EMC® VPLEX™ Host Multipathing
Implementation Planning and Best Practices

- ESXi NMP Policy
- PowerPath
- Native MPIO

Abstract
This document describes multipath options and settings for use with VPLEX in Local, Metro, Geo and host cross-cluster configurations.

September 2014
Copyright © 2014 EMC Corporation. All rights reserved. Published in the USA.

Published September 2014

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

The information in this publication is provided as is. EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose. Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

EMC², EMC, and the EMC logo are registered trademarks or trademarks of EMC Corporation in the United States and other countries. All other trademarks used herein are the property of their respective owners.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

**EMC VPLEX Host Multipathing Implementation Planning and Best Practices Technical Notes**

Part Number H13547
# Contents

## Chapter 1  Introduction

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audience</td>
<td>6</td>
</tr>
<tr>
<td>Document Organization</td>
<td>6</td>
</tr>
<tr>
<td>Host Multipath for VPLEX Overview</td>
<td>7</td>
</tr>
<tr>
<td>Multipath Software</td>
<td>8</td>
</tr>
</tbody>
</table>

## Chapter 2  ESXi NMP Policy

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round Robin - New Policy Recommendation</td>
<td>11</td>
</tr>
<tr>
<td>Round Robin</td>
<td>11</td>
</tr>
<tr>
<td>Changing Policy from Fixed Path to Round Robin</td>
<td>11</td>
</tr>
<tr>
<td>Changing Policy</td>
<td>11</td>
</tr>
<tr>
<td>Additional Information</td>
<td>12</td>
</tr>
</tbody>
</table>

## Chapter 3  PowerPath

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>15</td>
</tr>
<tr>
<td>Path management</td>
<td>15</td>
</tr>
<tr>
<td>PowerPath Features</td>
<td>15</td>
</tr>
<tr>
<td>Supported Versions and Settings</td>
<td>16</td>
</tr>
</tbody>
</table>

## Chapter 4  Native MPIO

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>18</td>
</tr>
<tr>
<td>AIX 19</td>
<td></td>
</tr>
<tr>
<td>HP-UX</td>
<td>21</td>
</tr>
<tr>
<td>Linux</td>
<td>22</td>
</tr>
<tr>
<td>Red Hat Enterprise Linux (RHEL)</td>
<td>22</td>
</tr>
<tr>
<td>SuSE Linux Enterprise Server (SLES)</td>
<td>27</td>
</tr>
<tr>
<td>Microsoft Windows</td>
<td>31</td>
</tr>
<tr>
<td>Windows 2008 Server Core and Windows 2008 R2</td>
<td>31</td>
</tr>
<tr>
<td>Server Core</td>
<td>31</td>
</tr>
<tr>
<td>Path management in Multipath I/O for VPLEX</td>
<td>32</td>
</tr>
<tr>
<td>Enabling Native MPIO</td>
<td>32</td>
</tr>
<tr>
<td>Native MPIO with Windows Server 2012</td>
<td>33</td>
</tr>
<tr>
<td>Configuring Native MPIO for Windows Server 2012</td>
<td>33</td>
</tr>
<tr>
<td>Device discovery and claiming devices for MPIO</td>
<td>34</td>
</tr>
<tr>
<td>Path management in Multipath I/O for VPLEX</td>
<td>35</td>
</tr>
</tbody>
</table>
### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solaris</td>
<td>37</td>
</tr>
<tr>
<td>VPLEX and MultiPath software</td>
<td>37</td>
</tr>
<tr>
<td>VPLEX and MPxIO</td>
<td>39</td>
</tr>
<tr>
<td>VPLEX with Oracle Solaris Cluster support</td>
<td>40</td>
</tr>
</tbody>
</table>
This chapter presents the following topics:

- **Audience** .......................................................... 6
- **Document Organization** ....................................... 6
- **Host Multipath for VPLEX Overview** .................. 7
Chapter 1: Introduction

**Audience**

These technical notes are for EMC field personnel and partners and customers who will be configuring, installing, and supporting VPLEX. An understanding of these technical notes requires an understanding of the following:

- SAN technology and network design
- Fiber Channel block storage concepts
- VPLEX concepts and components

**Document Organization**

This technical note is one of a set of documents that supersede the monolithic Implementation Planning and Best Practices for EMC VPLEX Technical Notes document that had previously been available. It is intended to provide more concise, per-topic information that will be easier to maintain and keep up-to-date.

All documents are available on [http://support.emc.com](http://support.emc.com).

The following list represents the technical note best practice documents that will be made available:

- EMC VPLEX Overview and General Best Practices
- EMC VPLEX SAN Connectivity
- EMC VPLEX Host Multipathing
- EMC VPLEX IP Networking
- EMC VPLEX and RecoverPoint Interoperability
- EMC VPLEX Data Mobility and Migrations
- EMC VPLEX REST API
- EMC VPLEX Upgrades
- EMC VPLEX Virtual Edition
Host Multipath for VPLEX Overview

This document is designed to address host multipath configurations as they apply to VPLEX.

The recommendations and requirements are based on general SAN best practices and rigorous testing through E-Lab, QE and the VPLEX Performance Engineering Team.

Changes to software and hardware architecture both within VPLEX and within the overall environment cause these recommendations and requirements to be adjusted occasionally.

Please refer to the ESSM (EMC Simple Support Matrix) for recommendations and requirements for multipath support statements.

**Important** VPLEX uses a Distributed Cache Model which is significantly different for the purpose of multipath configurations than most arrays which typically use a Global Cache Model. Additionally, from a host perspective, VPLEX is viewed as an active/active architecture.

The Distributed Cache Model is designed so that each VPLEX Director has its own independent cache. This means that if a host writes to one Director and then performs a read against that write from another Director then the result is a cache read miss. An array with Global Cache would result is a cache read hit. To compensate for these differences, the host multipath configuration setting will differ when connected to VPLEX from an array so that the chance of a cache read hit can be optimized.

Also as mentioned, VPLEX is an active/active architecture from the host perspective. This is completely independent from the backend array. VPLEX will handle the load balancing and path access for the backend array and will present the device to the host as active/active. This means that there is absolutely no consideration for active/passive or ALUA configuration settings within the host multipath configuration for the devices accessed through VPLEX even if the backend array where that device resides is active/passive or ALUA.
Chapter 1: Introduction

Multipath Software

There are several choices when selecting the host multipath software. These can be broken down into two basic categories:

1. Path Failover
2. Load Balancing

Path Failover

Path Failover is designed as single active path and the other available paths are for failover access only. This design has advantages with the VPLEX Distributed Cache model as it has the highest possible chance of a cache read hit. All data will be written and read from the same VPLEX Director for that host and if the data is in cache on VPLEX then it will be on that one Director that the active path is configured to and there will be a 100% chance of a cache read hit.

While this may be a distinct advantage when considering performance, there are a few disadvantages. First, this type of configuration requires manual configuration to assign the active path to the least utilized VPLEX Director and may create a lot of additional work for the administrator. Second, this type of software may not have the same responsiveness and adaptability in the event of various environment failures as compared to load balancing software. Third, VPLEX offers a very unique type of configuration called “Host Cross-Cluster Connect Configuration”. This is a configuration where the host has access to the device across two different VPLEX Clusters in a Metro Configuration across data centers. The initial configuration would dictate that the active path be defined to the local VPLEX Cluster however if the local VPLEX Cluster were to fail then there would be a high possibility that the first available failover paths would be to the same Director on the remote VPLEX Cluster for all hosts which would overwhelm that director and create a serious workload for the administrator to try and rebalance the active paths. All this work would then have to be redone when the original VPLEX Cluster came back online to rebalance the active paths again on the local cluster.

