Veritas™ Cluster Server Hardware Replication Agent for Hitachi TrueCopy Configuration Guide

Windows 2000, Windows Server 2003

5.0

11293738

Veritas Cluster Server Hardware Replication Agent for Hitachi TrueCopy Configuration Guide

Veritas Cluster Server 5.0

THIS DOCUMENTATION IS PROVIDED “AS IS” AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID, SYMANTEC CORPORATION SHALL NOT BE LIABLE FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH THE FURNISHING PERFORMANCE, OR USE OF THIS DOCUMENTATION. THE INFORMATION CONTAINED IN THIS DOCUMENTATION IS SUBJECT TO CHANGE WITHOUT NOTICE.

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

Chapter 2 Installing the Hitachi TrueCopy agent

Chapter 3 Configuring the Hitachi TrueCopy agent

Chapter 4 Managing and testing clustering support for Hitachi TrueCopy

Chapter 1

Introduction

This chapter contains the following topics:

- About the Hitachi TrueCopy agent
- Supported software and hardware
- Typical setup
- Agent functions

About the Hitachi TrueCopy agent

The VCS enterprise agent for Hitachi TrueCopy monitors and manages the state of replicated devices attached to local hosts. The agent ensures that the system on which the TrueCopy resource is online has safe and 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 supports TrueCopy in all fence levels that are supported on a particular array.

When replicating between Lightning arrays, the agent supports the following fence levels: *data*, *never*, and *async*.

When replicating between Thunder arrays, the agent supports the following fence levels: *data* and *never*.

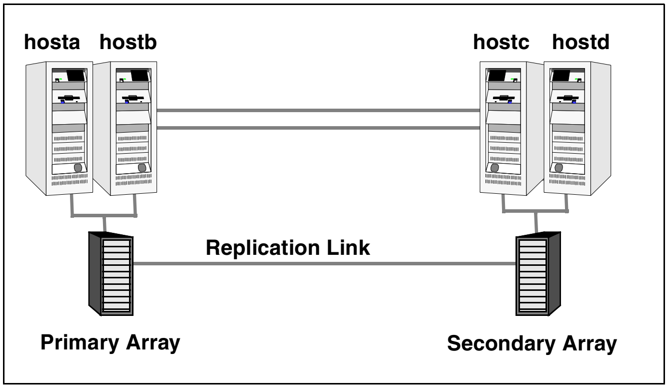
Supported software and hardware

The agent supports all versions of the Hitachi RAID Manager. It supports TrueCopy on all microcode levels on all Lightning arrays, provided the host/HBA/array combination is in Hitachi's hardware compatibility list. The agent supports Sun StorEdge 9900 and Hewlett-Packard XP arrays with TrueCopy rebranded as Continuous Access. The agent supports all fence levels on 9900 arrays and supports synchronous replication on the 9500 series.

The agent does not support other Hewlett-Packard replication solutions under the Continuous Access umbrella such as Continuous Access Storage Appliance (CASA); it only supports Continuous Access XP.

Typical setup

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



- The primary array, comprising P-VOL hosts directly attached by SCSI or Fibre Channel to a Hitachi array containing TrueCopy P-VOL volumes.
- The secondary array, comprising S-VOL hosts directly attached by SCSI or Fibre Channel to a second Lightning array containing TrueCopy S-VOL devices. These devices pair with the P-VOL devices in the primary array.

  These hosts and the array must be at a significant distance apart from the primary side to survive a disaster that may occur there.
- 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 by 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 Veritas agent for Hitachi TrueCopy monitors and manages the state of replicated devices that are attached to VCS nodes.

The agent performs the following operations:

online

If the state of all local devices is read-write enabled, the agent creates a lock file on the local host to indicate that the resource is online. This action makes the devices writable for the application.

If one or more devices are not in a writable state, the agent runs the `horctakeover` command to enable read-write access to the devices.

See "About the Hitachi TrueCopy agent's online function" on page 11.

offline

The agent removes the lock file on the device. The agent does not run any TrueCopy commands because taking the resource offline is not indicative of an intention to give up the devices.

monitor

Verifies the existence of the lock file to determine the resource status. If the lock file exists, the agent reports the status of the resource as online. If the lock file does not exist, the agent reports the status of the resource as offline.

The monitor entry point does not examine the state of the devices or the state of the replication link between the arrays.

open

Removes the lock file on the system on which this entry point is called. This prevents potential concurrency violation if the group fails over to another node.

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.

info

Reports the current role and status of the devices in the device group. This entry point can be used to verify the device state and to monitor dirty track trends.

action

Resynchronizes the devices from the VCS command line after connectivity failures are detected and corrected.

The agent supports the following actions:

- pairdisplay—Displays information about all devices.
- pairresync—Resynchronizes the S-VOLs.
- pairresync-swaps—Promotes the S-VOLs to P-VOLs and resynchronizes the original P-VOLs.
- localtakeover—Makes the local devices write-enabled.

