VERITAS™

VERITAS Cluster Server 4.1 Enterprise Agent for EMC SRDF

Installation and Configuration Guide

HP-UX

N16818G

June 2005

Disclaimer

The information contained in this publication is subject to change without notice. VERITAS Software Corporation makes no warranty of any kind with regard to this manual, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. VERITAS Software Corporation shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance, or use of this manual.

VERITAS Legal Notice

Copyright © 2005 VERITAS Software Corporation. All rights reserved. VERITAS, the VERITAS Logo, and all other VERITAS product names and slogans are trademarks or registered trademarks of VERITAS Software Corporation. VERITAS and the VERITAS logo, Reg. U.S. Pat. & Tm. Off. Other product names and/or slogans mentioned herein may be trademarks or registered trademarks of their respective companies.

VERITAS Software Corporation
350 Ellis Street
Mountain View, CA 94043
USA
Phone 650–527–8000 Fax 650–527–2908
www.veritas.com

Contents

Preface . . . . . **v**
How This Guide is Organized v
VERITAS Cluster Server Documentation vi
Conventions vi
Getting Help vii
Documentation Feedback vii

Chapter 1. Introduction **1**
About the EMC SRDF Agent 1
Supported Software and Hardware 1
Typical Setup 2
Agent Operations 3

Chapter 2. Installing the EMC SRDF Agent **5**

Chapter 3. Configuring the EMC SRDF Agent **7**
Before Configuring the SRDF Agent 7
Resource Type Definition 8
Attribute Definitions 9
Cluster Heartbeats 11
Configuration Concepts 12
Individual Component Failure 12
All Host or All Application Failure 13
Total Site Disaster 14



Replication Link Failure 14
Split-brain 15
Dynamic Swap 15
Configuring the Agent Using the Wizard 16
Configuring the Agent Manually 20
Configuring the Agent in a Global Cluster 20
Configuring the Agent in a Replicated Data Cluster 22

Chapter 4. Managing and Testing Clustering Support for EMC SRDF 23
Service Group Migration 24
Host Failure 25
Disaster Test 26
Failback Test 26
Removing the Agent 27

Chapter 5. Setting Up a Fire Drill 29
Fire Drill Configurations 30
SRDFSnap Agent 31
Agent Operations 31
Resource Type Definition 32
Attribute Definitions 33
Configuring the Snapshot Attributes 34
Sample Configuration 35
Configuring the Fire Drill Service Group 36
Prerequisites 36
Configuration Instructions 37
Verifying a Successful Fire Drill 39

Index 41

Preface

This document describes how to install and configure the VERITAS Cluster Server (VCS) enterprise agent for EMC SRDF on the HP-UX operating system.

If this document is dated more than six months prior to the date you are installing the enterprise agent, contact VERITAS Technical Support to confirm you have the latest supported versions of the application and operating

How This Guide is Organized

- Chapter 1. "Introduction" on page 1 introduces the VCS enterprise agent for EMC SRDF and describes the tasks performed by the agent.
- Chapter 2. "Installing the EMC SRDF Agent" on page 5 provides instructions on installing the EMC SRDF agent.
- Chapter 3. "Configuring the EMC SRDF Agent" on page 7 describes key configuration concepts and provides instructions on configuring the agent.
- Chapter 4. "Managing and Testing Clustering Support for EMC SRDF" on page 23 provides test scenarios and expected outcomes. It also describes how to remove the agent.
- Chapter 5. "Setting Up a Fire Drill" on page 29 describes how you can test the fault-readiness of the disaster recovery environment by running a fire drill.



VERITAS Cluster Server Documentation

The following documents, along with the online help and the Release Notes, comprise the VCS documentation for this release:

Title	File Name
VERITAS Cluster Server Installation Guide	`vcs_install.pdf`
VERITAS Cluster Server User's Guide	`vcs_users.pdf`
VERITAS Cluster Server Bundled Agents Reference Guide	`vcs_bundled_agents.pdf`
VERITAS Cluster Server Agent Developer's Guide	`vcs_agent_dev.pdf`

See the Release Notes for a complete list of documents, including VCS enterprise agent guides.

Conventions

The following conventions apply throughout the documentation set.

Typeface/Font	Usage
bold	names of screens, windows, tabs, dialog boxes, options, buttons
italic	new terms, book titles, emphasis, variables in tables or body text
`Courier`	computer output, command references within text
`Courier` (bold)	command-line user input, keywords in grammar syntax
`Courier` (bold, italic)	variables in a command
#	UNIX superuser prompt (all shells)

Getting Help

For technical assistance, visit http://support.veritas.com and select phone or email support. This site also provides access to resources such as TechNotes, product alerts, software downloads, hardware compatibility lists, and our customer email notification service. Use the Knowledge Base Search feature to access additional product information, including current and past releases of VERITAS documentation.

For license information, software updates and sales contacts, visit https://my.veritas.com/productcenter/ContactVeritas.jsp. For information on purchasing product documentation, visit http://webstore.veritas.com.

Documentation Feedback

Your feedback on product documentation is important to us. Send suggestions for improvements and reports on errors or omissions to clusteringdocs@veritas.com. Include the title and part number of the document (located in the lower left corner of the title page), and chapter and section titles of the text on which you are reporting. Our goal is to ensure customer satisfaction by providing effective, quality documentation. For assistance with topics other than documentation, visit http://support.veritas.com.



Introduction 1

The VCS enterprise agent for EMC SRDF provides failover support and recovery in environments employing SRDF to replicate data between EMC Symmetrix arrays.

About the EMC SRDF Agent

The VCS enterprise agent for EMC SRDF monitors and manages the state of replicated Symmetrix devices attached to VCS nodes. The agent ensures that the system on which the SRDF resource is online has safe exclusive access to the configured devices.

The agent can be used in single VCS replicated data clusters and multi-cluster environments set up using the VCS Global Cluster Option. The agent also supports parallel applications, such as VERITAS Storage Foundation for Oracle RAC.

The agent supports SRDF in the synchronous mode only; the agent does not support semi-synchronous, Adaptive Copy, and SRDF/A.