Load Balancing

Load Balancing software is designed to take advantage of the additional available paths by utilizing additional resources to move data. This increases the total bandwidth available to move larger amounts of data faster.

Based on the specific load balancing software chosen, there are typically several choices for configuration settings. In basic terms, these settings fall into a few different categories:

1. Path Policy - i.e. Round Robin, Adaptive...
2. How much I/O is sent before moving on to next path
3. Path timeouts and retries

Path Policy is a setting within the multipath software and is configured at the host level. The Path Policies vary from software vendor to software vendor and is important to reference the ESSM (EMC Simple Support Matrix) before committing to a product or setting.
Most products available today were designed to provide multiple path access within the datacenter. VPLEX offers additional capabilities and therefore additional challenges for host access.

As mentioned with the Path Failover software, VPLEX uses a Distributed Cache model and requires multipath software to be configured to optimize the VPLEX caching algorithms which differ from most arrays.

Not all host multipath software is created equal. Basic configurations within the datacenter may offer the option of selecting host based native multipath solutions or third party software but there will be circumstances where we recommend PowerPath. PowerPath has added additional support specifically for VPLEX Metro which is not available with any other product on the market.

Specific recommendations and settings will be discussed in the following chapters.
This chapter presents the following topics:

- **Round Robin - New Policy Recommendation** ................................................................. 11
- **Changing Policy from Fixed Path to Round Robin** ......................................................... 11
- **Additional Information** .................................................................................................... 12
Round Robin - New Policy Recommendation

Round Robin

This section provides the recommendations and steps to achieve optimal performance and continuous availability with EMC® VPLEX™

- The recommended multipathing setting is Round Robin for VPLEX Local, VPLEX Metro (non-cross-connect), and VPLEX Geo. The I/O Limit value should be left at the default setting of 1000.
  - For VPLEX Metro cross-connect with VMware, PowerPath/VE is highly recommended.
    - PowerPath/VE 5.8 includes the auto-standby feature which allows each ESXi host to automatically prefer to send I/O to its local VPLEX cluster over the remote cluster. The host paths connected to the local VPLEX Cluster will be the active paths whereas those connected to the remote VPLEX Cluster will be the standby paths.
    - For more information on PowerPath/VE and the auto-standby feature, see the support page: https://support.emc.com/products/1800_PowerPath-VE-for-VMware
  - Additional background: There are two problems with using NMP for VPLEX Metro cross-connect environments:
    - Round-robin path policy for a host connected to both VPLEX clusters will incur extra read and write latency for I/O operations to the remote cluster. Roughly half of the I/O will be local and half will be remote. WAN bandwidth for front-end host traffic will be consumed. Additional VPLEX inter-cluster cache-coherency traffic will be sent between clusters.
    - Fixed path policy requires a lot of manual administrative work to have all ESXi hosts and all volumes on both clusters to prefer their local cluster. For a handful of hosts and only a few volumes this might be acceptable. But for hundreds of hosts and thousands of volumes this is too onerous.
    - In addition, should the single preferred path fail for whatever reason, the new path chosen by a host might be at the remote cluster. And it’s entirely possible that multiple hosts could by the luck of the draw unfortunately all choose the same new remote director and thus overload that one director. A manual re-balancing of paths would be required at the new cluster, and then when the old cluster is back online, the exercise has to be repeated all over again.

Changing Policy from Fixed Path to Round Robin

Changing Policy

To change the default policy setting from Fixed to Round Robin for EMC VPLEX (Invista) devices, complete the following steps:

1. Open the vSphere CLI (recommended) or the service console.
2. Run the following command:
Chapter 2: ESXi NMP Policy

- For ESXi 5.x:
  
  # esxcli storage nmp satp set --default-psp=VMW_PSP_RR --satp=VMW_SATP_INV

- For ESXi/ESX 4.x:
  
  # esxcli nmp satp setdefaultpsp --psp VMW_PSP_RR --satp VMW_SATP_INV

- Set the Multipath Round Robin policy for the I/O operation limit to a value of 1000 for optimal performance with VPLEX. Currently, the VMware default value for this setting is 1000.

  1. Open the vSphere CLI (recommended) or the service console.
  2. Run the following commands:

     - For ESXi 5.x:
       
       To check the I/O operations limit:
       
       # esxcli storage nmp psp roundrobin deviceconfig get --device=device_NAA

       To set the I/O operations limit:
       
       #esxcli storage nmp psp roundrobin deviceconfig set --device=device_NAA --iops=1000 --type iops

     - For ESXi/ESX 4.x:

       To check the I/O operations limit:
       
       # esxcli nmp roundrobin getconfig --device=device_NAA

       To set the I/O operations limit:
       
       # esxcli nmp roundrobin setconfig --device=device_NAA --iops 1000 --type iops

**Additional Information**

For information on VMware and VPLEX versions recommended settings, refer to VMware Compatibility Guide for EMC VPLEX:


For related information, refer to the following VMware Knowledgebase article:

Changing the default pathing policy for new/existing LUNs (1017760) located at [http://kb.vmware.com/kb/1017760](http://kb.vmware.com/kb/1017760)
For information on VPLEX host cross-cluster connect configuration using PowerPath and autostandby when using an ICL instead of ISL with the latency to each VPLEX cluster being nearly identical please see the Knowledgebase article:

**Solution: How to modify esxi powerpath paths to standby alive at boot time on esxi host**

Specific recommendations and settings will be discussed in the following chapters.
Chapter 3: PowerPath

This chapter presents the following topics:

- **Introduction**................................................................. 15
- **Supported Versions and Settings**................................. 16
**Introduction**

PowerPath is a host-based software that provides path management. PowerPath operates with several storage systems including VPLEX, on several operating systems, with Fibre Channel.

**Path management**

PowerPath works with VPLEX to intelligently manage the frontend I/O paths.

Path refers to the physical route between a host and VPLEX Virtual Volume or Logical Unit (LU) as referred to by PowerPath.

This includes the host bus adapter (HBA) port, cables, a switch, VPLEX interface and port, and an LU. LU refers to a virtual volume addressable as a single storage volume behind a VPLEX target. For the iSCSI standard, path is the Initiator-Target-LU, or ITL nexus and encompasses the connection between the HBA, Storage Port and LUN.

Bus refers to two connected SAN edge points (for example, Fibre Channel fabric N-port addresses) in the storage configuration: an HBA port on the server on one end and an array port on the other. For the iSCSI standard, bus is the Initiator-Target, or IT nexus. This differs from a storage path, which refers to a host’s end-to-end storage connection with an LU. Typically, multiple storage paths traverse a single bus.

PowerPath supports multiple paths to a logical device, enabling PowerPath to provide:

- Automatic failover in the event of a hardware failure. PowerPath automatically detects path failure and redirects I/O to another path.
- Dynamic multipath load balancing. PowerPath distributes I/O requests to a logical device across all available paths, thus improving I/O performance and reducing management time and downtime by eliminating the need to configure paths statically across logical devices.

