Implementing vSphere Metro Storage Cluster using HPE 3PAR Peer Persistence
## Contents

- Introduction .................................................................................................................. 4
- Terminology .................................................................................................................. 4
- Features and benefits ................................................................................................. 4
- Requirements ............................................................................................................... 5
  - Host Persona 11 ......................................................................................................... 5
    - Verify each VV has same WWN on both arrays ...................................................... 6
    - Verify VLUN status on both arrays ...................................................................... 6
    - Quorum Witness ..................................................................................................... 6
  - Management port on HPE 3PAR StoreServ .............................................................. 7
- VMware vSphere Metro Storage Cluster (vMSC) Configuration .................................... 7
  - vMSC uniform ......................................................................................................... 7
  - vMSC non-uniform ................................................................................................. 7
- Planned switchover ...................................................................................................... 8
- HPE 3PAR Peer Persistence failures handling .............................................................. 9
  - Array to Array communication failure .................................................................... 9
  - Single site to QW communication failure ............................................................... 9
  - Site 1 to QW and Array to Array communication failure ....................................... 9
  - Site 2 to QW and Array to Array communication failure ....................................... 9
  - Site 1 and site 2 both lose communications to the QW .......................................... 9
- VMware vMSC and HPE 3PAR Peer Persistence best practices .................................. 10
  - VMware DRS Affinity Group Usage ..................................................................... 10
  - Managing VMware Storage DRS settings .............................................................. 10
  - Workload management across sites ...................................................................... 10
  - Multi-VM application placement ........................................................................... 11
  - Heartbeat Setting .................................................................................................... 11
  - VMware HA Admission Control ............................................................................ 11
  - Isolation addresses & false positive ....................................................................... 11
  - HPE 3PAR Quorum Witness Hosting .................................................................... 11
  - Remote Copy Replication links .............................................................................. 11
  - HPE 3PAR Peer Persistence maximums .................................................................. 11
- HPE OneView for VMware vCenter ............................................................................... 12
  - Features and benefits ............................................................................................. 12
- HPE 3PAR Recovery Manager for VMware vSphere ................................................... 13
  - Features and benefits ............................................................................................. 13
Introduction

HPE 3PAR StoreServ Peer Persistence software enables HPE 3PAR StoreServ systems located in different data centers at metropolitan distances to act as peers to each other, presenting continuous storage system access to hosts connected to them. This capability allows you to configure a high-availability solution between two sites or data centers where storage access failover and failback remains completely transparent to the hosts and applications running on those hosts. Compared to traditional cluster failover models where upon failover, the applications must be restarted, Peer Persistence software allows hosts to remain online serving their business applications even when they switch storage access from their original site to the storage array in the remote site, resulting in a much improved recovery time.

This paper provides information about deploying a vSphere Metro Storage Cluster (vMSC) across two data centers or sites using HPE 3PAR StoreServ Peer Persistence on StoreServ Storage systems. A vMSC configuration is designed to maintain data availability beyond a single physical or logical site.

Terminology

Throughout this paper, we identify the volume that is part of a primary remote copy group as source volume. The replicated volume in the secondary remote copy group is called the target volume. Volumes created on an HPE 3PAR StoreServ are called Virtual Volumes, or VVs. A VV exported to the host is known as a VLUN. A zoning operation creates a logical connection in the Fibre Channel (FC) SAN between an FC host bus adapter (HBA) on a server and one on a storage system.

Features and benefits

Peer Persistence is a high availability configuration between two data centers in which the ESXi hosts are setup in a metro cluster configuration with access to storage arrays in both sites. Storage volumes created on one storage array are replicated to the other array using synchronous remote copy to ensure that the volumes are in sync at all times. Peer Persistence software takes advantage of the Asymmetric Logical Unit Access (ALUA) capability that allows paths to a SCSI device to be marked as having different characteristics. With ALUA the same LUN can be exported from both arrays simultaneously, but only the paths to the side accepting write to the volume will be marked as active. The paths to the secondary side volume will be marked as standby preventing the host from performing any I/O using those paths. In the event of a non-disruptive array volume migration scenario, the standby paths are marked as active and host traffic to the primary storage array is redirected to the secondary storage array without impact to the hosts. Figure 1 shows a recommended Peer Persistence configuration.

![Figure 1. Peer Persistence setup with a Uniform vMSC configuration](image-url)
The Peer Persistence software supports bi-directional replication which enables each site to perform as a disaster recovery center for the other. It enables you to move your applications from one site to another based on your business and performance needs without impacting the applications running on those hosts.

