WHITE PAPER: I/O FENCING DEPLOYMENT CONSIDERATIONS

SYMANTEC CLUSTER SERVER 6.2 I/O FENCING DEPLOYMENT CONSIDERATIONS

Anthony Herr, Technical Product Manager

Table of Contents

Executive Summary

I/O Fencing provides protection against data corruption and can guarantee data consistency in a clustered environment. Data is the most valuable component in today's enterprises. Having data protected and therefore consistent at all times is a number one priority.

This White Paper describes the different deployment methods and strategies available for I/O Fencing in a Symantec Cluster Server (VCS) environment. It is designed to illustrate configuration options and provide examples where they are appropriate.

Symantec has led the way in solving the potential data corruption issues that are associated with clusters. We have developed and adopted industry standards (SCSI3 Persistent Reservations [PR]) that leverage modern disk-array controllers that integrate tightly into the overall cluster communications framework.

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 **Symantec Cluster Server Release Notes**.

Licensing and registration

Symantec Cluster Server is a licensed product. See the **Symantec Cluster Server Installation Guide** for license installation instructions.

Technical support

For technical assistance, visit:
http://www.symantec.com/enterprise/support/assistance_care.jsp.
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.

Scope of document

This document is intended to explain and clarify I/O Fencing for Symantec Cluster Server (VCS) Clusters. It will provide information to assist with adoption and architecture of I/O Fencing when using VCS. Note that VCS is included in several product bundles from Symantec, including but not limited to, Storage Foundation High Availability, Storage Foundation Cluster File System, and Storage Foundation for Oracle RAC.

The document describes how I/O Fencing operates as well as providing an outline of the available functionality. Installation and Administration procedures are well covered in publicly available documentation that is unique to each software release.

This document focuses on the I/O Fencing functionality provided in VCS 6.2. Information may or may not be applicable to earlier and later releases. Where possible, we will mention in which version a specific feature was introduced.

Audience

This document is targeted for technical users and architects who wish to deploy VCS with I/O Fencing. The reader should have a basic understanding of VCS. More information around VCS can be found here: **http://www.symantec.com/business/cluster-server**.

Background

Providing data high availability naturally exposes the risk to protect that data because independent nodes have access to the same storage device. In its infancy, this technology caused data corruptions. As availability evolved, different technologies have been developed to prevent data corruption while continuing to protect services.

In short, the problem arises when two or more servers are accessing the same data independently of each other. This is termed "split brain" and is outlined in the **Introduction chapter** of this document.

Preventing data corruption during a split brain scenario is relatively easy. However, some cluster solutions handle this situation by forcing downtime to the applications. This is not acceptable in today's computing environments that require constant uptime and tighter SLAs.

As VCS evolved, several methods of avoiding split brain and data corruption have been put into place. Since VCS 3.5 (Released in 2001), I/O Fencing has been available. I/O Fencing can eliminate the risk of data corruption in a split brain scenario by ensuring that a single set of clustered nodes remains online and continues to access shared storage.

This document focuses on the various options for I/O Fencing including implementation considerations, architecture, and guidelines. It also provides a comparison among the different deployment methods and options.

NOTE: VCS for Windows (known as Storage Foundation for Windows HA or SFWHA) uses another method to prevent data corruptions in split brain scenarios and will not be covered in this document. Please refer to public documentation for the Storage Foundation for Windows HA release for more information.

Introduction

Introduction to VCS topology

Cluster communication network, also called a "heartbeat link", is an essential factor in availability design. A single VCS cluster can consist of multiple systems that are all connected via heartbeat networks. In most cases, two independent heartbeat networks are used to ensure that communication can still occur if a single network port or switch was to go offline. Protecting the heartbeat networks is crucial for cluster stability. In some instances, the heartbeat networks are referred to as "private networks" as they generally are not used for public network traffic.

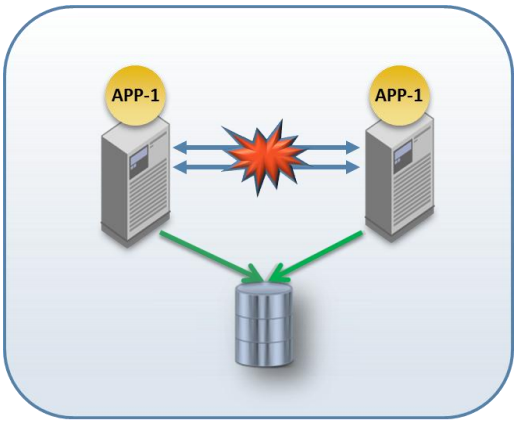
VCS replicates the current state of all managed resources from each cluster node to all other nodes in the cluster. State information is transferred over the heartbeat networks; hence all nodes have the same information about all cluster resource states and activities. VCS also recognizes active nodes, nodes joining and leaving the cluster, and faulted nodes over the heartbeat networks.

Shared storage isn't required when using VCS. However, VCS is most commonly configured with shared storage. With the 6.1 release, Symantec has introduced functionality to share data across cluster nodes even if the local node does not have access to the shared storage. This functionality is called FSS or Flexible Shared Storage. In these configurations, ensuring all cluster nodes can communicate is even more important. I/O Fencing is recommended to ensure only nodes that can communicate are included in the cluster. For more information, see the Cluster File System Documentation on SORT.

Heartbeat networks are also used to transfer lock information when configured with a cluster file system or in a configuration using parallel applications like Oracle RAC or Sybase ASE CE. In this case, data may also be transmitted across the heartbeat network.

Split brain outlined – what is the problem?

A split brain condition occurs when two or more nodes in a cluster act independently, without coordinating the activities with the other nodes. If all cluster heartbeat links fail simultaneously, it is possible for one cluster to separate into two or more subclusters of nodes. Subclusters contain one or more nodes that continue to communicate when heartbeat links fail. In this situation, each individual subcluster would not be aware of the status for the other subcluster. Each subcluster could carry out recovery actions for the departed systems. For example, a passive node can bring online an instance of an application, despite that the application instance is already online on another node. This concept is known as split brain.



Example of a split-brain condition

There are three types of split brain conditions:

Traditional split brain

Given a local cluster, with possible mirroring of shared storage, if no protection is in place, it's very likely that split brain will likely lead to corruption of data. The issue in this scenario is that since shared storage is being used, all cluster nodes can access the data. If an application is running on the primary node, when heartbeats get disconnected a secondary node may attempt to startup that application. In this case more than one node will have the same disk enabled and active at the same time. I/O Fencing can provide data protection against this scenario. Beyond implementing I/O Fencing, other options are available in VCS to avoid traditional split brain such as adding low-priority heartbeat links over a public network segment. These links are not used to transmit cluster data unless all high-priority or standard heartbeat links are unavailable.

Serial Split brain

Given a cluster whose nodes span across separate facilities (campus, buildings), there is a need to ensure that multiple locations accessing the same storage, with a local copy in each facility, continue to function as if they are the only application instance running. A serial split brain occurs when each location uses the local copy of the data and changes made are not copied to the remote storage causing the data to be inconsistent throughout the environment.

This condition can occur when the cluster is configured across two or more sites in a campus cluster configuration with disk mirroring configured between the sites. In this situation, if heartbeat networks and storage connectivity are unavailable between the two sites; the application will be able to go online on both sites simultaneously.

Though it may not corrupt data, it could invalidate the data as each site will have a copy of the data updated by duplicate applications writing local data to separate storage devices which would otherwise be a single unified storage device (mirrored). Since this issue is faced within a single cluster that is distributed, I/O Fencing can determine which subcluster nodes should stay online. The data would be protected and protect against splitting the mirror as only one system would be online at a time so all changes would be on one storage device.

