be used to stop the initial sync of the Primary and Secondary volumes if they already contain the same data.

Setting up third-party replication

Replication can be set up using VVR or various third-party replication technologies.

Set up replication according to the following best practices:

- Set up the replication. If you plan to use third-party software replication technology (Oracle Dataguard, IBM DB2HADR, etc.) or third-party hardware replication technology (EMC SRDF, Hitachi TrueCopy, IBM SVCCopyServices, etc.), review the vendor documentation for planning and configuring replication.
- Once the replication is configured, verify that the replicated storage is visible to all the nodes in the cluster(s).
- Verify that the storage vendor tools (CLIs or APIs) are installed on the VCS nodes as recommended in the respective replication agent Installation and Configuration Guide. For example, install RAID Manager CCI on each of the cluster nodes if you are using HITACHI TrueCopy replication.
- Download and install the required VCS replication agent from https://sort.symantec.com.
- Configure the replication agent resources using guidelines from the replication agent's Installation and Configuration Guide.
- Using the scenarios provided in the Install and Configuration Guide, test failover for the configuration.

Fire drill in global clusters

A fire drill procedure verifies the fault-readiness of a disaster recovery configuration. This procedure is performed without stopping the application at the Primary site and disrupting user access.

A fire drill is performed at the Secondary site using a special service group for fire drills. The fire drill service group is identical to the application service group, but uses a fire drill resource in place of the replication agent resource. The fire drill service group uses a copy of the data that is used by the application service group.

Almost all of the replication agents have a corresponding fire drill agent bundled with them. These fire drill agents are responsible for taking a snapshot inside the storage array and importing the snapshot disks on the hosts. The fire drill agents are named similar to the replication agents with the suffix “Snap” appended to the agent name. For example, the fire drill agent for EMC SRDF is SRDFSnap and the fire drill agent for Hitachi TrueCopy (HTC) agent is HTCSnap.

For more information about configuring fire drills, please refer the Installation and Configuration Guide of the appropriate replication agent.

Chapter 8

Configuring a global cluster with Storage Foundation Cluster File System or Storage Foundation for Oracle RAC

This chapter includes the following topics:

- About global clusters
- About replication for parallel global clusters using Storage Foundation and High Availability (SFHA) Solutions
- About setting up a global cluster environment for parallel clusters
- Configuring the primary site
- Configuring the secondary site
- Setting up replication between parallel global cluster sites
- Testing a parallel global cluster configuration

About global clusters

Global clusters provide the ability to fail over applications between geographically distributed clusters when disaster occurs. This type of clustering involves migrating

applications between clusters over a considerable distance. You can set up HA/DR using hardware-based or software-based replication technologies.

About replication for parallel global clusters using Storage Foundation and High Availability (SFHA) Solutions

Global clusters provide the ability to fail over applications between geographically distributed clusters when disaster occurs. This type of clustering involves migrating applications between clusters over a considerable distance. You can set up HA/DR using hardware-based or software-based replication technologies.

You can set up a primary cluster for replication to a secondary cluster by configuring global VCS service groups and using a replication technology. The database cluster at the secondary site can be a single node cluster. For example, you can have a two-node cluster on the primary site and a two-node or single-node cluster on the secondary site.

You can use one of the following replication technologies:

- Veritas Volume Replicator (VVR), which provides host-based volume replication. Using VVR you can replicate data volumes on a shared disk group in Veritas Storage Foundation Cluster File System High Availablity (SFCFS HA) or Veritas Storage Foundation for Oracle RAC (SF Oracle RAC).
- Supported hardware-based replication technologies. Using hardware-based replication you can replicate data from a primary array to a secondary array.

To verify your configuration is supported, review the product requirements and licensing information:

See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*.

See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

To confirm the compatibility of your hardware, see the current compatibility list in the Symantec Technical Support website:

http://www.symantec.com/docs/TECH170013

SFCFSHA and SF Oacle RAC support multiple third-party replication options.

For current information on third-party replication support:

See: https://sort.symantec.com/agents and select **Replication Agents** under **Agent type**.

Notes:

- Check your vendor's compatibility list for the supported software versions. The support listed above only exists if the host, HBA, and array combination is in your vendor's hardware compatibility list. Check your array documentation.
- All arrays must support SCSI-3 persistent reservations.

The Veritas replication agents provide application failover and recovery support to your replication configuration in environments where data is replicated between clusters.

VCS replication agents control the direction of replication. They do not monitor the progress or status of replication. The replication agents manage the state of replicated devices that are attached to global cluster nodes. The agents make sure that the system which has the resource online also has safe and exclusive access to the configured devices.

For more current information on the replicated agents:

See the *Veritas Cluster Server Bundled Agents Guide*

Technical Support TechNote for the latest updates or software issues for replication agents:

http://www.symantec.com/docs/TECH46455

About setting up a global cluster environment for parallel clusters

Configuring a global cluster for environment with parallel clusters requires the coordination of many component setup tasks. The procedures provided here are guidelines. Refer to your product installation guide for your product installation and basic configuration for each cluster. You will need this guide to configure a global cluster environment and replication between the two configured clusters.

Table 8-1 Tasks for setting up a global cluster environment with parallel clusters

Task	Description
Configure a parallel cluster at the primary site	See "Configuring the primary site" on page 102.
Configure an parallel cluster at the secondary site	See "Configuring the secondary site" on page 105.

Table 8-1 Tasks for setting up a global cluster environment with parallel clusters *(continued)*

Task	Description
Configure a global cluster environment	See "Setting up replication between parallel global cluster sites" on page 110.
Test the HA/DR configuration	See "Testing a parallel global cluster configuration" on page 117.

Upon successful testing, you can bring the environment into production

Some configuration tasks may require adjustments depending upon your particular starting point, environment, and configuration: details of your configuration may differ from the examples given in the procedures. Review the installation requirements and sample cluster configuration files for primary and secondary clusters.

For requirements, installation instructions, and sample configuration files:

See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*.

See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

Configuring the primary site

Table 8-2 Tasks for setting up a parallel global cluster at the primary site

Task	Description
Set up the cluster	See "To set up the cluster at the primary site" on page 103.
Set up the database	See "To set up the database at the primary site" on page 105.

Consult your product installation guide for planning information as well as specific configuration guidance for the steps below.

See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*.

See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

You can use an existing parallel cluster or you can install a new cluster for your primary site. If you are using an existing cluster as the primary and you want to set up a global cluster, skip the steps below and proceed to configure your secondary cluster.

See "Configuring the secondary site" on page 105.

Note: You must have a Global Cluster Option (GCO) license enabled for a global cluster. If you are using Veritas Volume Replicator (VVR) for replication, you must have a VVR license enabled.

If you do not have an existing cluster and you are setting up two new sites for a global cluster, follow the steps below.

To set up the cluster at the primary site

1. Install and configure servers and storage.
2. If you are using hardware-based replication, install the sofware for managing your array.
3. Verify that you have the correct installation options enabled, whether you are using keyless licensing or installing keys manually. You must have the GCO option enabled for a global cluster. If you are using VVR for replication, you must have it enabled.
4. Prepare, install, and configure your Storage Foundation and High Availability (SFHA) Solutions product according to the directions in your product's installation guide.

   For a multi-node cluster, configure I/O fencing.
5. Verify the CVM group is online on all nodes in the primary cluster.

   ```
   # hagrp -state cvm
   ```
6. Set storage connectivity for volume asymetry and I/O shiping for DCO logs policies for the disk group.
7. Prepare systems and storage for a global cluster. Identify the hardware and storage requirements before installing your database software.

   For SFCFS HA, you will need to set up:

   - Local storage for database software
   - Shared storage for resources which are not replicated as part of the hardware-based or host-based replication

- Replicated storage for database files

For SF Oracle RAC, you will need to set up:

- Local storage for Oracle RAC and CRS/GRID binaries
- Shared storage for OCR and Vote disk which is not replicated as part of the hardware-based or host-based replication
- Replicated shared storage for database files

8 For SFCFS HA, install and configure your database binaries. Consult your database documentation.

Note: Resources which will not be replicated must be on non-replicated shared storage.

After successful database installation and configuration, verify that database resources are up on all nodes.

9 For Oracle RAC, see the instructions in the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide* for installing, configuring, and linking Oracle database binaries:

- Oracle Clusterware/Grid Infrastructure software
- Oracle RAC database software
- The Oracle RAC binary versions must be exactly same on both sites.

Note: OCR and Vote disk must be on non-replicated shared storage.

After successful Oracle RAC installation and configuration, verify that CRS daemons and resources are up on all nodes.

For Oracle RAC 10gR2, use:

```
$ CRS_HOME/bin/crs_stat -t
```

For Oracle RAC 11gR1, use:

```
$ CRS_HOME/bin/crs_stat -t
```

For Oracle RAC 11gR2, use:

```
$ GRID_HOME/bin/crsctl stat res -t
```

To set up the database at the primary site

1. Identify the disks that will be replicated, create the required CVM disk group, volume, and file system.
2. Create the database on the file system you created in the previous step.
3. Configure the VCS service groups for the database.
4. Verify that all VCS service groups are online.

Configuring the secondary site

The setup requirements for the secondary site parallel the requirements for the primary site with a few additions or exceptions as noted below.

Table 8-3 Tasks for setting up a parallel global cluster at the secondary site

Task	Description
Set up the cluster	See "To set up the cluster on secondary site" on page 106.
Set up the database	See "To set up the SFCFS HA database for the secondary site" on page 108. See "To set up the Oracle RAC database for the secondary site" on page 108.

Important requirements for parallel global clustering:

- Cluster names on the primary and secondary sites must be unique.
- You must use the same OS user and group IDs for your database for installation and configuration on both the primary and secondary clusters.
- For Oracle RAC, you must use the same directory structure, name, permissions for the CRS/GRID and database binaries.
- For Sybase ASE CE, the binary versions must be exactly same on both sites, including the ESD versions.

You can use an existing parallel cluster or you can install a new cluster for your secondary site.

Consult your product installation guide for planning information as well as specific configuration guidance for the steps below.

See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*.

See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

To set up the cluster on secondary site

1. Install and configure servers and storage.
2. If you are using hardware-based replication, install the sofware for managing your array.
3. Verify that you have the correct installation options enabled, whether you are using keyless licensing or installing keys manually. You must have the GCO option for a global cluster. If you are using VVR for replication, you must have it enabled.
4. Prepare, install, and configure your Storage Foundation and High Availability (SFHA) Solutions product according to the directions in your product's installation guide.

   For a multi-node cluster, configure I/O fencing.
5. For a single-node cluster, do not enable I/O fencing. Fencing will run in disabled mode.
6. Prepare systems and storage for a global cluster. Identify the hardware and storage requirements before installing your database software.

   For SFCFS HA, you will need to set up:

   - Local storage for database software
   - Shared storage for resources which are not replicated as part of the hardware-based or host-based replication
   - Replicated storage for database files
   - You must use the same directory structure, name, permissions for the quorum and database binaries as on the primary.

   For SF Oracle RAC, you will need to set up:

   - Local storage for Oracle RAC and CRS binaries
   - Shared storage for OCR and Vote disk which is not replicated as part of the hardware-based or host-based replication
   - Replicated shared storage for database files
   - You must use the same directory structure, name, permissions for the CRS/GRID and database binaries as on the primary.

   Note: You must use the same directory structure, name, permissions for the CRS/GRID and database binaries.