**PowerPath Features**

PowerPath features include:

- Multiple paths, for higher availability and performance — PowerPath supports multiple paths between a logical device and a host. Having multiple paths enables the host to access a logical device even if a specific path is unavailable. Also, multiple paths can share the I/O workload to a given logical device.
- Path management insight capabilities—PowerPath characterizes I/O patterns and aids in diagnosing I/O problems due to flaky paths or unexpected latency values.

Metrics are provided on:

- Read, write MB/seconds per LUN
- Latency distribution: high and low water marks per path
- Retries: number of I/Os that did not succeed down a path

PowerPath also defines and measures performance on I/O throughput, fault detection, and path restore. Three new CLI commands (powermt set perfmon, powermt display perf, and powermt display perf bus) provide this information. The
EMC PowerPath Family CLI and System Messages Reference Guide provides more information. The commands associated with path management insight may cause some performance overhead. However, EMC qualifications and in-house testing have displayed no measurable impact to performance.

- Expanded support for standby: autostandby—An autostandby feature has been added to the standby feature to automatically put paths into autostandby that have intermittent I/O failures (also known as flaky paths) and/or to automatically select autostandby for high-latency paths in VPLEX cross-connected Metro configurations. The EMC PowerPath Family CLI and System Messages Reference Guide provides more information.

**Supported Versions and Settings**

Please refer to the EMC Simple Support Matrix for latest supported versions required.

Before configuring VPLEX, complete the following on each host:

- Confirm that all necessary remediation has been completed.

This ensures that OS-specific patches and software on all hosts in the VPLEX environment are at supported levels according to the *EMC Support Matrix*.

- Confirm that each host is running VPLEX-supported PowerPath version and has at least one available path to each VPLEX fabric.

- Confirm that the load-balancing and failover policy is set to **Adaptive**.

PowerPath is highly recommended for the host cross-cluster connect configuration.

The load-balancing and failover policy apply to:

- ESXi
- Windows
- HP-UX
- AIX
- Linux

**Note:** Windows 2008 is no longer supported with VPLEX
This chapter presents the following topics:

**Introduction** .......................... 18

**AIX**  19

**HP-UX** ..................................... 21

**Linux** ........................................ 22

**Microsoft Windows** ......................... 31

**Solaris** ...................................... 37
Introduction

Most host operating systems provide a native form of multipath software known as MPIO. This allows utilization of additional paths for failover or multipath access while lowering the overhead cost involved with implementing a solution. These solutions are relatively basic and designed for solutions within the data center.

VPLEX has tested and qualified these solutions and provide guidance found in the ELAB Host Configuration Guides.

The following sections summarize the MPIO configuration settings specific to VPLEX.
AIX

Before configuring VPLEX in the IBM AIX environment, complete the following on each host:

- Confirm that all necessary remediation has been completed. This ensures that OS-specific patches and software on all hosts in the VPLEX environment are at supported levels according to the EMC Support Matrix.
- Confirm that each host is running VPLEX-supported failover software and has at least one available path to each VPLEX fabric.

**Note:** Always refer to the EMC Support Matrix for the most up-to-date support information and prerequisites.

- If a host is running EMC PowerPath, confirm that the load-balancing and failover policy is set to Adaptive.
- Support requires minimum AIX 7100-00-02-1041, AIX 6100-04-03-1009, and AIX 5300-11-02-1009.
- AIX attach to VPLEX requires minimum EMC ODM 5.3.0.3.
- PowerPath, Veritas DMP, and AIX MPIO are the only multipath solutions supported with VPLEX and AIX hosts.
- AIX MPIO minimum requirements are as follows:
  - Use ODM files on each host. This includes the reset_delay parameter set to 0. The ODM versions are posted at [ftp://ftp.emc.com/pub/elab/aix/ODM_DEFINITIONS](ftp://ftp.emc.com/pub/elab/aix/ODM_DEFINITIONS)
    - AIX 6.1 Technology Level 6100-08-00-1241, 6100-07-06-1241 and later. Requires minimum ODM version 5.3.0.8
    - AIX 7.1 Technology Level 7100-02-00-1241, 7100-01-06-1241 and later. Requires minimum ODM version 6.0.0.3
    - VIOS 2.2.2.0 and later. Requires minimum ODM version 5.3.0.8
    - VIOS 2.2.2.1-SP1 FP26 and later. Requires minimum ODM version 5.3.0.8

  The host must be rebooted after updating ODM file for the settings to take effect.

- The following configurations with AIX 6.1 and 7.1 are supported with MPIO:
  - AIX LVM must be used
  - IBM PowerHA, GPFS filesystem
  - VIOS, VIOC, LPAR, NPIV
- The following configurations are not supported at this time:
  - Oracle RAC
  - Veritas VxVM, VCS, SFRAC, VxCFS
Chapter 4: Native MPIO

- VPLEX 5.1.x is the minimum version supported

IMPORTANT

For optimal performance in an application or database environment, ensure alignment of your host's operating system partitions to a 32 KB block boundary.

Veritas DMP settings with VPLEX

- Veritas DMP 5.1 SP1 requires the asl package 5.1.100.100 to correctly detect VPLEX
- If a host attached to VPLEX is running Veritas DMP multipathing, change the following values of the DMP tunable parameters on the host to improve the way DMP handles transient errors at VPLEX in certain failure scenarios:
  - `dmp_lun_retry_timeout` for the VPLEX array to 60 seconds using the following command:
    "vxdmpadm setattr enclosure emc-vplex0 dmp_lun_retry_timeout=60"
  - `recoveryoption` to throttle and `iotimeout` to 30 using the following command:
    "vxdmpadm setattr enclosure emc-vplex0 recoveryoption=throttle iotimeout=30"
Before configuring VPLEX in the HP-UX environment, complete the following on each host:

- Confirm that all necessary remediation has been completed

This ensures that OS-specific patches and software on all hosts in the VPLEX environment are at supported levels according to the EMC Support Matrix.

- Confirm that each host is running VPLEX-supported failover software and has at least one available path to each VPLEX fabric.

Note: Always refer to the EMC Support Matrix for the most up-to-date support information and prerequisites.

- If a host is running EMC PowerPath, confirm that the load-balancing and failover policy is set to Adaptive
- The following are supported with VPLEX:
  - HP-UX 11iV3 with VxVM 5.1 SP1
  - HP-UX 11iV3 with VxMP 5.0 - 6.0
  - HP-UX 11iV2 with VxMP 4.1 - 5.0

**IMPORTANT**

For optimal performance in an application or database environment, ensure alignment of your host's operating system partitions to a 32 KB block boundary.

