Veritas™ Cluster Server Hardware Replication Agent for EMC SRDF Configuration Guide

Windows 2000, Windows Server 2003

5.0

11293737

symantec™

Veritas Cluster Server Hardware Replication Agent for EMC SRDF Configuration Guide

Veritas Cluster Server 5.0

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

Third-party legal notices

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

Windows is a registered trademark of Microsoft Corporation.

Licensing and registration

Veritas Cluster Server is a licensed product.

Technical support

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

Contents

Chapter 1

Introduction

This chapter contains the following topics:

About the EMC SRDF agent

The VCS enterprise agent for EMC SRDF monitors and manages the state of replicated Symmetrix devices that are 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 replicated data clusters and global clusters, and the agent also supports parallel applications.

The agent supports SRDF in the synchronous and asynchronous modes; the agent does not support semi-synchronous and Adaptive Copy. The agent does not require special configuration for SRDF/A support; the agent detects SRDF/A backed devices and manages their failover accordingly.

The agent also supports dynamic SRDF (role swap). If all devices in a given device group are configured for dynamic SRDF, the agent attempts a role swap.

Supported software and hardware

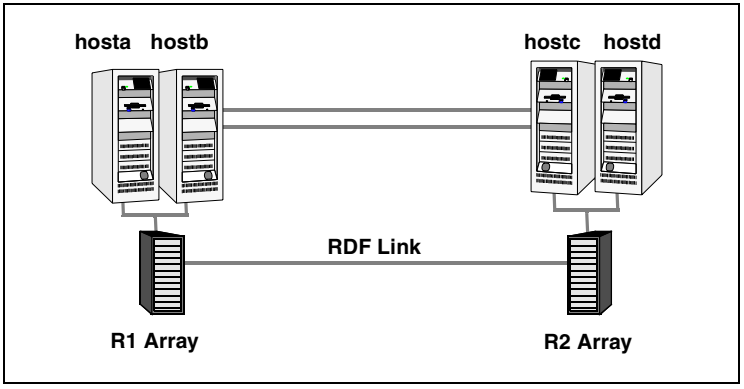
The EMC SRDF agent supports SFW HA 5.0.

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. Contact EMC for details if necessary.

Typical EMC SRDF setup in a VCS cluster

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



- The R1 array, comprising one or more R1 hosts directly attached by SCSI or Fibre Channel to a Symmetrix array containing SRDF R1 devices.
- The R2 array, comprising one or more R2 hosts directly attached by 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. The R2 hosts and arrays must be at a significant distance 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 “About cluster heartbeats” on page 19 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.

SRDF agent functions

The VCS enterprise agent for EMC SRDF monitors and manages the state of replicated Symmetrix devices that are attached to VCS nodes.

The agent performs the following functions:

Function (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 operation makes the devices writable for the application. If one or more devices are write-disabled (WD), the agent runs an `symrdf` command to enable READ-WRITE access to the devices. See "About the EMC SRDF agent's online function" on page 9.
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 operation 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 the `hastop<-all | -local> -force` command.
clean	Determines 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.

About the EMC SRDF agent's online function

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.

If one or more devices are in the write-disabled (WD) state, 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 agent waits for the devices to transition to the R1 UPDATED state before the agent runs a `symrdf` command.
- For R2 devices in the SYNCINPROG state, the agent waits for the devices to transition to the SYNCHRONIZED state before the agent runs a `symrdf` command.

The agent does not run any commands if there is not enough time remaining for the entry point to complete the command.

See "To set the OnlineTimeout attribute" on page 33.

About dynamic swap

The agent supports the SRDF dynamic swap capability. The agent performs a swap for healthy arrays that are configured for dynamic swap when a service group fails over between the arrays. If one array is down, a unilateral read-write enable occurs. The agent fails over the 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 occurs:

- 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.

Chapter 2

Installing the EMC SRDF agent

This chapter contains the following topics:

Before you install the SRDF agent

Set up your cluster. For information about installing and configuring VCS, see the *Veritas Storage Foundation and High Availability Solutions Installation and Upgrade Guide*.

Set up replication and the required hardware infrastructure.

See "Typical EMC SRDF setup in a VCS cluster" on page 8.

Installing the agent for SRDF

If you did not install the EMC SRDF agent when you installed Veritas Storage Foundation for Windows High Availability (SFW HA), follow these instructions to install the agent.

