Installation and Reference Guide Device Mapper Multipath Enablement Kit for HP StorageWorks Disk Arrays

Version 4.0.0

Part number: AA-RW8RD-TE
First edition: February 2008

Contents

Tables

About this guide

HP Device Mapper Multipath (HPDM Multipath) software kit is an HP released Device Mapper multipath binary based on Device Mapper included in Linux operating system distributions with enablement for HP StorageWorks Disk Arrays. This guide describes the Device Mapper multipath enablement for HP StorageWorks Disk Arrays (HPDM Multipath) and provides information to help you:

- Determine hardware and software prerequisites
- Install HPDM Multipath
- Configure HPDM Multipath
- Use HPDM Multipath
- Troubleshoot HPDM Multipath

NOTE:

This version of HPDM Multipath supports all the features available in the earlier versions. It also supports a range of new features and OS. For more information, see What's New.

Intended audience

This document is intended for users who install, configure, and manage HPDM Multipath in their Linux server environment. Readers must be familiar with Linux system administration, including hardware and software installation.

Document conventions and symbols

Table 1 Document conventions

Convention	Element
Medium blue, underlined text (http://www.hp.com)	Website addresses
Bold font	• Key names • Text typed into a GUI element, such as into a box • GUI elements that are clicked or selected, such as menu and list items, buttons, and check boxes
Italic font	Text emphasis
`Monospace font`	• File and directory names • System output • Code • Text typed at the command line
`Monospace, italic font`	• Code variables • Command-line variables
`Monospace, bold font`	Emphasis of file and directory names, system output, code, and text typed at the command line

IMPORTANT:

Provides clarifying information or specific instructions.

NOTE:

Provides additional information.

HP technical support

Telephone numbers for worldwide technical support are listed on the HP support website: http://www.hp.com/support/

Collect the following information before calling:

- Technical support registration number (if applicable)
- Product serial numbers
- Product model names and numbers
- Applicable error messages
- Operating system type and revision level
- Detailed, specific questions

For continuous quality improvement, calls may be recorded or monitored.

HP recommends that customers sign up online using the Subscriber's choice website: http://www.hp.com/go/e-updates

- Subscribing to this service provides you with e-mail updates on the latest product enhancements, newer versions of drivers, and firmware documentation updates as well as instant access to numerous other product resources.
- After signing up, you can locate your products by selecting **Business support** and then **Storage** under Product Category.

Helpful websites

For additional product information, see the following websites:

- http://www.hp.com/country/us/eng/prodserv/storage.html
- http://www.hp.com/support/
- http://www.docs.hp.com
- http://sourceware.org/dm/
- http://download.qlogic.com/drivers/56660/readme_driver_80107.html#driver_parameters

1 Introduction to HPDM Multipath

This chapter addresses the following topics:

Overview

HPDM Multipath software kit is an HP released Device Mapper multipath binary, based on Device Mapper Multipath included in Linux Operating System distributions, with enablement for HP StorageWorks Disk Arrays. Device Mapper is an infrastructure in the Linux kernel. It provides a generic way to create virtual layers of block devices. It supports striping, mirroring, snapshots, concatenation, and multipathing. The multipath feature is provided with combination of DM Multipath kernel modules and multipath-tools user-space package.

Features

HPDM Multipath supports the following features:

- Detects path failure and automatically reroutes (failover) I/O to an available alternate path, if an existing connection fails.
- Provides an option to perform failback of the LUN to the repaired paths.
- Implements failover or failback actions transparently without disrupting applications.
- Monitors each path and notifies if the path status changes.
- Facilitates I/O load balancing among the multiple paths.
- Provides CLI with display options to configure and manage Multipath features.
- Provides an option to customize names for the HPDM Multipath devices.
- Provides device name persistence across reboots or configuration SAN changes.
- Provides policy-based path grouping for the user to customize the I/O flow through specific set of paths.
- Provides online LUN addition and deletion support.
- Provides support for the HP StorageWorks Arrays (EVA-AA, MSA-AA, and XP) with latest firmware revisions.
- Provides multipathing in clustered environment.
- Provides support for multipathing devices to be configured under Logical Volume Manager.

What's New

HPDM Multipath provides the following additional features as part of this version:

- Provides support for RHEL4.5/RHEL4.6/RHEL5.1/SLES9 SP4 and SLES10 SP1 along with RHEL5/SLES9 SP3/SLES10 OS distributions.
- Provides support for HP StorageWorks EVA4400 Disk Arrays.
- Provides support for HP StorageWorks MSA1000/1500 Disk Arrays.
- Provides support for HP StorageWorks MSA2000 Product family Disk Arrays (MSA2012fc/MSA2212fc/MSA2012i).