About the Hitachi TrueCopy agent’s online function

If the state of all local devices is read-write enabled, the agent makes the devices writable by creating a lock file on the local host.

If one or more devices are not in a writable state, the agent runs the `horctakeover` command to enable read-write access to the devices.

For S-VOL devices in any state other than SSWS or SSUS, the agent runs the `horctakeover` command and makes the devices writable. The time required for failover depends on:

- The health of the original primary.
- The RAID Manager timeouts as defined in the horcm configuration file for the device group.

The agent considers P-VOL devices writable and takes no action other than going online, regardless of their status.

If the S-VOL devices are in the COPY state, the agent runs the `horctakeover` command after one of the following:

- The synchronization from the primary completes.
- The OnlineTimeout period of the entry point expires, in which case the resource faults.

Chapter 2

Installing the Hitachi TrueCopy agent

This chapter contains the following topics:

Before you install the TrueCopy 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 9.

Installing the agent for TrueCopy

If you did not install the Hitachi TrueCopy 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 TrueCopy 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, hgihlight 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 TrueCopy 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 Hitachi TrueCopy agent

This chapter contains the following topics:

Configuration concepts

Review the configuration concepts and failure scenarios for the agent..

Resource type definition for the Hitachi TrueCopy agent

The Hitachi TrueCopy agent is represented by the HTC resource type in VCS.

```
type HTC (
    static str ArgList[] = { BaseDir, GroupName, Instance }
    static int NumThreads = 1
    static keylist SupportedActions = { localtakeover,
    pairresync, pairresync-swaps, pairdisplay }
    NameRule = resource.GroupName
    str BaseDir = "C:\\HORCM\\etc"
    str GroupName
    int Instance
    int SplitTakeover = 1
    int LinkMonitor = 0
    )
```

Attribute definitions

Review the description of the agent attributes.

BaseDir

Path to the RAID Manager Command Line interface.

Type-dimension: string-scalar

Default is `C:\\HORCM\\etc`.Default: `C:\\HORCM\\etc`.

GroupName

Name of the device group managed by the agent.

Type-dimension: string-scalar

Instance

The Instance number of the device group that the agent manages. Multiple device groups may have the same instance number.

Do not define the attribute if the instance number is zero.

Type-dimension: integer-scalar

SplitTakeover

A flag that determines whether the agent permits a failover to S-VOL devices if the replication link is disconnected; that is, if P-VOL devices are in the PSUE state.

See “About the SplitTakeover attribute” on page 19.

Type-dimension: integer-scalar

Default: 1.

LinkMonitor

A flag that defines whether the agent perodically attempts to to resynchronize the S-VOL side if the replication link is disconnected. The agent uses the `pairresync` command to resynchronize arrays.

The value 1 indicates that when the replication link is disconnected, the agent periodically attempts to resynchronize the S-VOL side using the `pairresync` command.

Setting LinkMonitor does not affect the SplitTakeoverbehavior. However, you can minimize the time during which the P-VOL is in the PSUE state by setting the LinkMonitor attribute.

Type-dimension: integer-scalar

Default: 0.

About the SplitTakeover attribute

The SplitTakeover attribute determines whether the agent permits a failover to S-VOL devices if the replication link is disconnected; that is, if P-VOL devices are in the PSUE state.

The default value for this attribute is 1. If you set the value to 0, the agent does not permit a failover to S-VOL devices if the P-VOL devices are in the PSUE state. If a failover occurs when the the replication link is disconnected, there is a possibility of data loss because the S-VOL devices may not be in synch.

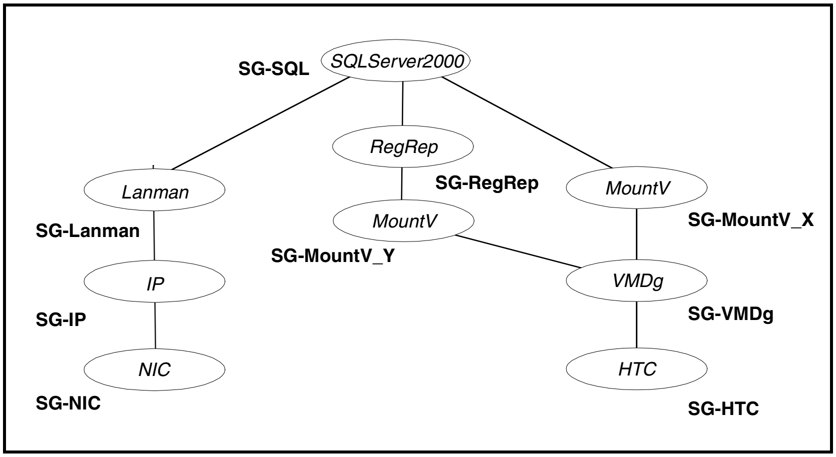
In this scenario, the agent attempts to contact the RAID manager at the P-VOL side to determine the status of the arrays. If the P-VOL side is down, the agent attempts to go online.