7 For SFCFS HA, install and configure your database binaries. Consult your database documentation.

Note: Resources which will not be replicated must be on non-replicated shared storage.

After successful database installation and configuration, verify that database resources are up on all nodes.

8 For Oracle RAC, see the instructions in the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide* for installing and configuring:

- Oracle Clusterware/Grid Infrastructure software
- Oracle RAC database software
- The Oracle RAC binary versions must be exactly same on both sites.

Note: OCR and Vote disk must be on non-replicated shared storage.

After successful Oracle RAC installation and configuration, verify that CRS daemons and resources are up on all nodes.

For Oracle RAC 10g R2, use:

```
$ CRS_HOME/bin/crs_stat -t
```

For Oracle RAC 11gR1, use:

```
$ CRS_HOME/bin/crs_stat -t
```

For Oracle RAC 11gR2, use:

```
$ GRID_HOME/bin/crsctl stat res -t
```

Do not create the database. The database will be replicated from the primary site.

To set up the SFCFS HA database for the secondary site

1. If you are using hardware-based replication, the database, disk group, and volumes will be replicated from the primary site.

   Create the directory for the CFS mount point which will host the database data and control files.

2. If you are using VVR for replication, create an identical disk group and volumes for the replicated content with the same names and size as listed on the primary site.

   Create the directories for the CFS mount points as they are on the primary site. These will be used to host the database and control files when the failover occurs and the secondary is promoted to become the primary site.

3. Create subdirectories for the database as you did on the primary site.

To set up the Oracle RAC database for the secondary site

1. If you are using hardware-based replication, the database, disk group, and volumes will be replicated from the primary site.

   Create the directory for the CFS mount point which will host the database data and control files.

2. If you are using VVR for replication, create an identical disk group and volumes for the replicated content with the same names and size as listed on the primary site.

   Create the directories for the CFS mount points as they are on the primary site. These will be used to host the database and control files when the failover occurs and the secondary is promoted to become the primary site.

3. On each node in the cluster, copy the initialization files (pfiles,spfiles) from the primary cluster to the secondary cluster maintaining the same directory path.

   For example, copy init$ORACLE_SID.ora and orapw$ORACLE_SID.ora from $ORACLE_HOME/dbs at the primary to $ORACLE_HOME/dbs at the secondary.

4. For Oracle 11g R2 databases, modify the init$ORACLE_SID.ora file to add the following entries:

   ```
   SPFILE= <spfile location>
   remote_listener='<SCAN_LISTENER>:1521'
   ```

5 As Oracle user, ceate the following subdirectories on the secondary site to parallel the directories on the primary site:

For Oracle RAC 10g:

```
$ mkdir -p /$ORACLE_BASE/admin/database_name/adump
$ mkdir -p /$ORACLE_BASE/admin/database_name/bdump
$ mkdir -p /$ORACLE_BASE/admin/database_name/cdump
$ mkdir -p /$ORACLE_BASE/admin/database_name/dpdump
$ mkdir -p /$ORACLE_BASE/admin/database_name/hdump
$ mkdir -p /$ORACLE_BASE/admin/database_name/udump
$ mkdir -p /$ORACLE_BASE/admin/database_name/pfile
```

For Oracle 11gR1:

```
$ mkdir -p $ORACLE_BASE/admin/$database_name
$ cd $ORACLE_BASE/admin/$database_name
$ mkdir adump dpdump hdump pfile
```

For 11gR2 create following directory structure:

```
$ mkdir -p $ORACLE_BASE/diag
$ mkdir -p $ORACLE_BASE/admin/database_name/adump
```

6 Configure listeners on the secondary site with same name as on primary. You can do this by one of the following methods:

- Copy the listener.ora and tnsnames.ora files from the primary site and update the names as appropriate for the secondary site.
- Use Oracle's netca utility to to configure the listener.ora and tnsnames.ora files on the secondary site.

7 On the secondary site, register the database using the `srvctl` command as the database software owner.

Registering the database only has to be done once from any node in the secondary cluster.Use the following command as the Oracle database software owner

```
$ $ORACLE_HOME/bin/srvctl add database -d database_name -o oracle_home
```

8 To prevent automatic database instance restart, change the Management policy for the database (automatic, manual) to MANUAL using the `srvctl` command:

```
$ $ORACLE_HOME/bin/srvctl modify database -d database_name -y manual
```

You need only perform this change once from any node in the cluster.

9 Register the instances using srvctl command. Execute the following command on each node:

```
$ $ORACLE_HOME/bin/srvctl add instance -d database_name \
-i instance_name -n node-name
```

If the secondary cluster has more than one node, you must add instances using the `srvctl` command.

For example, if the database instance name is racdb, the instance name on sys3 is racdb1 and on sys4 is racdb2.

```
$ $ORACLE_HOME/bin/srvctl add instance -d racdb -i racdb1 -n sys3
```

```
$ $ORACLE_HOME/bin/srvctl add instance -d racdb -i racdb2 -n sys4
```

10 Register all other resources (for example listener, ASM, service) present in cluster/GRID at the primary site to the secondary site using the `srvctl` command or `crs_register`. For command details, see Oracle documentation at Metalink.

Setting up replication between parallel global cluster sites

You have configured Veritas Cluster Server (VCS) service groups for the database on each cluster. Each cluster requires an additional virtual IP address associated with the cluster for cross-cluster communication. The VCS installation and creation of the ClusterService group typically involves defining this IP address.

Configure a global cluster by setting:

- Heartbeat
- Wide area cluster (wac)

- GCO IP (gcoip)
- remote cluster resources

Table 8-4 Tasks for configuring a parallel global cluster

Task	Description
Prepare to configure global parallel clusters	Before you configure a global cluster, 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 your SFHA Solutions product, 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 database user and group IDs must be the same on all nodes. ■ The database, which is replicated 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.
Configure a global cluster using the global clustering wizard.	See "To modify the ClusterService group for global clusters using the global clustering wizard" on page 112.
Define the remote global cluster and heartbeat objects	See "To define the remote cluster and heartbeat" on page 113.
Configure global service groups for database resources	See "To configure global service groups for database resources" on page 117.

Table 8-4 Tasks for configuring a parallel global cluster *(continued)*

Task	Description
Start replication between the sites.	For software-based replication using Veritas Volume Replicator (VVR): See "About configuring a parallel global cluster using Veritas Volume Replicator (VVR) for replication" on page 120. For replication using Oracle Data Guard see the Data Guard documentation by Oracle. For replication using hardware-based replication see the replicated agent guide for your hardware. See the *Veritas Cluster Server Bundled Agents Guide*
Test the HA/DR configuration before putting it into production	See "Testing a parallel global cluster configuration" on page 117.

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 modify the ClusterService group for global clusters using the global clustering wizard

1 On the primary cluster, start the GCO Configuration wizard:

```
# /opt/VRTSvcs/bin/gcoconfig
```

2 The wizard discovers the NIC devices on the local system and prompts you to enter the device to be used for the global cluster. Specify the name of the device and press Enter.

3 If you do not have NIC resources in your configuration, the wizard asks you whether the specified NIC will be the public NIC used by all the systems. Enter **y** if it is the public NIC; otherwise enter **n**. If you entered **n**, the wizard prompts you to enter the names of NICs on all systems.

4 Enter the virtual IP address for the local cluster.

5 If you do not have IP resources in your configuration, the wizard prompts you for the netmask associated with the virtual IP. The wizard detects the netmask; you can accept the suggested value or enter another one.

The wizard starts running commands to create or update the ClusterService group. Various messages indicate the status of these commands. After running these commands, the wizard brings the ClusterService failover group online on any one of the nodes in the cluster.

After configuring global clustering, add the remote 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 On the primary site, define the remote cluster and its virtual IP address.

In this example, the remote cluster is clus2 and its IP address is 10.11.10.102:

```
# haclus -add clus2 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 clus1 and its IP address is 10.10.10.101:

```
# haclus -add clus1 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 lists the remote cluster.
- Arguments enable you to define the virtual IP address for the remote cluster.

For example:

```
# hahb -modify Icmp ClusterList clus2
# hahb -modify Icmp Arguments 10.11.10.102 -clus clus2
```

6 Save the configuration and change the access to read-only on the local cluster:

```
# haconf -dump -makero
```

7 Complete step 4-6 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 It is advisible to modify "OnlineRetryLimit" & "OfflineWaitLimit" attribute of IP resource type to 1 on both the clusters:

```
# hatype -modify IP  OnlineRetryLimit  1

# hatype -modify IP  OfflineWaitLimit  1
```

9 Verify cluster status with the `hastatus -sum` command on both clusters.

```
# hastatus -sum
```

For example, for SF Oracle RAC, the final output should resemble the output displayed below, from rac_clus101 (primary):

For example, the final output should resemble the output displayed below, from rac_clus101 (primary):

```
# hastatus -sum
.........
-- WAN HEARTBEAT STATE
-- Heartbeat       To                   State

L  Icmp            clus2          ALIVE

-- REMOTE CLUSTER STATE
-- Cluster         State

M  clus2     RUNNING

-- REMOTE SYSTEM STATE
-- cluster:system       State               Frozen

N  clus2:sys3  RUNNING              0
N  clus2:sys4  RUNNING              0
```

10 Display the global setup by executing haclus -list command.

```
# haclus -list
    clus1
    clus2
```

Example of heartbeat additions to the main.cf file on the primary site:

```
.
.
remotecluster clus2 (
Cluster Address = "10.11.10.102"
)
heartbeat Icmp (
    ClusterList = { clus2 }
    Arguments @clus2 = { "10.11.10.102" }
    )

system sys1 (
    )
.
.
```

Example heartbeat additions to the main.cf file on the secondary site:

```
.
.
remotecluster clus1 (
    Cluster Address = "10.10.10.101"
    )

heartbeat Icmp (
    ClusterList = { clus1 }
    Arguments @clus1 = { "10.10.10.101" }
    )

system sys3 (
    )
.
.
```

See the *Veritas Cluster Server Administrator's Guide* for more details for configuring the required and optional attributes of the heartbeat object.

To configure global service groups for database resources

1. Configure and enable global groups for databases and resources.
   - Configure VCS service groups at both sites.
   - Configure the replication agent at both sites.
   - For SF Oracle RAC, make the Oracle RAC service group a global service group, enabling failover across clusters.
   - For example:
     See "Modifying the Veritas Cluster Server (VCS) configuration on the primary site" on page 137.
2. To test real data in an environment where HA/DR has been configured, schedule a planned migration to the secondary site for testing purposes.

   For example:

   See "To migrate the role of primary site to the remote site" on page 149.

   See "To migrate the role of new primary site back to the original primary site" on page 150.
3. Upon successful testing, bring the environment into production.

For more information about VCS replication agents:

See the *Veritas Cluster Server Bundled Agents Guide*

For complete details on using VVR in a shared disk environment:

See the *Veritas Storage Foundation and High Availability Solutions Replication Administrator's Guide*.

Testing a parallel global cluster configuration

Symantec recommends testing before putting a global cluster configuration into production.

To test a planned failover

1. Offline the VCS service group for the database on the cluster at the primary site.
2. Online the VCS service group for the database on the cluster at the secondary site.

To test distaster recovery at the recovery site

1. Plan downtime to test the disaster recovery configuration.
2. Simulate a disaster at the primary site.

   For example:

   Shut down the hosts and storage arrays at the primary. If you can not shut down the storage arrays, disconnect the replication link between the sites.
3. Use VCS to fail over the database to the cluster at the secondary site.