An example would be the use of vMotion within a VMware® vSphere environment. Peer Persistence provides the same stretched cluster configuration across data centers that are used with VMware's vSphere Metro Storage Cluster certification to enable multi-site storage high availability. In figure 1, a few virtual machines (VMs) are being serviced by a HPE 3PAR StoreServ storage system on site 1 while other VMs are being serviced by another HPE 3PAR StoreServ storage system at site 2 located within metropolitan distances from one another. vMotion allows customers to move VMs across sites. Peer Persistence software enables the movement of VMs across data centers to be completely transparent to the applications those VMs are running.

Peer Persistence achieves automatic transparent failover in the event of a complete array or data center failure with the help of the Quorum Witness software. The Quorum Witness software needs to be installed at a third site and will act as an independent witness should one of the arrays or data centers become unavailable.

**Requirements**

- Firmware version on both HPE 3PAR StoreServ arrays must be 3.12 MU2 or newer.
- The sites must be set up in a Remote Copy 1-to-1 configuration in synchronous mode.
- The round trip latency between the two sites must be 2.6 milliseconds or less.
- Quorum Witness software must be installed on a virtual machine at a third site.
- The HPE 3PAR StoreServ arrays communicate with the Quorum Witness via the management port over TCP/IP. The Quorum witness needs to be reachable from both sites.
- This non-disruptive failover configuration is supported with VMware ESXi 5.x or newer.
- All source and target volumes exported from the two HPE 3PAR StoreServ arrays must have the same volume WWN.
- Hosts have to be connected to the HPE StoreServ storage system using FC, iSCSI or FCoE. Support for the iSCSI and FCoE host connectivity in HPE 3PAR Peer Persistence has been added in later version of HPE 3PAR StoreServ OS. Please refer to the Single Point of Connectivity Knowledge (SPOCK) for more details.
- All associated hosts are connected to both arrays (Uniform vMSC).
- The ESXi hosts that these volumes are exported to must be configured with host persona 11 (VMware), which supports host capability Asymmetric Logical Unit Access (ALUA) configuration considerations.

For the latest supported configuration and requirement for HPE 3PAR Peer Persistence in a vSphere environment please refer to HPE Single Point of Connectivity Knowledge (SPOCK).

**Host Persona 11**

Host personas are a set of behaviors that permit hosts connected to FC or iSCSI ports on the system to deviate from the default host behavior. By assigning a persona to a host, multiple host types that require distinct customized responses can share a single system port. For example, hosts running Windows®, Linux®, and AIX operating systems can all connect to the same system port. This simplifies connecting hosts to the system and reduces management costs related to complex host connections.

Personal 11 supports ALUA path management and it must be used for host configurations that have ESXi hosts connected to both the primary and the secondary. Existing ESXi hosts defined with host persona 6 should be changed to persona 11 if these hosts are connected to both primary and secondary array.
For host configurations that have independent ESXi hosts attached to the primary or secondary arrays, host persona 6 (Generic-Legacy) should be configured. This persona does not support Asymmetric Logical Unit Access (ALUA) path management.

For more information on host persona migration, please refer to the HPE 3PAR StoreServ Storage VMware ESX Host Persona Migration Guide (HPE Part Number: QL226-97875) hp.com/go/storage/docs.

**Verify each VV has same WWN on both arrays**

It's important that the source and the target volumes on the HPE 3PAR StoreServ arrays have the same identity when both volumes are exported to the same ESXi host. Hence both volumes should have the same WWN. User can create volumes with the same WWN automatically using the new option in `admitrcopyvv` command. See the HPE 3PAR StoreServ Copy Software User’s Guide for instructions on how to use the `admitrcopyvv` command. The HPE 3PAR Management Console (MC) also give the users option to create VVs with the same WWN when configuring RC groups.

The replicated volume WWN can also be changed later if the volumes were created using the management console or any method other than the admitrcopyvv. See the HPE 3PAR StoreServ Command Line Interface Administrator’s Manual or the HPE 3PAR StoreServ Management Console Online Help for instructions on how to edit a VV's WWN.

**Verify VLUN status on both arrays**

Make sure that volumes are exported using the same LUN ID from both HPE 3PAR StoreServ arrays. Once the volumes are exported from both arrays, the showvlun command will show the state as “active” for VLUNs exported from the source side and “standby” for the VLUNs exported from the target side. ESXi hosts will send I/O requests only to the LUNs reported as active.

**Quorum Witness**

The HPE 3PAR StoreServ Quorum Witness (QW) software enables transparent automatic failover between the two sites in a vMSC cluster. The Quorum Witness software gets regularly updated with status information from the HPE 3PAR StoreServ arrays. In the event of a failure of one of the HPE 3PAR StoreServ arrays, the surviving HPE 3PAR StoreServ array detects that the Quorum Witness is not getting updated by the failed array and initiates a failover operation on any secondary groups associated with that failed array.

