# Veritas™ Cluster Server Hardware Replication Agent for IBM PPRC, Configuration Guide

Windows 2000, Windows Server 2003

5.0

11293740

symantec.

# Veritas Cluster Server Hardware Replication Agent for IBM PPRC 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.

Oracle is a registered trademark of Oracle 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 IBM PPRC agent

The VCS enterprise agent for IBM PPRC provides failover support and recovery in environments employing PPRC to replicate data between arrays. It monitors replicated devices attached to the local hosts and transitions those devices to a writable state if necessary.

The agent ensures that a VCS service group having a PPRC volume set configured is online on only one node in the cluster.

The agent supports parallel applications. The agent serializes parallel management, with one agent performing the failover, and the others realizing that it has occurred successfully. In global cluster environments, VCS ensures that the application runs only on hosts attached to the same array. This eliminates the chance, other than a “split-brain,” that instances are running on both source and target volumes simultaneously.

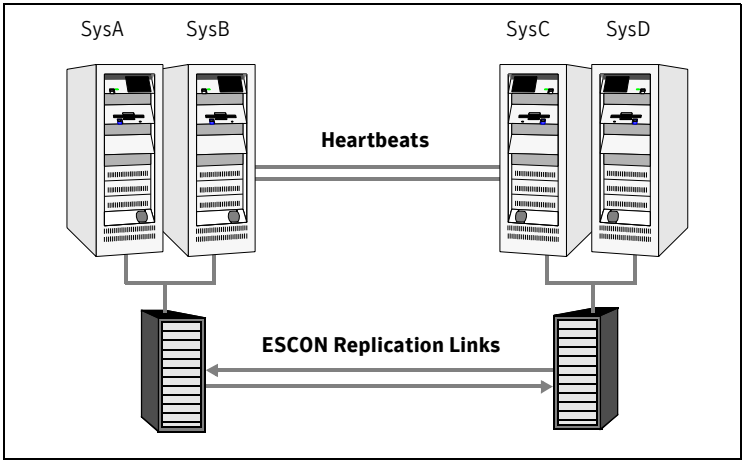
**Note:** The VCS enterprise agent for IBM PPRC does not support PPRC-XD.

## Supported software and hardware

The agent supports all versions of the ESSCLI and any Shark array model 800. Contact IBM for details if necessary.

# Typical setup

Clustering in a PPRC environment typically consists of the following hardware infrastructure:




- The *source array*, comprising one or more hosts directly attached via SCSI or Fibre Channel to a TotalStorage 800 (“Shark”) array containing PPRC source volumes.
- The *target array*, comprising one or more hosts directly attached via SCSI or Fibre Channel to a second Shark array containing PPRC target devices. The target devices are paired with the source devices in the source array. The hosts and the array must be at a significant distance from the source to survive a disaster that may occur at the source. Set up these relationships before configuring the cluster.
- Network heartbeats, using LLT or TCP/IP, between the two data centers to determine their health.
- 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 array.

# Agent functions

The following table lists agent functions. The behavior of some entry points depends on the current PPRC state (SOURCE or TARGET) and status (FULLCOPY, SUSPENDED, or COPY_PENDING) of the PPRC devices that the resource manages.

| Function (Entry Point) | PPRC State and Status | Description |
|---|---|---|
| online | State: SOURCE<br>Status: ANY STATUS | If the PPRC state of all local devices is SOURCE, the agent touches a lock file to indicate that the host has read-write access to the volumes. |
| | State: TARGET<br>Status: ANY STATUS | If the PPRC state of all of the local devices is TARGET and the status of all of devices is FULLCOPY, the agent runs the failover task for those volumes. If the state transitions to SOURCE, the agent touches the lock file to indicate that the resource can come online. The agent then issues the failback task to convert the original source volumes into targets.<br>If the status of any target device is COPY_IN_PROGRESS, the agent waits until all devices are FULLCOPY, otherwise the entry point times out.<br>If the status of any target device is SUSPENDED or SIMPLEX, the entry point times out and the resource faults. |
| | State: SIMPLEX<br>Status: NONE | The resource goes online without initiating any action. The administrator should decide whether these volumes should be part of a PPRC pair or should not be included in the VCS PPRC management. |
| offline | | The agent removes the lock file regardless of the device state and the next monitor cycle returns the resource state as OFFLINE. No PPRC commands are required during an offline because offlining is not indicative of the intention to give up the devices. |
| monitor | State: SOURCE or TARGET<br>Status: ALL FULLCOPY | Verifies the existence of the lock file to determine the resource status. If the lock file exists, the agent reports the PPRC resource as being online. If the lock file does not exist, it reports the resource as being offline. |