You must install the agent for SRDF on each node in the cluster. In global cluster environments, install the agent on each node in each cluster. These instructions assume that you have already installed SFW HA.

To install the agent

1. Open the Windows Control Panel and click **Add or Remove Programs**.
2. Click the SFW HA Server Components entry and click **Change**.
3. On the installer screen, click **Add or Remove** and click **Next**.
4. In the Option Selection dialog box, select the agent and click **Next**.
5. The installer validates the system for installation.
   If a system is rejected, the Comments column displays the cause of rejection. Highlight the system to view detailed information about the failure in the Details box. Resolve the error, highlight the node in the selected systems list, and click **Validate Again**.
   After all the systems are accepted, click **Next**.
6. An informational message appears if you selected the DMP option. Review the information and click **OK** to continue.
7. Review the summary of your selections and click **Next**.
8. Click **Update** to start the installation.
9. The installer displays the status of installation. After the installation is complete, review the installation report and click **Next**.
10. Click **Finish**.

Removing the agent

This section describes steps for uninstalling the agent. Do not attempt to remove the agent if service groups accessing the shared storage are online.

To remove the agent

1. Open the Windows Control Panel and click **Add or Remove Programs**.
2. Click the VSFW HA Server Components entry and click **Remove**.
3. Review the Welcome page and click **Next**.
4. In the Option Selection dialog box, select the SRDF agent and click **Next**.
5. The installer validates the system for uninstallation.
   If a system is rejected, the Comments column displays the cause of rejection. Highlight the system to view detailed information about the failure in the Details box. Resolve the error, highlight the node in the selected systems list, and click **Validate Again**.
   After all the systems are accepted, click **Next**.
6. Review the summary of your selections and click **Uninstall**.
7. The installer displays the status of uninstallation.
8. After the uninstallation is complete, review the report and click **Next**.
9. Click **Finish**.

Note: For Win IA64 and Win x64 architectures, you will have to manually delete the agent directory if it is not removed after the uninstallation.

Chapter 3

Configuring the agent

This chapter contains the following topics:

Configuration concepts

Review the description of the agent attributes.

Resource type definition

The EMC SRDF agent is represented by the SRDF resource type in VCS.

```
type SRDF (
    static str ArgList[] = { SymHome, GrpName, DevFOTime,
    AutoTakeover, SplitTakeover }
        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 = "C:\\Program Files\\EMC\\SYMCLI\\bin"
        str GrpName
        int DevFOTime = 2
        int AutoTakeover = 1
        int SplitTakeover = 1
        temp str VCSResLock
)
```

Attribute definitions for the SRDF agent

Review the description of the agent attributes.

Required attributes

You must assign values to required attributes.

GrpName — Name of the Symmetrix Device Group that the agent manages.

Type-dimension: string-scalar

Optional attributes

Configuring these attributes is optional.

SymHome — Path to the Symmetrix command line interface.

Type-dimension: string-scalar

Default is C:\Program Files\EMC\SMYCLI\bin.

DevFOTime — Average time in seconds that is required for each device in the group to fail over. This value helps the agent to determine whether it has adequate time for the online operation 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.

Type-dimension: integer-scalar

Default is 2 seconds per device.

AutoTakeover — A flag that determines whether the agent performs a read-write enable on partitioned devices in the write-disabled state during a failover.

Type-dimension: integer-scalar

Default is 1, which means that the agent performs a read-write enable if devices are consistent.

SplitTakeover — A flag that determines whether the agent permits a failover to R2 devices in the Split state. The value 0 indicates that the agent does not permit a failover to R2 devices in the Split state. The value 1 indicates the agent brings service groups online on the R2 side even if the devices are in the split state because they are read-write enabled. The attribute has no effect on failing over to a host attached to R1 devices.

Set the attribute to 0 to minimize the risk of data loss on a failover to devices that may not be in synch.

Type-dimension: integer-scalar

Default is 1.