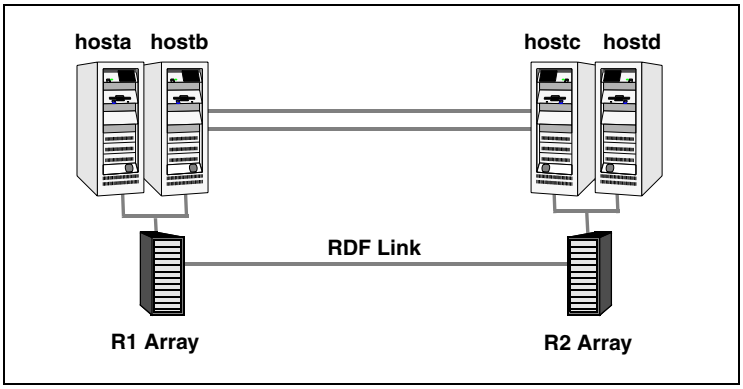
Supported Software and Hardware

The agent supports all versions of SYMCLI, including WideSky. The agent supports SRDF on all microcode levels on all Symmetrix arrays, provided the host/HBA/array combination is in EMC's hardware compatibility list. When using VERITAS Storage Foundation for Oracle RAC, the microcode level of both arrays must be at a level that supports SCSI-III persistent reservations with SRDF devices. Contact EMC for details if necessary.



Typical Setup

Clustering in an SRDF environment typically consists of the following hardware infrastructure:



- ✔ The *R1 array,* comprising one or more *R1 hosts* directly attached via SCSI or Fibre Channel to a Symmetrix array containing SRDF R1 devices.
- ✔ The *R2 array,* comprising one or more *R2 hosts* directly attached via SCSI or Fibre Channel to a Symmetrix array containing SRDF R2 devices. The R2 devices are paired with the R1 devices in the R1 array. These R2 hosts and the array must be at a significant distance from the R1 side to survive a disaster that may occur at the R1 side.
- ✔ Network heartbeats, LLT or TCP/IP, between the two data centers to determine their health. See "Cluster Heartbeats" on page 11 for more information.
- ✔ In a replicated data cluster environment, all hosts are part of the same cluster. You must connect them with dual, dedicated networks that support LLT.
- ✔ In a global cluster environment, you must attach all hosts in a cluster to the same Symmetrix array.
- ✔ In parallel applications such as VERITAS Storage Foundation for Oracle RAC, all hosts attached to the same array must be exclusively part of the same GAB membership. VERITAS Storage Foundation for Oracle RAC is supported with SRDF only in a global cluster environment and not in a replicated data cluster environment.

Agent Operations

Operation (Entry Point)	Description
online	If the state of all local devices is read-write enabled (RW), the agent creates a lock file on the local host to indicate that the resource is online. This makes the devices writable for the application. If one or more devices are write-disabled (WD), the agent runs a `symrdf` command to enable read-write access to the devices. ◆ For R2 devices in the SYNCHRONIZED state, the agent runs the `symrdf failover` command to make the devices writable. ◆ For R1 devices in the FAILED OVER or R1 UPDATED state, the agent runs the `symrdf failback` command to make the devices writable. ◆ For all devices in the PARTITIONED state, the agent runs the `symrdf rw_enable` command to make the devices writable. The agent runs the command only if the AutoTakeover attribute is set to 1 and if there are no dirty tracks on the local device. Dirty tracks indicate that an out-of-order synchronization was in progress when the devices became partitioned, rendering them inconsistent and unusable. If dirty tracks exist, the online entry point faults on timeout. ◆ For R1 devices in the UPDINPROG state, the agents waits for the devices to transition to the R1 UPDATED state before running a `symrdf` command. ◆ For R2 devices in the SYNCINPROG state, the agent waits for the devices to transition to the SYNCHRONIZED state before running a `symrdf` command. **Note** The agent does not run any commands if it detects that there is not enough time remaining for the entry point to complete the command. See "To set the OnlineTimeout attribute" on page 13 for more information.
offline	Removes the lock file on the device. The agent does not run any SRDF commands because taking the resource offline is not indicative of the intention to give up the devices.
monitor	Verifies that the lock file exists. If the lock file exists, the monitor entry point reports the status of the resource as online. If the lock file does not exist, the monitor entry point reports the status of the resource as offline.
open	Removes the lock file on the host where the entry point is called. This prevents potential concurrency violation if the service group fails over to another node. **Note** The agent does not remove the lock file if the agent was started after an `hastop -force` command.



Operation (Entry Point)	Description
clean	Determines whether if it is safe to fault the resource if the online entry point fails or times out. The main consideration is whether a management operation was in progress when the online thread timed out and was killed, potentially leaving the devices in an unusable state.
info	Reports the device state to the VCS interface. This entry point can be used to verify the device state and to monitor dirty track trends.
action	Performs a `symrdf update` from the R2 side to merge any dirty tracks from the R2 to the R1.

Installing the EMC SRDF Agent 2

You must install the SRDF enterprise agent on each node in the cluster. In global cluster environments, install the agent on each node in each cluster.

▼ To install the agent

1. Insert the disc into a drive connected to the host.
2. Create a mount point directory, `/cdrom`, if it does not exist. The directory must have read-write permissions.
3. Determine the block device file for the disc drive:

   ```
   # ioscan -fnC disk
   ```

   For example, the listing may indicate the block device is `/dev/dsk/c1t2d0`.
4. Start the Portable File System (PFS).

   ```
   # nohup pfs_mountd &
   # nohup pfsd &
   ```
5. Mount the disc:

   ```
   # /usr/sbin/pfs_mount -t rrip /dev/dsk/c#t#d# /cdrom
   ```

   The variable */c#t#d#* represents the location of the drive.
6. Install the software:

   ```
   # swinstall -s /cdrom/depot VRTSvcse
   # swinstall -s /cdrom/depot VRTScsecw
   # swinstall -s /cdrom/depot VRTScsfdw
   ```




Configuring the EMC SRDF Agent 3

Most applications configured in VCS can be adapted to a disaster recovery environment by:

- Converting their devices to SRDF devices
- Synchronizing the devices
- Adding the EMC SRDF agent to the service group

Volumes of Symmetrix device groups are configured as resources of type SRDF.

Before Configuring the SRDF Agent

- ✔ Verify the agent is installed on all systems in the cluster.
- ✔ Verify the hardware setup for the agent. See "Typical Setup" on page 2 for more information.
- ✔ Make sure the cluster has an effective heartbeat mechanism in place. See "Cluster Heartbeats" on page 11 for more information.
- ✔ Review the agent's resource type definition and its attribute definitions.
- ✔ Review the section "Configuration Concepts" on page 12. This section presents information about how VCS behaves during failover and how you can set attributes to customize VCS behavior.