| Function (Entry Point) | PPRC State and Status | Description |
|---|---|---|
| open | | Removes the lock file on the host where this entry point is called. This prevents potential concurrency violation if the group fails over to another node.<br>**Note:** The agent does not remove the lock file if the agent was started after an `hastop<-all \| -local> -force` command. |
| 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. |
| attr_changed | | Updates the volume list maintained by the agent when the Volumes attribute is a list of volumes. |
| info | | Retrieves the current state and status of all volumes managed by the resource.<br>This entry point is not called automatically because it uses ESSCLI, which is resource-intensive. To schedule this operation at a regular interval, set the InfoInterval attribute to a non-zero value equalling the interval. For example, to run this operation every 180 seconds, set the InfoInterval attribute to 180. |

Chapter 2

# Installing the IBM PPRC agent

This chapter contains the following topics:



## Before you install the PPRC 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 setup" on page 8.

## Installing the PPRC agent

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

You must install the PPRC agent 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 SFW HA Server Components entry and click **Remove**.

3 Click **Next**.

4 In the Option Selection dialog box, select the PPRC 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**.

Chapter 3

# Configuring the IBM PPRC agent

This chapter contains the following topics:



## Configuration concepts

Review the description of the resource type definition for the agent, the agent attributes, and the configuration examples.

## Resource type definition

```
type PPRC (
    static keylist RegList = { Volumes }
    static int MonitorInterval = 300
    static int MonitorTimeout = 290
    static int NumThreads = 1
    static int OfflineMonitorInterval = 0
    static i18nstr ArgList[] = { BaseDir, LocalServers,
    SecurityFile,
    Volumes, Tasks, RemoteServers, RemoteSecurity }
    str BaseDir = "C:\\Program Files\\IBM 2105 CLI"
    str LocalServers[]
    str SecurityFile
    str Volumes[]
    str Tasks{}
    str RemoteServers[]
    str RemoteSecurity
    temp str VCSResLock
)
```

## Attribute definitions

| Required Attributes | Type-Dimension | Description |
|---|---|---|
| LocalServers | string-vector | List of IP addresses for the ESS management server of the array where local nodes are connected.<br>The first IP address is the primary address and a second address, if it exists, is the backup. |
| SecurityFile | string-vector | Fully qualified path name of a text file containing a valid username and password for executing ESS tasks. Each host in the cluster must see a valid SecurityFile, which you can share or copy to all hosts.<br>The format for the file is: *username password*. |

| Required Attributes | Type-Dimension | Description |
|---|---|---|
| Volumes | string-vector | List of volumes in the array where the local nodes are connected.<br>You can also enter the list of volumes in a file and set this attribute to the file path by entering a forward slash before the file path. In the file, make sure to separate each volume by a new line. The contents of the file must be identical on all hosts connected to the same array. |
| Tasks | string-association | An association specifying predefined PPRC tasks. The agent performs the following functions by invoking the appropriate tasks:<br>■ failover (FO)<br>■ failback (FB)<br>■ resync (RS)<br>Specify tasks with the following syntax:<br>`{ taskname1 = task1, taskname2 = task2, taskname3 = task3 }`<br>Define the task as: FO, FB, or RS in VCS.<br>For example:<br>Tasks = `{ FO306-219 = FO, FB219-306 = FB, RS219-306 = RS }`<br>You can set multiple tasks of the same type, but Symantec recommends that you group several related tasks into a single task in the ESS. |