- Provides support for XP20000/XP24000 Disk Arrays.
- Provides support for EVA iSCSI devices on RHEL5/RHEL5.1/SLES9 SP3/SLES9 SP4/SLES10/SLES10 SP1.

Software Components

This section discusses the following software components of Device Mapper Multipath:

- **dm-multipath kernel module** -- Routes I/O and provides `failover` to paths and path groups.
- **multipath configuration tool** -- Provides commands to configure, list, and flush Multipath devices.
- **multipathd daemon** -- Monitors path status. When paths revert, `multipathd` daemon may also initiate path group switches to ensure that the optimal path group is used.
- **kpartx utility** -- Reads partition tables on the specified device and creates device maps over the detected partitions. The **kpartx utility** is called from `hotplug` whenever device maps are created and deleted.
- **devmap-name** -- Provides a meaningful device name to `udev` for device maps (devmaps).

HPDM Multipath Operation

HPDM Multipath enables hosts to route I/O over the multiple paths available to an end storage unit (LUN). A path refers to the connection from an HBA port to a storage controller port. When an active path through which I/O happens fails, HPDM Multipath reroutes the I/O over other available paths. On a Linux host, when there are multiple paths to a storage controller, each path appears as a separate block device and hence results in multiple block devices for single LUN. HPDM Multipath creates a new Multipath block device for those devices that have the same LUN WWN.

For example, a host with two HBAs when attached to a storage controller with two ports through a single FC switch provides four block devices: `/dev/sda`, `/dev/sdb`, `/dev/sdc`, and `/dev/sdd`. HPDM Multipath creates a single block device, `/dev/mapper/mpath1`, that reroutes I/O through these four underlying block devices.

2 Installing HPDM Multipath

This chapter addresses the following topics:

- HPDM Multipath Support Matrix
- Installing HPDM Multipath Tools
- Setting Up HPDM Multipath

HPDM Multipath Support Matrix

Table 2 lists the hardware and software prerequisites for installing HPDM Multipath.

Table 2 Hardware and Software Prerequisites

System Feature	Supported hardware and software			
Operating system versions	Red Hat Enterprise Linux 4 Update 5 Red Hat Enterprise Linux 4 Update 6 Red Hat Enterprise Linux 5 errata 2.6.18–8.1.6.el5 or later Red Hat Enterprise Linux 5 Update 1 SuSE LINUX Enterprise Server 9 - Service Pack 3 errata 2.6.5–7.283 or later SuSE LINUX Enterprise Server 9 - Service Pack 4 SuSE LINUX Enterprise Server 10 SuSE LINUX Enterprise Server 10 - Service Pack 1			
Host Bus Adapters (HBA)	**Qlogic:**			
	FCA2214	PCI-X	1-port 2Gbps	Fibre Channel
	FCA2214DC	PCI-X	2-port 2Gbps	Fibre Channel
	A7538A	PCI-X	1-port 2Gbps	Fibre Channel
	A6826A	PCI-X	2-port 2Gbps	Fibre Channel
	FC1142SR	PCI-E	1-port 4Gbps	Fibre Channel
	FC1242SR	PCI-E	2-port 4Gbps	Fibre Channel
	FC1143	PCI-X	1-port 4Gbps	Fibre Channel
	FC1243	PCI-X	2-port 4Gbps	Fibre Channel
	AB379A	PCI-X	2-port 4Gbps	Fibre Channel
	QMH 2462	4 Gbps Mezzanine for HP Blade Servers		
	Emulex:			
	A8002A/LPe1150	PCI-E	1-port 4Gbps	Fibre Channel
	A8003A/LPe11002	PCI-E	2-port 4Gbps	Fibre Channel
	AD167A	PCI-X 2.0	1-port 4Gbps	Fibre Channel
	AD168A	PCI-X 2.0	2-port 4Gbps	Fibre Channel
Servers	HP BladeSystem c-Class Server Blades, ProLiant x86, ProLiant AMD64, ProLiant EM64T Servers, Integrity Servers			