Resource Type Definition

```
type SRDF (
  static str ArgList[] = { SymHome, GrpName, DevFOTime,
  AutoTakeover }
    static int NumThreads = 1
    static int ActionTimeout = 180
    static int OfflineMonitorInterval = 0
    static int MonitorInterval = 300
    static int RestartLimit = 1
    static keylist SupportedActions = { update }
    NameRule = resource.GrpName
    str SymHome = "/usr/symcli"
    str GrpName
    int DevFOTime = 2
    int AutoTakeover = 1
    temp str VCSResLock
)
```

Attribute Definitions

Required Attribute	Type-Dimension	Description
GrpName	string-scalar	Name of the Symmetrix Device Group managed by the agent.

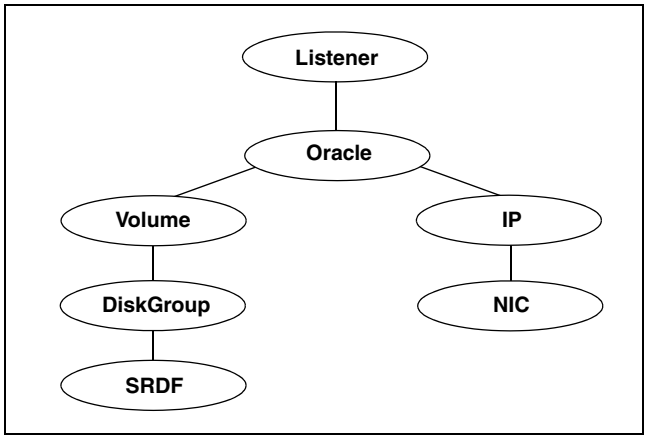
Optional Attributes	Type-Dimension	Description
SymHome	string-scalar	Path to the Symmetrix command line interface. Default is `/usr/symcli`.
DevFOTime	integer-scalar	Average time in seconds required for each device in the group to fail over. This value helps the agent to determine whether there is adequate time for the online operation to complete after waiting for other device groups to fail over. If the online operation cannot be completed in the remaining time, the failover does not proceed. See "All Host or All Application Failure" on page 13 for information on failover serialization and the recommended VCS restart settings. Default is 2 seconds per device.
AutoTakeover	integer-scalar	A flag that determines whether the agent performs a read-write enable on write-disabled partitioned devices during a failover. Default is 1, which means that the agent will perform a read-write enable if devices are consistent.

Internal Attribute	Type-Dimension	Description
VCSResLock	temporary string	This attribute is used by the agent to guarantee serialized management in case of a parallel application. *Do not modify this value.*



Sample Configuration

The following dependency graph shows a VCS service group that has a resource of type SRDF. The DiskGroup resource depends on the SRDF resource.



A resource of type SRDF may be configured as follows in `main.cf`:

```
SRDF oradf_rdf (
   GrpName = "oracle_grp"
)
```

Cluster Heartbeats

In a replicated data cluster, robust heartbeating is accomplished through dual, dedicated networks over which the Low Latency Transport (LLT) runs. Additionally, you can configure a low-priority heartbeat across public networks.

In a global cluster, network heartbeating is accomplished by sending ICMP pings over the public network between the two sites. VCS global clusters minimize the risk of split-brain by sending ICMP pings to highly available IP addresses and by notifying administrators when the sites cannot communicate.

In global clusters, the VCS Heartbeat agent sends heartbeats directly between the Symmetrix arrays, given the Symmetrix ID of each array. This heartbeat offers the following advantages:

- VCS does not mistakenly interpret the loss of ICMP heartbeats over the public network as a site failure because the Symmetrix heartbeat shows that the arrays are alive.
- If the loss of heartbeats occurs due to the failure of all hosts in the primary cluster, a failover may be required even if the array is alive. In any case, it is important to distinguish between a host-only crash and a complete site failure. In a host-only crash, only the ICMP heartbeat signals a failure via an SNMP trap. No cluster failure notification occurs because a surviving heartbeat exists. This trap is the only notification to fail over an application.
- The heartbeat is then managed completely by VCS and reports as being down only when the remote array is not visible by the `symrdf ping` command.

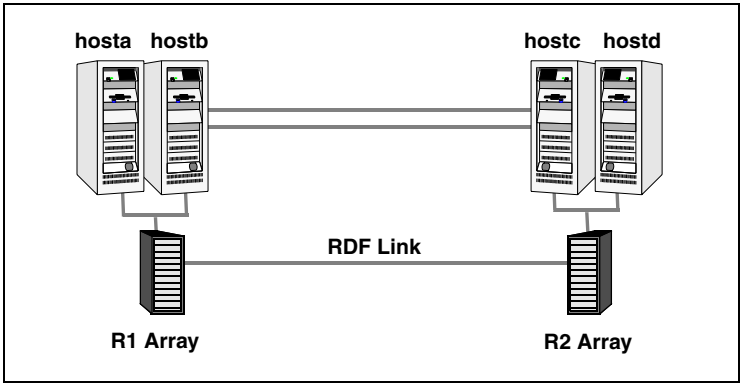


Configuration Concepts

This section describes some failure scenarios and provides guidelines on how to configure the agent.

Individual Component Failure

In a replicated data cluster, you can prevent unnecessary SRDF failover or failback by configuring hosts attached to an array as part of the same system zone. VCS attempts to fail over applications within the same system zone before failing them over across system zones.



In this sample, hosta and hostb are in one system zone, and hostc and hostd are in another system zone. The SystemZones attribute enables you to create these zones. You can modify the SystemZones attribute using the following command:

```
# hagrp -modify grpname SystemZones hosta 0 hostb 0 hostc 1 hostd 1
```

The variable *grpname* represents the service group in the cluster.

This command creates two system zones: zone 0 with hosta and hostb, zone 1 with hostc and hostd.

Global clusters do not require system zones because failover occurs on a remote cluster if all local targets have been exhausted.

While running on R2 devices, SRDF does not synchronize data back to the R1 automatically. You must update out-of-synch tracks manually. Monitor the number of out-of-synch tracks by viewing the ResourceInfo attribute of an online SRDF resource. If the value is too high, update tracks to the R1 using the update action, which is defined as a supported action in the SRDF resource type.

All Host or All Application Failure

If all hosts on the R1 side are disabled or if the application cannot start successfully on any R1 hosts, but both arrays are operational, the service group fails over.

In replicated data cluster environments, the failover can be automatic, whereas in global cluster environments, failover requires user confirmation by default. In both environments, multiple device groups may fail over in parallel. VCS serializes `symrdf` commands to ensure that SRDF does not lock out a command while another command is running.

Set the OnlineTimeout and RestartLimit attributes for the SRDF resource to make sure that its entry points do not time out, or that they are automatically restarted if they time out.

