

HP StorageWorks

Device Mapper Multipath Enablement Kit for HP StorageWorks Disk Arrays installation and reference guide



Legal and notice information

© Copyright 2009 Hewlett-Packard Development Company, L.P.

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Red Hat is the registered trademark of Red Hat, Inc.

SuSE is the registered trademark of Novell, Inc.

Contents

- About this guide 5
 - Intended audience 5
 - Related documentation 5
 - Document conventions and symbols 5
 - HP technical support 6
 - Subscription service 6
 - HP websites 6
 - Documentation feedback 7
 - Product feedback 7
- 1 Introduction 9
 - Overview 9
 - Features 9
 - Software components 10
 - DM Multipath operation 10
- 2 Installing HPDM Multipath 11
- 3 Configuring DM Multipath 13
 - Using /etc/multipath.conf File 13
 - Setting up Device Mapper Multipath daemon 14
- 4 Using DM Multipath 17
 - Creating an alias using /etc/multipath.conf file 17
 - Using Multipath map 17
 - Path grouping 19
 - Clustering Solutions with DM Multipath 20
 - Configuring HP Service Guard with DM Multipath devices 20
 - Configuring SteelEye LifeKeeper with DM Multipath devices 21
- 5 Removing HPDM Multipath tools 23
 - Removing HPDM Multipath tools 23
- 6 Troubleshooting 25
- 7 Abbreviations 27

Tables

1 Document conventions	5
2 HP Recommended configuration parameters	13
3 Basic operations of DM Multipath	20
4 Abbreviations	27

About this guide

This guide describes the Device Mapper Multipath enablement for HP StorageWorks Disk Arrays (DM Multipath) and provides information to help you:

- [Install DM Multipath](#)
- [Configure DM Multipath](#)
- [Use DM Multipath](#)
- [Troubleshoot DM Multipath](#)

Intended audience

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

Related documentation

The *Device Mapper Multipath Enablement Kit for HP StorageWorks Disk Arrays release notes* provides related information:

The document is available with the Device Mapper installation kit, and also in the Manual page of the HP Business Support Center website:

<http://www.hp.com/support/manuals>

In the Storage section, click **Storage Software** and then select **Multi-path Device Mapper for Linux Software** under **Storage Infrastructure Software** list.

Document conventions and symbols

Table 1 Document conventions

Convention	Element
Blue text: Table 1	Cross-reference links and e-mail addresses
Blue, underlined text: http://www.hp.com	Website addresses
Bold text	<ul style="list-style-type: none">• Keys that are pressed• Text typed into a GUI element, such as a box• GUI elements that are clicked or selected, such as menu and list items, buttons, tabs, and check boxes
<i>Italic</i> text	Text emphasis

Convention	Element
Monospace text	<ul style="list-style-type: none"> • File and directory names • System output • Code • Commands, their arguments, and argument values
<i>Monospace, italic</i> text	<ul style="list-style-type: none"> • Code variables • Command variables
Monospace, bold text	Emphasized monospace text

❗ **IMPORTANT:**

Provides clarifying information or specific instructions.

📝 **NOTE:**

Provides additional information.

HP technical support

For worldwide technical support information, see the HP support website:

<http://www.hp.com/support>

Before contacting HP, collect the following information:

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

Subscription service

HP recommends that you register your product at the Subscriber's Choice for Business website:

<http://www.hp.com/go/e-updates>

After registering, you will receive e-mail notification of product enhancements, new driver versions, firmware updates, and other product resources.

HP websites

For additional information, see the following HP websites:

- <http://www.hp.com>

- <http://www.hp.com/go/devicemapper>
- <http://www.hp.com/support/manuals>
- <http://www.hp.com/support/downloads>

Documentation feedback

HP welcomes your feedback.

To make comments and suggestions about product documentation, please send a message to storagedocsFeedback@hp.com. All submissions become the property of HP.

Product feedback

To make comments and suggestions about Device Mapper Multipath Enablement Kit for HP StorageWorks Disk Arrays, please send a message to multipathfeedback@hp.com.