Wide Area Split brain

Having two or more clusters configured for a site-to-site failover scenario is called a wide-area or global cluster. In this setup, individual clusters are configured locally on two or more sites. Manual operations are typically how switch-over functions are initiated in global clusters, although some solutions can be configured to perform automatic failovers. Data is replicated across sites and each site has access to a full copy. Replication can be accomplished from either the server where the application is running or from the storage array itself. Regardless of the method, VCS will be responsible for controlling which site is the source of the data and which site is the replication target. There is an added consideration of how out-of-date the replication state is (known as synchronous and asynchronous replication), but we will not get into this discussion in this paper.

There are two ways for a wide area split brain to occur:

- If using manual failover: The wide-area-heartbeat link between two or more clusters is down, and a System Administrator performs manual operations to bring up applications that actually are already online in another global cluster location concurrently.
- If using automatic failovers: When the wide-area-heartbeat links between two or more clusters go down, the cluster automatically brings up applications on the remote cluster.

As global heartbeats or wide-area-heartbeats are typically deployed over a great distance, it's difficult to get appropriate reliability on the global heartbeat networks. In addition, global clusters are usually deployed for DR purposes. Many companies prefer to have manual DR operations.

In global cluster configurations with VCS and Global Cluster Option, a steward process can be utilized. The steward process is run on a server on a third site to be used when heartbeat communication is lost between the primary site and the DR site. The DR site checks with the steward process when the heartbeats are lost, which is located outside of both the primary and DR sites, to determine if the primary site is down. It does this by ensuring the machine where the steward process is being run can determine the state of the primary sites cluster. The use of the steward process is only needed in configurations limited to 2 sites. For Global Clusters with more than 2 sites, a third site will act as an arbitrator in the same capacity as the steward process.

NOTE: Wide area split brains are not handled by I/O Fencing. I/O Fencing operates at the individual cluster level, and is only for local and campus clusters. This example is not a use case covered in this document but is included to show all possible split brain examples.

What are the most common cases for a split brain to happen?

- Heartbeat networks simultaneously disconnected, dividing cluster nodes into subclusters
- A cluster node hangs. Other nodes expect that the hanging node is down, and start action to prevent downtime (bringing up services).
- Operating System Break/Pause and Resume. If the break feature of an OS is used, the cluster expects that this node is down, and will start actions to prevent downtime. If the OS is resumed soon after, this can introduce risk of a data corruption.
- Some virtualization technologies also support the ability to Pause a running Virtual Machine. If a paused VM was resumed, just like a paused physical machine or partition, it would continue from the point it was paused and it could potentially assume that it is the active cluster node running an application while writing to active disk partitions causing data corruption. Some virtualization technologies help prevent data corruption in this scenario. See the **I/O Fencing protecting virtual environments** section for more information.

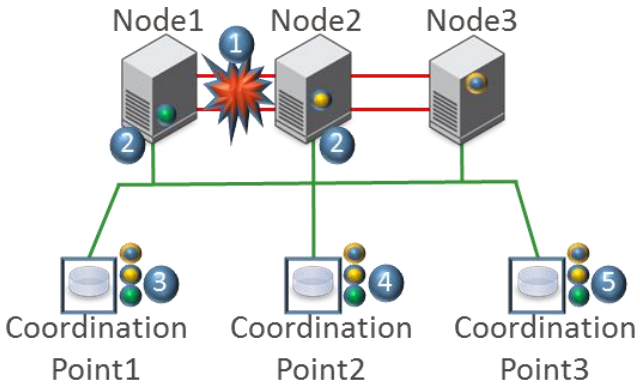
I/O Fencing under VCS should be deployed to protect from all scenarios described above.

General notes on I/O Fencing

I/O Fencing is a core component of VCS focused on properly handling a cluster partition event due to the loss of cluster communication. I/O Fencing consists of two distinct components, Membership Arbitration and Data Protection; together they are able to deliver maximum data integrity in a cluster environment.

Detailed Fencing Race Condition Example

To begin with, let's review how VCS decides which node wins a race. Here is a picture of our environment:



This is a 3-node cluster with 3 coordination points. It does not matter if they are SCSI3 coordinator disks or coordination point servers for our example. When nodes form a cluster they put reservations on each of the coordination points. In our example, the colored balls in the node represent their individual reservation. You will notice that each coordination point has a unique reservation (colored ball) next to them as well. This shows that when the cluster formed, all nodes registered with each coordination point. To be very clear, though it is not shown in this example, when a node registers with a coordination point, they have to do so down all possible paths. This comes into play when disks are used as coordination points and multipathing is being used.
Step1: A fencing event occurs when cluster nodes stop communicating. In our example next to the #1, we see that Node1 is separated from Node2 and Node3. This causes a race condition.

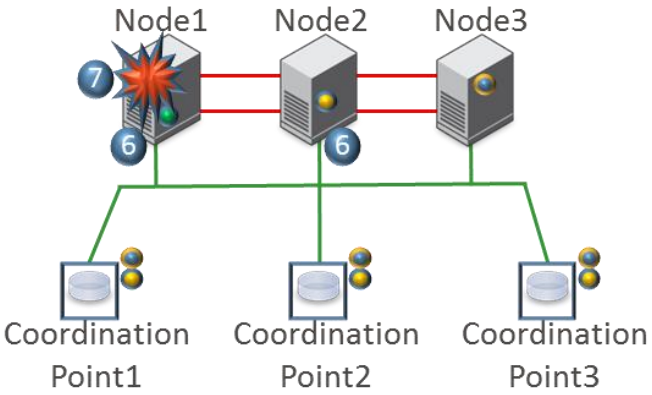
Step2: At this point, each subcluster determines who can still communicate. Node1 cannot talk to any other nodes, so it is a subcluster of 1. Node2 and Node3 can still talk over the heartbeat network so they form a subcluster. When a subcluster of multiple nodes are formed, a racer node is needed. The node with the lowest cluster node number is the racer node, which in our example is Node2.

Step3: Once the racer nodes are determined, then they race for the coordination points. They connect with the coordination point and first determine if they have their reservation still active. If they do, then they remove the reservations from all nodes not in their subcluster. If their reservation is not available then the subcluster lost the race to the coordination point.

Step4: After the first coordination point is completed, the race occurs for the second coordination point.

Step5: Finally, the last coordination point is raced for. There are a few things to note:

1) In a majority of the cases, if a racer node wins the first coordination point, they will win the rest of the coordination points.
2) Preferred Fencing, mentioned later in this document, can influence the race outcome.



Step6: Once the race is complete, the subclusters will both determine if they won at least half of the coordination points. If they did, as is the case for Node2 and Node3, then they continue running and attempt to recover services on the losing subclusters.

Step7: If they did not win at least half of the coordination points then the vxfen driver will cause the node to panic and reboot. Once the heartbeat connections are restored, the node will join the cluster.

Membership Arbitration

Membership arbitration is necessary to ensure that when cluster members are unable to communicate over the cluster heartbeat network, only a single subcluster should continue to remain active. In the event of a cluster split, it also determines which nodes should panic to ensure cluster integrity. The ultimate goal is to have a process that guarantees multiple servers in the same high availability cluster are not attempting to startup the same application at the same time on more than one node in a failover cluster. It should also be done in a timely fashion, so as to limit any chance of data corruption.