**Note**

QW details have to be setup at the target level, and transparent failover will occur only for Remote Copy groups that use the target, are started and have the “auto_failover” policy set. Failover is not automatic for remote copy groups between the same two HPE 3PAR StoreServ systems which do not have the “auto_failover” policy set. For HPE 3PAR OS 3.1.3 or later, the setting of path management is required as well. A volume group must have the path management policy set if it is to be valid for automated and transparent failover. If this policy is not set, then no automatic failover will occur if a disaster strikes.

The Quorum Witness is typically set up on a third site where events that may impact site 1 or site 2 cannot impact the Quorum Witness site at the same time. Additionally, the QW connects to arrays on both sites using non-RC links. With the above two configuration characteristics (site and link independence), QW helps determine the following nature of failures:

- **Link failure**: The QW can detect if the two arrays are alive but not communicating because of a link failure. QW would still be receiving updates from both of the arrays.

- **Array/Site failure**: The QW can detect when one of the arrays/sites fails. QW would not be receiving updates from one of the array that has failed.

The “Handling failures” section covers the various failure conditions handled by Peer Persistence.
The Quorum Witness software is packaged as an Open Virtualization Format OVF package for deployment in a vSphere or a Hyper-V environment. The Quorum Witness virtual machine should not be stored on a data store allocated from either of the arrays that are part of the Peer Persistence configuration. If the datastore is located on one of the arrays that is part of the Peer Persistence configuration failure of the array will make the data store unavailable which will result in the shutdown of the Quorum Witness virtual machine. The Quorum Witness software uses port 8080 to communicate with the two HPE 3PAR StoreServ systems in the quorum. Hence firewall rules need to be changed to open port 8080 at the QW site.

The Quorum witness is a passive component of the configuration and the software itself does not provide any high availability, however the virtual machine can be made highly available by using VMware—High Availability or Fault Tolerance, or Microsoft® Failover Clustering for the Hyper-V based QW package. For the latest supported configuration and requirement for HPE 3PAR Peer Persistence QW, refer to the Single Point of Connectivity Knowledge (SPOCK).

Management port on HPE 3PAR StoreServ
The HPE 3PAR StoreServ’s communicate with the Quorum Witness using the HPE 3PAR StoreServ management interface. The network should be setup such that the QW server has network connectivity to the admin interface on both arrays. It is required to connect the administration port from a minimum of two controller nodes to the network. However if the array has more than 2 controllers the best practice is to connect all controller nodes to the network.

VMware vSphere Metro Storage Cluster (vMSC) Configuration
Two vMSC configurations are possible based on how the hosts are connected to the storage arrays.

vMSC uniform
In this configuration each host can access LUNs exported from both the sites. Figure 1 shows a typical Uniform configuration. This setup helps in protecting against a storage failure at a site. So, if the array on site 1 fails storage will failover to the HPE 3PAR StoreServ array on site 2 transparently and automatically allowing VMs on site 1 continued access to storage on site 2 without any interruption. Of course this implies that customers need to have good bandwidth between their sites. Outside of what happens in sync replication (ask only after write to remote array is completed), even host to storage array traffic will be routed across the FC connections to the remote site (in synchronous replication, hosts are writing only to the local array).

Note that with Uniform, the use case is typically of active load balancing being done where secondary paths/volumes are made active on site 2 array as a means to achieve load balancing. This is in addition to what Uniform provides which protection against storage failure.

vMSC non-uniform
In this configuration each host can access storage resources available only on its local site. Figure 2 shows a typical non-uniform configuration. This is geared towards protecting against a complete site failure similar to what VMware vCenter Site Recovery Manager (SRM) does.

The concept here is that if a complete site fails (storage and server cluster), Peer Persistence will be able to transparently and automatically failover the storage to site 2 while the host application fails over in parallel. Consider this really as a DR solution. Non-uniform configurations will not allow an application to transparently failover. If a switchover is executed the local ESXi servers will lose connectivity, the VMs will go offline and will be switched according to the HA setup, if they are switched to the remote site, they will restart automatically.