Internal attributes

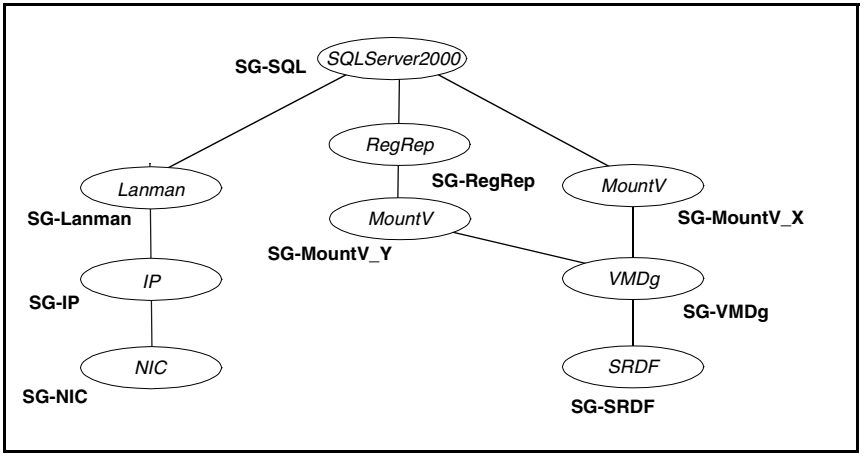
These attributes are for internal use only. Do not modify their values.

VCSResLock — The agent uses the VCSResLock attribute to guarantee serialized management in case of a parallel application.

Type-dimension: temporary string

Sample configuration

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



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

```
SRDF SG-SRDF (
        GrpName = "SQLDG"
)
```

Before you configure the SRDF agent

- Review the configuration concepts, which describe the agent's type definition and attributes.
  See "Configuration concepts" on page 16.
- Verify that the SRDF agent is installed on all systems in the cluster.
- Verify the hardware setup for the agent.
  See "Typical EMC SRDF setup in a VCS cluster" on page 8.
- Make sure the cluster has an effective heartbeat mechanism in place.
  See "About cluster heartbeats" on page 19.
  See "About preventing split-brain" on page 19.
- Set up system zones in replicated data clusters.
  See "About configuring system zones in replicated data clusters" on page 20.

About 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, VCS sends ICMP pings over the public network between the two sites for network heartbeating. To minimize the risk of split-brain, VCS sends ICMP pings to highly available IP addresses. VCS global clusters also notify the administrators when the sites cannot communicate.

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

- The Symmetrix heartbeat shows that the arrays are alive even if the ICMP heartbeats over the public network are lost. So, VCS does not mistakenly interpret this loss of heartbeats as a site failure.
- Heartbeat loss may occur due to the failure of all hosts in the primary cluster. In such a scenario, a failover may be required even if the array is alive. In any case, a host-only crash and a complete site failure must be distinguished. In a host-only crash, only the ICMP heartbeat signals a failure by 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. VCS reports that the site is down only when the remote array is not visible by the `symrdf ping` command.

About preventing split-brain

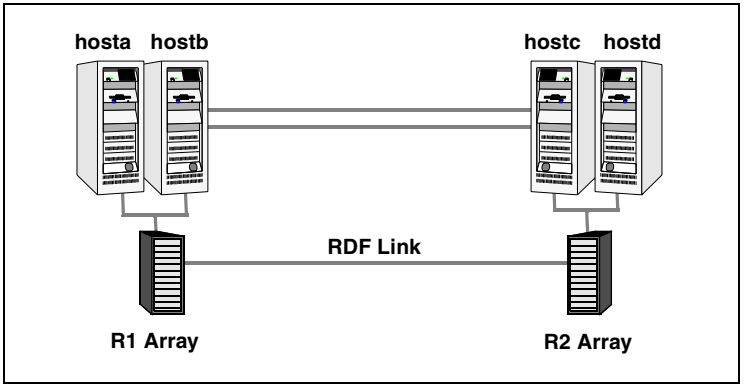
Split-brain occurs when all heartbeat links between the primary and secondary hosts are cut. In this situation, each side mistakenly assumes that the other side is down. Minimize the effects of split-brain by ensuring the cluster heartbeat links pass through similar physical infrastructure as the replication links so that if one breaks, so does the other.

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 R1 to R2 and vice-versa. In this case, the application faults because its underlying volumes become write-disabled, causing the service group to fault. VCS tries to fail it over to another host, causing the same consequence in the reverse direction. This phenomenon continues until the group comes online on the final node. You can avoid this situation by setting up your infrastructure such that loss of heartbeat links also mean the loss of replication links.

About configuring system zones in replicated data clusters

In a replicated data cluster, you can prevent unnecessary failover or failback. Cluster Server attempts to fail over applications within the same system zone before failing them over across system zones. Configure the hosts that are attached to an array as part of the same system zone to avoid unnecessary failover.

The following example depicts a sample configuration where hosta and hostb are in one system zone, and hostc and hostd are in another system zone. Use the SystemZones attribute to create these zones.