To test failback on the primary site

1. Offline the VCS service group for the database on the cluster at the secondary site.
2. If the nodes and storage are down, restart the nodes and the storage array at the primary site.
3. Reconnect the replication link if it was broken.
4. Resynchronize the data from the secondary to make sure the data at the primary site is current.
5. Bring the VCS service group online at the primary site.

Chapter 9

Configuring a global cluster with Veritas Volume Replicator and Storage Foundation Cluster File System or Storage Foundation for Oracle RAC

This chapter includes the following topics:

- About configuring a parallel global cluster using Veritas Volume Replicator (VVR) for replication
- Setting up replication on the primary site using VVR
- Setting up replication on the secondary site using VVR
- Starting replication of the primary site database volume to the secondary site using VVR
- Configuring Veritas Cluster Server to replicate the database volume using VVR
- Replication use cases for global parallel clusters

About configuring a parallel global cluster using Veritas Volume Replicator (VVR) for replication

Configuring a global cluster for environment with parallel clusters using Veritas Volume Replicator requires the coordination of many component setup tasks. The tasks listed below are guidelines.

Before configuring two clusters for global clustering, you must verify that:

- You have the correct installation options enabled for your Storage Foundations and High Availability (SFHA) Solutions product, whether you are using keyless licensing or installing keys manually. You must have the GCO option for a global cluster and VVR enabled.
  Review your product requirements and licensing information.
- Both clusters have your SFHA Solutions product software installed and configured.
  See "Configuring the primary site" on page 102.
  See "Configuring the secondary site" on page 105.

Note: You can install and configure both clusters at the same time, or you can configure the second cluster at a later time than the first.

You you can use this guide to configure a global cluster environment and replication between the two clusters. For installation and configuration information:

See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*.

See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

With two clusters installed and configured , you are ready to configure a global cluster environment using VVR. You must perform the following tasks to modify both cluster configurations to support replication in the global cluster environment.

Table 9-1 Tasks for configuring a parallel global cluster with VVR

Task	Description
Setting up replication on the primary site	■ Create the Storage Replicator Log (SRL) in the disk group for the database. ■ Create the Replicated Volume Group (RVG) on the primary site. See " Setting up replication on the primary site using VVR" on page 122.
Setting up replication on the secondary site	■ Create a disk group to hold the data volume, SRL, and RVG on the storage on the secondary site. You must match the names and sizes of these volumes with the names and sizes of the volumes on the primary site. ■ Edit the /etc/vx/vras/.rdg file on the secondary site. ■ Use resolvable virtual IP addresses that set network RLINK connections as host names of the primary and secondary sites. ■ Create the replication objects on the secondary site. See "Setting up replication on the secondary site using VVR" on page 125.
Starting replication of the database.	You can use either of the following methods to start replication: ■ Automatic synchronization ■ Full synchronization with Storage Checkpoint See "Starting replication of the primary site database volume to the secondary site using VVR" on page 130.
Configuring VCS for replication on clusters at both sites.	Configure Veritas Cluster Server (VCS) to provide high availability for the database: ■ Modify the VCS configuration on the primary site ■ Modify the VCS configuration on the secondary site See "Configuring Veritas Cluster Server to replicate the database volume using VVR" on page 132.

Once the global clusters and replication with VVR are configured, the following replication use cases are supported for it:

- Migration of the role of the primary site to the remote site
- Takeover of the primary site role by the secondary site

- Migrate the role of primary site to the secondary site
- Migrate the role of new primary site back to the original primary site
- Take over after an outage
- Resynchronize after an outage
- Update the rlink to reflect changes

See "Replication use cases for global parallel clusters" on page 147.

Setting up replication on the primary site using VVR

If you have not already done so, create a disk group to hold data volume, Storage Replicator Log (SRL), and Replicated Volume Group (RVG) on the storage on the primary site. For example, create the storage for your database.

See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*.

See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*.

Table 9-2 Tasks for setting up replication on the primary site using VVR

Task	Description
Create the Storage Replicator Log (SRL) in the disk group for the database.	See "Creating the data and SRL volumes on the primary site" on page 122.
Create the Replicated Volume Group (RVG) on the primary site.	See "Setting up the Replicated Volume Group on the primary site" on page 124.

Creating the data and SRL volumes on the primary site

Create the data volume if you do not have one already.

- The data volume on the secondary site has the same name and the same size as the data volume on the primary site.
  See "Configuring the secondary site" on page 105.
- The data volume and Storage Replicator Log (SRL) volume should exist in the same disk group.
- Mirror the data volume in the absence of hardware-based mirroring.

To create the data volume on the primary site

- In the disk group created for the database, create a data volume for the data on primary site. In the examples below, the dbdata_vol volume on the primary site is 12 GB:

```
# vxassist -g dbdatadg make dbdata_vol 12000M disk1 disk2
```

Create the SRL. The SRL is a volume in the Replicated Volume Group (RVG). The RVG also holds the data volumes for replication.

- The SRL on the secondary site has the same name and the same size as the SRL on the primary site.
- You must create SRLs on disks without other volumes.
- Mirror SRLs and in the absence of hardware-based mirroring.

In the example procedure below, dbdatadg is the disk group and dbdatavol is the data volume to be replicated.

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 amount of use.

   See the *Veritas Storage Foundation and High Availability Solutions Replication Administrator's Guide* for details.

2. Issue the following command:

```
# vxassist -g dbdatadg make dbdata_srl 6800M nmirror=2 disk4
   disk5
```

   Note: Assume that for the example setup that *disk4* and *disk5* are already added and are part of the same disk group. They are used in this step for mirroring and creation of the SRL.

3. If the SRL volume is not already started, start the SRL volume by starting all volumes in the disk group:

```
# vxvol -g dbdatadg startall
```

Setting up the Replicated Volume Group on the primary site

Before creating the Replicated Volume Group (RVG) on the primary site, make sure the volumes and Cluster Volume Manager (CVM) group 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 to verify that the CVM group is online.
3. On each site, verify vradmin is running:

```
#  ps -ef |grep vradmin
    root  536594  598036   0 12:31:25      0  0:00 grep vradmin
```

If vradmin is not running start it:

```
# vxstart_vvr
VxVM VVR INFO V-5-2-3935 Using following ports:
heartbeat: 4145
vradmind: 8199
vxrsyncd: 8989
data: Anonymous-Ports
To change, see vrport(1M) command
#  ps -ef |grep vradmin
    root  536782       1   0 12:32:47      -  0:00 /usr/sbin/vradmind
    root 1048622  598036   0 12:32:55      0  0:00 grep vradmin
# netstat -an |grep 4145
tcp4       0      0  *.4145                 *.*          LISTEN
udp4       0      0  *.4145                 *.*
```

After reviewing the status of replication objects on the primary site, you can create the primary RVG.

The command to create the primary RVG takes the form:

```
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 Storage Replicator Log (SRL)

To create the primary RVG

1 Determine which node is the CVM master node by entering:

```
# vxdctl -c mode
```

2 To create the dbdata_rvg RVG, you must run the following on the master node:

```
# vradmin -g dbdatadg createpri dbdata_rvg
  dbdata_vol
  dbdata_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 *dbdata_vol*.

Setting up replication on the secondary site using VVR

To create objects for replication on the secondary site, use the `vradmin` command with the `addsec` option. To set up replication on the secondary site, perform the following tasks:

Table 9-3 Tasks for setting up replication on the secondary site using VVR

Task	Description
Create a disk group to hold the data volume, Storage Replicator Log (SRL), and Replicated Volume Group (RVG) on the storage on the secondary site. You must match the names and sizes of these volumes with the names and sizes of the volumes on the primary site. For example, you must set up the data files. For example, create the disk group, volume, and mount point for the database datafiles.	See the *Veritas Storage Foundation Cluster File System High Availability Installation Guide*. See the *Veritas Storage Foundation for Oracle RAC Installation and Configuration Guide*. See "Creating the data and SRL volumes on the secondary site" on page 126.
Edit the /etc/vx/vras/.rdg file on the each site.	See "Editing the /etc/vx/vras/.rdg files" on page 127.

Table 9-3 Tasks for setting up replication on the secondary site using VVR *(continued)*

Task	Description
Use 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 127.
Create the replication objects on the secondary site.	See "Setting up the disk group on secondary site for replication" on page 128.

Creating the data and SRL volumes on the secondary site

Note the following when creating volumes for the data and Storage Replicator Log (SRL):

- The sizes and names of the volumes must match the sizes and names of the corresponding volumes in the primary site before you create the disk group.
- The disk group must match the size and name of the disk group at 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 database, create a data volume of same size as that in primary for data; in this case, the dbdata_vol volume on the primary site is 12 GB:

```
# vxassist -g dbdatadg make dbdata_vol 12000M nmirror=2 disk11 disk12
```

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 dbdatadg make dbdata_srl 6800M nmirror=2 disk14 disk16
```

See the *Veritas Storage Foundation and High Availability Solutions Replication Administrator's Guide* for details.

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 Replicated Volume Group (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 (dgid).

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 /etc/vx/vras/.rdg file and enter the secondary disk group ID on a single line.

Setting up IP addresses for RLINKs on each cluster

Creating objects with the vradmin 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. Using the following command, determine whether a node is the CVM master or the slave:

```
# vxdctl -c mode
```

   You must configure RLINKs on the CVM master node.

2. For each Replicated Volume Group (RVG) running on each cluster, set up a virtual IP address on one of the nodes of the cluster. These IP addresses are part of the RLINK.

   The example assumes for the cluster on the primary site:

   - The public network interface is public_NIC1:2
   - The virtual IP address is 10.10.9.101
   - The net mask is 255.255.255.0

```
# ifconfig en0 10.10.9.101 netmask 255.255.255.0 \
  broadcast 10.180.95.255 alias
  # ifconfig en0 up
```

3 Use the same commands with appropriate values for the interface, IP address, and net mask on the secondary site.

The example assumes for the secondary site:

- The public network interface is public_NIC1:1
- virtual IP address is 10.11.9.102
- net mask is 255.255.255.0

4 Define the virtual IP addresses to correspond to a host name in the virtual cluster on the primary site and a host name in the virtual cluster on the secondary site.

Update the /etc/hosts file on all the nodes on both the primary and secondary sites.

The examples assume:

- clus1 has IP address 10.10.9.101
- clus2 has IP address 10.11.9.102

5 Use the ping command to verify the links are functional.

Setting up the disk group on secondary site for replication

Create the replication objects on the secondary site from the master node of the primary site, using the `vradmin` 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:

```
# vradmin -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: dbdata_vol
- rvg_pri is the Replicated Volume Group (RVG) on the primary site. For example: dbdata_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 clus1

- 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 clus2

For example, the command to add the cluster on the primary site to the Replicated Data Set (RDS) is:

```
vradmin -g dbdatadg addsec dbdata_rvg clus1 clus2
```

or

```
vradmin -g dbdatadg addsec dbdata_rvg 10.10.9.101 10.11.9.102
```

On the secondary site, the above command performs the following tasks:

- Creates an RVG within the specified disk group using the same name as the one for the primary site
- Associates the data and Storage Replicator Log (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_dbdata_clus2_dbdata_rvg and the "secondary" RLINK created is rlk_dbdata_clus1_dbdata_rvg.
  If you use 10.10.9.101 and 10.11.9.102, creates cluster RLINKs for the primary and secondary sites with the default names; for example, the "primary" RLINK created for this example is rlk_10.11.9.102_dbdata_rvg and the "secondary" RLINK created is rlk_10.10.9.101_dbdata__rvg.