Another reason membership arbitration is necessary is because systems may falsely appear to be down. If the cluster heartbeat network fails, a cluster node can appear to be faulted when it actually is not. Limited bandwidth, the OS hanging, driver bugs, improper configuration, power outage or network issues can cause heartbeat networks to fail. Even if no SPOF (Single Points of Failure) exist in the heartbeat configuration, human mistakes are still possible. Therefore, the membership arbitration functionality in the I/O Fencing feature is critical to ensure cluster integrity and the prevention of a split brain.

Membership arbitration with I/O Fencing protects against such split brain conditions. The key components for membership arbitration in VCS are coordination points. Coordination points can be serviced in multiple ways which will be expanded upon in the upcoming chapters. This mechanism is used to determine which nodes are entitled to stay online and which will be forced to leave the cluster in the event of a loss of communication. Typically, the required number of coordination points is an odd number to ensure a clear majority. Most commonly, three coordination points are utilized. This is the case because the winning subcluster must remain active and retain keys or reservations on a majority of coordination points. If a customer was to configure 2 or 4 coordination points, it is possible that both racer nodes could obtain ½ of the coordination points and would result in all nodes to lose the race and panic. Although you can configure a single coordination point, it would be a single-point-of-failure (SPOF) and is only recommended in non-production environments like test and development.
Single coordination points are not supported in Production Environments.

When a cluster node starts up, a component called the vxfen kernel module will register keys or reservations on all coordination points. VCS I/O fencing technologies, called coordination points can be disk devices (coordinator disks), distributed server nodes (coordination point servers, CP Servers or just CPS) or both as the combination of disk and server based coordination is supported.

Data Protection

I/O Fencing uses SCSI3 Persistent Reservations (PR) for data protection. SCSI3-PR allows access to a device from multiple clustered systems down single or multiple paths. At the same time it blocks access to the device from other systems that are not part of the cluster. It also ensures persistent reservations across SCSI bus resets.

Using SCSI3-PR eliminates the risk of data corruption in a split brain scenario by fencing off nodes from the protected data disks. If a node has been fenced off from the data disks, there is no possibility for that node to write data to the disks. The term being fenced off means that a blocker through the SCSI3 protocol has been erected to ensure that access to the device has been prevented.

SCSI3-PR also protects against accidental use of LUNs. For example, if a LUN is used on one system, and is unintentionally provisioned to another server, there is no possibility for corruption; the LUN will simply not be readable or writable. Also, both the coordinator disks and the data disks need to be SCSI3-PR compliant for this protection.

Note: SCSI3-PR protection requires Storage Foundation. It is not supported with native volume managers. SCSI3-PR needs to be supported by the disk array, node architecture and multipathing solution. To determine if an operating system and array vendor are supported with Symantec SFHA, check the Hardware Compatibility List (HCL) for version being implemented. SCSI3 protection is supported with Dynamic Multipathing (DMP) included within Storage Foundation and only other utilities mentioned in the HCL.

Non-SCSI3 Based Fencing

In some environments, SCSI3-PR support is not available. This can be due to multiple reasons: lack of support from the disk array, lack of support from a HBA driver or from the architecture as is the case with some virtual machine technologies. There is also a data protection aspect of Non-SCSI3 Based Fencing. This is implemented through the use of judicious timing. When a fencing race occurs, a minimal time gap is put in place before attempting to bring the application online on the subcluster that wins the race. This is to ensure there is enough time for the losing node to panic and reboot. In environments that do not support SCSI3-PR, Symantec recommends deployment of Non-SCSI3 Based Fencing. For the latest status on SCSI3-PR support, refer to the HCL found here: **https://sort.symantec.com/documents**

NOTE: Although Non-SCSI3 based Fencing greatly reduces the risk of data corruption during split brain scenarios in a VCS environment, it's not 100% eliminated. There is still a small risk of having data corruption. Symantec advises customers to use SCSI3-PR based Fencing if you require 100% data guarantee in a split brain scenario. The use of pause and resume features included in some virtualization technology can lead to data corruption and is not supported with Non-SCSI3 Based Fencing.

Majority Fencing

In VCS 6.2, an additional type of fencing was introduced called Majority Fencing. This was added to support high availability in appliances where there is no option for external dependencies. In all other cases, Symantec recommends customers implement SCSI3 or Non-SCSI3 fencing in their environments. The major drawback with Majority Fencing is that without external coordination points, the process can only determine if a subcluster should remain online based on the number of nodes within it or if it is the leader node.

If the cluster has an odd number of nodes, then the subcluster with more than half will remain online. If a subcluster has less than half of the nodes, then it will panic itself.

If the cluster has an even number of nodes, then the lowest cluster node number that is active is called the leader node. If the subcluster has more than half of the nodes, then it acts the same way as the odd number cluster; the subcluster with less than half will panic itself. If there is a tie and both subclusters have exactly half of the nodes, then the subcluster with the leader node survives.

To understand this option, we should work through a few examples:

1) 3-node cluster with a single node disconnected heartbeats
2) 4-node cluster splitting into 2 equal subclusters
3) 3-node cluster with all nodes unable to communicate

Scenario 1: 3-node cluster with a single node disconnected heartbeats

In this scenario there are an odd number of cluster nodes so a subcluster would need more than half to remain online. Looking at the figure, node 1 cannot communicate with any other node. Because it does not have more than half it will panic itself. Nodes 2 and Node 3 combine into a subcluster because they can continue to communicate over the heartbeat link. Since they have 2 members, which is more than half of 3, they remain online. Because there are an odd number of cluster nodes active at the start, the leader node concept does not apply.

Scenario 2: 4-node cluster splitting into 2 equal subclusters

In this scenario there is an even number of cluster nodes. We begin by determining if either subcluster has more than half. In this instance, there are two equal subclusters so we look to the lowest cluster node. If the cluster number is 1 through 4 just like the names then the lowest node would be Node 1. In this scenario, the subcluster with Node 1 and Node 2 would remain online while the subcluster with Node 3 and Node 4 will panic themselves.

Scenario 3: 3-node cluster with all nodes unable to communicate

In this scenario there are an odd number of cluster nodes so a subcluster would need more than half to remain online. Looking at the figure, all nodes form their own subcluster and cannot communicate to each other. Because there is no subcluster with more than half of the node, all nodes will panic themselves. The leader node concept is not used as there are an odd number of cluster nodes to begin with.

Note: If Scenario number 2 had Node 4 shutdown prior to the communication issue, although there are four nodes in cluster, it would have acted in the same manner as Scenario 1 because the number of active nodes in the cluster at the start of fencing event was an odd number.

Protection Examples

Membership arbitration and data protection offer different levels of protection. Arbitration alone doesn't give a 100% guarantee against corruptions in split brain scenarios. Consider these scenarios:

1. System Hangs. If a system is hung, VCS will interpret this as Node Faulted condition and will take action on another cluster node to ensure application availability
2. Operating System break/resume used
3. Very busy cluster nodes can start dropping heartbeats

All of these scenarios are rare, but they do happen. Let's consider a sample scenario in the following table to see what happens when I/O Fencing is disabled, I/O Fencing with SCSI3-PR is enabled and Non-SCSI3 Fencing is enabled:

	Node 1	Node 2	Result
System Hang with no I/O Fencing Enabled	Node 1 is Hung	Node 2 detects the loss of communication and starts up application	Application is online on both nodes. The risk is data corruption as both Node 1 and Node 2 have the disks mounted. If Node 1 flushes its buffers while Node 2 is writing to the disk then the data could be corrupted.
System Hang with SCSI3 Fencing Enabled	Node 1 is Hung and is fenced out of the cluster. With SCSI3 protection, once the node is out of the cluster and the SCSI3 keys are removed, it can no longer flush its buffers to the disk.	Another cluster node detects the loss of communication and begins a fencing race. As the Node 1 is hung, Node 2 wins the race and brings the application up	Since Node 1 lost the race, the box is panicked and the application is running on Node 2. Once Node 1 loses the race and its keys are removed it cannot access the data disks without rebooting. Once Node 1 comes back online and has its heartbeat communication reestablished, it can join the cluster and access the disks.