Modify the SystemZones attribute using the following command:

```
C:\> 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.

Configuring the SRDF agent

Most clustered applications 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.

Configuring the agent in a global cluster

Configuring the agent manually in a global cluster involves the following tasks.

To configure the agent in a global cluster

1. Start Cluster Manager and log on to the cluster.
2. If the agent resource type (SRDF) is not added to your configuration, add it. From the Cluster Explorer **File** menu, choose **Import Types** and select `C:\Program Files\Veritas\Cluster Server\conf\config\SRDFTypes.cf`
3. Click **Import**.
4. Save the configuration.
5. Add a resource of type SRDF at the bottom of the service group.
6. Configure the attributes of the SRDF resource.
7. If the service group is not configured as a global group, configure the service group using the Global Group Configuration Wizard.
8. Change the ClusterFailOverPolicy from the default, if necessary. Symantec recommends keeping the default, which is Manual, to minimize the chance of failing over on a split-brain.
   Repeat step 5 through step 8 for each service group in each cluster that uses replicated data.
9. Configure the Symm heartbeat at each cluster.
   - From Cluster Explorer **Edit** menu, choose **Configure Heartbeats**.
   - On the Heartbeats Configuration dialog box, enter the name of the heartbeat (Symm).
   - Select the check box next to the name of the cluster to add it to the cluster list for the heartbeat.
   - Click the icon in the **Configure** column to open the **Heartbeat Settings** dialog box.

- 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.
- Click **OK**.

Configuring the agent in a replicated data cluster

Configuring the agent manually in a replicated data cluster involves the following tasks.

To configure the agent in the replicated data cluster

1. Start Cluster Manager and log on to the cluster.
2. If the agent resource type (SRDF) is not added to your configuration, add it. From the Cluster Explorer **File** menu, choose **Import Types** and select `C:\Program Files\Veritas\Cluster Server\conf\config\SRDFTypes.cf`
3. Click **Import**.
4. Save the configuration.
5. In each service group that uses replicated data, add a resource of type SRDF at the bottom of the service group.
6. Configure the attributes of the SRDF resource. Note that some attributes must be localized to reflect values for hosts that are attached to different arrays.
7. Set the SystemZones attribute for the service group to reflect which hosts are attached to the same array.

Setting the OnlineTimeout attribute

Set the OnlineTimeout attribute 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.

Set the OnlineTimeout attribute to at least the amount of time the largest device group takes to fail over.

To set the OnlineTimeout attribute

1. 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 that might fail over simultaneously.
- The additional epsilon is for the command instantiation overhead.

2 If the resulting value seems excessive, divide it by two for every increment in the value of the RestartLimit attribute.

To set the OnlineTimeout attribute using the script

- Run the perl script `\Program Files\Veritas\cluster server\bin\SRDF\sigmal.pl` 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 appropriate `symrdf` command. Specify another value to the sigma script if the instantiation takes shorter or longer.

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

Additional configuration considerations

- 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.
- In global clusters, the AYARetryLimit for the Symm heartbeat must be shorter than the ICMP retry limit. This value ensures that VCS does not confuse a site failure with an all-host failure and that VCS detects array failure first.

Chapter 4

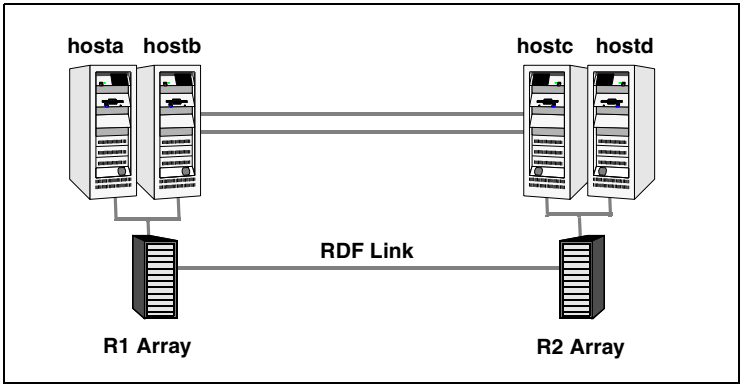
Managing and testing clustering support for EMC SRDF

This chapter contains the following topics:

Typical test setup

A typical test environment includes:

- Two hosts (hosta and hostb) attached to the R1 array.
- Two 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.



Testing 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. In the Service Groups tab of the Cluster Explorer configuration tree, right-click the service group.
2. 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.
3. Migrate the service group to a host that is attached to a different array. In the Service Groups tab of the Cluster Explorer configuration tree, right-click the service group.

4 Click **Switch To**, and click **Remote switch**.

5 Select a system attached to the another array (hostc) and click **OK**.
The service group comes online on hostc and volumes there transition to the RW/FAILED OVER state.

6 Accumulate dirty tracks on the R2 side and update them back on the R1:
`C:\> hares -action srdf_res_name update -sys hostc`
The variable *srdf_res_name* represents the name of the SRDF resource.

7 After the devices transition to R1 UPDATED state, migrate the service group back to its original host. In the Service Groups tab of the Cluster Explorer configuration tree, right-click the service group.

8 Click **Switch To**, and click **Remote switch**. 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.

Testing host failure

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

To perform the host failure test

1 Shut down the host where the application runs.
The service group fails over to hostb and devices are in the RW/SYNCHRONIZED state.

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

5 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.

Performing a disaster test

Test how robust your cluster is in case of a disaster.

To perform a disaster test

1. Shut down all hosts on the source side and shut down the source array. If you can not shut down the R1 Symmetrix, disconnect the ESCON link between the two arrays and simultaneously shut down the hosts. This action mimics a disaster scenario from the point of view of the R2 side.
2. In a replicated data cluster, the service group fails over to hostc or hostd if:
   - All devices were originally SYNCHRONIZED.
   - No synchronization was in progress at the time of disaster.
3. In a global cluster, the administrator is notified of the failure. The administrator can then initiate the failover.

Performing the failback test

You can set up your cluster for a failback test.

The failback test verifies the application can fail back to its original host after a failover to a remote site.

To perform the failback test

1. Reconnect the ESCON cable and reboot the original R1 hosts.
2. Take the service group offline. If you are running this test in a replicated data cluster, run the following command from any host:
   ```
   C:\> 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:
   ```
   C:\> 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:

```
C:\> hagrp -online grpname -sys hosta
```

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

Failure scenarios

Review the failure scenarios and agent behavior in response to failure.

Site disaster

In a total site failure, all hosts and the array are completely disabled, either temporarily or permanently.

In a replicated data cluster, site failure is detected the same way as a total host failure, that is, the loss of all LLT heartbeats.

In a global cluster, VCS detects site failure by the loss of both the ICMP and Symm heartbeats. Make sure that a site failure is not confused with an all-host failure. Set the AYARetryLimit for the Symm heartbeat to be shorter than the ICMP retry limit. With such a setting, 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 the target devices are inconsistent until the synchronization completes, the agent does not write-enable the devices, but it times out and faults. You must restore consistent data from a snapshot or tape backup.

All host or all application failure

Even if both arrays are operational, the service group fails over in the following conditions:

- All hosts on the *PrimarySite* side are disabled.
- The application cannot start successfully on any *PrimaryHost* host.

In replicated data cluster environments, the failover can be automatic, whereas in global cluster environments failover requires user confirmation by default.

VCS serializes `symrdf` commands to ensure that SRDF does not lock out a command while another command is running.

Make sure that SRDF agent's entry points do not time out. Set the OnlineTimeout and RestartLimit attributes for the SRDF resource to restart automatically if the agent entry points are timed out.

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.

Before the SRDF takes any action it waits for the synchronization to complete in the following situations:

- The two arrays are healthy and the link that failed is restored.
- A failover is initiated while synchronization is in progress.

After the synchronization completes, the SRDF runs the `symrdf failover` command.

If the agent times out before the synchronization completes, the resource faults.

The R2 devices are rendered inconsistent and unusable in the following conditions:

- A failover is initiated due to a disaster at the R1 site, and
- A synchronization was in progress

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. Instead, the agent faults. You must restore consistent data to these devices, either from BCV or from a tape backup. Then, you must 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 primary and secondary hosts are cut. In this situation, each side mistakenly assumes that the other side is down. Minimize the effects of split-brain by ensuring the cluster heartbeat links pass through similar physical infrastructure as the replication links so that if one breaks, so does the other.

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 *PrimaryHost* to *SecondaryHost* and vice-versa. In this case, the application faults because its underlying volumes become write-disabled, causing the service group to fault.

VCS tries to fail it over to another host, causing the same consequence in the reverse direction. This phenomenon continues until the group comes online on the final node. You can avoid this situation by setting up your infrastructure such that loss of heartbeat links also mean the loss of replication links.

Index

T

V