▼ **To set the OnlineTimeout attribute**

Use the following formula to calculate an appropriate value for the OnlineTimeout attribute:

$$\text{OnlineTimeout} = \sum_{1}^{n_{devicegroups}} ((n_{devices} \times d_{failovertime}) + \varepsilon)$$

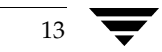
- $n_{devices}$ represents the number of devices in a device group.
- $d_{failovertime}$ represents the value of the DevFOTime attribute.
- $n_{devicegroups}$ represents the total number of device groups managed by VCS that might fail over simultaneously.
- The additional epsilon is for the command instantiation overhead.

If the resulting value seems excessive, divide it by two for every increment in the value of the RestartLimit. However, the OnlineTimeout must be at least the time taken for the largest device group to fail over, otherwise the group will never be able to complete its operation, regardless of the value of RestartLimit.

Run the perl script `/opt/VRTSvcs/bin/SRDF/sigma` to get recommendations for VCS attribute values.

Run the script on a node where VCS is running and has the SRDF agent configured. Note that the sigma calculator adds 10 seconds to the value for each device group to compensate for the overhead of launching an `symrdf` command. Specify another value to the sigma script if you feel the instantiation takes shorter or longer.

The script assumes that all devices in the Symmetrix array are managed by VCS. Other operations outside of VCS that hold the array lock might delay the online operation unexpectedly.



Total Site Disaster

In a total site failure, all hosts and the Symmetrix array are completely disabled.

In a replicated data cluster, VCS detects site failure and total host failure by the loss of all LLT heartbeats.

In a global cluster, VCS detects site failure by the loss of both the ICMP and Symm heartbeats. In order not to confuse a site failure with an all-host failure, the AYARetryLimit for the Symm heartbeat must be shorter than the ICMP retry limit, so that the failure of the Symmetrix array is detected first.

A total disaster renders the devices on the surviving array in the PARTITIONED state. If the AutoTakeover attribute is set to its default value of 1, the online entry point runs the `symrdf_rw` command. If the attribute is set to 0, no takeover occurs and the online entry point times out and faults.

The online entry point detects whether any synchronization was in progress when the source array was lost. Since synchronization renders the target SRDF devices inconsistent until the synchronization completes, write-enabling the devices would be futile since the data stored on them is unusable. In this case, the agent does not enable the devices and instead times out and faults. In such a scenario, you must restore consistent data from a BCV or tape backup.

Replication Link Failure

SRDF detects link failures, monitors changed tracks on devices, and resynchronizes R2 devices if the R1 was active at the time of the link failure.

If the two arrays are healthy and the link fails and is restored, and if a failover is initiated while one or more devices are in the SYNCINPROG state, the SRDF agent waits for the synchronization to complete before running the `symrdf failover` command. If the agent times out before the synchronization completes, the resource faults.

If a failover is initiated due to a disaster at the R1 site, and if a synchronization was in progress, the R2 devices are rendered inconsistent and unusable. In this case, even if the AutoTakeover attribute of the agent is set to 1, the agent does not enable read-write access to the devices and instead it faults. You must restore consistent data to these devices, either from BCV or from a tape backup, and then enable read-write access to the devices manually before they can be used.

If the AutoTakeover attribute is set to 0, the agent does not attempt a `symrdf rw_enable`, but it times out and faults. If you write-enable the devices manually, the agent can come online after it is cleared.

Split-brain

Split-brain occurs when all heartbeat links between the R1 and R2 hosts are cut and each side mistakenly thinks the other side is down. To minimize the effects of split-brain, it is best if the cluster heartbeat links pass through similar physical infrastructure as the replication links so that if one breaks, so does the other.

In a replicated data cluster, VCS attempts to start the application assuming a total disaster because the R1 hosts and array are unreachable. Once the heartbeats are restored, VCS stops the applications on one side and restarts the VCS engine (HAD) to eliminate concurrency violation of the same group being online at two places simultaneously. You must resynchronize the volumes manually using the `symrdf merge` or `symrdf restore` commands.

In a global cluster, you can confirm the failure before failing over the service groups. You can check with the site administrator to identify the cause of the failure. If you do mistakenly fail over, the situation is similar to the replicated data cluster case; however, when the heartbeat is restored, VCS does not stop HAD at either site. VCS forces you to choose which group to take offline. You must resynchronize data manually.

If it is physically impossible to place the heartbeats alongside the replication links, there is a possibility that the cluster heartbeats are disabled, but the replication link is not. A failover transitions the original R2 volumes to R1 volumes and vice-versa. In this case, the application faults because its underlying volumes become write-disabled. VCS tries to fail the application over to another host, causing the same consequence in the reverse direction. This phenomenon continues until the group comes online on the final node. This situation can be avoided by setting up your infrastructure such that the loss of heartbeat links also means the loss of replication links.

Dynamic Swap

The agent supports the SRDF dynamic swap capability. If all devices are configured for dynamic swap, when a service group fails over between arrays, the agent performs a swap if both arrays are healthy. If one array is down, a unilateral read-write enable occurs. The agent fails over device groups that are not configured for dynamic swap using the `symrdf` failover command, which read-write enables the R2.

The agent checks the following criteria before determining if a swap will occur:

- ✔ All devices in the device group are configured as dynamic devices.
- ✔ Dynamic RDF is configured on the local Symmetrix array.
- ✔ The SYMCLI version is 5.4 or above.
- ✔ The microcode is level 5567 or above.

Dynamic swap does not affect the ability to perform fire drills.



Configuring the Agent Using the Wizard

This section describes how to use the wizard to configure the SRDF agent in an application service group.

1. Run the wizard on a system attached to the Symmetrix array. Make sure the command line package (SYMCLI) for the Symmetrix array is installed on the system.

2. Set the *DISPLAY* variable and start the SRDF Configuration wizard as `root`.

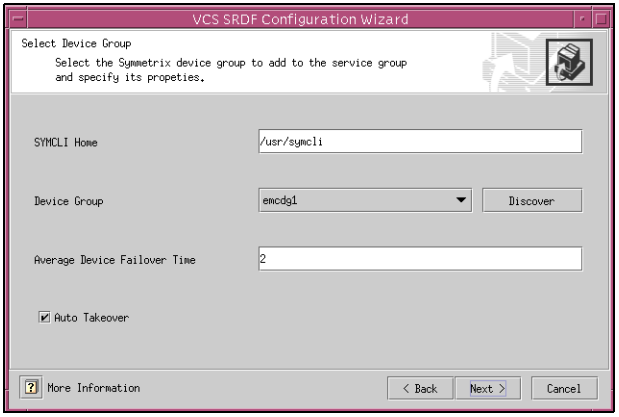
   ```
   # hawizard srdf
   ```

3. Read the information on the Welcome screen and click **Next**.

4. In the Wizard Options dialog box, select the application service group to which you want to add an SRDF resource.

   Note The wizard displays service groups having disk group resources; it does not display service groups having SRDF resources.