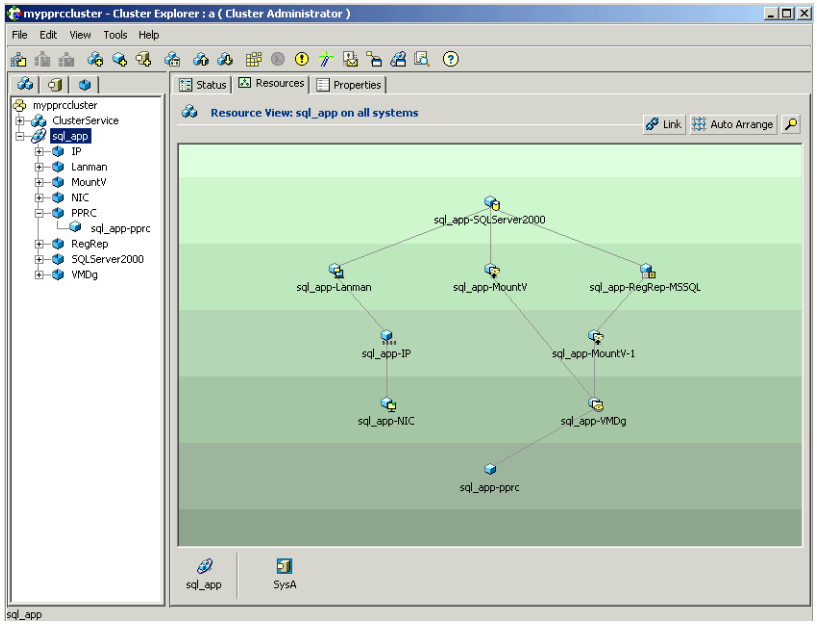
| Optional Attributes | Type-Dimension | Description |
|---|---|---|
| BaseDir | string-scalar | The base path to the Enterprise Storage Server Command Line Interface (ESSCLI).<br>Default is "`C:\\Program Files\\IBM 2105 CLI`" |

| Optional Attributes | Type-Dimension | Description |
|---|---|---|
| RemoteServers | string-vector | List of IP addresses of the ESS management server of the remote array where local nodes are connected.<br>The first IP address is the primary address and a second address, if it exists, is the backup.<br>Use this attribute in conjunction with the RemoteSecurity attribute. |
| RemoteSecurity | string-scalar | Fully qualified path name of a text file containing a valid username and password for executing ESS tasks on the remote ESS. Each host in the cluster attached to an ESS must see a valid file, which you can share or copy to all hosts.<br>The format for the text file is: *username password*. |

| Internal Attributes | Type-Dimension | Description |
|---|---|---|
| VCSResLock | temporary-string | This attribute is used by the agent to guarantee serialized management in parallel applications.<br>*Do not modify this attribute.* |

# Configuration examples

The following dependency graph—applicable for both replicated data and global clusters—shows a VCS service group that has a resource of type PPRC. The VMDg resource depends on the PPRC resource.




## Replicated data cluster

This configuration has a resource of type PPRC configured in a two-node replicated data cluster. This cluster consists of nodes SysA and SysB.

```
PPRC sql_app-pprc (
        LocalServers@SysA = { "10.162.25.139" }
        LocalServers@SysB = { "10.162.25.167" }
        SecurityFile = "C:\temp\shark-local.pwd"
        Volumes @SysA = { 22222634, 22322634, 22422634, 22522634
}
        Volumes @SysB = { 30722634, 30822634, 30922634, 30A22634
}
        Tasks @SysA = { FOsql22 = FO, FBsql22 = FB, RSsql22 = RS
}
        Tasks @SysB = { FOsql23 = FO, FBsql23 = FB, RSsql23 = RS
}
        RemoteServers@SysA = { "10.162.25.167" }
        RemoteServers@SysB = { "10.162.25.139" }
        RemoteSecurity = "C:\temp\shark-remote.pwd"
)
```

In this configuration:

- A group of four PPRC volumes 22222634, 22322634, 22422634, 22522634 on the array attached to SysA.
- PPRC volumes 30722634, 30822634, 30922634, 30A22634 on the array attached to SysB.
- Host SysA is attached to the ESS managed via 10.162.25.139; host SysB connects to 10.162.25.167 as its local array.
- No backup server is specified for either host.
- Tasks FOsql22, FBsql22, and RSsql22 are configured at the array attached to SysA; these can failover, failback, and copy out-of-synch tracks to the four volumes 22222634, 22322634, 22422634, and 22522634 and their PPRC pairs 30722634, 30822634, 30922634, 30A22634 respectively.
- All hosts access the username and password for the local array in the file `C:\temp\shark-local.pwd` and username and password to access the remote array at `C:\temp\shark-remote.pwd`.