In a global cluster environment, if the agent at the P-VOL side detects the PSUE state locally, it freezes the service group at the S-VOL side to prevent a failover. The agent unfreezes the service group when the devices are resynchronized after the link is restored.

Note: Setting LinkMonitor does not affect the SplitTakeover behavior. However, you can minimize the time during which the P-VOL is in the PSUE state by setting the LinkMonitor attribute.

Sample configuration

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



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

```
HTC SQLDG (
        GroupName = SQLDG
        Instance = 1
)
```

Before you configure the Hitachi TrueCopy agent

Before you configure the agent, review the following information:

- Review the configuration concepts, which describe the agent's type definition and attributes.
  See "Configuration concepts" on page 18.
- Verify that the agent is installed on all systems in the cluster.
- Verify the hardware setup for the agent.
  See "Typical setup" on page 9.
- Make sure the cluster has an effective heartbeat mechanism in place.
  See "About cluster heartbeats" on page 21.
  See "About preventing split-brain" on page 21.
- Verify that the clustering infrastructure is in place. If you plan to configure the agent in a global cluster, make sure the global service group for the application is configured. If you plan to configure the agent in a replicated

data cluster, make sure the required replication infrastructure is in place and that the application is configured. See the *Veritas Cluster Server Administrator's Guide* for more information.

- Set up system zones in replicated data clusters.
  See "About configuring system zones in replicated data clusters" on page 22.

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.

Hitachi arrays do not support a native heartbeating mechanism between the arrays. The arrays send a support message on detecting replication link failure. You can take appropriate action to recover from the failure and to keep the devices in a synchronized state. The TrueCopy agent supports actions that can automate the resynchronization of devices after a replication link outage is corrected.

About preventing split-brain

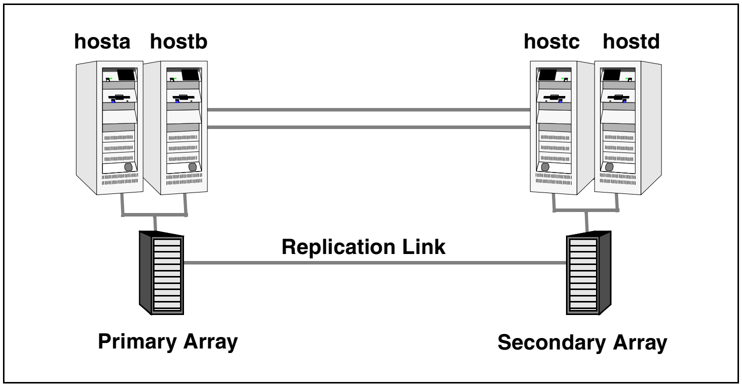
Split-brain occurs when all heartbeat links between the primary and secondary hosts are cut. In this situation, each side mistakenly assumes that the other side is down. Minimize the effects of split-brain by ensuring the cluster heartbeat links pass through similar physical infrastructure as the replication 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 primary host to secondary host 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. VCS Hardware Replication Agent for Hitachi TrueCopy 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.



You can 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 host b, 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.

Configuring the Hitachi TrueCopy agent

You can adapt most clustered applications to a disaster recovery environment by:

- Converting their devices to TrueCopy devices
- Synchronizing the devices
- Adding the Hitachi TrueCopy agent to the service group

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's resource type HTC 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\HTCTypes.cf.`
3. Click **Import**.
4. Save the configuration.
5. Add a resource of type HTC at the bottom of the service group.
6. Configure the attributes of the HTC resource.
7. 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 Administrator's Guide* for more information.
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.

Configuring the agent manually in a replicated data cluster

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

To configure the agent in a replicated data cluster

1. Start Cluster Manager and log on to the cluster.
2. If the agent resource type (HTC) 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\HTCTypes.cf.`
3. Click **Import**.
4. Save the configuration.
5. In each service group that uses replicated data, add a resource of type HTC at the bottom of the service group.
6. Configure the attributes of the HTC 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

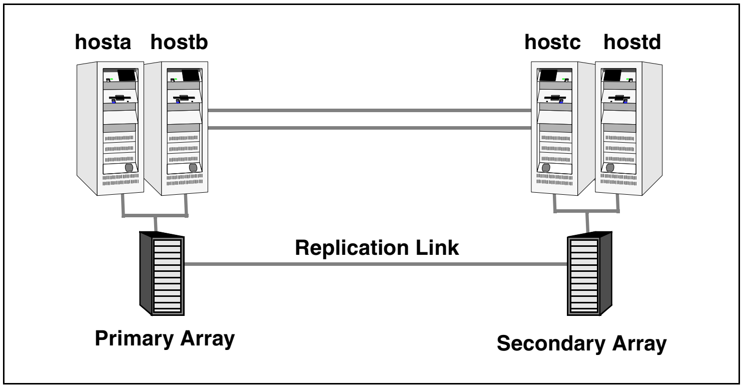
Managing and testing clustering support for Hitachi TrueCopy