	Node 1	Node 2	Result
System Hang with Non-SCSI3 Fencing Enabled	Node 1 is Hung and is fenced out of the cluster. When the cluster recognizes the loss of communication it attempts to race and determines that it lost the race and will panic.	Another cluster node detects the loss of communication and begins a fencing race. As the Node 1 is hung, Node 2 wins the race and brings the application up	Since Node 1 lost the race, the box is panicked and the application is running on Node 2. Non-SCSI3 Fencing does not put a lock on the disk like SCSI3-PR does.

Comparison between Coordinator Disk and Coordination Point Server

We will discuss reasons for implementing each technology and sample architectures more thoroughly later in the document. This chart includes information to assist in deciding which I/O Fencing technology to implement:

	Coordinator Disk	Coordination Point Server
Communication	SAN Connection using SCSI3-PR	Network to a CP Server
Benefits	• SCSI3-PR based data Protection • Disk based Membership Arbitration • SANs are generally more reliable than IP Networks • 100% Guaranteed Data Protection • Only cluster members can access data disks • During a fencing race, the losing subcluster has disk access prevented	• Basis for Non-SCSI3 Fencing • Network Membership Arbitration • I/O fencing for environments that do not support SCSI3-PR (e.g. some virtualization platforms or some storage arrays) • Used in conjunction with SCSI3-PR disks to help with Campus Clusters split site concern • CP Servers can serve as a coordination point for up to 2048 individual cluster nodes
Drawbacks	• Dedicated LUN per Coordinator Disk • Some low-end storage arrays do not support SCSI3-PR • Some Virtualization technologies do not support SCSI3-PR	• Requires an additional server to run the CPS process • When only CP Servers are used, it does not provide the guaranteed data protection in SCSI3-PR
Primary use case	• Need Guaranteed Data Protection	• Campus Cluster across two sites • In virtual environments where SCSI3-PR is not supported

VCS Agents related to SCSI3 Protection

SCSI3 Persistent Reservations for failover DiskGroups:

The VCS DiskGroup agent is responsible of setting the SCSI3 Persistent Reservations on all disks in the managed failover diskgroup.

NOTE: Do not import the DiskGroup manually, and then enable the DiskGroup resource. If the MonitorReservation attribute is set to false (default), the DiskGroup resource will be reported as online; however no Persistent Reservations are present to protect the DiskGroup. If the MonitorReservation attribute is set to true, the DiskGroup resource will be faulted.

SCSI3 Persistent Group Reservations Shared DiskGroups (CVM/CFS)

Protection works slightly different for shared DiskGroups. Shared DGs are imported during the cluster join process for each node and reservations are set at that time. The difference between the two is when the Persistent Reservations are set and if the DG resource is responsible for placing keys on the disks. Persistent Reservations are set on shared DGs. This is necessary to control concurrent access to the DiskGroup. Regular persistent reservations cannot be used for this purpose. However, this is nothing you need to configure. VCS will set appropriate reservations based on the agent being used. If a new shared DiskGroup is created, reservations will be set when the DiskGroup is imported.

VCS Agents and Attributes related to I/O Fencing

DiskGroup Agent notes and attributes

The DiskGroup VCS agent sets reservations on all disks in the diskgroup during online process/import. When a DiskGroup resource is brought online by VCS, and SCSI3-PR is enabled (UseFence=SCSI3 in the VCS main.cf configuration file), Persistent Reservations will be set on the disks.

MonitorReservation attribute

Symantec has noted that some array operations, for example online firmware upgrades, have removed the reservations. This attribute enables monitoring of the reservations. If the value is 1, and SCSI3 Based Fencing is configured, the agent monitors the SCSI reservations on the disks in the disk group. If a reservation is missing, the monitor agent function takes the resource offline. This attribute is set to 0 by default.

Reservation attribute

The Reservation attribute determines if you want to enable SCSI3 reservation. This attribute was added in VCS 5.1 SP1 release to enable granular reservation configuration, for individual disk groups. This attribute can have one of the following three values:

ClusterDefault (Default) - The disk group is imported with SCSI3 reservation if the value of the cluster-level UseFence attribute is SCSI3. If the value of the cluster-level UseFence attribute is NONE, the disk group is imported without reservation.

SCSI3 - The disk group is imported with SCSI3 reservation if the value of the cluster-level UseFence attribute is SCSI3.

NONE - The disk group is imported without SCSI3 reservation.

CoordPoint Agent notes

The CoordPoint Agent is used to monitor the state of your coordination points, regardless if they are Disk or CP Server based. Customers typically configure this agent within their cluster to ensure that the coordination points are currently active. Any issue with the coordination points will be logged in the engine_A.log and if notification is enabled a message will be sent. This agent will be automatically included in the customer's configuration during fencing setup based on customer preferences. To enable fencing using the command: # */opt/VRTS/install/installvcs<Version> -fencing*

The FaultTolerance attribute determines when the CoordPoint agent declares that the registrations on the coordination points are missing or connectivity between the nodes and the coordination points is lost. See the VCS bundled agent guide for more information on implementation.

Here is an example of the agent configured within a main.cf cluster configuration:

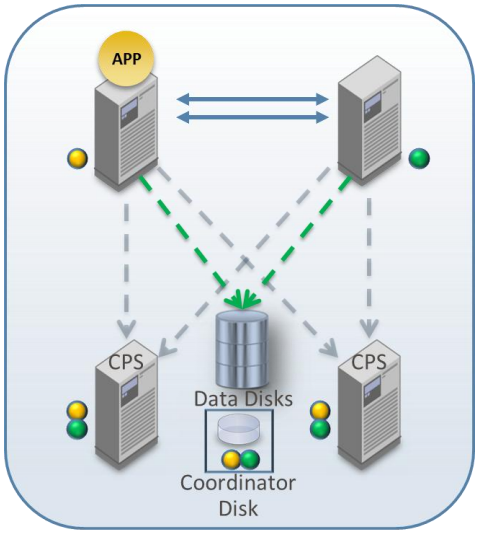
```
group vxfen (
        SystemList = { sysA = 0, sysB = 1 }
        Parallel = 1
        AutoStartList = { sysA, sysB }
        )

CoordPoint coordpoint (
        FaultTolerance=0
        )
```

Recent Enhancements to I/O Fencing

With the release of VCS 6.1, customers can choose the order in which coordination points are accessed. This allows customers to determine if a specific 3rd site CPS is used prior to other coordination point or to choose the coordinator disk prior to racing for the CP Server. Another change in the 6.1 version is that all communication to the CPS is secure using the HTTPS protocol automatically. A CPS installed at version 6.1 and above can communicate to both pre-6.1 clusters using the older IMP protocol and using https for clusters at VCS version 6.1 and beyond.