**Veritas DMP settings with VPLEX**

- Veritas DMP 5.1 SP1 requires the asl package 5.1.100.100 to correctly detect the VPLEX Virtual Volumes
- If a host attached to VPLEX is running Veritas DMP multipathing, change the following values of the DMP tunable parameters on the host to improve the way DMP handles transient errors at the VPLEX array in certain failure scenarios:
  - `dmp_lun_retry_timeout` for the VPLEX array to 60 seconds using the following command:

```
vxdmpadm setattr enclosure emc-vplex0 dmp_lun_retry_timeout=60
```
  - `recoveryoption` to throttle and `iotimeout` to 30 using the following command:

```
vxdmpadm setattr enclosure emc-vplex0 recoveryoption=throttle iotimeout=30
```
Linux

Native MPIO

DM-MPIO, native multipathing, provides a mechanism to address device names persistently through the use of udev and scsi-id. The names used to address multipath names rely on the properties of the physical device, and are thus both unique and consistent across reboots.

Red Hat Enterprise Linux (RHEL)

The following steps detail the procedure for configuring native multipath failover for Invista or VPLEX virtualized storage in a RHEL 4 host.

RHEL 4

To configure MPIO on a RHEL 4 host, complete the following steps:

1. Install the multipath-tools rpm package.

The device-mapper-multipath-tools package is not installed as part of a default installation of the operating system. Select the package as part of the OS install or install the rpm later either from the install CD or by downloading the rpm from the Red Hat website.

2. Verify that the required version of udev has been installed for your environment. To query the udev version, execute the following command:

```
# rpm -q udev
```

If required, upgrade the udev package and then execute the command to create a new device under the '/dev' directory.

```
# udevstart
```

3. Load the dm_multipath kernel module if it is not already loaded, as follows:

```
# modprobe dm_multipath
```

4. Replace the default /etc/multipath.conf with the following multipath.conf file recommended by EMC for attach to EMC storage.

To find the file on a RHEL system, issue the following command:

```
# rpm -ql `rpm -qa | grep multipath` | grep synt
```

The following is sample output from RHEL:

```
/usr/share/doc/device-mapper-multipath-0.4.5/multipath.conf.synthetic
```

IMPORTANT

The following is only an example. Consult Red Hat’s documentation to ensure the correct syntax is followed for your release. MPIO continues to evolve with each release of RHEL.

Note: VPLEX require its own specific configuration stanza in /etc/multipath.conf file.
## This is the /etc/multipath.conf file recommended for EMC storage devices.
## OS: RHEL 4
## Arrays: Invista/VPLEX
## The blacklist is the enumeration of all devices that are to be excluded from multipath control
## Replace the wwid with the output of the command MPIO
## 'scsi id -g -u -s /block/[internal scsi disk name]'
## Enumerate the wwid for all internal scsi disks.
## wwid 35005076718 d4224d
devnode ""(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]""
devnode ""hd[a-z][0-9]""
devnode ""cciss!c[0-9]d[0-9]*[p[0-9]*]"
## Use user friendly names, instead of using WWIDs as names.
defaults {
  user friendly names yes
} devices {
  # Device attributes requirements for Invista/VPLEX
  device {
    vendor "EMC"
    product "Invista"
    path checker tur
    no path retry 5
    product blacklist "LUNZ"
  }
}

5. Perform a dry run and evaluate the setup by running the `multipath` command.

    # multipath -v2 -d

create: mpath42 (360060480000190100501533030383644)
  [size=5 GB][features="0"] [hwhandler="0"]
  \ round-robin 0
  \ 11:0:1:52 sdao 66:128
  \ 11:0:2:52 sdaz 67:48
  \ 10:0:1:52 sdm 8:192
  \ 10:0:2:52 sdx 65:112
create: mpath43 (360060480000190100501533030383645)
  [size=5 GB][features="0"] [hwhandler="0"]
  \ round-robin 0
  \ 11:0:1:53 sdap 66:144
  \ 11:0:2:53 sdba 67:64
  \ 10:0:1:53 sdn 8:208
  \ 10:0:2:53 sdy 65:128

6. If the listing is appropriate, commit the configuration as follows:
   a. Start the required multipath processes.
      
      # /etc/init.d/multipathd start
   b. Execute the `multipath` command.
      
      # multipath -v2
   c. Perform an `lsmod` and verify that the processes are running.
      
      # lsmod |grep dm
      dm round robin 4929 1
Chapter 4: Native MPIO

To get a listing of the current setup do:

```bash
# multipath -ll
```

Integrate the startup of the appropriate daemons in the boot sequence as follows:

```bash
# chkconfig --add multipath
# chkconfig multipathd on
```

You may optionally reboot the host to verify that the required processes automatically startup.

Additional documentation is available in the directory `/usr/share/doc/device-mapper-multipath-x.x.x/`, the manpages on your system, and from Red Hat's website.

**RHEL 5 and RHEL 6**

The following steps detail the procedure for configuring native multipath failover on a RHEL 5 or RHEL 6 host.

The `/etc/multipath.conf` file is installed by default when you install the device-mapper-multipath package.

To configure MPIO on a RHEL 5 or RHEL 6 host, complete the following steps:

1. Replace the default `/etc/multipath.conf` with the following `multipath.conf` file recommended by EMC for attach to EMC storage.

To find the file on a RHEL system to modify, issue the following command:

```bash
# rpm -ql `rpm -qa | grep multipath` | grep synt
```

The following is a sample output from RHEL:

`/usr/share/doc/device-mapper-multipath-0.4.5/multipath.conf.synthetic`

**IMPORTANT**

The following is only an example. Consult Red Hat documentation to ensure the correct syntax is followed for your release. MPIO continues to evolve with each release of RHEL.

The Linux native MPIO in RHEL 5.5 and later and RHEL 6.0 and later already contain default configuration parameters for EMC VPLEX virtualized storage to provide optimal performance in most environments. There is no need to create a device stanza for VPLEX unless you wish to modify the default behavior.

If you wish to modify the default behavior the default stanza is contained in the `multipath.conf.defaults` located in `/usr/share/doc/device-mapper-multipath-x.x.x`. The device stanza begins with the following:

```bash
vendor "EMC"
product "Invista"
```
Copy this stanza into your `/etc/multipath.conf` file and modify it as desired.

```
# This is an example configuration file for device mapper multipath.
# For a complete list of the default configuration values, see
# /usr/share/doc/device-mapper-multipath-0.4.5/multipath.conf.defaults
# For a list of configuration options with descriptions, see
# /usr/share/doc/device-mapper-multipath-0.4.5/multipath.conf.annotated
# Blacklist all devices by default. Remove this to enable multipathing
# on the default devices.

# Note: Insert # to disable the all device blacklist and enable multipathing on your system.

#blacklist {
#  devnode "**"
#}
## By default, devices with vendor = "IBM" and product = "S/390.*" are
## blacklisted. To enable multipathing on these devices, uncomment the
## following lines.
#blacklist exceptions {
#  device {
#    vendor "IBM"
#    product "S/390.*"
#  }
#}
## Use user friendly names, instead of using WWIDs as names.
defaults {
  user friendly names yes
}
##
## This is a template multipath-tools configuration file
## Uncomment the lines relevant to your environment
230
Native Multipath Failover
##
defaults {
  udev dir /dev
  polling interval 10
  selector "round-robin 0"
  path grouping policy multibus
  getuid callout "/sbin/scsi_id -g -u -s /block/"%
  prio callout /bin/true
  path checker readsector0
  rr min io 100
  rr weight priorities
  fallback immediate
  no path retry fail
  user friendly name yes
}
##
## The wwid line in the following blacklist section is shown as an
## example
## of how to blacklist devices by wwid. The 3 devnode lines are the
## compiled in default blacklist. If you want to blacklist entire types
## of devices, such as all scsi devices, you should use a devnode line.
## However, if you want to blacklist specific devices, you should use
## a wwid line. Since there is no guarantee that a specific device will
## not change names on reboot (from /dev/sda to /dev/sdb for example)
## devnode lines are not recommended for blacklisting specific devices.
##
#blacklist {
wwid 360060480000019010196553303033230
devnode "\(ram|raw|loop|fd|md|dm-|sr|scd|st\) [0-9]*"
devnode "\(hd[a-z]\)"
devnode "\(cciss[c0-9]d[0-9]*\)"
}

#multipaths {
# multipath {
#  wwid 3600508b4000156d700012000000b0000
#  alias yellow
#  path grouping policy multibus
#  path checker readsector0
```
Chapter 4: Native MPIO