5. In the Select Device Group dialog box, specify the device group from the Symmetrix array for which the SRDF resource is to be added.



   a. In the **SYMCLI Home** field, specify the path where the CLI package for the Symmetrix array is installed. The default location is `/usr/symcli`.

   b. Select the device group to monitored. If the device group does not appear in the list, click **Discover**.

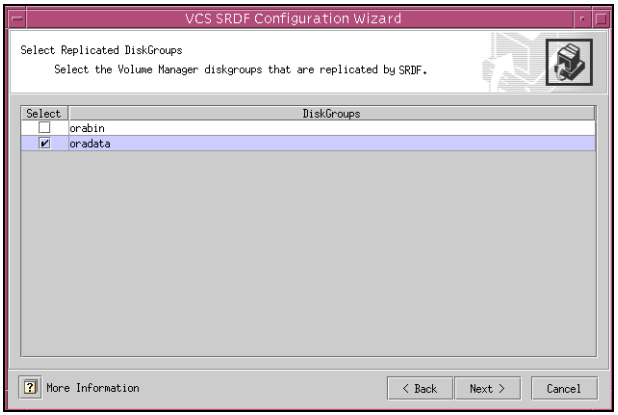
   c. In the **Average Device Failover Time** field, specify the average time in seconds for each device group in the service group to fail over. Default is 2 seconds per device.

   d. Select the **Auto Takeover** check box if you want the SRDF resource to perform a read-write enable on write-disabled partitioned devices during a failover.

   e. Click **Next**.

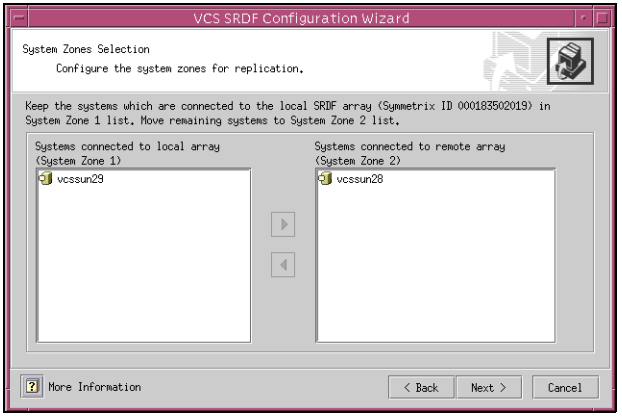


6. Select the replicated diskgroups and click **Next**.



If you are adding an SRDF resource in a service group configured in a replicated data cluster, proceed to the next step. Otherwise, proceed to step 8 on page 19.

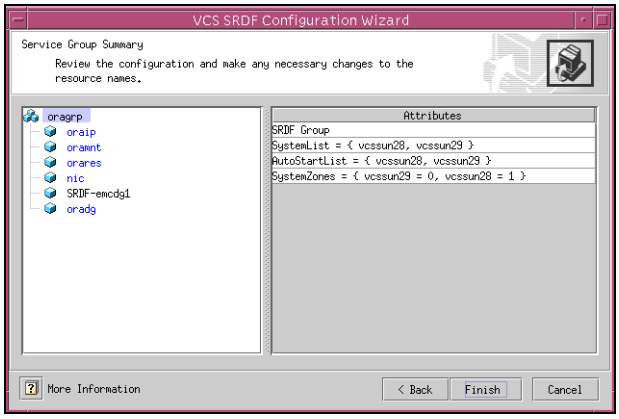
7. In the System Zones Selection dialog box, specify the systems for each zone in a replicated data cluster.



   a. If you had configured SystemZones in the application service group, verify the configuration. Use the arrows to move systems to their respective zones.

   b. Click **Next**.

8. In the Service Group Summary dialog box, review the service group configuration and change the name of the SRDF resource, if desired.



   a. To change the name of the SRDF resource, select the resource name and either click it or press the F2 key. Press Enter after editing the resource name. To cancel editing a resource name, press Esc.

   b. Click **Finish**.

   The wizard starts running commands to add the SRDF resource to the service group. Various messages indicate the status of these commands.

9. In the Completing the SRDF Configuration Wizard dialog box, select the check box to bring the service group online on the local system.

10. Click **Close**.



Configuring the Agent Manually

This section describes how to configure the agent using the Java Console in global and replicated data clusters.

Configuring the Agent in a Global Cluster

1. If the SRDF resource type is not added to your configuration, add it.

   a. Start Cluster Manager and log on to the cluster.

   b. From the Cluster Explorer **File** menu, choose **Import Types** and select `/etc/VRTSvcs/conf/SRDFTypes.cf`.

   c. Click **Import**.

   d. Save the configuration.

Note You can also add the resource type by running the following command: **`/etc/VRTSvcs/conf/sample_srdf/addSRDFType.sh`.**

2. Configure the Symm heartbeat at each cluster:

   a. From Cluster Explorer **Edit** menu, choose **Configure Heartbeats**.

   b. On the Heartbeats Configuration dialog box, enter the name of the heartbeat.

   c. Select the check box next to the name of the cluster to add it to the cluster list for the heartbeat.

   d. Click the icon in the **Configure** column to open the **Heartbeat Settings** dialog box.

   e. Specify as the value of the Arguments attribute the Symmetrix ID of the array in the other cluster. Set the value of the AYARetryLimit attribute for this heartbeat to 1 less than the value for the ICMP heartbeat.

   f. Click **OK**.

3. Perform the following tasks for each service group in each cluster that uses replicated data:

   a. Add a resource of type SRDF at the bottom of the service group. See "Sample Configuration" on page 10 for more information.

   b. Configure the attributes of the SRDF resource. See "Attribute Definitions" on page 9 for more information.

   c. If the service group is not configured as a global group, configure the service group using the Global Group Configuration Wizard. See the *VERITAS Cluster Server User's Guide* for more information.

   d. Change the ClusterFailOverPolicy from the default, if necessary. VERITAS recommends keeping the default, which is Manual, to minimize the chance of failing over on a split-brain.

   e. To configure the agent to manage volumes used by VERITAS Storage Foundation for Oracle RAC, configure the SupportedActions attribute for the CVMVolDg resource and add *import* and *deport* as keys to the list. Note that SupportedActions is a resource type attribute and defines a list of action tokens for the resource.