## Global cluster

The following configuration has a resource of type PPRC configured in a global cluster.

```
PPRC sql_app-pprc (
    BaseDir = "C:\\Program Files\\IBM 2105 CLI"
    LocalServers = { "10.162.25.139" }
    SecurityFile = "C:\temp\shark.pwd"
    Volumes = { 23222634 }
    Tasks = { FOSrcTo232_12 = FO, FB316_13toTarget = FB,
    rsync_232_to_316 = RS }
    RemoteServers = { "10.162.25.140" }
    RemoteSecurity = "C:\temp\shark.pwd"
    )
```

In this configuration:

- All hosts within the cluster are using the PPRC volume 23222634.
- All hosts are attached to the ESS managed via 10.162.25.139.
- No backup server is specified for any host.
- Tasks FOSrcTo232_12, FB316_13toTarget, and rsync_232_to_316 are configured at the array. These can failover, failback, and copy out-of-synch tracks.
- All hosts find the username and password to access the local array in the file `C:\temp\shark-local.pwd` and username and password to access the remote array at `C:\temp\shark-remote.pwd`.

# Before you configure the PPRC agent

- Review the configuration concepts, which describe the agent's type definition and attributes.
  See "Configuration concepts" on page 13.
- Verify that the PPRC agent is installed on all systems in the cluster.
- Verify the hardware setup for the agent.
  See "Typical setup" 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.

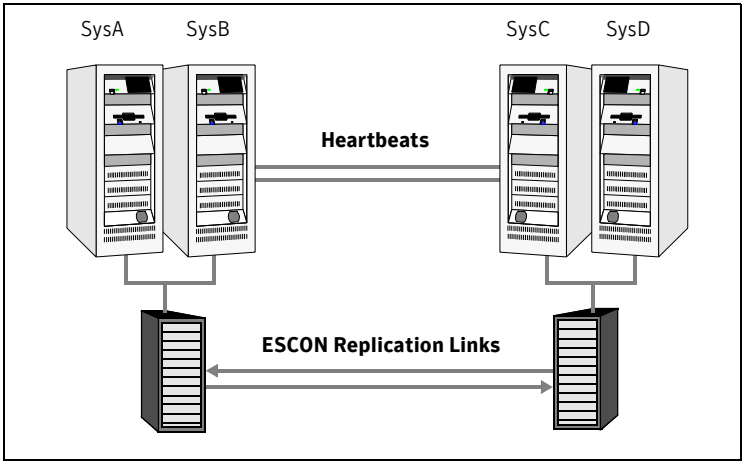
In global clusters, the VCS Heartbeat agent sends heartbeats directly between the Symmetrix arrays if the Symmetrix ID of each array is known. This heartbeat offers the following advantages:

- 
- 

## 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.

# About configuring system zones in replicated data clusters




This example depicts a sample configuration where SysA and SysB are in one system zone, and SysC and SysD are in another system zone. Use the SystemZone attribute to create these zones.

Modify the SystemZones attribute using the following command:

```
C:\> hagrp -modify grpname SystemZones SysA 0 SysB 0 SysC 1 SysD 1
```

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

This command creates two system zones: zone 0 with SysA and SysB and zone 1 with SysC and SysD.

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

# Configuring the PPRC agent

You can adapt most applications configured in VCS to a disaster recovery environment by:

- Converting their devices to PPRC devices
- Synchronizing the devices
- Adding the PPRC agent to the service group
- Configure volumes of PPRC arrays as resources of type PPRC.

## 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 (PPRC) 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\PPRCTypes.cf`.
3. Click **Import**.
4. Save the configuration.