2 Verify the list of RVGs in the RDS by executing the following command.

```
# vradmin -g dbdatadg -l printrvg
```

For example:

```
Replicated Data Set: dbdata_rvg
Primary:
HostName: 10.180.88.187 <localhost>
RvgName: dbdata_rvg
DgName: dbdata_vol
datavol_cnt: 1
vset_cnt: 0
srl: dbdata_srl
RLinks:
```

```
name=rlk_clus2_dbdata_rvg, detached=on,
synchronous=off
Secondary:
HostName: 10.190.99.197
RvgName:dbdata_rvg
DgName: dbdatadg
datavol_cnt: 1
vset_cnt: 0
srl: dbdata_srl
RLinks:
name=rlk_clus1_dbdata_rvg, detached=on,
synchronous=off
```

Note: Once the replication is started the value of the detached flag will change the status from ON to OFF.

Starting replication of the primary site database volume to the secondary site using VVR

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
- Storage Replicator Log (SRL) overflow protection: srlprot=autodcm
- Packet size: `packet_size=8400`
- Network protocol: `protocol=TCP`

Method of initial synchronization:

- Automatic synchronization
- Full synchronization with Storage Checkpoint

For guidelines on modifying these settings and information on choosing the method of replication for the initial synchronization:

See the *Veritas Storage Foundation and High Availability Solutions Replication Administrator's Guide*

Use the `vradmin` command to start replication or the transfer of data from the primary site to the secondary site over the network. Because the cluster on the secondary site uses only one host name, the command does not require the *sec_host* argument.

To start replication using automatic synchronization

- From the primary site, use the following command to automatically synchronize the Replicated Volume Group (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 Replicated Volume Group (RVG) on the primary site
  - sec_host is the virtual host name for the secondary site

  For example:

  ```
  # vradmin -g dbdatadg -a startrep dbdata_rvg clus2
  ```

Use the `vradmin` command with the Storage Checkpoint option to start replication using full synchronization with Storage Checkpoint.

To start replication using full synchronization with Storage 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 Storage 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 dbdatadg -c dbdata_ckpt syncrvg dbdata_rvg clus2
```

2 To start replication after full synchronization, enter the following command:

```
# vradmin -g dbdatadg -c dbdata_ckpt startrep dbdata_rvg clus2
```

Verify that replication is properly functioning.

To verify replication status

1 Check the status of VVR replication:

```
# vradmin -g disk_group_name repstatus rvg_name
```

2 Review the `flags` output for the status. The output may appear as `connected` and `consistent`. For example:

```
# vxprint -g dbdatadg -l rlk_clus2_dbdata_rvg
  Rlink: rlk_clus2_dbdata_rvg
  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: dbdata_ckpt
  flags: write enabled attached consistent connected
asynchronous
```

Configuring Veritas Cluster Server to replicate the database volume using VVR

After configuring both clusters for global clustering and setting up the database for replication, configure Veritas Cluster Server (VCS) to provide high availability for the database. Specifically, configure VCS agents to control the cluster resources, including the replication resources.

Table 9-4 Tasks for configuring VCS to replicate the database volume using VVR

Task	Description
Modify the VCS configuration on the primary site	See "Modifying the Veritas Cluster Server (VCS) configuration on the primary site" on page 137.
Modifying the VCS configuration on the secondary site	See "Modifying the VCS configuration on the secondary site" on page 142.

The following resources must be configured or modified for replication:

Table 9-5 VCS resource modifications for replication with VVR

Resource	Modification
Log owner service group	Create a log owner service group including the RVGLogowner resources. The RVGLogowner resources are used by: ■ RLINKs for the RVG ■ RVGLogowner resource. The RVG and its associated disk group are defined as attributes for the RVGLogowner resource.
Replicated Volume Group (RVG) service group	Create an RVG group that includes the RVGShared resource replication objects. 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 RVG log owner service group has an online local firm dependency on the service group containing the RVG. VCS uses the following agents to control the following resources: ■ RVGLogowner agent to control the RVGLogowner resource ■ RVGShared agent to control the RVGShared resource
CVMVolDg resource	The CVMVolDg resource does not have replicated service 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.

Table 9-5 VCS resource modifications for replication with VVR *(continued)*

Resource	Modification
RVGSharedPri resource	Add the RVGSharedPri resource to the existing database service group. The CVMVolDg resource must be removed from the existing database service group.
Database service group	The existing database service group is a parallel group consisting of the database resource, CVMVolDg resource, and CFSMount resource (if the database resides in a cluster file system). Define the database service group as a global group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute.

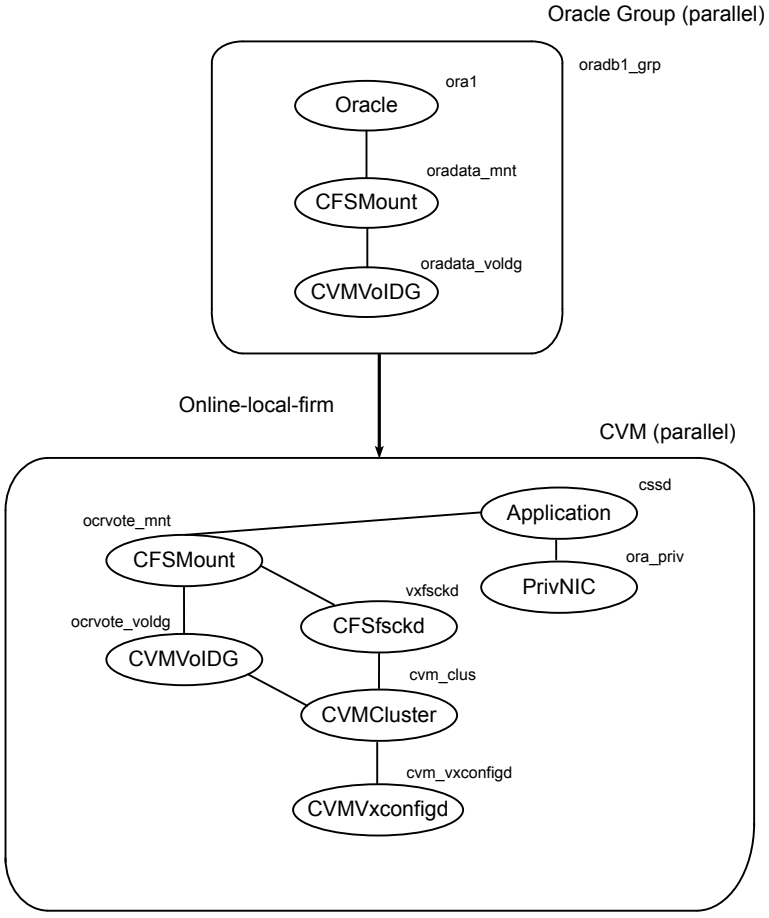
For more information on service replication resources:

See the *Veritas Cluster Server Bundled Agents Guide*

Review the following illustrations that display the changes to the VCS configuration, after setting up replication on the existing database. All of the dependencies between parent and child groups are online local firm. The CVM service group is the same in all illustrations because its definition requires no changes.

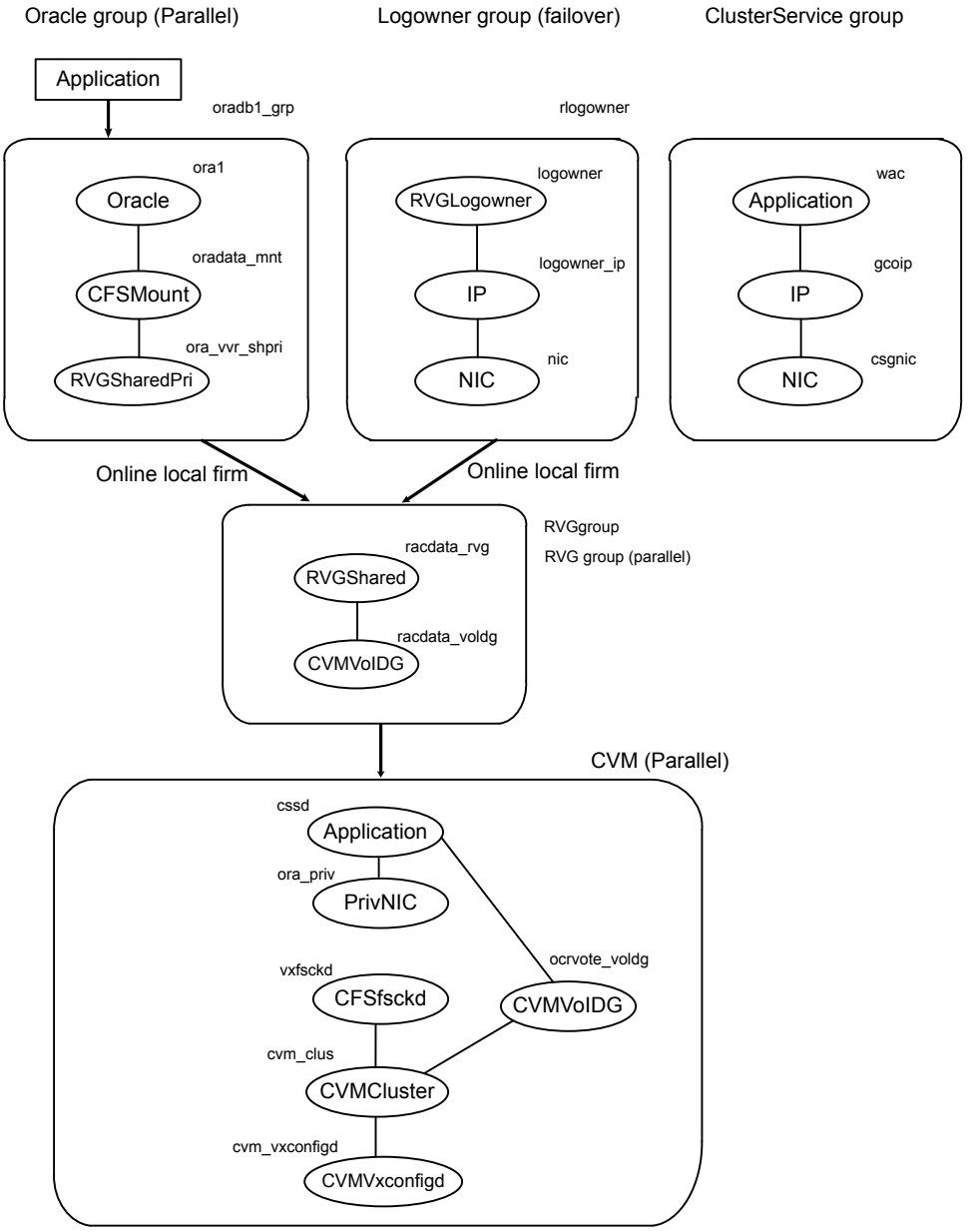
Configuration before modification for replication:

Figure 9-1 Dependencies before modification for replication of Oracle RAC



Configuration after modification for replication:

Figure 9-2 Dependencies after modification for replication of Oracle RAC



Modifying the Veritas Cluster Server (VCS) configuration on the primary site

The following are the tasks required to modify the existing VCS configuration on the primary site:

Table 9-6 Tasks for modifying the VCS configuration on the primary site

Task	Description
Configure two service groups: ■ A log owner group including the RVGLogowner resource. ■ A Replicated Volume Group (RVG) group including the RVGShared resource replication objects.	See "To modify VCS on the primary site" on page 138. See "Modifying the VCS configuration on the secondary site" on page 142.
Add the RVGSharedPri resource to the existing database service group and define this group as a global group by setting the ClusterList and ClusterFailOverPolicy attributes.	See "To modify VCS on the primary site" on page 138. See "Modifying the VCS configuration on the secondary site" on page 142.
Move the CVMVolDg resource from the existing database service group to the newly created RVG group.	See "To modify VCS on the primary site" on page 138. See "Modifying the VCS configuration on the secondary site" on page 142.
To view the sample main.cf files on your system:	See "To view sample configuration files for SF Oracle RAC" on page 137.

To view sample configuration files for SF Oracle RAC

1 Change directories to the find the sample main.cfs:

```
# cd /etc/VRTSvcs/conf/sample_rac
```

2 Enter:

```
# ls *sfrac*
sfrac07_main.cf sfrac08_main.cf
```

The following files include CVM/VVR configuration examples:

- For the primary: sfrac_07_main.cf
- For the secondary: sfrac_08_main.cf

See "Sample Storage Foundation for Oracle RAC configuration files" on page 156.

To modify VCS on the primary site

1. Log into one of the nodes on the primary cluster.
2. Use the following command to save the existing configuration to disk, and make the configuration read-only while you make changes:

```
# haconf -dump -makero
```

3. Use the following command to make a backup copy of the main.cf file:

```
# cd /etc/VRTSvcs/conf/config
# cp main.cf main.orig
```

4. Use vi or another text editor to edit the main.cf file. Review the sample configuration file after your product installation.

   Add a failover service group using the appropriate values for your cluster and nodes. Include the following resources:

   - 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

   The following are examples of RVGLogowner service group for the different platforms.

   Example for Oracle RAC:

```
group rlogowner (
    SystemList = { sys1 = 0, sys2 = 1 }
    AutoStartList = { sys1, sys2 }
    )

    IP logowner_ip (
        Device = en0
        Address = "10.10.9.101"
        NetMask = "255.255.255.0"
        )
    NIC nic (
        Device = en0
        NetworkType = ether
```

```
        NetworkHosts = "10.10.8.1"
        )
RVGLogowner logowner (
     RVG = dbdata_rvg
     DiskGroup = dbdatadg
     )
requires group RVGgroup online local firm
logowner requires logowner_ip
logowner_ip requires nic
```

5 Add the RVG service group using the appropriate values for your cluster and nodes.

Example `RVGgroup` service group:

```
group RVGgroup (
     SystemList = { sys1 = 0, sys2 = 1 }
     Parallel = 1
     AutoStartList = { sys1, sys2 }
     )
RVGShared dbdata_rvg (
     RVG = dbdata_rvg
     DiskGroup = dbdatadg
     )
     CVMVolDg dbdata_voldg (
         CVMDiskGroup = dbdatadg
         CVMActivation = sw
         CVMVolume = { dbvol, dbdata_srl }
         )
requires group cvm online local firm
dbdata_rvg requires dbdata_voldg
```

6 Modify the database service group using the appropriate values for your cluster and nodes:

- Define the database 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 (online, local, firm) to depend on the RVG service group.
- Remove the existing dependency of the Database service group on the CVM service group. Remove the line:

  ```
  requires group CVM online local firm
  ```

- Remove the existing dependency between the CFSMount for the database and the CVMVoldg for the database. Remove the line:

  ```
  dbdata_mnt requires dbdata_voldg
  ```

See configuration examples below.

7 Save and close the main.cf file.

8 It is advisible to modify "OnlineRetryLimit" & "OfflineWaitLimit" attribute of IP resource type to 1 on both the clusters:

```
# hatype -modify IP  OnlineRetryLimit  1

# hatype -modify IP  OfflineWaitLimit  1
```

9 Use the following command to verify the syntax of the /etc/VRTSvcs/conf/config/main.cf file:

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

10 Stop and restart VCS.

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

Wait for port h to stop on all nodes, and then restart VCS with the new configuration on all primary nodes:

```
# hastart
```

Storage Foundation Cluster File System High Availability (SFCFS HA) example of a database service group configured for replication:

```
group database_grp (
    SystemList = { sys1 = 0, sys2 = 1 }
    ClusterList = { clus1 = 0, clus2 = 1 }
    Parallel = 1
    ClusterFailOverPolicy = Manual
    Authority = 1
```

```
    AutoStartList = { sys1,sys2 }
    OnlineRetryLimit = 3
    TriggerResStateChange = 1
    OnlineRetryInterval = 120
    )

    CFSMount dbdata_mnt (
        MountPoint = "/dbdata"
        BlockDevice = "/dev/vx/dsk/dbdatadg/dbdata_vol"
        )

    Process vxfend (
    PathName = "/sbin/vxfend"
    Arguments = "-m sybase -k /tmp/vcmp_socket"
    )

    RVGSharedPri dbdata_vvr_shpri (
       RvgResourceName = dbdata_rvg
       OnlineRetryLimit = 0
       )

requires group RVGgroup online local firm
oradata_mnt requires dbdata_vvr_shpri
```

Storage Foundation (SF) for Oracle RAC example of a database service group configured for replication:

```
group database_grp (
    SystemList = { sys1 = 0, sys2 = 1 }
    ClusterList = { clus1 = 0, clus2 = 1 }
    Parallel = 1
    ClusterFailOverPolicy = Manual
    Authority = 1
    AutoStartList = { sys1,sys2 }
    )

    CFSMount oradata_mnt (
        MountPoint = "/oradata"
        BlockDevice = "/dev/vx/dsk/oradatadg/racdbdata_vol"
        )

    RVGSharedPri ora_vvr_shpri (
       RvgResourceName = racdata_rvg
```

```
        OnlineRetryLimit = 0
        )

Oracle rac_db (
    Sid @sys1 = vrts1
    Sid @sys2 = vrts2
    Owner = Oracle
    Home = "/oracle/orahome"
    StartUpOpt = SRVCTLSTART
    ShutDownOpt = SRVCTLSTOP
    )
requires group RVGgroup online local firm
oradata_mnt requires ora_vvr_shpri
rac_db requires oradata_mnt
```

Modifying the VCS configuration on the secondary site

The following are highlights of the procedure to modify the existing VCS configuration on the secondary site:

- Add the log owner and Replicated Volume Group (RVG) service groups.
- Add a service group to manage the 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 following steps are similar to those performed on the primary site.

To modify VCS on the secondary site

1. Log into one of the nodes on the secondary site as root.
2. Use the following command to save the existing configuration to disk, and make the configuration read-only while making changes:

   ```
   # haconf -dump -makero
   ```

3. Use the following command to make a backup copy of the main.cf file:

   ```
   # cd /etc/VRTSvcs/conf/config
   # cp main.cf main.orig
   ```

4 Use vi or another text editor to edit the main.cf file. Edit the CVM group on the secondary site.

Review the sample configuration file after the VCS installation to see the CVM configuration.

See "To view sample configuration files for SF Oracle RAC" on page 137.

In our example, the secondary site has clus2 consisting of the nodes sys3 and sys4. To modify the CVM service group on the secondary site, use the CVM group on the primary site as your guide.

5 Add a failover service group using the appropriate values for your cluster and nodes. Include the following resources:

- 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 = { sys3 = 0, sys4 = 1 }
    AutoStartList = { sys3, sys4 }
    )

    IP logowner_ip (
       Device = en0
       Address = "10.11.9.102"
       NetMask = "255.255.255.0"
       )
NIC nic (
       Device = en0
       NetworkHosts = { "10.10.8.1" }
       NetworkType = ether
       )
RVGLogowner logowner (
       RVG = dbdata_rvg
       DiskGroup = dbdatadg
       )
requires group RVGgroup online local firm
logowner requires logowner_ip
logowner_ip requires nic
```

6 Add the RVG service group using the appropriate values for your cluster and nodes.

The following is an example `RVGgroup` service group:

```
group RVGgroup (
    SystemList = { sys3 = 0, sys4 = 1 }
    Parallel = 1
    AutoStartList = { sys3, sys4 }
    )

RVGShared dbdata_rvg (
    RVG = dbdata_rvg
    DiskGroup = dbdatadg
    )

    CVMVolDg dbdata_voldg (
        CVMDiskGroup = dbdatadg
        CVMActivation = sw
        )

requires group cvm online local firm
dbdata_rvg requires dbdata_voldg
```

7 It is advisible to modify "OnlineRetryLimit" & "OfflineWaitLimit" attribute of IP resource type to 1 on both the clusters:

```
# hatype -modify IP  OnlineRetryLimit  1

# hatype -modify IP  OfflineWaitLimit  1
```

8 Add an database service group. Use the database service group on the primary site as a model for the database service group on the secondary site.

- Define the database 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, database_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.

See configuration examples below.

9 Save and close the `main.cf` file.

10 Use the following command to verify the syntax of the `/etc/VRTSvcs/conf/config/main.cf` file:

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

11 Stop and restart VCS.

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

Wait for port h to stop on all nodes, and then restart VCS with the new configuration on all primary nodes one at a time.

```
# hastart
```

12 Verify that VCS brings all resources online. On one node, enter the following command:

```
# hagrp -display
```

The database, RVG, and CVM groups are online on both nodes of the primary site. The RVGLogOwner and ClusterService groups are 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 `hagrp -online` command. This information applies to the secondary site, except for the database group which must be offline.

13 For AIX, on the primary site, enter the following commands:

```
# hagrp -online rlogowner -sys sys1

# hagrp -online database_grp -sys sys1

VCS WARNING V-16-1-50817 Please use hagrp -online -force to
online a global group for the first time

# hagrp -online -force database_grp -sys sys1
```

On the secondary site, enter the following command:

```
# hagrp -online rlogowner -sys sys3
```

14 Verify the service groups and their resources that are brought online. On one node, enter the following command:

```
# hagrp -display
```

The database service group is offline on the secondary site, but the ClusterService, CVM, RVG log owner, and RVG groups are online.

This completes the setup for a global cluster using VVR for replication. Symantec recommends testing a global cluster before putting it into production.

Example of the Oracle RAC database group on the secondary site:

```
group database_grp (
    SystemList = { sys3 = 0, sys3 = 1 }
    ClusterList = { clus2 = 0, clus1 = 1 }
    Parallel = 1
    OnlineRetryInterval = 300
    ClusterFailOverPolicy = Manual
    Authority = 1
    AutoStartList = { sys3, sys4 }
    )

    RVGSharedPri dbdata_vvr_shpri (
        RvgResourceName = rdbdata_rvg
        OnlineRetryLimit = 0
        )

    CFSMount dbdata_mnt (
        MountPoint = "/dbdata"
        BlockDevice = "/dev/vx/dsk/dbdatadg/dbdata_vol"
        Critical = 0
```

```
        )
    RVGSharedPri dbdata_vvr_shpri (
        RvgResourceName = dbdata_rvg
        OnlineRetryLimit = 0
        )

Oracle rac_db (
     Sid @sys3 = vrts1
     Sid @sys4 = vrts2
     Owner = Oracle
     Home = "/oracle/orahome"
     Pfile @sys3 = "/oracle/orahome/dbs/initvrts1.ora"
     Pfile @sys4 = "/oracle/orahome/dbs/initvrts2.ora"
     StartUpOpt = SRVCTLSTART
     ShutDownOpt = SRVCTLSTOP
     )

requires group RVGgroup online local firm
dbdata_mnt requires dbdata_vvr_shpri
rac_db requires dbdata_mnt

RVGSharedPri dbdata_vvr_shpri (
        RvgResourceName = dbdata_rvg
        OnlineRetryLimit = 0
        )

requires group RVGgroup online local firm
dbdata_mnt requires dbdata_vvr_shpri
```

Replication use cases for global parallel clusters

For information on the VCS commands for global clusters:

See the *Veritas Cluster Server Administrator's Guide*.