Supported arrays	EVA 3000 (HSV101) VCS 4.007 or later EVA 5000 (HSV111) VCS 4.007 or later EVA 4000 (HSV200) XCS 5.110/6.100 or later EVA 6000 (HSV200) XCS 5.110/6.100 or later EVA 8000 (HSV210) XCS 5.110/6.100 or later EVA 4100 (HSV200) XCS 6.110 or later EVA 6100 (HSV200) XCS 6.110 or later EVA 8100 (HSV210) XCS 6.110 or later EVA 4400 (HSV300) XCS 0900 or later XP128* fw rev 21-14-18-00/00 or later XP1024* fw rev 21-14-18-00/00 or later XP10000 fw rev 50-07-30-00/00 or later XP12000 fw rev 50-09-34-00/00 or later XP 20000 fw rev 60-02-04-00/00 or later XP 24000 fw rev 60-02-04-00/00 or later MSA 1000/MSA 1500 fw rev 7.0.0 or later MSA 2000 Storage product family (MSA2012fc/MSA2212fc/MSA2012i) *If you use XP128/1024, the system modes 140 and 293 must be switched on, otherwise the XP does not respond to a SCSI inquiry to code page 0x83 with a unique serial number (scsi_id tool). This prevents the HPDM Multipath from detecting the redundant paths. Provides support for iSCSI devices presented from supported EVA disk arrays.
HBA drivers	Emulex: 8.0.16.32 or later (for SLES9/RHEL 4.5/RHEL 4.6), 8.1.6.8 or later (for SLES10), 8.1.10.11 or later (for RHEL5/RHEL5.1/SLES10 SP1) available at: http://h18006.www1.hp.com/products/storageworks/4gbpciehba/index.html Qlogic: Version 8.01.07.25 or later for all OS distributions available at: http://h18006.www1.hp.com/products/storageworks/fca2214/index.html

NOTE:

- On any single host, the HBAs must be from the same HBA vendor.
- Does not support coexistence with other multipath products.
- Does not support Active-Passive Storage Arrays.
- EVA iSCSI devices are not supported on RHEL 4U5 and RHEL 4U6.

Prerequisites for HPDM Multipath Installation

Ensure the following RPMs are installed on the system before proceeding with HPDM Multipath installation:

- For RHEL 5/RHEL 5.1:
  `libsysfs-devel, device-mapper >= 1.02.02-2, libselinux-devel, libsepol-devel, readline-devel, ncurses-devel`
- For RHEL 4.6:
  `sysfsutils-devel, device-mapper, libaio-devel`
- For RHEL 4.5:
  `sysfsutils-devel, device-mapper`
- For SLES10 SP1:
  `device-mapper, device-mapper-devel, readline-devel, sysfsutils`
- For SLES10:
  `device-mapper, readline-devel, sysfsutils`
- For SLES9 SP3 /SLES9 SP4:
  `readline-devel, device-mapper, ncurses-devel`

- For all OS:
  ```
  aaa_base acl attr bash bind-utils bison bzip2 coreutils cpio cpp
  cracklib cvs cyrus-sasl db devs diffutils e2fsprogs file filesystem
  fillup findutils flex gawk gdbm-devel glibc glibc-devel glibc-locale
  gpm grep groff gzip info insserv kbd less libacl libattr libgcc
  libselinux libstdc++ libxcrypt m4 make man mktemp module-init-tools
  ncurses net-tools netcfg openldap2-client openssl pam pam-modules patch
  permissions popt procinfo procps psmisc pwdutils rcs readline sed strace
  syslogd sysvinit tar tcpd texinfo timezone unzip util-linux vim zlib
  zlib-devel autoconf automake binutils gcc gdbm gettext libtool perl rpm
  termcap udev
  ```

Installing/Upgrading HPDM Multipath Tools

To install or upgrade HPDM Multipath tools, complete the following steps:

1. Download the HPDM Multipath tools installation package from the following HP website at: http://www.hp.com
2. Log in as `root` to the host system.
3. Copy the installation `tar` package to a temporary directory (for example, `/tmp/HPDMmultipath`).
4. To unbundle the package, enter the following commands:
   ```
   # cd /tmp/HPDMmultipath
   # tar -xvzf HPDMmultipath-4.0.0.tar.gz
   # cd HPDMmultipath-4.0.0
   ```
5. Verify that the directory contains the `INSTALL.sh` shell script, the `SRPMS`, README.txt, and the `docs` directories.
6. To install or upgrade HPDM Multipath tools software on the server, enter the following command:
   ```
   # ./INSTALL.sh
   ```
7. Follow the on-screen instructions to complete the installation.
8. When the installation is complete, enter the following command:
   ```
   #rpm -qa |grep multipath
   ```
9. Verify that the output lists the `HPDMmultipath-tools-4.0.0-<OS>` package.