Note
vMSC non-uniform is not supported with HPE 3PAR Peer Persistence.
Planned switchover

A switchover operation migrates the remote copy group role from source to target without impacting host I/O. This operation requires that the associated hosts are connected to the arrays on both sites (Uniform vMSC). This is a planned operation and the remote copy groups need to already be started and synced for this command to work. This operation is useful when you want have the least possible latency for accessing storage when using VMs from a host. For example, if vMotion is used to move all VMs on a host from site 1 to site 2, one can use the switchover command to reverse the replication resulting in site 2 hosts being able to access the LUNs locally from site 2 array. The system performs the following action as part of a switchover operation:

- I/O from the Virtual Machines to the volumes in the source remote copy group is blocked and in flight I/O is allowed to drain. The remote copy group is stopped and snapshots are taken on the primary array.
- The primary array target port group is changed to transition state and it sends a remote failover request to the secondary remote copy group.
- The secondary array target port group is then changed to transition state and it takes a recovery point snapshot.
- The secondary remote copy group changes state to become primary-reversed and makes the volumes read/write.
- The pri-rev target port group is changed to active state and the array returns a failover complete message to the primary remote copy group.
- The primary array target port group is changed to standby state and any blocked I/O is returned to the host with a sense error: NOT READY, LOGICAL UNIT NOT ACCESSIBLE, TARGET PORT IN STANDBY STATE.
- The host will then perform SCSI inquiry requests to detect what target port groups have changed and which paths are now active and I/O will now be serviced on the active path to the primary-reverse remote copy group.
- All operations to this point should complete within 30 seconds to ensure that host I/O does not timeout.
- The primary remote copy group will then send a remote recover request to the primary-reverse remote copy group. The primary-reverse remote copy group will perform a recover operation and will change the primary remote copy group state to secondary-reverse state. The remote copy group will then be started from primary-reverse to secondary-reverse.
- When the primary-reverse and secondary-reverse volumes are back in sync the snapshots that were taken on both sides will be removed. The remote copy group will then undergo a reverse(- natural) operation. This operation will change the primary-reserve group to primary and secondary-reverse group to secondary. It is possible for this operation to time out or fail if the target links go down during processing. If this happens, issue the “setcopygroup reverse - natural” command to complete the process manually. The system is now fully reversed and ready for another switchover request.
HPE 3PAR Peer Persistence failures handling

Peer Persistence with Quorum Witness can identify different types of failures and perform transparent failover automatically. Peer Persistence will perform transparent failover only for remote copy groups that have the auto_failover and path_management enabled.

Array to Array communication failure

In the case of an array to array communication failure, the HPE 3PAR StoreServ continues to send/receive communication to/from the Quorum Witness. Automatic transparent failover does not occur and host I/Os continue to go to their primary volumes; however replication of I/O across RC links will stop due to the failure.

Single site to QW communication failure

In the case where the arrays at either site 1 or site 2 lose connectivity with the Quorum Witness, an automatic failover will not occur. Host I/O will continue as before and replication of I/O across RC links will continue as normal. An automatic failover does not need to occur because the two HPE 3PAR StoreServ arrays can still communicate with one another via the replication links they share.

Site 1 to QW and Array to Array communication failure

If the array in site 1 becomes isolated due to a dual network failure (communication with the QW fails and the replication links to the array in site 2 fails) Remote copy groups that are in primary state on the array in site 1 will be transitioned to failsafe mode and will stop serving I/O to the ESXi hosts in site 1. The array in site 2 (which is still communicating with the QW) will perform an automatic failover and host I/O will be redirected to the new primary volumes in site 2.

Site 2 to QW and Array to Array communication failure

This is the same as the previous case with the difference that site 2 will become isolated due to a dual network failure.

Site 1 and site 2 both lose communications to the QW

In the case of a communication failure with the QW from both sites, but where the Remote Copy links remain operational both arrays will be aware of each other’s operational state. Automatic failover does not happen and host I/Os continue to go to the primary volumes; replication of I/O across RC links will continue as normal.

Site 1 and site 2 to QW and Array to Array communication failure.

In the case where all network communications between the arrays and the QW and the communication between both arrays also fails, both arrays will be isolated as they cannot communicate with each other over the RC links nor can they communicate with the QW. In this case remote copy groups that are primary will go into failsafe mode and stop serving I/O, and failover actions will not be performed. This will result in host I/O failure and replication of I/O across RC links will stop due to the failure.
### Table 1. HPE 3PAR Peer Persistence failures handling

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Replication Stopped</th>
<th>Automatic Failover</th>
<th>Host I/O Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Array to Array remote copy links failure</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Single site to QW network failure</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Single Site to QW network and Array to Array remote copy link failure</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Both sites to QW network failure</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Both sites to QW network and Array to Array remote copy link failure</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### VMware vMSC and HPE 3PAR Peer Persistence best practices

The implementation of the VMware vSphere Metro Storage Cluster requires certain considerations for optimal deployment. VMware has published a technical white paper titled: “VMware vSphere Metro Storage Cluster Case Study.” The paper discusses vMSC design considerations and operational procedures. It is a recommended document to review for better outcome of the deployment of HPE 3PAR StoreServ Peer Persistence.