# 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 (PPRC) 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\PPRCTypes.cf`.
3. Click **Import**.
4. Save the configuration.
5. In each service group that uses replicated data, add a resource of type PPRC at the bottom of the service group.
6. Configure the attributes of the PPRC 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.

Chapter 4

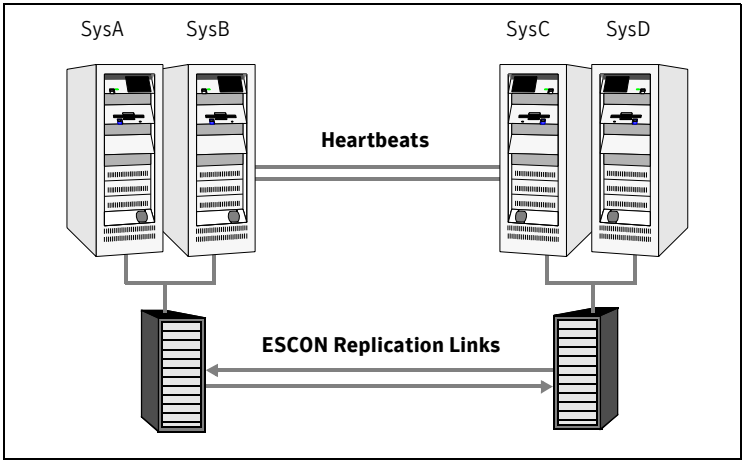
# Managing and testing clustering support for IBM PPRC

This chapter contains the following topics:

# Typical test setup

A typical test environment includes:

- Two hosts (SysA and SysB) attached to the source array.
- Two hosts (SysC and SysD) are attached to the target array.
- 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 (SysB) from the menu.
   The service group comes online on SysB and local volumes remain in the SOURCE 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 In the Switch Global Group dialog box, select a system attached to the another array (SysC) and click **OK**.

6 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.

7 Click **Switch To** and click **Remote switch**.

8 In the Switch Global Group dialog box, select the original system on which the group was initially online (SysA) and click **OK**.
The group comes online on SysA. The devices return to the SOURCE state at the array attached to SysA and SysB. The status of all volumes is FULLCOPY at all times, except during the PPRC device failover.

# 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 SysB and devices are in the SOURCE state.

2 Shut down SysB.
In a replicated data cluster, the group fails over to SysC or SysD 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 SOURCE 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 where the service group was initially online (SysA).
The service group comes online on SysA and devices transition to the SOURCE 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 cannot shut down the ESS, disconnect the ESCON link between the two arrays and simultaneously shut down the hosts. This action mimics an actual disaster scenario from the point of view of the target side.
2. In a replicated data cluster, the service group fails over to SysC or SysD if all volumes were originally in the FULLCOPY state.
   The target devices become source volumes, and the status is SUSPENDED.
3. In a global cluster, the administrator is notified of the failure. The administrator can then initiate the failover.

# 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.

Bringing the service group online at the target side causes the failover tasks to be run. The success of the failover tasks promotes the target volumes to source, which can be mounted and written to.

The online operation attempts to perform a failback to demote the original source volumes. Since the source volumes are inaccessible, the failback does not succeed. However the ability to mount and write the volumes is not predicated on failback succeeding, so the failure is logged and the resource comes online.

If the RemoteServer and RemoteSecurity attributes are set, the monitor entry point of the PPRC agent periodically attempts the failback tasks

In replicated data clusters, the failover can be automatic, while in global clusters, failover requires user confirmation. In a global cluster, you can *declare* a cluster as being down due to a disaster or an outage. The PPRC agent will not attempt failback tasks on a cluster that has been declared a disaster. It attempts failback on a cluster that is declared faulted due to an outage.

## 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.

## Split-brain

When split-brain occurs in a replicated database cluster, VCS assumes a total disaster because the *PrimarySite* hosts and array are unreachable. VCS attempts to start the application. Once the heartbeats are restored, VCS stops the applications on one side and restarts the VCS engine (HAD). This action eliminates 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.

# Index

## R



## S



## T



## V