Setting up HPDM Multipath

Setting up HPDM Multipath involves the following:

- Configuring QLogic HBA Parameters
- Configuring Emulex HBA parameters

Configuring QLogic HBA Parameters

You must configure the QLogic HBA parameters for HPDM Multipath. The HBA time outs are typically set up for non HPDM Multipath environments, where longer time outs are necessary. This is because the only alternative is to send an error to the application. However, with Multipath, errors such as cable failures must be intimated at the earliest so that the Multipath layer can quickly take action and redirect the I/O to another path.

To configure the QLogic HBA parameter, complete the following steps:

1. For QLogic `2xxx` family of HBAs, edit the `/etc/modprobe.conf` file in RHEL4 or RHEL5 hosts and `/etc/modprobe.conf.local` file in SLES9/SLES10 hosts with the following values:

   ```
   options qla2xxx qlport_down_retry=10 ql2xfailover=0
   ```
2. Rebuild the `initrd` by executing the following script:

   ```
   /opt/hp/src/hp_qla2x00src/make_initrd
   ```
3. Reboot the host.

Configuring Emulex HBA Parameters

To configure the Emulex HBA parameter, complete the following steps:

1. For Emulex `lpfc` family of HBAs:
   - In RHEL4 hosts, edit the `/etc/modprobe.conf` file with the following values:

     ```
     options lpfc
               lpfc_nodev_tmo=14
               lpfc_lun_queue_depth=16
               lpfc_discovery_threads=32
     ```

   - In SLES9 hosts, edit the `/etc/modprobe.conf.local` file with the following values:

     ```
     options lpfc
               lpfc_nodev_tmo=14
               lpfc_lun_queue_depth=16
               lpfc_discovery_threads=32
     ```

   - In SLES10 hosts, edit the `/etc/modprobe.conf` file with the following values:

     ```
     options lpfc
               lpfc_nodev_tmo=28
               lpfc_lun_queue_depth=16
               lpfc_discovery_threads=32
     ```

   - In RHEL5 hosts edit the `/etc/modprobe.conf` file with the following values:

     ```
     options lpfc
                lpfc_nodev_tmo=28
                lpfc_lun_queue_depth=16
                lpfc_discovery_threads=32
     ```

2. Rebuild the `initrd` by executing the following script:

   ```
   /opt/hp/hp-lpfc/make_initrd
   ```

3. Reboot the host.

NOTE:

Ensure you have a HBA driver without multipathing capability installed. For more information, see the installation and reference guides for the respective HBA drivers.

3 Configuring HPDM Multipath

This chapter addresses the following topics:

Using `/etc/multipath.conf` File

The first step of configuring HPDM Multipath involves editing the `/etc/multipath.conf` configuration file, which is distributed as part of the HPDM Multipath tools user space package. This file also enables you to customize HPDM Multipath for a variety of SAN storage subsystems .

The `multipath.conf` file provides the following options to configure the attributes of a `multipath` device:

- I/O grouping policy
- polling interval for subsequent path status verification
- path selector algorithm
- program used by multipath to obtain a unique path identifier

The HPDM Multipath package comes with a template file with the recommended configuration for HP supported arrays. To use this template file:

For **RHEL** systems:

1. Copy the `/usr/share/doc/HPDMmultipath-tools-4.0.0/multipath.conf.HPTemplate` file as `/etc/multipath.conf`

For **SLES** systems:

1. Copy the `/usr/share/doc/packages/HPDMmultipath-tools-4.0.0/multipath.conf.HPTemplate` file as `/etc/multipath.conf`

The `/etc/multipath.conf` file consists of the following sections, to configure the attributes of a multipath device:

- System defaults (`defaults`)
- Black-listed devices (`devnode_blacklist/blacklist`)
- Storage array model settings (`devices`)
- Multipath device settings (`multipaths`)
- Blacklist exceptions (`blacklist_exceptions`)

The `defaults` section defines default values for attributes which are used whenever required setting is unavailable. The `blacklist` section defines which devices should be excluded from the multipath topology discovery. The `blacklist_exceptions` section defines which devices should be included in the multipath topology discovery, despite being listed in the `blacklist` section. The `multipaths` section defines the multipath topologies. They are indexed by a World Wide Identifier (wwid). The `devices` section defines the device-specific settings based on vendor and product values.