In the next section, we will discuss few of the considerations and recommendations that customer should consider when deploying HPE 3PAR Peer Persistence.

### VMware DRS Affinity Group Usage

The VMs should be grouped together to the preferred site during normal operation. This will guarantee that they use local storage in their preferred site. However, during certain failure scenarios these VMs may be running on the storage in the other data center. The user should consider using the “should rule” rather than the “must rule” with affinity group. Making this choice will allow VMs to failover to the other site as a result of the VMware HA initiating those moves. The result of this implementation is to simply employ affinity group to create site awareness leveraging the DRS “should rule.” For more information about those setting refer to the HA & DRS best practices.

### Managing VMware Storage DRS settings

VMware vSphere 5.0 introduced a new feature called Storage DRS. The feature provides smart VM placement and load balancing methods based on space capacity utilization and I/O latency. It offers an efficient way for provisioning of VMs and monitoring of storage used by the vSphere cluster.

Since the goal during normal operation is to have VMs running on their preferred site, the setup of Storage DRS should be “Manual Mode.” In this mode, Storage DRS will make recommendations when thresholds for space utilization or I/O latency are exceeded. This will help avoid having Storage DRS moving VMs to a LUN on remote storage. One also needs to take into consideration the extra load on the links between the two sites due to storage vMotion between LUNs from the two sites. Its best to schedule these moves for non-peak hours and implement the recommendations manually if needed. Even with manual mode, the user should still take advantage of the Storage DRS on the initial placement of the VMs during the provisioning process along with the monitoring of the environment.

### Workload management across sites

Users should consider the impact that vMotion and Storage vMotion will have on the inter-site links. Making an effort to avoid having server I/O cross sites, vMotion and Storage vMotion can be used to ensure that the inter-site bandwidth is dedicated to replicating production data.
Multi-VM application placement
User should consider grouping or placing multi-VM applications together on the same site. Such placement would ensure that VMs which have dependencies on other VMs and could have significant traffic between them are not consuming resources on replication links. There is also the added latency due to the distance between the two sites in the case where VMs sharing a data store are split between the two sites. The added latency can impact the application user’s experience.

Heartbeat Setting
VMware vSphere 5.0 introduced Datastore Heartbeating. The datastore heartbeat mechanism is used to check if a host has failed in the VMware HA cluster or simply if it’s isolated due to management network issues. The master host in the HA cluster relies on this mechanism to assess if a host has failed and as a result if the VMs hosted on that host would need to be restarted on another host in the HA cluster. The default behavior of the VMware HA cluster is to use two datastores for the heartbeat. vCenter uses an algorithm to pick the two datastores. It is recommended that the user consider using a total of 4 with 2 at each site in case of the loss of communication between the two sites.

VMware HA Admission Control
Users should consider the resources required at each site to handle the workload in case of a failure of one site and the subsequent failover for the workload to the surviving site. VMware vCenter uses the Admission Control to guarantee that enough resources are available in the cluster to protect user workload in case of a failover. By enabling admission control and configuring its policy to be able to handle the all workload when restarted by VMware HA.

Isolation addresses & false positive
Since the network heartbeat is key in the indicating the state of a host, the user should ensure that the management network is resilient. Using the VMware “das.isolationaddress” setting the user should assign one isolation address per site as minimum to avoid having an isolation incident due to the communication loss between the two sites that might result in unnecessary failover and subsequent restart of VMs.

HPE 3PAR Quorum Witness Hosting
Given how critical the OW is to operation of HPE 3PAR Peer Persistence, it is highly recommended that the OW VM should be hosted outside the stretched cluster environment and should not be on the HPE 3PAR arrays used in the HPE 3PAR Peer Persistence deployment it arbitrates or the vSphere hosts that are part of the Metro Cluster.

Remote Copy Replication links
Since a Uniform vMSC configuration requires fiber channel connectivity between the data centers containing the HPE 3PAR StoreServ arrays the ESXi servers for host to array connectivity, it is recommended that RCFC be used for the synchronous replication of data between the arrays in a Peer Persistence configuration. Use of RCFC will ensure a high bandwidth low latency connection for the replication of data between the arrays and it uses a patented protocol for synchronous replication that only requires a single round trip across the replication links to replicate an I/O smaller than 80K in size.

HPE 3PAR Peer Persistence maximums
It is highly recommended to consolidate VVs inside few RC Groups in a HPE 3PAR Peer Persistence configuration. Such configuration will guarantee the host I/Os will be failed over within the timing window in case of a transparent failover.