### Multipath Configuration

```
# path selector "round-robin 0"
# failback manual
# rr weight priorities
# no path retry 5
#
# multipath {
# wwid 1DEC  321816758474
# alias red 
# }
#}
#devices {
# device {
# vendor "COMPAQ "
# product "HSV110 (C)COMPAQ"
# path grouping policy multibus
# getuid callout "/sbin/scsi id -g -u -s /block/%n"
# path checker readsector0
# path selector "round-robin 0"
# hardware handler "0"
# failback 15
# rr weight priorities
# no path retry queue
# }
# device {
# vendor "COMPAQ"
# product "MSA1000"
# path grouping policy multibus
# }
#}
devices {
 Device attributes for EMC Invista/VPLEX
device {
 vendor "EMC"
 product "Invista"
 path checker tur
 no path retry 5
 product blacklist "LUNZ"
 }
```

2. Perform a dry run and evaluate the setup by running the multipath command:

```
# multipath -v2 -d [RHEL 5 & 6]
create: mpath15 (360060480000190101965533030423744) EMC,INVISTA
 [size=8.4G][features=0][hwhandler=0]
 \
 round-robin 0 [pric=2][undef]
 \ 2:0:0:49 sdp 8:240 [undef][ready]
 \ 3:0:0:49 sds 65:32 [undef][ready]
create: mpath16 (360060480000190101965533030423745) EMC,INVISTA
 [size=8.4G][features=0][hwhandler=0]
 \
 round-robin 0 [pric=2][undef]
 \ 2:0:0:50 sdp 65:0 [undef][ready]
 \ 3:0:0:50 sdt 65:48 [undef][ready]
```

3. If the listing is appropriate, commit the configuration as follows:

   a. Start the required multipath processes.
      
      ```
      # /etc/init.d/multipathd start
      ```

   b. Execute the multipath command:
      
      ```
      # multipath -v2
      ```

   c. Perform an `lsmod` and verify that the processes are running.

      With VPLEX attached, the following modules will be listed:
      
      ```
      # lsmod |grep dm
      dm round robin 4929 1
dm multipath 22097 2 dm round robin
dm mod 66433 1 dm multipath
      ```

4. To get a listing of the current setup issue the following command:
# multipath -ll
5. Integrate the startup of the appropriate daemons in the boot sequence as follows:

```
# chkconfig --add multipathd
# chkconfig multipathd on
```

You may optionally reboot the host to verify that the required processes automatically startup.

---

**Note:** Additional documentation is available in the directory `/usr/share/doc/device-mapper-multipath-x.x.x/`.

---

This section details the procedure for configuring native multipath failover (MPIO) for EMC VPLEX virtualized storage on a SuSE Linux Enterprise Server (SLES) host.

### SLES 9 and SLES 10

The following steps apply for SLES 9 and SLES 10 servers.

To configure MPIO on a SLES host, complete the following steps:

1. Install the multipath-tools rpm package.

   The multipath-tools package is not installed as part of a default installation of the operating system. The recommended multipath-tools rpm package is available in the installation medium for the operating system or from the SuSE website.

2. Verify that the required version of udev has been installed for your environment. To query the udev version, execute the following command:

   ```
   # rpm -q udev
   ```

   If required, upgrade the udev package and then recreate the devices.

   - On SLES 9 execute:
     ```
     # udevstart
     ```

   - On SLES 10 execute:
     ```
     # /etc/init.d/boot.udev restart
     ```

3. On SLES 9, SuSE recommends that 'subfs' auto-mount be disabled when using MPIO. This is achieved by editing the `/etc/sysconfig/hotplug` and setting the value, as follows:

   ```
   HOTPLUG USE SUBFS=no
   ```

   Also, to prevent significant boot delays, edit `/etc/sysconfig/boot` and set the value as follows:

   ```
   DISABLE BLKID=yes
   ```

4. Load the `dm_multipath` kernel module if it is not already loaded, as follows:

   ```
   # modprobe dm_multipath
   ```

5. The following is the `/etc/multipath.conf` file recommended for attach to EMC storage.
Chapter 4: Native MPIO

If /etc/multipath should not exist on your server, you can locate a copy to use by using the following Linux command:

```shell
# rpm -ql `rpm -qa | grep multipath` | grep synt
```

The following is sample output from SLES 9:

```
/usr/share/doc/packages/multipath-tools/multipath.conf.synthetic
```

**IMPORTANT**

The following is only an example. Consult SuSE documentation to ensure the correct syntax is followed for your release. MPIO continues to evolve with each release of SLES.

**Note:** EMC VPLEX require its own specific configuration stanza in /etc/multipath.conf file.

- On a SLES 9 system:

```plaintext
## This is the /etc/multipath.conf file recommended for
## EMC storage devices.
##
## OS : SLES 9
## Arrays : INVISTA/VPLEX
##
## The blacklist is the enumeration of all devices that are to be
## excluded from multipath control
blacklist {
## Replace the wwid with the output of the command
## `scsi id -g -u -s /block/[internal scsi disk name]'`
## Enumerate the wwid for all internal scsi disks.
##
## wwid 35005076718d4224d
##
## devnode "(^ram[0-9]*|loop[0-9]*|md[0-9]*)"
## devnode "^hd[a-z][0-9]*"
## devnode "^cissd[0-9]*[p[0-9]*]"
}
## Use user friendly names, instead of using WWIDs as names.
defaults {
  user friendly names yes
}
devices {
## Device attributes for EMC SYMMETRIX
device {
  vendor "EMC "
  product "INVISTA"
  path grouping policy multibus
  path checker tur
  no path retry 5
  product blacklist "LUNZ"
}
```

- On a SLES 10 SP3 and later and SLES 11 and later systems, the Linux native MPIO already contains default configuration parameters for EMC VPLEX virtualized storage to provide optimal performance in most environments. There is no need to create a device stanza for VPLEX unless you wish to modify the default behavior.