Table 3 lists the important attributes of HP supported arrays.

Table 3 HP Recommended configuration parameters

Attribute	Description	Supported values
path_grouping_policy	Used for applying the policy to the multipath device hosted by this storage controller	multibus: All valid paths in one priority group group_by_prio: One priority group per path priority value
path_checker	Used for determining the state of the path	tur
path_selector	Used to select the path selector algorithm to be used for mpath. These algorithms are offered by the kernel mpath target	round-robin 0
failback	Used to manage the time during path group failback	immediate
prio_callout	Executable to obtain a path weight for a block device. Weights are summed for each path group to determine the next path group to be used in case of path failure	/bin/true /sbin/mpath_prio_alua %n
rr_weight	Used to assign weights to the path	uniform
no_path_retry	(n =12) indicates the number of retries until queuing is disabled (queues till n number of polling), *fail* indicates immediate failure (no queuing), or *queue* indicates never stop queuing (queue forever till the path comes alive)	n=12

Example 1. Recommended Values for HP Arrays

The product string of the storage controller contains 16 characters (including spaces) and the vendor string contains 8 characters (including spaces) in the /etc/multipath.conf file. To ease the task of manually editing the devices section of the multipath.conf file, the wild card, regular expression support for the product and vendor string is provided. For example, in /etc/multipath.conf file, the product string for the HSV101 arrays is given the value of HSV101*, instead of HSV101 \(C\)COMPAQ.

For EVA3000-(HSV101)

```
device
{
            vendor                  "HP"
            product                 "HSV101 \(C\)COMPAQ"
            path_grouping_policy    group_by_prio
            getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
            path_checker            tur
            path_selector           "round-robin 0"
            prio_callout            "/sbin/mpath_prio_alua %n"
            rr_weight               uniform
            rr_min_io               100
            failback                immediate
            no_path_retry           12
}
```

For EVA4000/6000 - (HSV200)

```
device
{
            vendor                  "HP"
            product                 "HSV200"
            path_grouping_policy    group_by_prio
            getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
            path_checker            tur
            path_selector           "round-robin 0"
            prio_callout            "/sbin/mpath_prio_alua %n"
            rr_weight               uniform
            rr_min_io               100
            failback                immediate
            no_path_retry           12
}
```

For EVA5000-(HSV111)

```
device
{
            vendor                  "COMPAQ"
            product                 "HSV111 \(C\)COMPAQ"
            path_grouping_policy    group_by_prio
            getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
            path_checker            tur
            path_selector           "round-robin 0"
            prio_callout            "/sbin/mpath_prio_alua /dev/%n"
            rr_weight               uniform
            rr_min_io               100
            failback                immediate
            no_path_retry           12
}
```

For EVA8000 (HSV210)

```
device
{
           vendor                  "HP”
           product                 "HSV210"
           path_grouping_policy    group_by_prio
           getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
           path_checker            tur
           path_selector           "round-robin 0"
           prio_callout            "/sbin/mpath_prio_alua /dev/%n"
           rr_weight               uniform
           rr_min_io               100
           failback                immediate
           no_path_retry           12
}
```

For EVA4400 (HSV300)

```
device
{
            vendor                  "HP"
            product                 "HSV300"
            path_grouping_policy    group_by_prio
            getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
            path_checker            tur
            path_selector           "round-robin 0"
            prio_callout            "/sbin/mpath_prio_alua /dev/%n"
            rr_weight               uniform
            rr_min_io               100
            failback                immediate
            no_path_retry           12
}
```

For MSA A/A arrays

```
device
{
            vendor                  "HP"
            product                 "MSA VOLUME*"
            path_grouping_policy    group_by_prio
            getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
            path_checker            tur
            path_selector           "round-robin 0"
            prio_callout            "/sbin/mpath_prio_alua /dev/%n"
            rr_weight               uniform
            failback                immediate
            hardware_handler        "0"
            no_path_retry           12
            rr_min_io               100

}
```

For MSA2000 Storage Product Family arrays

```
device
{
            vendor                  "HP"
            product                 "MSA2[02]*"
            path_grouping_policy    multibus
            getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
            path_selector           "round-robin 0"
            rr_weight               uniform
            prio_callout            "/bin/true"
            path_checker            tur
            hardware_handler        "0"
            failback                immediate
            no_path_retry           12
            rr_min_io               100

}
```

For XP arrays

```
device
{
            vendor                  "HP"
            product                 "OPEN-*"
            path_grouping_policy    multibus
            getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
            rr_weight               uniform
            rr_min_io               100
            path_checker            tur
            failback                immediate
            no_path_retry           12
}
```

NOTE:

For MSA 2000 Storage Product Family arrays (MSA2012fc/MSA2212fc/MSA2012i, add the device section in `/etc/multipath.conf` file. For HP recommended device parameters, see For MSA2000 Storage Product Family arrays.