In VCS 6.0, CPS was enhanced to provide the ability to use multiple networks. This functionality was included to allow the network based coordination point to have the same multipathing capabilities currently enjoyed by the Coordinator Disks. The Coordinator Disks can use Dynamic Multipathing (DMP) to ensure that the loss of a single path would not prevent a node from using I/O Fencing. As most customers have multiple networks in their environment (backup, maintenance, heartbeat, public, etc.), connecting the CPS to these networks to provide redundancy and reducing the single points-of-failure is an advantage.

Coordinator disks and coordination point servers can be used together to provide I/O Fencing for the same cluster. This can only occur when using SCSI3-PR based fencing. We can see an example of this configuration in the following diagram.



This diagram is an example of a 2-node VCS cluster configured with 3 coordination points. The green and yellow balls on the Coordinator Disk each represent the SCSI3-PR Fencing Key for each individual node while the green and yellow balls on the CP Servers represent each nodes registration. When there are keys and registrations on the coordination points, then a node is permitted to join the cluster, which is represented by the specific colored ball next to each of the cluster nodes. When at least one Coordinator Disk is used, SCSI3-PR based fencing is in active. Customers can choose to enable this protection only for membership arbitration if desired. To enable SCSI3 protection, the data disks will also need to be SCSI3 compliant.

2-node cluster using SCSI3 Fencing disk along with coordination point servers

I/O Fencing can be enabled for all environments

Protecting data and ensuring application availability is a main concern for all customers. Data availability can be compromised in several ways.

With the introduction of Symantec Cluster Server 5.1, clusters can be protected using a server-based I/O Fencing mechanism. While VCS can be configured to run without I/O Fencing, Symantec strongly recommends that I/O Fencing is configured for all VCS clusters and specifically all parallel applications like CFS and Oracle RAC to prevent split brain conditions and data corruption.

Preferred Fencing

At the time of a heartbeat network interruption, the fencing driver for each subcluster (or groups of cluster nodes that can still communicate) races for each of the coordination points. The subcluster that obtains a majority of the coordination points keys or reservations survives whereas the fencing driver causes a system panic on nodes from all other subclusters whose racer node lost the race.

By default, the fencing driver favors the subcluster with maximum number of nodes during the race for coordination points. If the subclusters are equal in number, then VCS will decide based on the order in which the nodes form and join the cluster, which may appear to be arbitrary.
Note that this behavior doesn't take VCS service groups or applications into consideration. It is possible that a passive node survives, and that an active node is fenced off and will panic, leaving the passive node to possibly go to an active state. This would cause application downtime unnecessarily.

With the introduction of Preferred Fencing in VCS version 5.1SP1, VCS introduced the ability to be configured to favor one of the subclusters using predefined policies. If the preferred node does not have access to the Coordination Points, it will lose the race regardless of the Preferred Fencing settings.
Preferred Fencing is controlled by the PreferredFencingPolicy attribute, found at the cluster level. The following values are possible for the PreferredFencingPolicy attribute:

Disabled – (default) Use the standard node count and node number based fencing policy as described above. Preferred fencing is essentially disabled.

Group – Enables Preferred Fencing based on Service Groups. Preferred Fencing using Service Group priority will favor the subcluster running with most critical service groups in an online state. Criticality of a Service Group is configured by weight using the Priority attribute of the SG.

System – Enables Preferred Fencing based on Cluster Members. Preferred Fencing using System priority will prioritize cluster nodes based on their Fencing Weight. For example, if one cluster node is more powerful in terms of CPU/Memory than others or if its location makes it higher priority, then we can give it a higher priority from a fencing perspective. The FencingWeight attribute is used to set the priority for each individual node. The calculation for determining the winner of the fencing race is done at racing time and would be combining all the FencingWeight values for the subcluster and comparing the racer nodes to determine who should win the race.

Site – Enables Preferred Fencing based on Sites. This policy is enabled in version 6.1 with the introduction of the Multi Site Management capability within VOM.

Note: Giving a Node, Service Group or Site priority does not mean that it will win the race. If the subcluster loses access to the SAN or the network and is unable to obtain the fencing reservations, it will not be able to win the race. Preferred Fencing is giving the subcluster with the preferred systems, service groups or nodes on the preferred site a head start in the race but it does not guarantee a winner.

Deploy I/O Fencing

Considerations for I/O Fencing in your environment

1. Choose coordination point technology or if you will use multiple technologies.
2. Decide where to put your coordination points in your environment.
3. Determine if you will use SCSI3 Persistent Reservations to fully protect your data from corruption. This option is recommended where possible, however there are cases when SCSI3-PR cannot be deployed, or where it doesn't make sense. In this case deploy Non-SCSI3 Fencing. Basically, if your environment supports SCSI3-PR, you should have it enabled.
4. (Optional) Determine if any applications (service groups) or any cluster nodes will have priority over others in a racing condition. This point is determining implementation of Preferred Fencing.

Choosing Coordination Point technology

This section contains general notes and guidelines regarding coordination points.

Disk-based coordination points

Fault tolerance for the coordination disks is vital to cluster stability and integrity. Coordination disks should be placed on enterprise-class storage and have an appropriate level of failure protection, with the use of hardware or software RAID, to ensure their ability to withstand disk failure. Each Coordinator disk is only required to have at least 64MB of storage, though some customers have limits on the size of the LUNs presented to the cluster nodes so there is no upper limit on the LUN size. Generally customers choose to have LUNs around 128 MB to ensure there is enough space for a 64MB header on the LUN, but this is not required for proper functionality as the header size can be reduced.

From disk performance point of view, you don't need high-performance disks. Availability is much more important that performance.

Pros:

- SAN networks are usually more stable than IP networks.
- No need to purchase or manage hardware & software for the CPS.
- Guaranteed Data Protection against corruption due to split brain.
- Proven technology as it has been available with VCS for over 10 years.

Cons:

- Each cluster requires three unique coordination LUNs. These LUNs can be small; however they cannot be shared between clusters. If multiple clusters are deployed, this can be quite expensive with the amount of disk consumed.
- Some environments have a single LUN size for an array. With these LUN sizes in the gigabytes, a lot of capacity is wasted as a very small amount of storage is needed for the coordinator disk.
- Required SCSI3-PR supported Storage and infrastructure. Though most enterprise arrays have this support today, not all do.
- Some virtual environments do not support SCSI3-PR or there may be limitations.
- In a Campus Cluster configuration there is an issue with one site having 2 Coordinator Disks and the second site having just one. The site with just one Coordinator Disk will be unable to come online if a full site failure occurs on the primary site. This can be overcome with the use of a single CPS on a third site for the third coordination point.

CP Server based coordination points

The CP Server process addresses some of the negatives for the Disk based coordination points as it can be utilized by more than one cluster at a time. The recommendation with CPS is to implement the server on reliable hardware and over multiple reliable networks.

Pros:

- Up to 2048 Cluster Nodes can share the same CP server.
- No waste of disk space from the disk array.
- Can be mixed with SCSI3 disks or used exclusively in Non-SCSI3 based Fencing.
- Clusters running in virtual environments.
- Replicated Data Clusters and Campus Clusters are fully supported.
- The installation of CPS comes with a single-node VCS license to protect its functionality. This is for non-failover environments only.
- CPS supports multiple IPs to withstand a network segment failure

Cons:

- Additional hardware/software to manage (CP servers), though it can be run on virtual machines to limit the amount of required resources.
- IP networks are usually less stable than SAN networks.
- Does not have the guaranteed data protection mechanism that is in SCSI3-PR if using CPS exclusively with Non-SCSI3 Fencing.