**IMPORTANT**

The following is an example of an edited file that would support VPLEX. Consult SuSE documentation to ensure the correct syntax is followed for your release. MPIO continues to evolve with each release of SLES.
If you wish to modify the default behavior the default stanza is contained in the multipath.conf.defaults located in /usr/share/doc/device-mapper-multipath-x.x.x. The device stanza begins with the following:

```plaintext
vendor "EMC"
product="INVISTA"
```

Copy this stanza into your /etc/multipath.conf file and modify it as desired.

```plaintext
## This is the /etc/multipath.conf file recommended for ## EMC storage devices

##
## OS : SLES 10
## Arrays: INVIsta/VPLEX
##
## defaults {
## user friendly names yes
##}
## blacklist {
## Replace the wwid with the output of the command MPIO
## 'scsi id -g -u -s /block/[internal scsi disk name]'
## Enumerate the wwid for all internal scsi disks.
##
## wwid 35005076718d4224
## devnode "(ram|raw|loop|fd|md|dm-|sr|scd|st) [0-9]*"
## devnode "^hd[a-z][0-9]*"
## devnode "cciss:c[0-9]d[0-9]*[p[0-9]*]"
##
## devices {
## Device attributes for EMC SYMMETRIX
##
## device {
## vendor "EMC"
## product "INVIsta"
## path grouping policy multibus
## path checker tur
## no path retry 5
## product blacklist "LUNZ"
##}
##}
```

6. Perform a dry run and evaluate the setup by running the `multipath` command:

- With Invista or VPLEX virtualized storage logical units on a SLES 9 host, the output will look similar to the following:

```
# multipath -v2 -d
```

```plaintext
create: 360060480000190100501533030383842
[size=898 MB] [features="0"] [hwhandler="0"]
\ round-robin 0 [ prio=4]
\ 2:0:2:105 sdab 65:160 [ready]
\ 2:0:3:105 sdds 66:192 [ready]
\ 3:0:2:105 sdcj 69:112 [ready]
\ 3:0:3:105 sddc 70:144 [ready]
create: 360060480000190100501533030383843
[size=898 MB] [features="0"] [hwhandler="0"]
\ round-robin 0 [ prio=4]
\ 2:0:2:106 sdab 65:176 [ready]
\ 2:0:3:106 sdat 66:208 [ready]
\ 3:0:2:106 sdc 69:128 [ready]
\ 3:0:3:106 sddc 70:160 [ready]
```

- With Invista or VPLEX virtualized storage logical units on a SLES 10 host, the output will look similar to the following:

```
# multipath -v2 -d
```

```plaintext
create: mpath27 (360060480000190100501533031353831) EMC,INVISTA
```
Chapter 4: Native MPIO

7. If the listing is appropriate, commit the configuration as follows.
   
a. Start the required multipath processes.
   
   # /etc/init.d/boot.multipath start
   # /etc/init.d/multipathd start
   
b. Perform an `lsmod` and verify that the processes are running.
   
   # lsmod | grep dm
   
   dm round robin 20480 1
   dm multipath 38544 2 dm round robin
   dm mod 77536 10 dm multipath
   
8. To get a listing of the current setup, type:

   # multipath -11
   
9. Integrate the startup of the appropriate daemons in the boot sequence as follows:

   # inserv boot.device-mapper multipathd boot.multipath

You may optionally reboot the host to verify that the required processes automatically startup.

Note: Additional documentation is available in the directory
/usr/share/doc/packages/multipath-tools/x.x.x/.
Microsoft Windows

Important

VPLEX is dropping support for Windows 2003.

Always refer to the VPLEX ESM for support details located on http://support.emc.com

Note: Native MPIO is not supported with failover clusters when stretched, or cross-connected. PowerPath must be used. See EMC knowledgebase article emc000187975.

Windows 2008 Server Core and Windows 2008 R2 Server Core

Note: Refer to Microsoft documentation for installing the Microsoft Multipath I/O feature.

Native MPIO must be configured to manage VPLEX. This can be accomplished by opening Control Panel then selecting the MPIO applet.

The claiming of array/device families can be done in one of two ways as described in “Method 1”, and in “Method 2”.

Method 1

Manually enter the Vendor and Device IDs of VPLEX for native MPIO to claim and manage.

Note: This may be the preferred method if all arrays are not initially connected during configuration and subsequent reboots are to be avoided.

To manually enter the array vendor and product ID information:

1. Use the MPIO-ed Devices tab in the MPIO Properties control panel applet.
2. Select Add and enter the vendor and product IDs of the VPLEX devices to be claimed by native MPIO.

The vendor ID must be entered as a string of eight characters (padded with trailing spaces) and followed by the product ID entered as a string of sixteen characters (padded with trailing spaces).

For example, to claim a VNX series and CLARiiON RAID 1 device in MPIO, the string would be entered as:

EMC*****Invista********

The asterisk is representative of a space.

<table>
<thead>
<tr>
<th>Array type</th>
<th>LUN type</th>
<th>Vendor ID</th>
<th>Product ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP.EX VS1/VS2</td>
<td>Any</td>
<td>EMC</td>
<td>Invista</td>
</tr>
</tbody>
</table>

Figure 1 Array and device types
Chapter 4: Native MPIO

Method 2
Use the MPIO applet to discover, claim, and manage the arrays already connected during configuration.

Note: This may be the preferred method if ease-of-use is required and subsequent reboots are acceptable when each array is connected.

IMPORTANT

MPIO limits the number of paths per LUN to 32. Exceeding this number will result in the host crashing with a Blue Screen stop message. Do not exceed 32 paths per LUN when configuring MPIO on your system.

Automatic discovery is configured using the Discover Multi-Paths tab of the MPIO Properties control panel applet. Note that only arrays which are connected with at least two logical paths will be listed as available to be added in this tab, as follows:

- Devices from VPLEX arrays will be listed in the Others section of the applet

Select the array / device types to be claimed and managed by MPIO by selecting the Device Hardware ID, and clicking the Add button.

Note: The OS will prompt you to reboot for each device type added. A single reboot will suffice after multiple devices types are added.

Path management in Multipath I/O for VPLEX

Following reboot, after all device types have been claimed by MPIO, each VPLEX-based disk will be shown in Device Manager as a Multi-Path Disk Device. When managed by MPIO, a new tab, named MPIO, will be available under Properties of the selected disk device. Under the MPIO tab, the number of logical paths configured between the host and array should be reported.

The default Load Balance Policy (as reported in the MPIO tab) for each disk device depends upon the type of disk device presented.

VPLEX devices will report a default Load Balance Policy as "Round Robin" with all active paths as “Active/Optimized”. The default policy can be overridden by changing the Load Balance Policy to any available, except "Fail Over Only". See the Windows Server 2008 and Windows Server 2008 R2 documentation for a detailed description of available Load Balance policies.

Note: The default Load Balance Policy cannot be changed globally for all disk devices. The change must be done on a per-disk device basis.

Enabling Native MPIO

Windows Server 2008 Server Core and Windows Server 2008 R2 Server Core

MPIO and other features must be started from the command line since Windows Server 2008 Server Core and Windows Server 2008 R2 Server Core are minimal installations that do not have traditional GUI interfaces. Refer to http://technet.microsoft.com for more information on Windows Server Core installations.

To enable the native MPIO feature from the command line, type:
Native MPIO with Windows Server 2012

Configuring Native MPIO for Windows Server 2012

start /w ocsetup Multipathlo

After the system reboots, you can manage MPIO with the mpiocpl.exe utility. From the command prompt, type:

mpiocpl.exe

The MPIO Properties window displays. From here, arrays/devices can be claimed and managed as described in the section above for standard Windows installations.