NOTE:

In XP arrays, there are different LUNs, such as `OPEN-<x>`, `3390-3A`, `3390-3B`, `OP-C:3390-3C`, `3380KA`, `3380-KB`, and `OP-C:3380-KC` where `x` = {`3,8,9,K,T,E,V`}.

The `product` strings for XP LUNs are based on these emulation types. A new device section must be added for each emulation type, because each `product` string requires a new device subsection.

`OPEN-*` is sufficient for the `product` string for all the XP LUNs with different OPEN emulations, because regular expressions are supported in the `/etc/multipath.conf` file.

IMPORTANT:

After editing the `/etc/multipath.conf` file for HP StorageWorks Arrays, you must restart `multipathd` daemon and run multipath command. For more information, see Starting HPDM Multipath.

Starting HPDM Multipath

Whenever an attribute for the LUNs of any of the HP supported storage arrays changes, the `multipathd` daemon must be restarted and the configuration must be reloaded by executing the following commands:

```
# /etc/init.d/multipathd restart
# /sbin/multipath
```

These steps are required only if the attributes are modified.

To view the status of the multipath devices, complete the following steps:

1. Run `multipath` utility.
   Use the `# /sbin/multipath -ll` option.
   For information on status maps in `multipath -ll` for HPDM Multipath devices, see Using HPDM Multipath.

4 Using HPDM Multipath

This chapter addresses the following topics:

Creating an Alias using /etc/multipath.conf file

This section describes various parameters available for updating /etc/multipath.conf file.

The multipath devices are created in the /dev/mapper directory in the hosts. These devices are similar to any other block devices present in the host, and are used for any block or file level I/O operations, such as creating the file system.

You must use the devices under /dev/mapper/ . You can create a new device alias by using the alias and the WWID attributes of the multipath device present in the multipath subsection of the/etc/mutipath.conf file.

Example 2. Updating /etc/multipath.conf file

For example, when the following subsection is added in the /etc/multipath.conf file for the LUN with WWID 3600508b30090f5d0d2a9d64590490022, a multipath device, mydatadisk1, is created under /dev/mapper:

```
multipaths
{
multipath
            wwid                    3600508b30090f5d0d2a9d64590490022
            alias                   mydatadisk1
{           path_grouping_policy    multibus
            path_checker            tur
            path_selector           "round-robin 0"
}
}
```

You can use /dev/mapper/mydatadisk1 like any other block device.

Using Multipath Map

This section describes a multipath map obtained by executing the multipath–ll command, based on a sample multipath.conf file.

Following is a sample multipath.conf file (for RHEL4) with parameters to support EVA 4000 as a storage controller:

```
defaults
{
```

```
            udev_dir                /dev
            polling_interval        5
            selector                "round-robin 0"
            path_grouping_policy    failover
            getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
            prio_callout            "/bin/true"
            path_checker            tur
            rr_min_io               1000
            rr_weight               uniform
            failback                immediate
            no_path_retry           12
            user_friendly_names     no

}

multipaths

{

multipath

            wwid                    3600508b30090f5d0d2a9d64590490022
            path_grouping_policy    multibus
{           path_checker            tur
            path_selector           "round-robin 0"

}

multipath

{           …….

            ….....                  for other target

}

multipath

{           …….

            ….....                  for other target

}

}

devices

{

device

{

            vendor                  "HP”
            product                 "HSV200"
            path_grouping_policy    group_by_prio
            getuid_callout          "/sbin/scsi_id -g -u -s /block/%n"
            path_checker            tur
            path_selector           "round-robin 0"
            prio_callout            "/sbin/mpath_prio_alua %n"
            rr_weight               uniform
            rr_min_io               100
            failback                immediate
            no_path_retry           60

}

device

{           …….

            ….....                  for targets from other storage controllers

}

}
```