For the latest supported configuration, maximum and requirement for HPE 3PAR Peer Persistence please refer to HPE Single Point of Connectivity Knowledge (SPOCK).
HPE OneView for VMware vCenter

HPE OneView for VMware vCenter is software for VMware's vCenter management console which enables the vSphere administrator to quickly obtain context-aware information about HPE servers and HPE storage in their VMware vSphere environment directly from within vCenter. This enables the vSphere administrator to easily manage physical servers, storage, data stores, and virtual machines. By providing the ability to clearly view and directly manage relationships between virtual machines and HPE infrastructure, the VMware administrator’s productivity increases, as does the ability to ensure quality of service.

**Features and benefits**

- Simplify administration with integration of the physical with the virtual infrastructure
- Reduce downtime by automating responses to hardware events
- Take control by launching your trusted HPE management tools
- Proactively manage changes with detailed relationship dashboards of server, networking, and storage
- Maintain stability and reliability with firmware inventory and deployment
- On demand server and storage capacity provisioning
- Visualize complex relationships:
  - How virtual machines are mapped to underlying storage
  - How peer persistence volumes are being configured

---

![Figure 3. HPE OneView and HPE Management—Storage](image)

![Figure 4. HPE OneView and HPE Management—VMs to Storage mapping](image)
HPE OneView for VMware vCenter v 7.4 added support for new storage provisioning wizards for HPE 3PAR Peer Persistence. Using the wizard will allow for the setting of the “auto_failover” and “path_management.”

Figure 5. HPE OneView and HPE Management-Switch Peer Persistence

**HPE 3PAR Recovery Manager for VMware vSphere**

Recovery Manager for VMware vSphere supports Peer Persistence in its latest release. This allows user to create and manage consistent HPE 3PAR StoreServ snapshots across sites.

**Features and benefits**

- RMV collects all HPE 3PAR storage devices information from vCenter Server.
- If a storage device contains multiple paths from different StoreServ, it is participated in Peer Persistence setup.
- The active path determines which StoreServ is the primary site.
- User can create crash level or application level both local and remote sites.
- Virtual Copy can be recovered from either sites.

**Sample Peer Persistence configuration**

This section gives the steps needed to setup a single remote copy group in Peer Persistence configuration. Please refer to the “HPE 3PAR StoreServ Remote Copy Software User’s Guide” for more details and the complete set of options available for the commands used in this section.

In this example, we are deploying the HPE 3PAR Peer Persistence between two sites: Kirkland and Redmond. The vSphere cluster consists of 4 hosts running ESXi 5.5 with 2 hosts at each site.

**QW installation**

The Quorum Witness software provides the arbitrator functionality in the Peer Persistence setup. The detailed instructions for deploying Quorum Witness are giving in Appendix F in “HPE 3PAR StoreServ Remote Copy Software User’s Guide.” Here are the instructions from the guide, please refer to the User’s Guide for accurate and updated information. User would need to note the IP address displayed in the Summary tab after powering on the virtual machine. This IP address will be used when configuring the Quorum Witness later.
Creating remote copy targets

This is a software configuration guide only and assumes that the right hardware requirements for setting up remote copy are met. This guide will not cover how to configure ports for remote copy, how to do cabling, how to perform zoning etc. This guide also assumes that you have met all requirements for Peer Persistence.

Before you start replicating volumes from one array to another you need to create remote copy targets on both arrays. Connect to both the arrays using MC, then start with creating “New Configuration” in the Remote Copy Manager.

- Select 1-1 configuration and select the two systems that will be used in the RC setup.
• Select the ports that will be used for remote copy. RCFC ports are used in this example.

• Click on finish as we would like to create groups later.

Peer Persistence Configuration
Click on the Peer Persistence Configuration tab after selecting the Remote Copy Configuration navigation tree node. Use the IP address of the Quorum Witness that was noted down earlier.

Do not click on the “Start Witness” checkbox at time of creating the configuration. Press OK after specifying the IP address.

Figure 7. HPE 3PAR Peer Persistence Configuration

After a while the Quorum Status will change to Not-started. Select both targets in the list and click on the Start button to start the Quorum Witness.
The status should now change to Started.