For more information on Microsoft Native MPIO, refer to:


Support for Native MPIO in Windows Server 2012

Windows Server 2012 include native multipathing (MPIO) support as a feature of the OS.

Native MPIO is supported with EMC CX4 Series, DMX-4, and VMAX storage array models.

Note the following:

- To use the Microsoft default DSM, storage must be compliant with SCSI Primary Commands-3 (SPC-3)
- Default Microsoft MPIO Timer Counters are supported
- VPLEX is active/active and does not support mode 4 [ALUA]

This section explains how to configure native MPIO for EMC storage arrays. Native MPIO is installed as an optional feature of the Windows Server 2012.

Note: Refer to Microsoft documentation for installing the Microsoft Multipath I/O feature.

Configuring MPIO and installing DSM

When MPIO is installed, the Microsoft device-specific module (DSM) is also installed, as well as an MPIO control panel. The control panel can be used to do the following:

- Configure MPIO functionality
- Install additional storage DSMs
- Create MPIO configuration reports

Opening the MPIO control panel

Open the MPIO control panel either by using the Windows Server 2012 control panel or by using Administrative Tools.

To open the MPIO control panel using the Windows Server 2012 control panel, complete the following steps:

1. On the Windows Server 2012 desktop, move your mouse to the lower left corner and click Start.
2. Click MPIO.
Chapter 4: Native MPIO

To open the MPIO control panel using Administrative Tools, complete the following steps:

1. On the Windows Server 2012 desktop, move your mouse to the lower left corner and click **Start**.
2. Point to **Administrative Tools** and click **MPIO**.

The MPIO control panel opens to the **Properties** dialog box.

**Note:** To access the MPIO control panel on Server Core installations, open a command prompt and type `MPIOCPL.EXE`.

Once installed, native MPIO must be configured to manage VPLEX. Open **Control Panel**, then the **MPIO** applet.

**IMPORTANT**

MPIO limits the number of paths per LUN to 32. Exceeding this number will result in the host crashing with a Blue Screen stop message. Do not exceed 32 paths per LUN when configuring MPIO on your system.

Automatic discovery is configured using the **Discover Multi-Paths** tab of the **MPIO Properties** control panel applet. Note that only arrays which are connected with at least two logical paths will be listed as available to be added in this tab, as follows:

- Devices from VPLEX arrays will be listed in the **Others** section of the applet

Select the array and device types to be claimed and managed by MPIO by selecting the Device Hardware ID, and clicking **Add**.

![Figure 2 MPIO Properties dialog box](image)

**Device discovery and claiming devices for MPIO**
Path management in Multipath I/O for VPLEX

Following reboot, after all device types have been claimed by MPIO, each VPLEX-based disk will be shown in Device Manager as a Multi-Path Disk Device.

When managed by MPIO, a new tab, MPIO, will be available under Properties of the selected disk device. Under the MPIO tab, the number of logical paths configured between the host and array should be reported.

This tab will also allow you to change the MPIO load balancing policy for a disk device.

**Note:** Some load balancing policies may not be available for specific array disk types. For example, the Round Robin policy is not available for VNX disk devices, but Round Robin with Subset is.

Options available for load balancing policies are as follows:

- **Fail Over Only** — Policy that does not perform load balancing. This policy uses a single active path, and the rest of the paths are standby paths. The active path is used for sending all I/O. If the active path fails, then one of the standby paths is used. When the path that failed is reactivated or reconnected, the standby path that was activated returns to standby.

- **Round Robin** — Load balancing policy that allows the Device Specific Module (DSM) to use all available paths for MPIO in a balanced way. This is the default policy that is chosen when the storage controller follows the active-active model and the management application does not specifically choose a load-balancing policy.

- **Round Robin with Subset** — Load balancing policy that allows the application to specify a set of paths to be used in a round robin fashion, and with a set of standby paths. The DSM uses paths from a primary pool of paths for processing requests as long as at least one of the paths is available. The DSM uses a standby path only when all the primary paths fail. For example, given 4 paths: A, B, C, and D, paths A, B, and C are listed as primary paths and D is the standby path. The DSM chooses a path from A, B, and C in round robin fashion as long as at least one of them is available. If all three paths fail, the DSM uses D, the standby path. If paths A, B, or C become available, the DSM stops using path D and switches to the available paths among A, B, and C.

- **Least Queue Depth** — Load balancing policy that sends I/O down the path with the fewest currently outstanding I/O requests. For example, consider that there is one I/O that is sent to LUN 1 on Path 1, and the other I/O is sent to LUN 2 on Path 1. The cumulative outstanding I/O on Path 1 is 2, and on Path 2, it is 0. Therefore, the next I/O for either LUN will process on Path 2.

- **Weighted Paths** - Load balancing policy that assigns a weight to each path. The weight indicates the relative priority of a given path. The larger the number, the lower ranked the priority. The DSM chooses the least-weighted path from among the available paths.

**Note:** The OS will prompt you to reboot for each device type added. A single reboot will suffice after multiple devices types are added.
Chapter 4: Native MPIO

- **Least Blocks** - Load balancing policy that sends I/O down the path with the least number of data blocks currently being processed. For example, consider that there are two I/Os: one is 10 bytes and the other is 20 bytes. Both are in process on Path 1, and both have completed Path 2. The cumulative outstanding amount of I/O on Path 1 is 30 bytes. On Path 2, it is 0. Therefore, the next I/O will process on Path 2.

The default **Load Balance Policy** (as reported in the MPIO tab) for each disk device depends upon the type of disk device presented:

- VPLEX devices will report a default **Load Balance Policy** as "Round Robin" with all active paths as "Active/Optimized."

Load balancing policies should be changed based on your particular environment. In most cases, the default policy will be suitable for your I/O load needs. However, some environments may require a change to the load balancing policy to improve performance or better spread I/O load across storage front-end ports. EMC does not require a specific load balancing policy for any environment, and our customers are free to make changes to their load balancing policies as they see fit to meet their environment's needs.

For more information on Microsoft Native MPIO, refer to:

Solaris

Before configuring VPLEX in the Solaris environment, complete the following on each host:

- Confirm that all necessary remediation has been completed.
- Confirm that each host is running VPLEX-supported failover software and has at least one available path to each VPLEX fabric.

**Note:** Always refer to the *EMC Support Matrix* for the most up-to-date support information and prerequisites.

- If a host is running EMC PowerPath, confirm that the load-balancing and failover policy is set to Adaptive.
- To run DMP, VPLEX 4.2 and Symantec 5.1 with the appropriate asl package are required.

To improve the way DMP handles transient errors at the VPLEX array in certain failure scenarios, the following attributes must be changed from the default value:

- `dmp_lun_retry_timeout` for the VPLEX array to 60 seconds using the following command:

  ```
  vxdmpadm setattr enclosure emc-vplex0 dmp_lun_retry_timeout=60
  ```

- `recoveryoption` to throttle and `iotimeout` to 30 using the following command:

  ```
  vxdmpadm setattr enclosure emc-vplex0 recoveryoption=throttle iotimeout=30
  ```

**IMPORTANT**

For optimal performance in an application or database environment, ensure alignment of your host’s operating system partitions to a 32 KB block boundary.

**VPLEX and MultiPath software**

VPLEX can work with PowerPath, DMP, or MPxIO as multipath software in Oracle Solaris environment, as described briefly in this section.