A combination of CP Servers and Coordinator using SCSI3-PR

CPS was originally designed to help with campus clusters site-bias. In an environment with 3 coordination points are required and only 2 sites, there will be a bias for one site. If the alternate site is the only one available, it still will not have ½ of the coordination points and would require manual intervention to be brought online. This is explained in more detail in further chapters.

Pros:

- Combines the Pros of both options
- Having both technologies allows for cluster to validate their access to the SAN and IP network in order to gain access to the coordination points.

Cons:

- Cannot be used in some virtual environments that do not support SCSI3-PR
- Requires storage array that supports SCSI3-PR

Majority Fencing

As majority fencing was designed for appliances where there is no possibility of external resources, Symantec recommends that customers only use it in this scenario.

Pros:

- Enables fencing without external coordination points
- Support HA in appliances

Cons:

- Can result in all nodes in the cluster panicing
- Without external coordination points, there is no way to ensure functional nodes stay online.

Choosing Coordination Point Placement

The first consideration in I/O Fencing is the placement of the coordination points. Where you place the coordination points will influence your choice of coordination point techniques (Disk based or CP Server based). Placement is dependent on the physical infrastructure and the number of physical sites available.

Analyze possible failure and disaster scenarios. If you have two disk arrays, the recommended configuration is to use a coordinator disk in each array and put a third coordination point on a CP Server to remove the requirement to have 2 coordination points in the same array. Remember that a majority of the coordination points need to be available during a failure scenario for a subcluster to remain online.

Deploying I/O Fencing

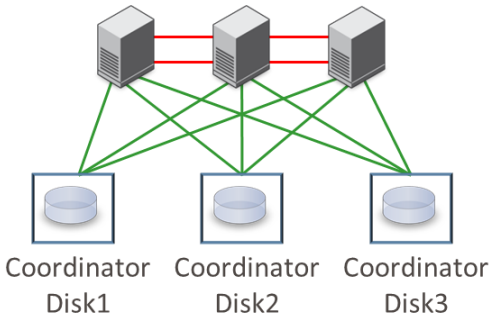
I/O Fencing is usually deployed using the CPI (Common Product Installer – installsf or installvcs scripts). This operation can be performed during the initial installation of the cluster or at a later stage using # */opt/VRTS/install/installvcs<version> -fencing*. When deploying I/O Fencing with only disk based coordination points, SCSI3 Persistent Reservations are enabled by default. If you have one or more CP servers available, the CPI script will ask if you want to enable SCSI3-PR. In most cases, it's recommended to have SCSI3-PR enabled if possible. The only time you should disable it is when your environment doesn't support SCSI3 ioctls. The CPI script asks explicitly whether the environment supports SCSI3-PR. To validate if a disk is SCSI3-PR compliant, use the vxfentsthdw utility. WARNING, this utility should only be used on blank LUNs or LUNs to be used as coordinator disks. This utility will destroy data on the disk so **be careful** when using it.

NOTE: Storage Foundation for Oracle RAC doesn't support Non-SCSI3 fencing.

Deploying Preferred Fencing (optional)

Preferred fencing provides three different levels of prioritization – System-based, Group-based and Site-based. In all cases, an internal value known as Node Weight is used. Depending on how Preferred Fencing is deployed, the Node Weight is calculated differently. Preferred Fencing is not required to implement SCSI3-PR fencing or Non-SCSI3 Fencing. It is optional in all cases.

To deploy Preferred Fencing modify the cluster-level attribute PreferredFencingPolicy based on the race policy previously discussed. If it is set to Disabled, then preferred fencing is disabled. If the value is set to System, VCS calculates node weight based on the system-level attribute FencingWeight. When the Policy is set to Group, VCS calculates node weight based on the group level attribute Priority for those service groups that are in the Online or Partial state and have their Priority set. This value will only be set to Site if the Multi-Site Management utility is used in VOM to define nodes and arrays belonging to a specific site.



3-node cluster using SCSI3-PR with 3 coordinator disks

Regardless if Preferred Fencing is using Group, System or Site, the fencing calculation works the same. The total fencing weight for the subcluster will be the combined fencing weight from all subcluster members. In the above example, the fencing weight from Node2 and Node3 will be combined for the fencing weight from the subcluster and used by the racer node. The subcluster with the larger fencing weight has the preference and should win the race.

CP Server considerations

This section contains general considerations for CP Server deployments.

CP Server scalability requirements

The maximum number of clusters for one single CP server is relative to the number of nodes within the clusters. During performance testing, a single CPS can comfortably maintain 2048 nodes. This could be (1024) 2-node clusters,(512) 4-node clusters or any combination of cluster nodes that equal 2048. Since the primary objective of the CPS is to respond to fencing race conditions as quickly as possible, if all nodes in all

clusters could not communicate over the heartbeat network, all nodes would be in a fencing race asking for immediate response from the CPS. With 2048 nodes racing at one time, all requests would be satisfied.

The CP Server can run on any UNIX and Linux OS supported by VCS as the software comes as part of the SFHA distribution. As a note, 128 4-node clusters will require a CP Server database of approximately 5 megabytes of storage space and with all nodes racing at once, only took up less than 1% of a single CPU on a virtual machine.

Clustering the CP-Server itself

Clustering of the CP Server itself is not required, however in an enterprise class environment, the availability of each CP Server is crucial to ensuring overall data integrity, application availability and the proper functioning of VCS with I/O Fencing.

In those situations, it makes sense to cluster the CP Server process. Using VCS to cluster the CP Server is free of charge in a one-node cluster non-failover configuration. One qualification for this free one-node VCS license is that no other applications are clustered other than Veritas Operations Manager (VOM) and/or CPS, as they can coexist on the same single-node cluster. If the CPS is in a failover configuration with 2 or more nodes or other applications are being protected, then a full VCS license is required.

Can coordination point servers be included in other VCS clusters that currently are hosting production applications? The answer is yes as long as there is only one instance of CPS per cluster. In this case, four individual CP clusters are recommended, as each cluster cannot use "itself" as a coordination point. This setup is covered in **Scenario 4 – Replacing Coordination Disks with CP Servers**.

I/O Fencing protecting virtual environments

With the continued migration in enterprise computing from physical to virtual, along with application uptime, data protection continues to be a concern. Virtual environments have their own challenges to both data protection as well as membership arbitration. In this section we will discuss specific virtualization concerns and examples for VMware and other UNIX/Linux virtualization technologies.

I/O Fencing considerations in VMware

There are multiple considerations when configuring I/O Fencing with VCS in VMware including:

1) What types of virtual disks are used (RDM-P, VMDK, iSCSI with the initiator in the VM)?
2) What is the desired configuration (CFS, failover VCS)?
3) If using VMDK disks, what is the Multi-Writer flag set to (On or Off)?

A large number of VMware customers utilize VMDK disks for their configurations. If the Multi-Writer Flag is turned off then VMware disk management will only allow one VM to access the VMDK at a time. The VCS VMDK Disk Agent controls the attach/detach process in moving disks between cluster nodes. In this scenario, VCS I/O Fencing can be used for membership arbitration to control which subcluster remains online when heartbeat communication is lost but is not needed for data protection as the architecture supplies this natively.

If the Multi-Writer flag is turned on in a cluster file system configuration, then VCS I/O Fencing is needed for data protection and membership arbitration. In this case, all nodes in the cluster can access the VMDK disks concurrently. If a split-brain were to

occur, then data corruption is possible without the fencing. In this scenario, SCSI3 protection is not available since VMware does not support SCSI3 on the VMDK disk format.