      The agent supports importing and deporting a VERITAS Volume Manager diskgroup when failing over Real Application Clusters across replicating arrays. Failing to do so might leave disk groups imported on hosts where the storage is read-only. While this is not an error, any attempted writes to the disk group will be rejected, causing the disk group to be disabled. You must deport and reimport the disk group to enable it.



Configuring the Agent in a Replicated Data Cluster

1. If the SRDF resource type is not added to your configuration, add it.

   a. Start Cluster Manager and log on to the cluster.

   b. From the Cluster Explorer **File** menu, choose **Import Types** and select `/etc/VRTSvcs/conf/SRDFTypes.cf`.

   c. Click **Import**.

   d. Save the configuration.

Note You can also add the resource type by running the following command: **`/etc/VRTSvcs/conf/sample_srdf/addSRDFType.sh`.**

2. Perform the following tasks for each service group that uses SRDF replicated data:

   a. Add a resource of type SRDF at the bottom of the service group.

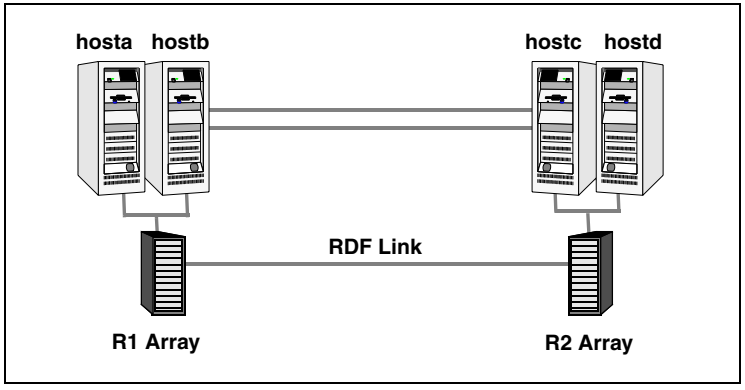
   b. Configure the attributes of the SRDF resource. See "Attribute Definitions" on page 9 for more information about these attributes. Note that some attributes must be localized to reflect values for hosts attached to different Symmetrix arrays.

   c. Set the SystemZones attribute for the service group to reflect which hosts are attached to the same array. See "Individual Component Failure" on page 12 for more information.

Managing and Testing Clustering Support for EMC SRDF 4

After configuring the SRDF agent in a VCS environment, you can perform some basic tests to verify the implementation. This chapter describes some test scenarios and expected behavior.

These tests assume the following environment:



Two hosts (hosta and hostb) are attached to the R1 array, and the other hosts are attached to the R2 array. The application is running on hosta and devices in the local array are read-write enabled, in the SYNCHRONIZED state.

A replicated data cluster has two dedicated heartbeat links; a global cluster has one network heartbeat and an optional SRDF replication link heartbeat. The test scenario is similar for both environments.



Service Group Migration

Verify the service group can migrate to different hosts in the cluster.

▼ To perform the service group migration test

1. Migrate the service group to a host attached to the same array.

   a. In the **Service Groups** tab of the Cluster Explorer configuration tree, right-click the service group.

   b. Click **Switch To**, and click the system attached to the same array (hostb) from the menu.

   The service group comes online on hostb and local volumes remain in the RW/SYNCHRONIZED state.

2. Migrate the service group to a host attached to a different array:

   a. In the **Service Groups** tab of the Cluster Explorer configuration tree, right-click the service group.

   b. Click **Switch To**, and click the system attached to the another array (hostc) from the menu.

   The service group comes online on hostc and volumes there transition to the RW/FAILED OVER state.

3. Accumulate dirty tracks on the R2 side and update them back on the R1:

   ```
   # hares -action srdf_res_name update -sys hostc
   ```

   The variable *srdf_res_name* represents the name of the SRDF resource.

4. After the devices transition to R1 UPDATED state, migrate the service group back to its original host:

   a. In the **Service Groups** tab of the Cluster Explorer configuration tree, right-click the service group.

   b. Click **Switch To**, and click the system on which the group was initially online (hosta).

   The group comes online on hosta. The devices return to the RW/SYNCINPROG state at the array attached to hosta and hostb, and then eventually transition to the SYNCHRONIZED state.

Host Failure

In this scenario, the host on which the application is running is lost and eventually all hosts in the system zone or cluster are lost.

▼ To perform the host failure test

1. Halt or shut down the host on which the application is running.

   The service group fails over to hostb and devices are in the RW/SYNCHRONIZED state.

2. Halt or shut down hostb.

   In a replicated data cluster, the group fails over to hostc or hostd depending on the FailOverPolicy in the cluster.

   In a global cluster, a cluster down alert appears and gives you the opportunity to fail over the service group manually.

   In both environments, the devices transition to the RW/FAILED OVER state and start on the target host.

3. Reboot the two hosts that were shut down.

4. Switch the service group to its original host when VCS starts.

   a. In the **Service Groups** tab of the Cluster Explorer configuration tree, right-click the service group.

   b. Click **Switch To**, and click the system on which the service group was initially online (hosta).

   The service group comes online on hosta and devices transition to the SYNCINPROG state and then to the SYNCHRONIZED state.



Disaster Test

Shut down all hosts on the source side and shut down the source array. If shutting down the R1 Symmetrix is not feasible, disconnect the ESCON link between the two arrays while simultaneously shutting down the hosts; this action mimics a disaster scenario from the point of view of the R2 side.

In a replicated data cluster, the service group fails over to hostc or hostd if all devices were originally SYNCHRONIZED, that is, no synchronization was in progress at the time of disaster.

In a global cluster, the administrator is notified of the failure. The administrator can then initiate the failover by declaring an outage.

Failback Test

▼ **To perform the failback test**

1. Reconnect the ESCON cable and reboot the original R1 hosts. You must manually resynchronize the device, which can be done only if both sides are write-disabled. This requires taking the service group offline.

2. If you are running this test in a replicated data cluster, run the following command from any host:

   ```
   # hagrp -offline grpname -any
   ```

   If you are running the test in a global cluster, run the command from hostc or hostd.

3. After the service group goes offline, run the following command:

   ```
   # symrdf -g device_group restore
   ```

   The variable *device_group* represents the name of the RDF device group at the R2 side. The restore command determines which tracks to merge between the R1 and R2 arrays and initiates the resynchronization. The operation of this command write disables both sides; use this command only when a brief downtime is acceptable.