**VPLEX and EMC PowerPath**

EMC PowerPath can work with VPLEX devices as a multipath software in Oracle Solaris host as it does with Symmetrix or CLARiiON devices in Active/Active mode.

Refer to *VPLEX EMC Simple Support Matrix* for supported PowerPath versions

Refer to https://support.emc.com for the PowerPath's user guide and other PowerPath support documentation.
VPLEX and DMP

Start with VPLEX 4.2 and Symantec VxVM 5.1. You can use DMP as multipath software in Oracle Solaris host to manage VPLEX devices.

IMPORTANT

Symantec DMP will be disabled automatically when PowerPath is installed. Uninstall PowerPath is required if you want to use DMP as host multipath management application.

To use DMP with VPLEX devices, It required that we must install an addition ASL package from Symantec.

Symantec ASL package for VPLEX can be downloaded from following location:
https://sort.symantec.com/asl

Follow the instruction to install ASL as provided by Symantec.

The default path policy is balanced.

The following are commands to verify ASL installation and the version of ASL.

1. Enable vxctl and then run the vxdladm list support command to ensure that the ASL is correctly installed.

bash-2.05# vxdladm listsupport libname=libvxInvista.so

bash-2.05# vxdladm listsupport libname=libvxInvista.so
ATTR NAME      ATTR VALUE
LIBNAME        libvxInvista.so
VID EMEC
PID Invista
ARRAY TYPE Inv-A/A, VPLEX-A/A
ARRAY NAME EMC Invista, EMC VPLEX

2. Verify the ASL version.

bash-2.05# vxdladm listversion libname=libvxInvista.so
LIB NAME ASL VERSION Min. VXVM version
libvxInvista.so.so vm-5.1-rev-1 5.1

3. After installing ASL package, a VPLEX device is recognized as follows:

# vxmpadm list dmpnode
dmpdev = emc-vpex0 00b6
state = enabled
enclosure = emc-vpex0
cab-sno = FNM00094900286
as1 = libvxInvista.so
vid = EMEC
pid = Invista
array-name = EMC-VPLEX
array-type = VPLEX-A/A
ipolicy = Balanced
av1d = 00b6
lun-sno = 60001440000000010A001CAAF895253B6
uid = EMC%5FInvista%5FFNM00094900286%20%20%5F60001440000000010A001CAAF895253B6
dev-attr = -
###path = name state type transport ctrlr hwpath aportID aportWWN attr
path = c2t500014425001CA10d2s2 enabled(a) - FC c2
/pci@e,60000/fibre-channel@2/fp@0,0 - 50:00:14:42:50:01:ca:10 -
path = c4t500014424001CA10d2s2 disabled - FC c4
/pci@e,60000/fibre-channel@3/fp@0,0 - 50:00:14:42:40:01:ca:10 -

**IMPORTANT**

ASL is tightly coupled with the VPLEX SCSI personality. If there is any change in the SCSI personality in the future VPLEX releases, then there should be a corresponding changes in the ASL as applicable.

**VPLEX and MPxIO**

Beginning with S10u11 and S11.1, you can use the MPxIO as multipath software in Oracle Solaris host to manage VPLEX devices.

**Required patches for MPxIO**

To use MPxIO with VPLEX devices, the following patches are required:

- For Solaris 11.1 (both Sparc and x86):
  - SRU 5.5
- For Solaris S10u11 (1/13) Sparc platform:
  - 148888-02 (kernel patch)
  - 142088-03 (fp/luxadm patch)
  - 150115-02 (sd patch)
- For Solaris S10u11 (1/13) x86 platform:
  - 148889-02 (kernel patch)
  - 142089-03 (fp/luxadm patch)
  - 150116-02 (sd patch)

Refer the README file that came with your SRU or patches for detail how to install those patches.

**Enabling MPxIO**

After the patches are installed, complete the following changes in the host to enable MPxIO:

- With S11.1 OS
  - Set the parameter mpxio_disable in the file
    `/etc/driver/drv/fp.conf` to `mpxio-disable="no"

- With S10u11 OS
  - Set the parameter mpxio_disable in the file
    `/kernel/drv/fp.conf` to `mpxio-disable="no"

- You must set the following parameter in the file
  `/kernel/drv/scsi_vhci.conf` as:

    load-balance="round-robin"
auto-failback="enable"

- If you are booting from the internal disk:, reboot the host.
- If you are booting from the SAN, issue the following command from the host:
  '/usr/sbin/stmsboot -e'

Stmsboot enables MPxIO and makes the appropriate changes in your system file (/etc/vfstab) so that you will be able to reboot after MPxIO is enabled since your boot disk's name will change to MPxIO name after the reboot.

If you do not use the stmsboot command, the system may fail to reboot after MPxIO is enabled.

There is a limited support of Solaris Cluster with VPLEX when using MPxIO as multipathing.

**IMPORTANT**

MPxIO is the only multipath software that can be used with Solaris Cluster and VPLEX. PowerPath is not supported.

To use Solaris Cluster with VPLEX and MPxIO the following restrictions must be strictly applied:

- Support only includes Solaris 10 1/13 (S10u11) with the Solaris Cluster SC3.3 3/13 (u2) or SC3.3, SC3.3 5/11 (u1) with the latest cluster core patch and agents patches.
- Only use cluster with local VPLEX configuration. (Stretch cluster over Metro configuration, Campus Cluster is not supported at this time.)
- Only use cluster with no more than two (2) nodes.
- Each VPLEX devices used by the cluster must not have more than two access paths cluster-wide.
- The quorum device must be a VPLEX device. No other shared storage device nor Quorum Server can be used.
- The cluster must use the SCSI-2 reservation.

The cluster must be configured with the following setting (to make sure they are using SCSI-2 reservation):

1. Cluster's global_fencing property is set to ‘pathcount’.
2. The quorum 'Access Mode' must be ‘scsi2’.
3. The property ‘default_fencing’ must set to ‘global’ for the quorum device and all data devices.

To check the current setting of the global_fencing property, use the following command:

`'/usr/cluster/bin/cluster show -t global'`
- To change the current setting of the `global_fencing` property to ‘pathcount’, use the following command:

  `'/usr/cluster/bin/cluster set --p global_fencing=pathcount'`

- To check the current setting of ‘default_fencing’ property and the number of device access paths for each device, use the following command:

  `'/usr/cluster/bin/cldevice show'` to show all devices

  or:

  `'/usr/cluster/bin/cldevice show d<N>'` to show a specific device d<N>

To change the ‘default_fencing’ property to ‘global’, use the following command:

`'/usr/cluster/bin/cldevice set --p default_fencing=global d<N>'`

(N is the device’s did number)

- After setting your `global_fencing` property to ‘pathcount’. The following steps must be made to guarantee your quorum device is using SCSI-2 reservation.

  - Make sure your quorum devices is using ‘Access Mode’ is scsi2, using the command

    `'/usr/cluster/bin/clquorum show'`

  - If the quorum device “Access Mode” is scsi3, the following steps must be done (at least step a and step b):

    a. Replace the quorum device with a new one.
    b. Change the ‘default_fencing’ property of the old quorum device to ‘global’.
    c. Replace the original quorum device back (in case you want to re-use the original quorum device).