SCSI3 protection is the only 100% guarantee against corruption in parallel disk access environments. Symantec Storage Foundation is required to implement SCSI3 protection. In VMware you have 2 options that support SCSI3:

1) Use RDM-P disks. There are multiple limitations associated with this format:
    a. There can only be one node per cluster on each ESXi. More than one node in the same cluster cannot run on the same physical hardware.
    b. vMotion is not supported.
2) iSCSI disks
    a. Both the coordinator disks and data disks need to use this format.
    b. vMotion is supported but as the storage is not under VMware control, Storage vMotion is not supported.
    c. All access is done through the virtual network interface.

I/O Fencing considerations in LDOM

I/O Fencing in this environment is slightly different from VMware. There is no native data protection mechanism built into the virtualization architecture so implementing I/O fencing provides even more benefit. With LDOM, SCSI3 is supported in layered fencing configurations.

Virtual devices backed by DMP devices in I/O domains can also support SCSI3. This support is with version 6.0.5 and higher along with VxVM version 6.1.1 or higher applied on the primary and alternate I/O domains.

Symantec recommends you disable I/O fencing if you exported the same physical device to multiple guest domains. If you use fencing on LDOMs, only 1 node of the cluster LDOM can be on each physical machine. Use coordination point servers if you need more than one node LDOM of the same cluster on the same physical machine.

Customers can install VCS inside of the control domain to manage the virtual machines and a separate cluster between guest domains to manage applications. A limitation of this configuration is that configuring I/O Fencing in both environments is not supported. In addition, it is required to have coordinator disks and data disks on different HBAs if more than one LDOM will run on the same physical host. These disks must be full disks and support SCSI3 on the Hardware Compatibility List (HCL).

If SCSI3 protection is enabled in a control domain cluster, then ensure the ReregPGR attribute is enabled. If customers use the `hagrp -migrate` command to live migrate an LDOM and the ReregPGR attribute is enabled then VCS will reregister the SCSI3 after the migration is complete.

I/O Fencing considerations in AIX DLPARs

I/O Fencing with VCS in on AIX DLPARs does not require additional considerations. DLPARS with NPIV have direct access to the disks and fully supports SCSI3-PR. If the storage is on the HCL showing support for SCSI3 then it can be used with VCS I/O Fencing SCSI3 protection.

Virtual SCSI (vSCSI) disks can be a SCSI disk, or a volume/file on a VIO Server. SCSI3 is not supported with vSCSI disks but I/O Fencing is supported using Coordination Point Servers.

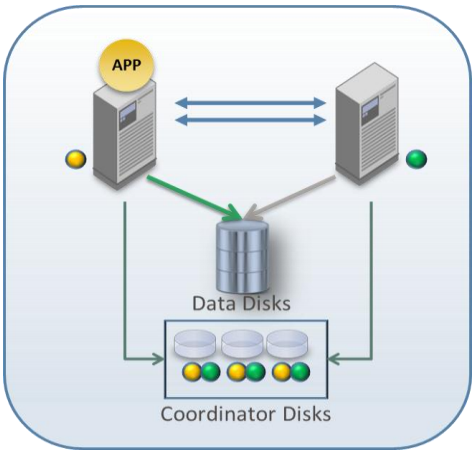
I/O Fencing considerations in linux virtualization

VCS fencing does not support SCSI3 for in-guest clusters in OVM but Non-SCSI3 I/O Fencing configuration using CPS is supported. SCSI3 disk-based fencing is also not supported in RHEV environments. With the support of SFCFSHA in RHEV on hypervisors, we support SCSI-3 fencing on the RHEL-H (and KVM host, too). For pure virt-to-virt clustering, we support fencing in KVM VMs and for RHEV VMs it's planned for 6.2.1

I/O Fencing Deployment Scenarios

To understand each scenario, we have developed diagrams relating to the example. Each picture is of a sample VCS cluster configured with 3 coordination points. The yellow and green circles on the CP Servers represent registrations and for the Coordinator Disks each represents SCSI3-PR keys. When there are registrations on all of the coordination points, then the node can join the cluster. When at least one Coordinator Disk is used, SCSI3-PR based fencing is in use.

Scenario 1: All nodes in the same Data Center using Disk based coordination points.

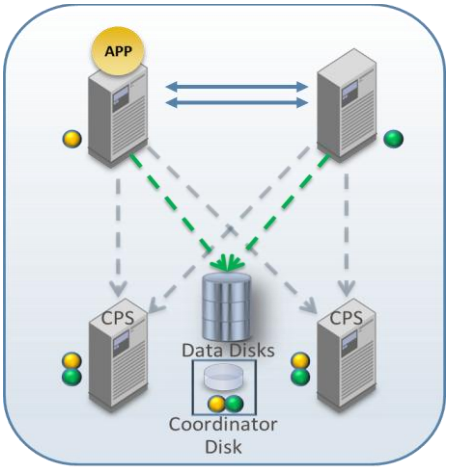


This is the simplest scenario to visualize, suitable for customers who are deploying clustering at the same location. In this scenario, you can place all coordination points on LUNs in the single array. This is the most common configuration as it was previously the only option customers had to receive SCSI3-PR protection prior to VCS 5.1.

Three LUNs are used to create a coordinator disk group after validating SCSI3-PR compatibility with the vxfentsthdw command. For more information, see the VCS Administrator's Guide.

2-node cluster utilizing SCSI3-PR Fencing with only SCSI3-PR Coordinator Disks

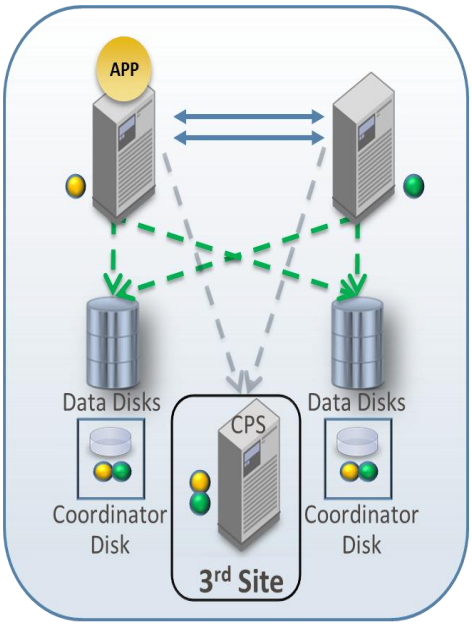
Scenario 2: All cluster nodes in the same datacenter, while reducing the amount of storage used for coordinator disks



In this scenario, the goal is to ensure that Single Points of Failure are reduced in the configuration while reducing the amount of storage used for coordination points. This configuration has the two of the three coordination points as CP Servers while it continues to provide SCSI3-PR data protection. Each CP Server can service up to 2048 cluster nodes. This configuration reduces the amount of disk space used for coordination points, while still providing data protection and membership arbitration with server based fencing. Each CPS can be protected using VCS in a single-node cluster to ensure the CPS process remains online.