If you have two clusters configured to use VVR for replication, the following replication use cases are supported:

Table 9-7 Replication use cases for global parallel clusters

Management option	Description
Migration of the role of the primary site to the remote site	Migration is a 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 of the primary site role by the secondary site	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.
Migrate the role of primary site to the secondary site	See "To migrate the role of primary site to the remote site" on page 149.
Migrate the role of new primary site back to the original primary site	See "To migrate the role of new primary site back to the original primary site" on page 150.
Take over after an outage	See "To take over after an outage" on page 152.
Resynchronize after an outage	See "To resynchronize after an outage" on page 153.
Update the rlink	See "To update the rlink" on page 154.

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 database service group, *database_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.

To migrate the role of primary site to the remote site

1. From the primary site, use the following command to take the database service group offline on all nodes.

   ```
   # hagrp -offline database_grp -any
   ```

   Wait for VCS to take all database 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. You can use the command from any node on the primary cluster.

   For example:

   ```
   # vxrlink -g data_disk_group status rlk_clus2_dbdata_rvg
   ```

   Where rlk_clus1_dbdata_rvg is the RLINK.

3. On the secondary site, which is now the new primary site, bring the database service group online on all nodes:

   ```
   # hagrp -online database_grp -any
   ```

After migrating the role of the primary site to the secondary site, you can use VCS commands to migrate the role of the cluster on the new primary site to the original primary site. In the procedure below, VCS takes the replicated database service group, database_grp, offline on the new primary (former secondary) site and brings it online on the original primary site; the original primary site now resumes 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.

To migrate the role of new primary site back to the original primary site

1. Make sure that all database resources are online, and switch back the group database_grp to the original primary site.

   Issue the following command on the remote site:

   ```
   # hagrp -offline database_grp -any
   ```

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. You can use the command from any node on the current primary cluster.

   For example:

   ```
   # vxrlink -g data_disk_group status rlk_clus1_dbdata_rvg
   ```

   Where rlk_clus1_dbdata_rvg is the RLINK.

3. Make sure that database_grp is offline on the new primary site. Then, execute the following command on the original primary site to bring the database_grp online:

   ```
   # hagrp -online database_grp -any
   ```

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). For a detailed description of concepts of taking over the primary role:

See the *Veritas Storage Foundation and High Availability Solutions Replication Administrator's Guide*.

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 and designate the failure type using one of the options for the `haclus` command.

Takeover options are:

Table 9-8 Takeover options on global parallel clusters

Takeover option	Description
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 of making 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.
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," which means 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.

Table 9-8 Takeover options on global parallel clusters *(continued)*

Takeover option	Description
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".

The examples illustrate the steps required for an outage takeover and resynchronization.

To take over after an outage

1 From any node of the secondary site, issue the `haclus` command:

```
# haclus -declare outage -clus clus1
```

2 After declaring the state of the remote cluster, bring the *database_grp* service group online on the secondary site. For example:

```
# hagrp -online -force database_grp -any
```

To resynchronize after an outage

1 On the original primary site, create a snapshot of the Replicated Volume Group (RVG) before resynchronizing it in case the current primary site fails during the resynchronization. Assuming the disk group is data_disk_group and the RVG is *dbdata1_rvg*, type:

```
# vxrvg -g data_disk_group -F snapshot dbdata_rvg
```

See the *Veritas Storage Foundation and High Availability Solutions Replication Administrator's Guide* for details on RVG snapshots.

2 Resynchronize the RVG. From any 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 dbdata_vvr_shpri fbsync -sys sys3

# 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 data_disk_group snapback dbdata_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 data_disk_group snaprestore dbdata_rvg
  ```

If the rlink is not up to date, use the `hares -action` command with the `resync` action token to synchronize the RVG.

To update the rlink

- The following command example is issued on any node (`sys1`, in this case) in the primary cluster, specifying the RVGSharedPri resource, `dbdata_vvr_shpri`:

```
# hares -action dbdata_vvr_shpri resync -sys sys1
```

Section 5

Reference

- Appendix A. Sample configuration files

Appendix A

Sample configuration files

This appendix includes the following topics:

- Sample Storage Foundation for Oracle RAC configuration files
- About sample main.cf files for Veritas Storage Foundation (SF) for Oracle RAC

Sample Storage Foundation for Oracle RAC configuration files

Storage Foundation for Oracle RAC provides several sample configuration files illustrating various scenarios. You may use the sample files as a guideline for setting up your cluster environment. These sample files are located at `/etc/VRTSvcs/conf/sample_rac/`.

This section briefly describes each of the sample files and illustrates the service group configuration for each of them. The section does not include a copy of the main.cf files.

The following sample files are useful references for disaster recovery use cases:

- sfrac02_main.cf file
- sfrac11_main.cf file
- See “Sample fire drill service group configuration” on page 166.

sfrac02_main.cf file

This sample file describes the following configuration:

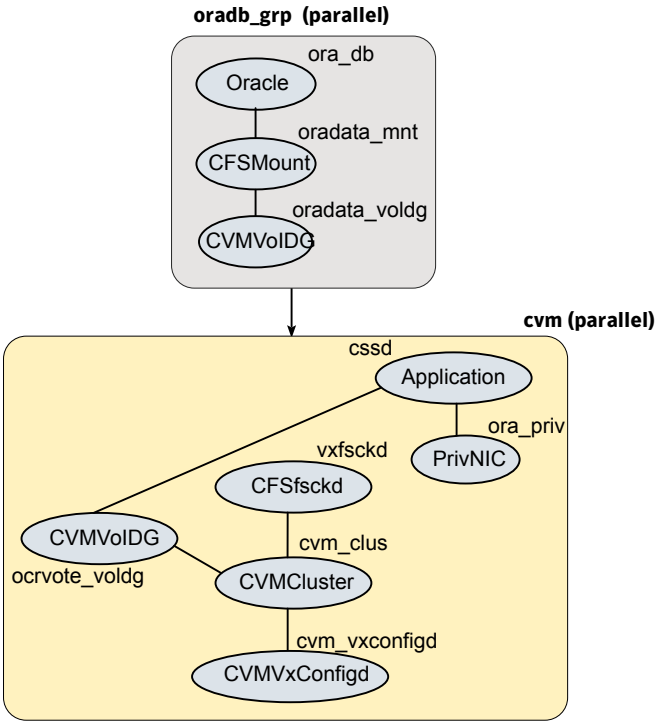
- A two node Storage Foundation for Oracle RAC cluster.
- The Oracle database is stored on CFS.
- The database is managed by the VCS agent for Oracle.

The agent starts, stops, and monitors the database.

- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion. The private IP address is managed by the PrivNIC agent for high availability.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CVM raw volumes.

Figure A-1 illustrates the configuration.

Figure A-1 Service group configuration for sfrac02_main.cf file



sfrac07_main.cf and sfrac08_main.cf files

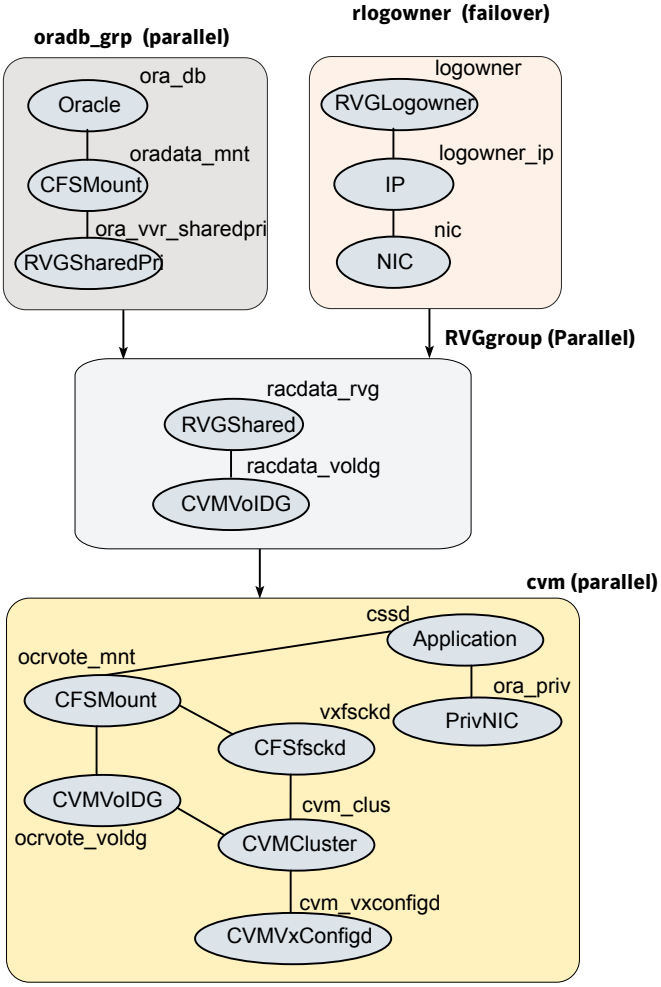
The sample configuration, sfrac07_main.cf, describes a disaster recovery configuration for the primary site. The sample configuration, sfrac08_main.cf, describes a disaster recovery configuration for the secondary site. The configuration uses VVR for replicating data between the sites.

This sample file describes the following configuration:

- Two Storage Foundation for Oracle RAC clusters, comprising two nodes each, hosted at different geographical locations.
- A single Oracle database that is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion. The private IP address is managed by the PrivNIC agent for high availability.
- One virtual IP address must be configured under the `ClusterService` group on each site for inter-cluster communication.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.
- Veritas Volume Replicator (VVR) is used to replicate data between the sites.
- The shared volumes replicated across the sites are configured under the RVG group.
- The replication link used by VVR for communicating log information between sites are configured under the `rlogowner` group. This is a failover group that will be online on only one of the nodes in the cluster at each site.
- The database group will be online on the primary cluster. The RVGSharedPri resource determines where the database group will be brought online.
- The database group is configured as a global group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute.

Figure A-2 illustrates the configuration. The service group configuration is the same on the primary and secondary site. The availability of groups (online/offline) differ between the sites.

Figure A-2 Service group configuration for sfrac07_main.cf and sfrac08_main.cf files



sfrac09_main.cf and sfrac10_main.cf files

The sample configuration, sfrac09_main.cf, describes a disaster recovery configuration for the primary site. The sample configuration, sfrac10_main.cf, describes a disaster recovery configuration for the secondary site. The sample configuration uses EMC SRDF technology for replicating data between the sites.

Note: You can use other supported hardware-based replication technologies with this configuration.

This sample file describes the following configuration:

- Two Storage Foundation for Oracle RAC clusters, comprising two nodes each, hosted at different geographical locations.
- A single Oracle database that is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion. The private IP address is managed by the PrivNIC agent for high availability.
- One virtual IP address must be configured under the `ClusterService` group on each site for inter-cluster communication.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.
- EMC SRDF is used to replicate data between the sites.
- The SRDF disk groups that are replicated across the sites using SRDF technology and the replication mode are specified under the SRDF resource in the database group. The CVM disk group that comprises the SRDF disk group must be configured under the CVMVolDg resource in the database group.
- The database group will be online on the primary cluster. The SRDF resource determines where the database group will be brought online.
- The database group is configured as a global group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute.

Figure A-3 illustrates the configuration on the primary site.

Figure A-3 Service group configuration for sfrac09_main.cf file

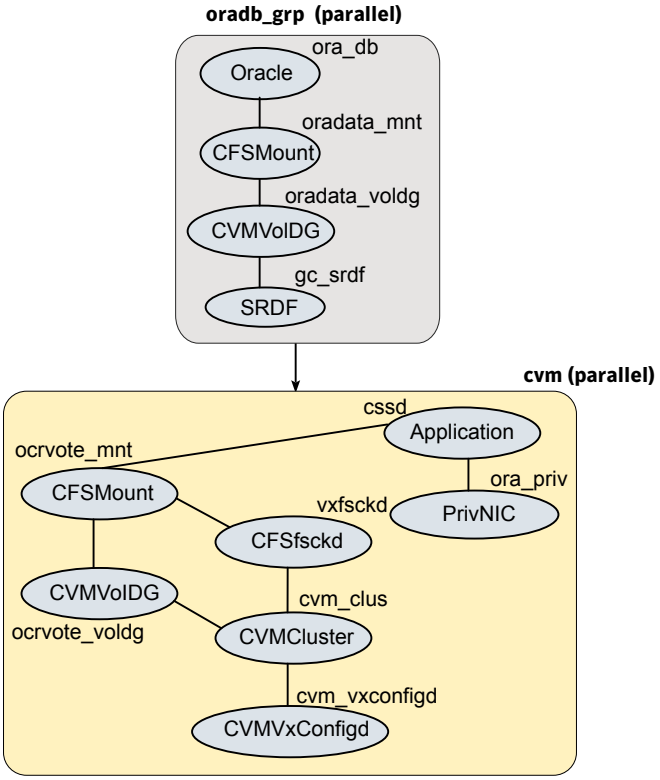
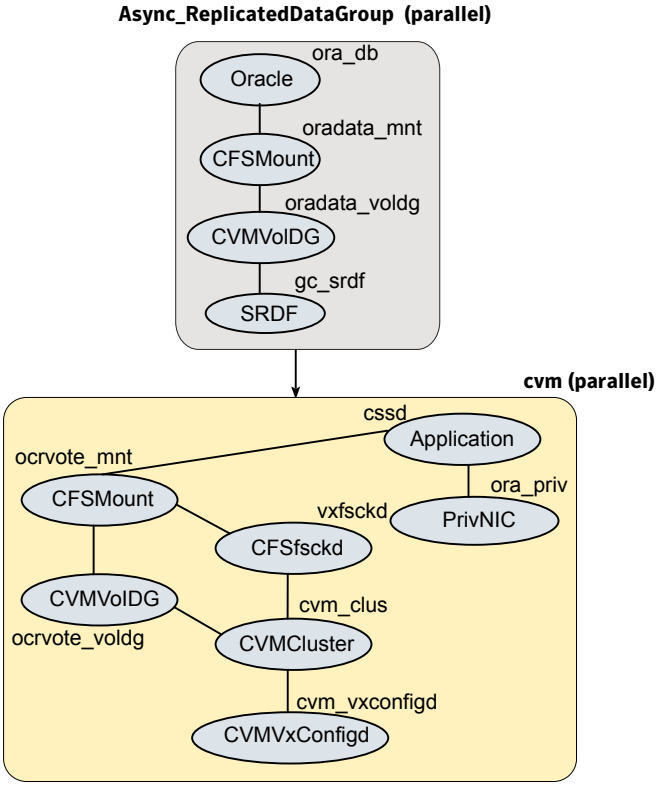


Figure A-4 illustrates the configuration on the secondary site.

Figure A-4 Service group configuration for sfrac10_main.cf file



sfrac11_main.cf file

This sample file describes the following configuration:

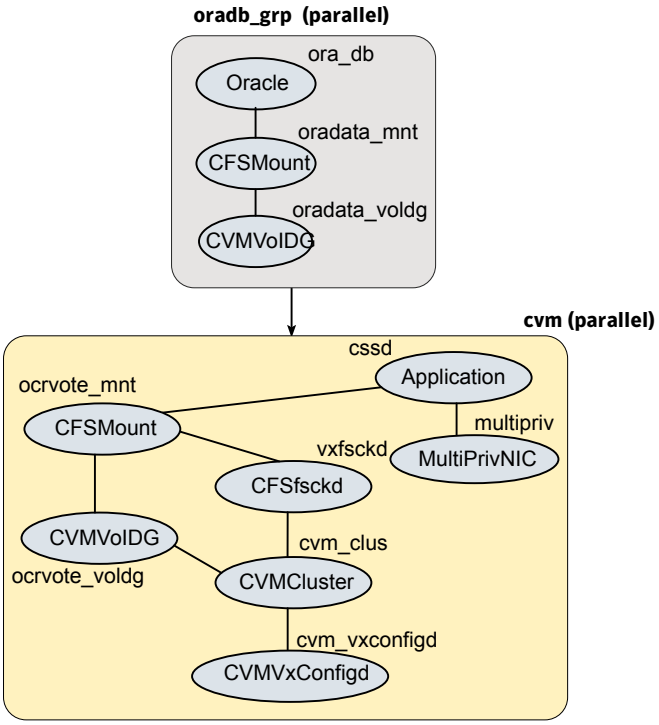
- An Storage Foundation for Oracle RAC campus cluster with four nodes hosted across two sites.
- Each site comprises two nodes of the cluster hosting a single database.
- The Oracle database is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- The IP address on NIC1 is used by Oracle Clusterware.
  The second IP address on NIC2 is used for Oracle database cache fusion.

The private IP addresses are managed by the MultiPrivNIC agent for high availability.

- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.
- Group the hosts at each physical site into separate logical system zones using the SystemZones attribute.

Figure A-5 illustrates the configuration.

Figure A-5 Service group configuration for sfrac11_main.cf file



sfrac12_main.cf and sfrac13_main.cf files

The sample configuration, sfrac12_main.cf, describes a disaster recovery configuration for the primary site. The sample configuration, sfrac13_main.cf, describes a disaster recovery configuration for the secondary site with fire-drill capability. The sample configuration uses Hitachi True Copy technology for replicating data between the sites.

Note: You can use other supported hardware-based replication technologies with this configuration.

This sample file describes the following configuration:

- Two Storage Foundation for Oracle RAC clusters, comprising two nodes each, hosted at different geographical locations.
- A single Oracle database that is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.
- A common IP address is used by Oracle Clusterware and database cache fusion. The private IP address is managed by the PrivNIC agent for high availability.
- One virtual IP address must be configured under the `ClusterService` group on each site for inter-cluster communication.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.
- Hitachi True Copy is used to replicate data between the sites.
- The HTC disk groups that are replicated across the sites using HTC technology and the replication mode are specified under the HTC resource in the database group. The CVM disk group that comprises the HTC disk group must be configured under the CVMVolDg resource in the database group.
- The database group will be online on the primary cluster. The HTC resource determines where the database group will be brought online.
- The database group is configured as a global group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute.
- The database group oradb_grp_fd on the secondary is configured for fire drill.
- When the group oradb_grp_fd is brought online, the HTCSnap creates a snapshot of the disk group configured under the HTC resource in the database group oradg_grp.
  Further, the Oracle database and the associated volumes and mount points configured under the service group oradb_grp_fd are brought online using the snapshots created by HTCSnap.

Figure A-6 illustrates the configuration on the primary site.

Figure A-6 Service group configuration for sfrac12_main.cf file

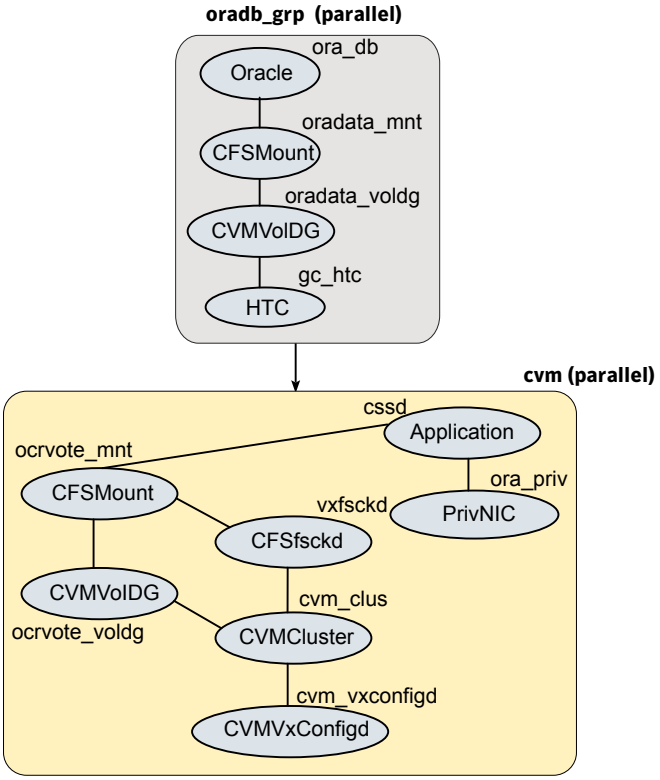
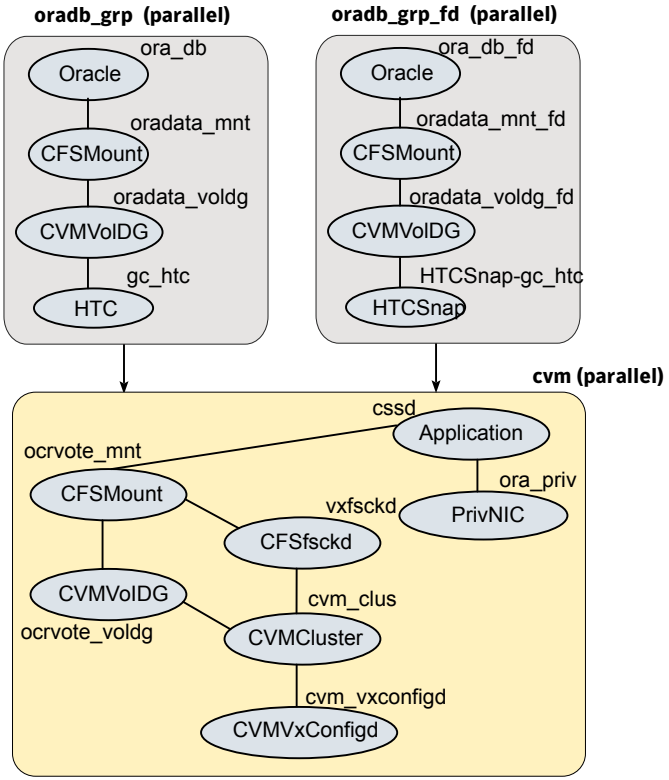


Figure A-7 illustrates the configuration on the secondary site.

Figure A-7 Service group configuration for sfrac13_main.cf file



Sample fire drill service group configuration

The sample configuration in this section describes a fire drill service group configuration on the secondary site. The configuration uses VVR for replicating data between the sites.

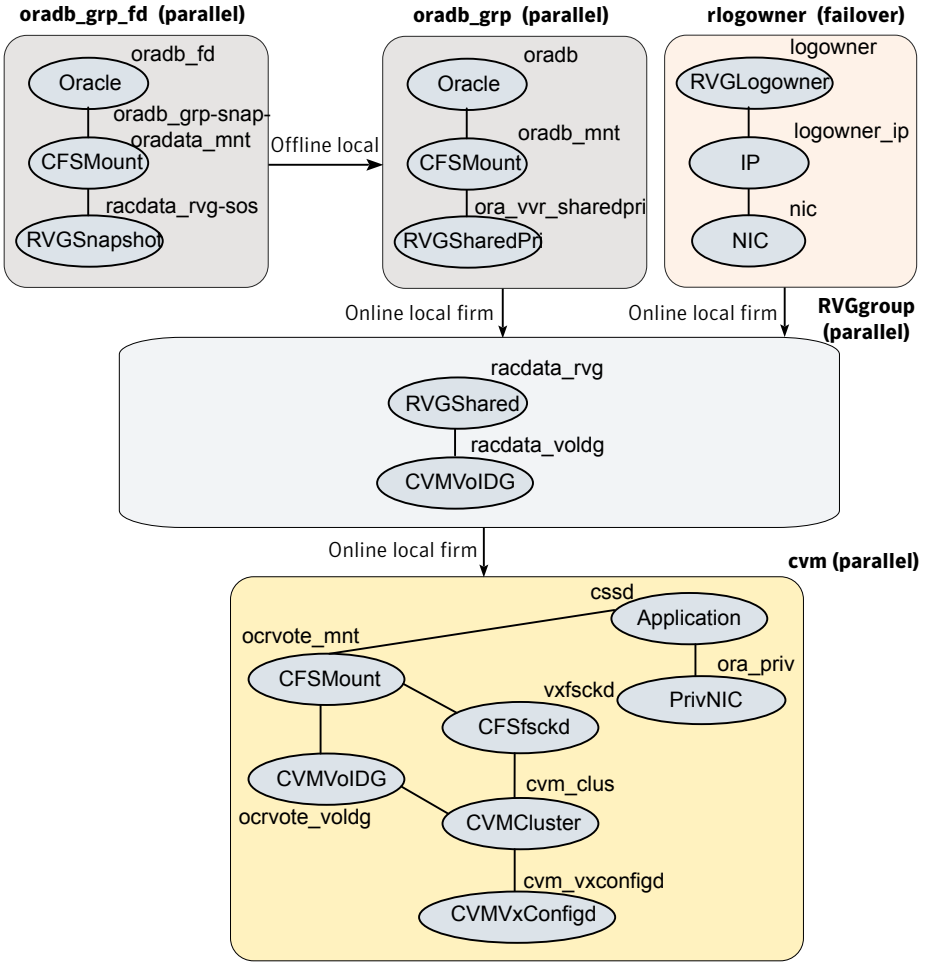
The sample service group describes the following configuration:

- Two Storage Foundation for Oracle RAC clusters, comprising two nodes each, hosted at different geographical locations.
- A single Oracle database that is stored on CFS.
- The database is managed by the VCS agent for Oracle.
  The agent starts, stops, and monitors the database.
- The database uses the Oracle UDP IPC for database cache fusion.

- A common IP address is used by Oracle Clusterware and database cache fusion. The private IP address is managed by the PrivNIC agent for high availability.
- One virtual IP address must be configured under the `ClusterService` group on each site for inter-cluster communication.
- The Oracle Cluster Registry (OCR) and voting disk are stored on CFS.
- Veritas Volume Replicator (VVR) is used to replicate data between the sites.
- The shared volumes replicated across the sites are configured under the RVG group.
- The replication link used by VVR for communicating log information between sites are configured under the `rlogowner` group. This is a failover group that will be online on only one of the nodes in the cluster at each site.
- The database group is configured as a global group by specifying the clusters on the primary and secondary sites as values for the ClusterList group attribute.
- The fire drill service group `oradb_grp_fd` creates a snapshot of the replicated data on the secondary site and starts the database using the snapshot. An offline local dependency is set between the fire drill service group and the application service group to make sure a fire drill does not block an application failover in case a disaster strikes the primary site.

Figure A-8 illustrates the configuration.

Figure A-8 Service group configuration for fire drill



About sample main.cf files for Veritas Storage Foundation (SF) for Oracle RAC

You can examine the VCS configuration file, main.cf, to verify SF Oracle RAC installation and configuration.

Sample main.cf file examples are provided for the following Oracle RAC configurations:

- Replicating data between two clusters

- For a primary site in a CVM VVR configuration
- For a secondary site in a CVM VVR configuration

Sample main.cf for Oracle 10g for CVM/VVR primary site

The following are the configuration details for this sample main.cf:

- Configuration file name: cvmvvr_primary_main.cf
- More general purpose, can have multiple Oracle databases