Figure 8. HPE 3PAR Peer Persistence Configuration—Start

Figure 9. HPE 3PAR Peer Persistence Configuration—Status
Troubleshooting Witness Quorum Communication issues

This operation needs to be done using CLI, so login to one of the HPE 3PAR StoreServ arrays and perform the following:

- Find the target name using showrcopy command:

  Redmond3PAR cli> showrcopy

  Remote Copy System Information
  Status: Started, Normal

  Target Information

  Name | ID | Type | Status | Options | Policy
  Kirkland3PAR | 1 | FC | ready | 2FF700C2AC01B575 | mirror_config

  Link Information

  Target | Node | Address | Status | Options
  Kirkland3PAR | 0:2:3 | 20230002AC01B575 | Up
  Kirkland3PAR | 1:2:4 | 21240002AC01B575 | Up
  receive | 0:2:3 | 20230002AC01B575 | Up
  receive | 1:2:4 | 21240002AC01B575 | Up

  Figure 10. HPE 3PAR Peer Persistence showrcopy output

- Use the witness check command to verify the communication between Quorum Witness and management interface of the two arrays. The two commands will not report any message on success. In earlier version of HPE 3PAR OS 3.2.1.n.n.n and 3.1.3.n.n.n. In version 3.2.1 a basic message is displayed if the verification is successful.

  Redmond3PAR cli> setrcopytarget witness check 192.168.0.34
  Connectivity check passed
  Redmond3PAR cli> setrcopytarget witness check --remote 192.168.0.34 Kirkland3PAR
  Connectivity check passed

  Figure 11. HPE 3PAR Peer Persistence QW check 1

- Following is an example if the Quorum Witness server is unreachable:

  Redmond3PAR cli> setrcopytarget witness check 192.168.0.36
  error: No route to Quorum Witness at 192.168.0.36

  Figure 12. HPE 3PAR Peer Persistence QW check 2
Creating remote copy groups

Start the Create Remote Copy group wizard and select the source and target system. Next specify the group name and select synchronous mode.

On the next screen, select the source volume you want to replicate and select the “Create new volume” on the target system. It’s preferred to create new volumes on the target, this way the volumes will have the same characteristics as the source volumes. If you manually create volume beforehand, then be sure that the volume size on target matches the volume size on the source and both volumes have the same WWN.

Add all volumes in the group and click Finish. The group will start and initial synchronization will be started.

More groups should be created. A typical uniform vMSC configuration will have remote copy groups replicating in both directions.
Changing Group Policy

It's required to edit the replication group and enable “auto_failover” policy to enable automatic transparent failover.

Click on the Enable/Disable Groups button in the Peer Persistence tab. Select the group that you want to enable automatic failover and press OK (Figure 15).

Enable the auto_failover and the path_management policies on the RC group.

Using showrcopy command to verify the auto failover and path management settings (Figure 16):

Redmond3PAR cttl showrcopy

Remote Copy System Information
Status: Started, Normal

Target Information
Name   ID  Type  Status Options  Policy
Redmond3PAR 1 FC  ready  3FF70002ACC018875 mirror_config

Link Information
Target  Mode  Address  Status Options
Redmond3PAR 01:2:0  2E2D002ACC018875 Up
receive  01:2:0  212D002ACC018875 Up
receive  01:2:3  212D002ACC018875 Up

Group Information
Name  Target  Status  Role  Mode Options
3par-pp-red-copy  Redmond3PAR Started Primary Sync auto_failover, path_management
LocalUV  ID  RemoteUV  ID  SyncStatus  Last3yrsTime
redmond-ds-2  64 redmond-ds-2  17 Synced  NA
redmond-ds-3  65 redmond-ds-3  19 Synced  NA
redmond-ds-3  66 redmond-ds-3  19 Synced  NA

Redmond3PAR cttl

Figure 15. HPE 3PAR Peer Persistence Auto Failover

Figure 16. HPE 3PAR Peer Persistence Auto Failover and path management setting (Site A)
From the other array on the remote site, the same command would yield the following:

```
FicklandSPAR c1ist showcopy
Remote Copy System Information
Status: Started, Normal

Target Information
Name       ID    Type Status Options       Policy
RedwoodSPAR 1 FC   ready 21770002A0018F76 active_config

Link Information
Target   Node Address Status Options
RedwoodSPAR 0:1:3 20210002A0018F76 Up
FicklandSPAR 1:2:1 21240002A0018F76 Up
remote    0:1:3 21240002A0018F76 Up
receive   1:2:1 21240002A0018F76 Up

Group Information
Name          Target Status Role Mode Options
Spar-cpg-cx5-cpg-1.090702 RedwoodSPAR Started Secondary Sync auto_failover,park_management
LocalVV       ID RemoteVV ID  SyncStatus LastSyncTime
Spar-cpg-cx5-cpg-1-1 17 Spar-cpg-cx5-cpg-1-1 65 Synchronized NA
Spar-cpg-cx5-cpg-1-2 18 Spar-cpg-cx5-cpg-1-2 65 Synchronized NA
Spar-cpg-cx5-cpg-1-3 19 Spar-cpg-cx5-cpg-1-3 65 Synchronized NA

FicklandSPAR c1ist
```

**Figure 17.** HPE 3PAR Peer Persistence Auto Failover and path management setting (Site B)

### Exporting LUNs to hosts

When creating ESXi hosts they need to be created using Personal 11, for best practice add all hosts in the cluster in a host set. And add all VVs in a remote copy group in a VV set.