4. Bring the service group online at the R1 side:

   ```
   # hagrp -online grpname -sys hosta
   ```

The devices synchronize, and the environment state will be the same as when the test began.

Removing the Agent

Type the following command on each system to remove the agent. Answer prompts accordingly:

```
# swremove VRTSvcse
# swremove VRTScsecw
# swremove VRTScsfdw
```




Setting Up a Fire Drill 5

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

A fire drill is performed at the secondary site. The initial steps involve configuring a service group identical to the application service group and replacing the SRDF resource with the SRDFSnap resource, a fire drill resource. The fire drill service group uses a copy of the data used by the application service group.

Bringing the fire drill service group online demonstrates the ability of the application service group to come online at the remote site when a failover occurs.

VCS supports several fire drill configurations and provides the SRDFSnap agent to manage the replication relationships during a fire drill. The Fire Drill Configuration wizard configures the fire drill service group.



Fire Drill Configurations

VCS supports the following fire drill configurations:

Fire Drill Configuration	Description
Gold	Runs the fire drill on a snapshot of the target array. Involves the following steps: ◆ Suspend replication to get a consistent snapshot. ◆ Take a snapshot of the target array on a BCV device. ◆ Modify the disk name and the disk group name in the snapshot. ◆ Bring the fire drill service group online using the snapshot data. **Note** For the Gold configuration, you must use VERITAS Volume Manager to import and deport the storage. For non-replicated devices: ◆ You must use VERITAS Volume Manager ◆ You must use the Gold configuration without the option to run in the Bronze mode.
Silver	Runs the fire drill on the target array after taking a snapshot. Involves the following steps: ◆ Suspend replication to get a consistent snapshot. ◆ Take a snapshot of the target array on a BCV device. ◆ Bring the fire drill service group online using the data on the target array.
Bronze	Runs the fire drill on the target array. No snapshots are taken. Involves the following steps: ◆ Suspend replication to get a consistent snapshot. ◆ Bring the fire drill service group online using the data on the 'target array.

SRDFSnap Agent

The SRDFSnap agent manages the replication relationship between the source and target arrays when running a fire drill. The agent is configured in the fire drill service group.

Agent Operations

Operation (Entry Point)	Description
Online	• **Gold Configuration** Suspends replication between the source and the target arrays, takes a local snapshot of the target LUN, resumes the replication between the arrays, and takes the fire drill service group online by mounting the snapshot. • **Silver Connfiguration** Suspends replication between the source and the target arrays, takes a local snapshot of the target LUN, and takes the fire drill service group online by mounting the target LUN. • **Bronze Configuration** Suspends replication between the source and the target arrays and takes the fire drill service group online using the target array. The operation also creates a lock file to indicate that the resource is online.
Offline	• **Gold Configuration** Destroys the snapshot by synchronizing data between the target array and the device on which snapshot was taken. • **Silver Configuration** Resumes replication between the source and the target arrays. Once the data is synchronized between the two arrays, the snapshot of the target array is destroyed by synchronizing data between the target array and the device where snapshot was taken. • **Bronze Configuration** Resumes the replication between the source and the target arrays. The operation also removes the lock file created by the Online operation.
Monitor	Verifies the existence of the lock file to make sure the resource is online.
Clean	Restores the state of the LUNs to their original state after a failed Online operation.
Action	For internal use.



Resource Type Definition

```
type SRDFSnap (
  static str ArgList[] = { TargetResName, MountSnapshot, UseSnapshot,
                           RequireSnapshot }
  static keylist RegList = { MountSnapshot, UseSnapshot }
  static int NumThreads = 1
  str TargetResName
  int MountSnapshot
  int UseSnapshot
  int RequireSnapshot
  temp str Responsibility
  temp str FDFile
)
```

Attribute Definitions

Required Attributes	Type-Dimension	Description
TargetResName	string-scalar	Name of the resource managing the LUNs to be snapshot. The target resource is of type SRDF if the data being snapshot is replicated; the resource is of type DiskGroup if the data is not replicated. For example, some applications like Oracle have data files and redo logs replicated, but temporary tablespace not replicated. The temporary tablespace must still exist at the DR site and may be part of its own disk group and is snapshot independently.
UseSnapshot	integer-scalar	Specifies whether the SRDFSnap resource takes a local snapshot of the target array. Set this attribute to 1 for Gold and Silver configurations. For Bronze, set this attribute to 0. See "Configuring the Snapshot Attributes" on page 34.
RequireSnapshot	integer-scalar	Specifies whether the SRDFSnap resource must take a snapshot before coming online. Set this attribute to 1 if you want the resource to come online only after it succeeds in taking a snapshot. Set this attribute to 0 if you do want the resource to come online even if it fails to take a snapshot. Setting this attribute to 0 creates the Bronze configuration. **Note** Set this attribute to 1 only if UseSnapshot is set to 1.
MountSnapshot	integer-scalar	Specifies whether the resource uses the snapshot to bring the service group online. Set this attribute to 1 for Gold configuration. For Silver and Bronze configurations, set the attribute to 0. **Note** Set this attribute to 1 only if UseSnapshot is set to 1.



Internal Attributes	Type-Dimension	Description
Responsibility	temporary string	For internal use only. Used by the agent to keep track of resynchonizing snapshots.
FDFile	temporary string	For internal use only. Used by the agent to locate the latest fire drill report.

Configuring the Snapshot Attributes

The UseSnapshot, MountSnapshot, and RequireSnapshot attributes define the fire drill configuration.

Attribute	Gold	Silver	Bronze
MountSnapshot	1	0	0
UseSnapshot	1	1	0

Setting the RequireSnapshot attribute to 0 enables a Gold or Silver configuration to run in the Bronze mode if the snapshot operation fails.

Sample Configuration

The sample configuration of a fire drill service group is identical to an application service group with a hardware replication resource. However, in a fire drill service group, the SRDF resource is replaced by the fire drill resource SRDFSnap.

The following configuration creates a Gold fire drill configuration, but allows VCS to run a Bronze fire drill if the snapshot does not complete successfully.

A resource of type SRDFSnap may be configured as follows in the main.cf:

```
SRDFSnap oradg_fd {
   TargetResName = "oradf_rdf"
   UseSnapshot = 1
   RequireSnapshot = 0
   MountSnapshot = 1
}
```




Configuring the Fire Drill Service Group

This section describes how to configure a fire drill service group using the Fire Drill Configuration wizard. Note that you can also use the text-based wizard, available at `/opt/VRTSvcs/bin/fdsetup-srdf`.