2-node cluster using SCSI3-PR Fencing with (2)CP Servers and (1)Coordinator Disk

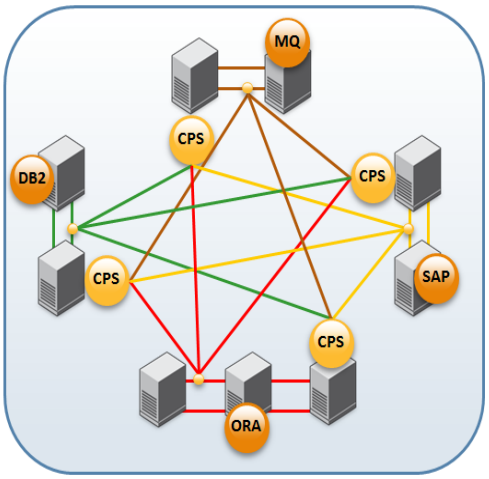
Scenario 3: Campus Cluster Configuration using three sites



Campus Cluster environment with two main sites and a third site to host a CP Server to act as an arbitrator in case there was a full site disconnection. In this scenario, the cluster is stretched between two main sites. Campus cluster requirements apply here, and those can be found in the VCS Administrator's Guide. Alternately, customers could put one coordination point in each of the two sites, and a third coordination point in a remote site. The downside to this change in config is the lack of SCSI3-PR data protection. The CP Server was originally designed to protect this exact scenario. When first introduced, campus cluster were configured to favor one site. The config would include 2 or a majority of the disks on the primary site. If there was a site failure, the secondary site would not be able to come online because it would not have access to more than half of the coordination points. The CPS on the 3rd site resolves this issue.

Campus cluster using SCSI3 Fencing (a Coordinator Disk on each site) and a CPS on a 3rd site

Scenario 4: Replacing all coordination disks with CP servers – Availability

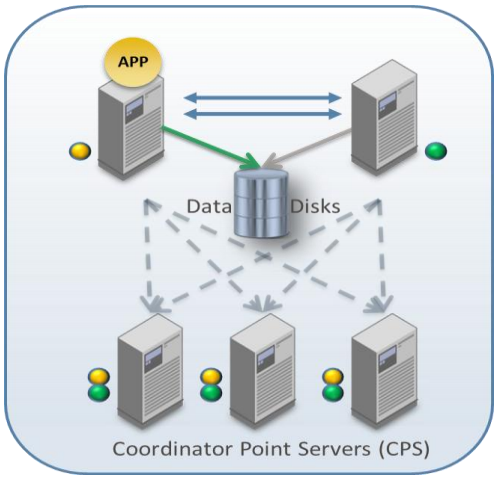


If a customer determines to utilize only CP servers for coordination points, availability of the CP servers is crucial. Guaranteeing availability for the CPS can be done with VCS as described earlier. In this scenario, the cluster environment is located on one site. Each coordination point server will be made highly available within a production failover cluster. In this configuration 4 CP Servers would be needed. Each cluster requires access to 3 CP Servers but they are unable to use the CPS contained within their own cluster configuration.

4 CP Servers throughout an environment to provide I/O Fencing to a distributed HA configuration

It is recommended to spread out the CP Servers throughout the environment to reduce Single Points of Failure (SPOF). Also, it is recommended to not have more than one of the CP Servers on the same VMware ESXi node in a VMware environment. To be clear, it is not a requirement to host the CPS within production environments. This is a suggestion to reduce the cost of clustering the CPS. It could be run on the failover node in an Active-Passive configuration or an N+1 cluster configuration or in using development or test clusters with high uptime levels to house the CP Servers.
Also, if you choose to have each CPS in a non-failover single-node cluster, that would also work just fine. For the purposes of reducing SPOFs, the suggestion is to have them included in failover clusters. To reduce the number of servers hosted in the CPS environment, see the next scenario on CPS Flexibility.

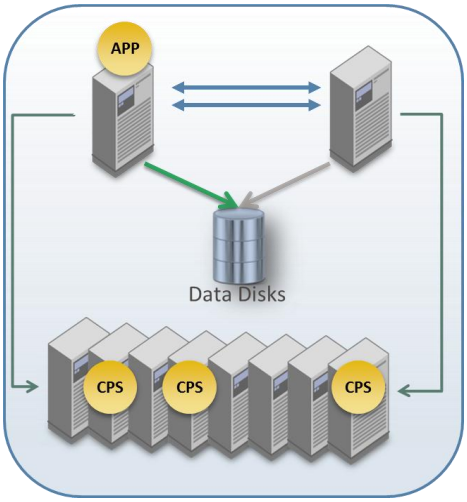
Scenario 5: Replacing all coordination disks with CP servers – Flexibility



If an enterprise determines it will replace all coordination disks with CP servers, and computing resources are scarce, each CPS can run in one-node cluster as described earlier. VCS is provided to ensure that the CPS application remains online, guaranteeing availability and access to the VCS cluster nodes. A Single-Node VCS license is supplied for free to protect the CP Server process in a Non-Failover environment as long as it and/or VOM are the only applications protected by VCS.

2-node cluster with 3 CP Servers as the coordination points

Scenario 6: Replacing all coordination disks with CP servers – Virtual Environment



With the continued adoption of virtualization many customers are interested in a fencing solution that works for both virtual and physical environments. Putting 3 CP Servers in a virtual environment allows for minimal resource usage while fulfilling their fencing requirements. There are a couple of considerations to keep in mind when architecting this configuration:

1) CP Servers should not run on the same physical host at the same time to avoid Single Points of Failure. This can be accomplished through Anti-Affinity rules wherever possible in virtual environments.

2) Each CPS should be protected by Single-Node VCS to ensure the process stays online.

2-node cluster with 3 CP Servers as the coordination points in a Virtual Environment

Coordination points availability considerations

Disk-based Fencing:

Coordinator disks should be placed on enterprise class storage, with appropriate RAID levels. Note that high performance is not required for the coordination LUNs, as no data resides on those. However, availability is crucial, so make sure to choose appropriate protection in the disk arrays for these LUNs.

Symantec recommends “the smallest possible LUNs” for the coordination disks. Note:

- With the vxfentsthdw command, 1MB LUNs are minimally required though 128 MB or more is recommended to ensure space for a 64MB disk header.
- For EMC arrays, the host based software may interpret smaller LUNs (smaller than 500mb) for command devices.

One Coordinator diskgroup per cluster is created regardless if using one or all three coordinator disks as coordination points. This diskgroup is deported and no volumes or mirrors should be created in the DG on the coordinator disks. Basically, one empty disk within a DG should be used for each coordination point. 3 LUNs in the Coordinator DG would equate to 3 disk-based coordination points.

It's a requirement to have Storage Foundation when using disk-based coordination points. If the diskgroup has been imported, make sure to deport it using the command
vxdg –t deport
When disk-based coordination points are used, if even used in combination with CPS, SCSI3-PR is enabled by default.

Server-based Fencing:

Coordination Point Servers cannot be located in a single VCS cluster. A single CPS instance can run on a server at a time, so more than one instance within a cluster is not supported.

CP Servers can run within a Virtual Machine. It is not recommended to house more than one CP Server on a single VMware ESXi host to prevent a Single Point of Failure (SPOF).

In Conclusion, using Disk-based I/O Fencing with SCSI3-PR, or Server-based I/O Fencing with Non-SCSI3 Fencing using CPS or a combination of both together, VCS enables Data Protection in your mission critical computing environments. Symantec recommends either disk-based or server-based fencing and advises them not to use Majority Fencing.

Last Updated December 2014.

About Symantec

Symantec is a global leader in providing security, storage and systems management solutions to help businesses and consumers secure and manage their information. Headquartered in Mountain View, Calif., Symantec has operations in 40 countries. More information is available at www.symantec.com.

For specific country offices and contact numbers, please visit our Web site. For product information in the U.S., call toll-free 1 (800) 745 6054.

Symantec Corporation
World Headquarters
350 Ellis Street
Mountain View, CA 94043 USA
+1 (408) 517 8000
1 (800) 721 3934
www.symantec.com

12/14