For example, in a typical configuration, a Linux host with a dual port HBA is connected to an EVA 4000 through two switches. In this case, if all the paths are available, the host has four I/O paths for any LUN presented from the EVA 4000. The `multipath.conf` file enables the host to view the multipath map, as follows:

```
3600508b30090f5d0cf46bd1a4c420023
               [size=1 GB][features="1 queue_if_no_path"][hwhandler="0"]
                       \_ round-robin  0      [prio=100][active]
                       \_ 3:0:1:3 sdad 65:208 [active][ready]
                       \_ 2:0:2:3 sdo  8:224  [active][ready]
                       \_ round-robin  0      [prio=20][enabled]
                       \_ 2:0:1:3 sdj  8:144  [active][ready]
                       \_ 3:0:0:3 sdy  65:128 [active][ready]

3600508b30090f5d0d2a9d64590490022
               [size=3 GB][features="1 queue_if_no_path"][hwhandler="0"]
                       \_ round-robin  0      [prio=120][active]
                       \_ 3:0:2:4 sdaj 66:48  [active][ready]
                       \_ 3:0:3:4 sdao 66:128 [active][ready]
                       \_ 2:0:0:4 sdf  8:80   [active][ready]
                       \_ 2:0:3:4 sdu  65:64  [active][ready]
```

Path Grouping

The information in the map is presented by grouping the paths for a LUN with unique identifiers, such as UID/WWN.The `size`, `features`, and the corresponding `hwhandlers` are displayed followed by the unique LUN identifier.

The grouping of paths is done and presented in a map based on the I/O load balancing policy. In the above example, for the `LUN 3600508b30090f5d0d2a9d64590490022`, the grouping is done with the policy as multibus and the devices, such as `sdaj`, `sdao`, `sdf`, and `sdu` belong to same path group.

The state of the group is either active or enabled depending on whether I/O is active or not.

For the LUN `3600508b30090f5d0cf46bd1a4c420023`, the grouping is done with the policy as `group_by_prio`.

The devices `sdad` and `sdo` belong to one path group, and the devices `sdj` and `sdy` belong to a different path group, because they are from two controllers with different serial numbers.

The I/O always happens in the path group which is active. If all paths in the active group fails, the failover occurs to the other path group that is enabled on changing that path group to active. When the paths are up again and the failback parameter is set as immediate, the failback occurs to the earlier group and I/O occurs through the earlier group.

Example 3. Failover and Failback

If group A has sdad and sdo, and group B has sdj and sdy, the failover occurs from group A to group B and the failback occurs from group B to group A .

The state of the path is given as [ready] [active] if the path is up, and ready for I/O. If the path is down, this state is shown as [faulty] [failed]. The path states are updated periodically based on the polling interval set in the /etc/multipath.conf file. The path states are updated by the multipathd daemon.

Table 4 lists the basic operations supported by the multipath CLI utility provided with HPDM Multipath.

Table 4 Basic operations of HPDM Multipath

Command	Description
# multipath -F	Deletes all HPDM Multipath devices.
# multipath -d	Displays potential paths, but does not create any device.
# multipath	Creates HPDM Multipath devices.
# multipath -l # multipath -ll	Displays the list of device status. Displays the detailed list of device status.
# multipath -v2 <device>	Configures and displays multipath map information.
# multipath -v3	Configures and displays multipath map information.

Clustering Solutions with HPDM Multipath

HPDM Multipath supports the following clustering solutions:

- HP Serviceguard for Linux version 11.18 or later
- SteelEye Lifekeeper for Linux version 6.1.4 or later

Configuring HP Service Guard with HPDM Multipath Devices

HPDM Multipath devices can be used for data storage under Serviceguard protection. To configure HPDM Multipath devices under Serviceguard, complete the following steps:

1. Configure the SAN such that each host has multiple paths to the array. Present the LUN's to all the hosts.
2. See *Configuring HPDM Multipath* for configuring the LUN's for HPDM Multipath
3. Use the HPDM Multipath device to create the logical volumes that are used by Serviceguard for Linux.

Please refer the HP ServiceGuard Compatibility matrix for the supported Operating Systems and the Arrays.