This chapter contains the following topics:

Typical test setup

A typical test environment includes:

- Two hosts (hosta and hostb) are attached to the primary array.
- Two hosts (hostc and hostd) attached to the secondary array.
- The application is running on hosta and devices in the local array are P-VOLs in the PAIR state.
- A replicated data cluster has two dedicated heartbeat links; a global cluster has one network heartbeat. The test scenario is similar for both environments.



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 that is 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 a system that is attached to the same array (hostb).

   The service group comes online on hostb and local volumes remain in the P-VOL/PAIR 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 the system that is attached to the another array (hostc) from the menu.

The service group comes online on hostc and volumes there transition to the P-VOL/PAIR state, changing the original P-VOLs to S-VOLs.

5 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.

6 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 original state in step 1.

Testing host failure

In this scenario, the host where 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 on which the application is running:

The service group fails over to hostb and devices are in the P-VOL/PAIR state.

2 Halt or shut down hostb.

In a replicated data cluster, the group fails over to hostc or hostd depending on the value of the FailOverPolicy attribute 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 on the target array remain S-VOLs. They do so because they cannot communicate with the original primary’s RAID manager, but they transition to the writable SSWS status. The failover can take some time as the RAID manager connection times out.

3 Reboot the two hosts that were shut down. A swap resynchronization is required to demote the original P-VOLs:

```
C:\> hares -action HTCRes pairresync-swaps -sys system
```

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 swap roles again.

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 primary array, disconnect the replication link between the two arrays and simultaneously shut down the hosts. This action mimics a disaster scenario to the secondary side.

2 In a replicated data cluster, the service group fails over to hostc or hostd if all devices were originally in the PAIR state; that is, 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 by declaring an outage.

Performing the failback test

You can set up your cluster for a failback test.

To perform a failback test

1 Reconnect the replication link and reboot the original P-VOL hosts.

2 Take the service group offline.

3 Write-disable both sides.

4 Manually resynchronize the device using the same steps as the preceding host failure test .

5 Once the resynchronization is complete, migrate the application back to the original primary side.

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

The devices swap roles again 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 environment, VCS detects the failure by the loss of the ICMP heartbeat between the clusters.

If a failover occurs, the online entry point of the TrueCopy agent runs the `horctakeover` command. The RAID manager waits for the timeout in trying to contact its peer RAID manager daemon before taking over the disks. This wait can cause delay in the failover. This timeout is defined in the device group's instance's configuration file. Make sure the value of the OnlineTimeout entry point of the HTC type is greater than the RAID manager timeout.

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 primary site side are disabled.
- The application cannot start successfully on any primary host.

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 service groups can fail over in parallel.

TrueCopy does not provide any serialization restrictions on simultaneous device group failover. However, the `horctakeover` command makes an attempt to contact the RAID manager on the original P-VOL when performing a failover. In such a case, if the RAID manager is inaccessible, failover is delayed until the surviving RAID manager's connect timeout expires. This timeout is defined in the configuration file for the particular instance.

Replication link failure

Hitachi arrays send an alert in the following situations:

- When the array detects a replication link failure

- When any P-VOLs, on which data has been written, transition from the PAIR state to the PSUE state

In fence levels never and async, a replication link failure does not compromise the application's ability to write to its local devices. The arrays start tracking changed regions on disk in preparation for resynchronization when the link is restored.

The devices do not automatically resynchronize when the link is restored , nor do they change state when the restoration is detected. An administrator can resynchronize the devices, either from the command line or by running a configured action using the agent's action entry point. The following situations require administrative action after you repair a link failure. These actions depend on the fence level and any events that occurred during the failure.

Table 4-1 Replication link failure and recommended action

Event	Fence Level	Recommended Action
Link fails and is restored, but application does not fail over.	never, async	Run the `pairresync` action to resynchronize the S-VOLs.
Link fails and application fails to the S-VOL side.	never, async, or data	Run the `pairresync-swaps` action to promote the S-VOLs to P-VOLs and resynchronize the original P-VOLs.
Application faults due to I/O errors.	data	Run the `localtakeover` action to write-enable the local devices. Clear faults and restart service group.

Split-brain

When split-brain occurs in a replicated database cluster, VCS assumes a total disaster because the primary site 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.

Administrators must resynchronize the volumes manually using the `pairresync` 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 a fail over mistakenly occurs, 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 the data manually.

Index