Prerequisites

- ✔ Make sure the application service group is configured with an SRDF resource.
- ✔ Make sure the infrastructure to take snapshots is properly configured between the source and target arrays. This involves associating and BCVs and synchronizing them with the source.
- ✔ When using Gold or Silver configuration, make sure you have TimeFinder for SRDF installed and configured at the target array.
- ✔ For the Gold configuration, you must use VERITAS Volume Manager to import and deport the storage.
- ✔ When taking snapshots of R2 devices, BCV's must be associated with the RDF2 device group and fully established with the devices.
- ✔ When taking snapshots of non-replicated devices, create a Symmetrix device group with the same name as the VxVM disk group, with the same devices as in the VxVM disk group and with BCVs associated.
- ✔ For non-replicated devices:
  - You must use VERITAS Volume Manager
  - You must use the Gold configuration without the option to run in the Bronze mode. This means the RequireSnapshot attribute must be set to 1.

Configuration Instructions

1. Set the *DISPLAY* variable and start the Fire Drill Configuration wizard as `root`.

   ```
   # hawizard firedrill
   ```

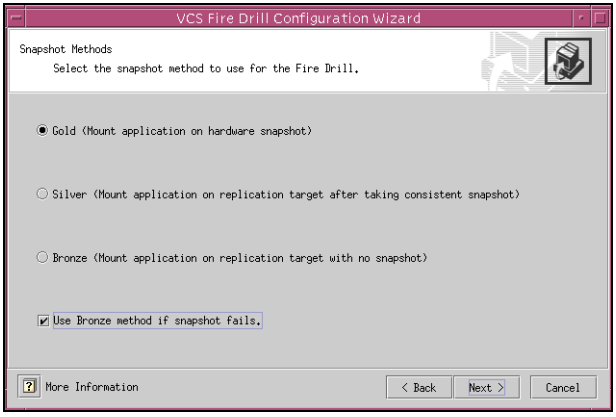
2. Read the information on the Welcome screen and click **Next**.

3. In the Wizard Options dialog box, select the application service group for which a fire drill service group is being configured.

   Note The wizard does not display service groups that do not have SRDF resources configured.

4. Verify the information presented in the Device Group Details dialog box and click **Next**.

5. In the Snapshot Methods dialog box, choose the configuration option for the fire drill service group.



   a. Choose either a **Gold**, **Silver**, or **Bronze** configuration option. For more information about these configuration options, refer to "Fire Drill Configurations" on page 30.

   b. Select the **Use Bronze method if snapshot fails** check box if you want the fire drill service group to come online even if the resource fails to take a snapshot. This check box is enabled only if you choose the Gold or Silver configuration.

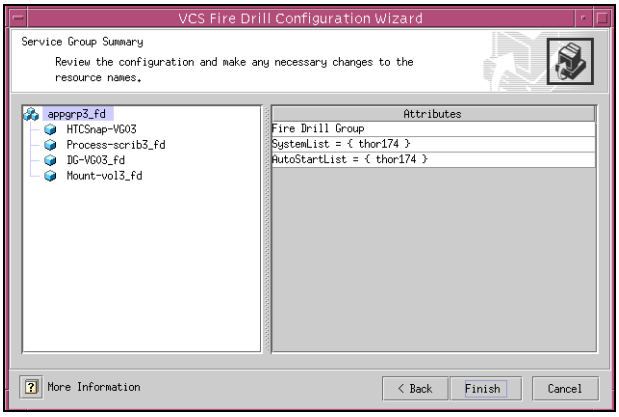
   c. Click **Next**.



6. In the Snapshot Details dialog box, the wizard informs whether the device group on the target array has synchronized BCV devices to take a snapshot. If the devices are synchronized, click **Next**.

   If the devices are not synchronized, quit the wizard, synchronize data between the target array and the BCV device, and rerun the wizard.

7. In the Service Group Summary dialog box, review the service group configuration and change the resource names if desired.



   a. To edit a resource name, select the resource name and either click it or press the F2 key. Press Enter after editing each resource name. To cancel editing a resource name, press Esc.

   b. Click **Finish**.

   The wizard starts running commands to create the fire drill service group. Various messages indicate the status of these commands.

8. In the Completing the Fire Drill Configuration Wizard dialog box, select the check box to bring the service group online on the local system.

9. Click **Close**.

Verifying a Successful Fire Drill

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

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. You can also view the fire drill log, located at `/tmp/fd-`*`servicegroup`*`.`

Caution Remember to take the fire drill offline once its functioning has been validated. Failing to take the fire drill offline could cause failures in your environment. For example, if the application service group were to fail over to the node hosting the fire drill service group, there would be resource conflicts, resulting in both service groups faulting.



Index

A

action entry point 4
agent operations 3
attribute definitions 9
AutoTakeover attribute 9

C

clean entry point 4

D

DevFOTime attribute 9
disaster test 26

E

EMC SRDF agent
  about 1
  attribute definitions 9
  configuration concepts 12
  configuring in a global cluster 20
  configuring in a replicated data cluster 20
  configuring using wizard 16
  installing 5
  operations 3
  removing 27
  testing 23
  type definition 8
EMC SRDF agent attributes
  AutoTakeover 9
  DevFOTime 9
  GrpName 9
  SymHome 9
  VCSResLock 9
entry points
  action 4
  clean 4
  monitor 3
  offline 3
  online 3
  open 3

F

failback Test 26
FDFile attribute 34
fire drill
  about 29
  configuration wizard 36
  running 39
  service group for 36
  SRDFSnap agent 31
  supported configurations 30

G

global cluster configuration 20
GrpName attribute 9

H

heartbeats 11

M

monitor entry point 3
MountSnapshot attribute 33

O

offline entry point 3
online entry point 3
OnlineTimeout attribute, setting 13
open entry point 3
operations 3

R

RDC configuration 20
RequireSnapshot attribute 33
resource type definition
  EMC SRDF agent 8
  SRDFSnap agent 32
Responsibility attribute 34

S

sample configuration 10
split-brain, handling in cluster 15
SRDF service group, migrating 24



SRDFSnap agent
about 31
attribute definitions 33
operations 31
type definition 32
SRDFSnap agent attributes
FDFile 34
MountSnapshot 33
RequireSnapshot 33
Responsibility 34
TargetResName 33
UseSnapshot 33
supported hardware 1
supported software 1
SymHome attribute 9

T

TargetResName attribute 33
type definition
EMC SRDF agent 8
SRDFSnap agent 32

U

UseSnapshot attribute 33

V

VCSResLock attribute 9