NOTE:

HPDM Multipath devices can be used with HP Serviceguard A.11.18 for Linux . The LVM volume groups must be configured on the HPDM Multipath devices. In HP Service Guard package configuration, the disk monitor must be configured using the HPDM Multipath device in the format /dev/dm-x. Please ensure that the device names (/dev/dm-x) are same across all the nodes. Configuration requirements HP Serviceguard for Linux are available in the certification matrix available at the following website:

http://www.hp.com/info/sglx

Configuring SteelEye LifeKeeper with HPDM Multipath Devices

HPDM Multipath devices in the clustered environment can be used for file systems having Lifekeeper as the clustering solution. To configure HPDM Multipath devices, complete the following steps:

1. Configure the SAN such that each host has multiple paths to the array. Present the LUN's to all the hosts.
2. Any file systems on the LUN's must be mounted using the HPDM Multipath device.
3. See *Configuring HPDM Multipath* for configuring the LUN's for HPDM Multipath
4. Install and start the LifeKeeper for Linux with HPDM Multipath Recovery kit installed on all the hosts.
5. Create a Resource Hierarchy on LifeKeeper using the HPDM Multipath device.

Please refer to SteelEye LifeKeeper documentation for the supported Operating Systems and the Arrays.

http://licensing.steeleye.com/support/docm.php

5 Removing HPDM Multipath Tools

This chapter addresses the following topic:

- Removing HPDM Multipath Tools

Removing HPDM Multipath Tools

You can remove HPDM Multipath tools by using the `INSTALL.sh` shell script.

To remove HPDM Multipath tools from your system, complete the following steps:

1. Log in to the system as a `root` user.
2. From the directory containing the script `INSTALL.sh`, execute the `# ./INSTALL.sh` command.
3. Follow the on-screen instructions.
4. Alternatively, you can also use `rpm -e HPDMmultipath-tools-4.0.0-<OS>` to remove Multipath tools.

6 Troubleshooting

Following are the troubleshooting steps for issues that you may encounter while using HPDM Multipath:

- Enter the following commands whenever a new LUN is added or deleted at the Linux host, to update the multipath maps in the kernel:
  1. `/etc/init.d/multipathd restart`
  2. `multipath -v <X>`
- Enter the following commands whenever multipath devices are deleted using `multipath -F`, to create the multipath devices and to create the same for the partitions on the multipath devices:
  1. `multipath -v <X>`
  2. `/sbin/dmsetup ls --target multipath --exec "/sbin/kpartx -a -p -part"` (SLES10/ SLES10 SP1)
     `/sbin/dmsetup ls --target multipath --exec "/sbin/kpartx -a"` (RHEL 4U5/RHEL 4U6/RHEL 5/RHEL 5U1)
- If **SELinux** is enabled, device maps cannot be created and **SELinux** denies access requested by `/sbin/multipathd`.

NOTE:

In some cases, labelling problems can cause **SELinux** denials. Try to restore the default system file context for `bin` by running the following commends:

```
restorecon -v bin
```

If this does not work, there is currently no automatic way to allow this access. Instead, you can generate a local policy module to allow this access - For more information, see the following website:http://fedora.redhat.com/docs/selinux-faq-fc5/#id2961385

You can also disable **SELinux** protection altogether. However, disabling **SELinux** protection is not recommended.

- If an existing LUN is deleted and a new LUN is presented in the same SCSI slot, LUN collision may occur. This results in the creation of a new LUN through old device special files. This may lead to data corruption. To recover from this error state, run following commands:
  1. multipath -F
  2. hp_rescan -a or echo "- - -" > /sys/class/scsi_host/<host instance>/scan
  3. /etc/init.d/multipathd restart
  4. /sbin/multipath -v3
- Path status may not be reflected immediately when it is recoved from FAILED to ACTIVE state. To reflect the correct path state run the following commands.
  1. /etc/init.d/multipathd restart
  2. /sbin/multipath -v3

- On hosts running SLES OS, if you do not enable boot.multipath to start at boot time, multipath maps are not created. To ensure the creation of device maps across reboots, enable boot.multipath using the following command:
  `# chkconfig boot.multipath on`

7 Abbreviations

Table 5 lists the abbreviations or terms used in this document.

Table 5 Abbreviations

Abbreviations/Acronyms	Definition
CLI	Command-Line Interpreter/Interface
HPDM	HP Device Mapper
GUI	Graphical User Interface
HBA	Host Bus Adapter
I/O	Input/Output
LUN	Logical Unit Number
LVM	Logical Volume Manager
OS	Operating System
RAID	Redundant Array of Independent (or Inexpensive) Disks
RHEL	Red Hat Enterprise Linux
SLES	SuSE LINUX Enterprise Server
UID	Unique Identifier
WWID	World-Wide Identifier
WWN	World-Wide Name
FC	Fibre Channel
SCSI	Small Computer System Interface
SAN	Storage Area Network