```
include "types.cf"
include "types.cf"
include "CFSTypes.cf"
include "CVMTypes.cf"
include "OracleTypes.cf"
include "PrivNIC.cf"
include "VVRTypes.cf"
include "VVRTypes.cf"
include "VVRTypes.cf"
include "/etc/VRTSvcs/conf/config/VVRTypes.cf"

cluster rac_cluster101 (
    UserNames = { admin = bopHo }
    ClusterAddress = "10.10.10.101"
    Administrators = { admin }
    UseFence = SCSI3
    )

remotecluster 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 = en0
        Address = "10.10.10.101"
        NetMask = "255.255.240.0"
        )

    NIC csgnic (
        NetworkHosts = { "10.10.12.2", "10.10.12.3" }
       Device = en0
       )

    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 (
        )

    CFSMount ocrvote_mnt (
        Critical = 0
        MountPoint = "/ocrvote"
        BlockDevice = "/dev/vx/dsk/ocrvotedg/ocrvotevol"
        )

CVMVolDg ocrvote_voldg (
    Critical = 0
    CVMDiskGroup = ocrvotedg
    CVMVolume = { ocrvotevol }
    CVMActivation = sw
    )

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

PrivNIC ora_priv (
    Critical = 0
    Device = { en1 = 0, en2 = 1}
    Address@galaxy = "192.168.12.1"
    Address@nebula = "192.168.12.2"
    NetMask = "255.255.240.0"
    )
```

```
    cssd requires ocrvote_mnt
    cssd requires ora_priv
    cssd requires ora_priv
    ocrvote_mnt requires ocrvote_voldg
    ocrvote_mnt requires vxfsckd
    ocrvote_voldg requires cvm_clus
    vxfsckd requires cvm_clus
    cvm_clus requires cvm_vxconfigd

group oradb1_grp (
    SystemList = { galaxy = 0, nebula = 1 }
    Parallel = 1
    ClusterList = { rac_cluster101 = 0, rac_cluster102 = 1 }
    OnlineRetryInterval = 300
    ClusterFailOverPolicy = Manual
    AutoStartList = { galaxy, nebula }
    Authority = 1
    )

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 = en0
        Address = "10.10.9.101"
        NetMask = "255.255.240.0"
        )

RVGLogowner logowner (
    RVG = rac1_rvg
    DiskGroup = oradatadg
    )

    requires group RVGgroup online local firm
    logowner requires logowner_ip
    logowner_ip requires nic

group VxSS (
SystemList = { north = 0, south = 1 }
Parallel = 1
OnlineRetryLimit = 3
OnlineRetryInterval = 120
)
Phantom phantom_vxss (
)
ProcessOnOnly vxatd (
IgnoreArgs = 1
PathName = "/opt/VRTSat/bin/vxatd"
)

group CMC (
SystemList = { north, south }
AutoStartList = { north, south }
OnlineRetryLimit = 3
OnlineRetryInterval = 120
)
ClusterConnectorConfig CMC_ClusterConfig (
MSAddress = "mgmtserver1.symantecexample.com"
MSPort = 14145
ClusterId = "1145613636"
ClusterType = "vcs"
```

```
ClusterPort = 14141
VCSLoggingLevel = "TAG_A"
Logging = "/opt/VRTScmccc/conf/cc_logging.properties"
ClusterConnectorVersion = "5.0.1000.0"
)
Process CMC_ClusterConnector (
PathName = "/bin/sh"
Arguments = "/opt/VRTScmccc/bin/cluster_connector.sh"
)
CMC_ClusterConnector requires CMC_ClusterConfig
```

Sample main.cf for Oracle 10g for CVM/VVR secondary site

The following are the configuration details for this sample main.cf:

- 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 "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 = en0
        Address = "10.11.10.102"
        NetMask = "255.255.240.0"
        )

    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 = 0, jupiter = 1 }
        CVMTransport = gab
        CVMTimeout = 200
        )

    CFSMount ocrvote_mnt (
        Critical = 0
        MountPoint = "/ocrvote"
        BlockDevice = "/dev/vx/dsk/ocrvotedg/ocrvotevol"
        )

    CVMVolDg ocrvote_voldg (
        Critical = 0
        CVMDiskGroup = ocrvotedg
        CVMVolume = { ocrvotevol }
        CVMActivation = sw
        )

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

```
    PrivNIC ora_privnic (
        Critical = 0
        Device = { en1 = 0, en2 = 1}
        Address@galaxy = "192.168.12.1"
        Address@nebula = "192.168.12.2"
        NetMask = "255.255.240.0"
        )

    cssd requires ocrvote_mnt
    cssd requires ora_priv
    ocrvote_mnt requires ocrvote_voldg
    ocrvote_mnt requires vxfsckd
    ocrvote_voldg requires cvm_clus
    vxfsckd requires cvm_clus
    cvm_clus requires cvm_vxconfigd

group oradb1_grp (
    SystemList = { mercury = 0, jupiter = 1 }
    Parallel = 1
    ClusterList = { rac_cluster101 = 0, rac_cluster102 = 1 }
    OnlineRetryInterval = 300
    ClusterFailOverPolicy = Manual
    Authority = 1
    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
    ora1 requires oradata_mnt
    oradata_mnt requires ora_vvr_sharedpri

group rlogowner (
    SystemList = { mercury = 0, jupiter = 1 }
    AutoStartList = { mercury, jupiter }
    OnlineRetryLimit = 2
    )

    RVGLogowner logowner (
        RVG = rac1_rvg
        DiskGroup = oradatadg
        )

requires group RVGgroup online local firm
    logowner requires logowner_ip
    logowner_ip requires nic

group VxSS (
SystemList = { north = 0, south = 1 }
Parallel = 1
OnlineRetryLimit = 3
OnlineRetryInterval = 120
)
Phantom phantom_vxss (
)
ProcessOnOnly vxatd (
IgnoreArgs = 1
PathName = "/opt/VRTSat/bin/vxatd"
)

group CMC (
SystemList = { north, south }
AutoStartList = { north, south }
OnlineRetryLimit = 3
OnlineRetryInterval = 120
)
ClusterConnectorConfig CMC_ClusterConfig (
MSAddress = "mgmtserver1.symantecexample.com"
MSPort = 14145
ClusterId = "1145613636"
ClusterType = "vcs"
ClusterPort = 14141
VCSLoggingLevel = "TAG_A"
```

```
Logging = "/opt/VRTScmccc/conf/cc_logging.properties"
ClusterConnectorVersion = "5.0.1000.0"
)
Process CMC_ClusterConnector (
PathName = "/bin/sh"
Arguments = "/opt/VRTScmccc/bin/cluster_connector.sh"
)
CMC_ClusterConnector requires CMC_ClusterConfig
```