After exporting the VVs from both arrays, press “Rescan All” in Configuration -> Storage Adapters for all ESXi hosts in the cluster.

If you look for path details for one of the LUNs you should see some Active paths and some Standby paths.

Now you are ready to create datastores on these LUNs and start using them in the vSphere cluster.

**Figure 18.** Exposing VLUNs to vSphere hosts
Planned switchover example

The CLI command has to be used to initiate a planned failover/switchover. The command has to be issued on the array with the primary role. The “Planned switchover” section provides details about the workings of this operation.

In this example the remote copy group is primary on Redmond3PAR.

```
Redmond3PAR cli> showcopy
Remote Copy System Information
Status: Started, Normal

Target Information
Name   ID Type Status Options    Policy
Kickland3PAR 5 PC ready 207F00002A018575 mirror_config

Link Information
Target Node Address Status Options
Kickland3PAR 01:23 20240002A018575 Up
Kickland3PAR 1:24 21240002A018575 Up
receive 01:23 20240002A018575 Up
receive 1:24 21240002A018575 Up

Group Information
Name   Target Status Role    Mode Options
Spar-cp-red-cp1 Kickland3PAR Started Primary Sync auto_failover,path_management
Loc1YY ID RemoteYY ID SyncStatus LastSyncTime
redmond-da-1 44 redmond-da-1 17 Synced NA
redmond-da-2 65 redmond-da-2 15 Synced NA
redmond-da-3 66 redmond-da-3 15 Synced NA
```

Figure 19. HPE 3PAR Peer Persistence status before switchover

Figure 20. HPE 3PAR Peer Persistence switchover using the HPE 3PAR MC
After the switchover command is executed successfully the primary role will move to Kirkland3PAR.

Recovering from automatic failover

If one array fails, then primary remote copy groups from that array that have “auto_failover” policy enabled will be failed over to the second array. Recovery process is needed to restart replication and reverse replication back in the original direction.
After the array is recovered and back online the showrcopy status will show.

```
Redmond3PAR cli> showcscopy
Remote Copy System Information
  Status: Started, Normal

Target Information
  Name   ID  Type  Status Options Policy
  Kirkland3PAR  1  FC  ready  2FF70002AC018878 mirror_config

Link Information
  Target   Mode  Address          Status Options
  Kirkland3PAR  0:2:3  20200002AC018878 Up
  Kirkland3PAR  1:2:4  21200002AC010575 Down
  receive     0:2:3  20200002AC018878 Up
  receive     1:2:4  21200002AC018878 Up

Group Information
  Name   Target   Status   Role   Node   Options
  Sync-cp-cod-rep1  Kirkland3PAR  Failover  Primary  Sync  auto_failover,pach_management
  LocalVV  ID  RemoteVV  ID  SyncStatus  LastSyncTime
  cedmond-de-1  54  cedmond-de-1  17  Failover  NA
  cedmond-dw-2  55  cedmond-dw-2  18  Failover  NA
  cedmond-he-3  56  cedmond-he-3  19  Failover  NA

Redmond3PAR cli>
```

**Figure 24.** HPE 3PAR Peer Persistence array status after is back online from an array failure

These steps should be used to recover the remote copy group after the array that had failed is back online.

Select the group that needs to be recovered under the Remote Copy Configuration node. Click on the Recover button and press OK. Select Yes at the next prompt to continue with the recovery.

```
Redmond3PAR cli>
```

**Figure 25.** HPE 3PAR Peer Persistence recover

The Restore button will be enabled after the recovery has completed. However, performing the restore operation might lead to ESXi hosts losing access to the LUNs for a short duration. Hence it’s recommended to use the CLI commands to complete the failback. Login to the array with the Primary-rev role and issue the following commands.
Figure 26. HPE 3PAR Peer Persistence reverse

The primary role should now change back to the original array and hosts will start performing IO to the VVs provisioned in that array.

For more information
HPE Enterprise Storage Information Library. [hp.com/Go/Storage/Docs](http://hp.com/Go/Storage/Docs)
HPE 3PAR StoreServ Storage: [hp.com/go/storeserv](http://hp.com/go/storeserv)
HPE 3PAR StoreServ VMware Implementation Guide
HPE 3PAR StoreServ Storage best practice guide
HPE 3PAR Remote Copy Software User Guide OS 3.2.2

Learn more at [hp.com/go/3parstoreserv](http://hp.com/go/3parstoreserv)