1 Introduction

This chapter addresses the following topics:

- [Overview](#)
- [Features](#)
- [Software components](#)
- [DM Multipath operation](#)

Overview

HPDM Multipath is an enablement kit for HP StorageWorks Disk Arrays based on Device Mapper Multipath, included in the Linux operating system distributions. 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

Device Mapper Multipath offers the following features:

- **I/O failover and failback:** Provides transparent failover and failback of I/Os by rerouting I/Os automatically to an alternative path when a path failure is sensed, and routing them back when the path is restored.
- **Path grouping policies:** Paths are coalesced based on the following path-grouping policies:
 - *Priority based path-grouping*
 - Provides priority to group paths based on Asymmetric Logical Unit Access (ALUA) state
 - Provides static load balancing policy by assigning higher priority to the preferred path
 - *Multibus* — All paths are grouped under a single path group
 - *Group by serial* — Paths are grouped together based on controller serial number
 - *Failover only* — Provides failover without load balancing by grouping the paths into individual path groups
- **I/O load balancing policies:** Provides the following load balancing policies within a path group:
 - *Weighted round robin* — This round-robin algorithm routes `rr_min_io` number of I/Os on a selected path before switching to the next path.
 - *Least pending I/O path* — This determines the number of non-serviced requests pending on a path and selects the path which has the least number of pending requests for service.
 - *DM service time* — This is a service time oriented dynamic load balancer, which selects a path to complete the incoming I/O with the shortest time.
- **Device name persistence:** Device names are persistent across reboots and Storage Area Network (SAN) reconfigurations. Device Mapper also provides configurable device name aliasing feature for easier management.

- **Persistent device settings:** All the device settings such as load balancing policies, path grouping policies are persistent across reboots and SAN reconfigurations.
- **Device exclusion:** Provides device exclusion feature through blacklisting of devices.
- **Path monitoring:** Periodically monitors each path for status and enables faster failover and failback.
- **Online device addition and deletion:** Devices can be added to or deleted from Device Mapper (DM) Multipath without rebooting the server or disrupting other devices or applications.
- **Management Utility:** Provides Command Line Interface (CLI) to manage Multipath devices.
- **Boot from SAN:** Provides multipathing for operating system installation partitions on SAN devices.
- **Cluster support:** Provides multipathing in HP Serviceguard and SteelEye LifeKeeper clustering environment.
- **Volume Manager support:** Provides support for multipathing devices to be configured under Logical Volume Manager.

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).

DM Multipath operation

DM 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, DM 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. DM 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`. DM Multipath creates a single block device, `/dev/mapper/mpath1`, that reroutes I/O through these four underlying block devices.

2 Installing HPDM Multipath

To install 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/go/devicemapper/>
2. Log in as root to the host system.
3. Copy the installation tar package to a temporary directory (for example, /tmp/HPDMmultipath).
4. Unbundle the package by executing the following commands:

```
# cd /tmp/HPDMmultipath
# tar -xvzf HPDMmultipath-<version>.tar.gz
# cd HPDMmultipath-<version>
```
5. Verify that the directory contains `INSTALL`, `README.txt`, `COPYING`, `bin`, `conf`, `SRPMS`, `RPMS`, and `docs` directories.
6. To install DM Multipath tools software on the server, execute the following command:

```
# ./INSTALL
```
7. Follow the on-screen instructions to complete the installation.
8. Restart Multipath services by executing the following command:

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


3 Configuring DM Multipath

This chapter addresses the following topics:

- [Using /etc/multipath.conf File](#)
- [Setting up Device Mapper Multipath daemon](#)

Using /etc/multipath.conf File

DM Multipath kit includes a template file with the recommended configuration for HP supported arrays. To configure DM Multipath, edit the /etc/multipath.conf configuration file. This file also enables you to customize DM Multipath for a variety of SAN storage subsystems.

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 2](#) lists the important attributes of HP supported arrays.

Table 2 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 weightedpath <hctl devname> <regex1> <prio1> <regex2> <prio2>: Provides static load balancing policy by assigning higher priority to the preferred path failover: One path per priority group
path_checker	Used for determining the state of the path	tur

Attribute	Description	Supported Values
<code>path_selector</code>	Used to select the path selector algorithm to be used for <code>mpath</code> . These algorithms are offered by the kernel <code>mpath</code> target	"round-robin 0" "least-pending 0" "service-time 0"
<code>failback</code>	Used to manage the time during path group failback	immediate $n > 0$ (user assigned values) manual
<code>prio_callout</code>	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	<code>/bin/true</code> <code>/sbin/mpath_prio_alua /dev/%n</code>
<code>prio</code>	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	const alua
<code>rr_weight</code>	Used to assign weights to the path	uniform priorities
<code>rr_min_io</code>	The number of IOs to route to a path before switching to the next in the same path group	100 1000
<code>no_path_retry</code>	($n=18$) indicates the number of retries until queuing is disabled (queues till n number of polling), <i>fail</i> indicates immediate failure (no queuing), or <i>queue</i> indicates never stop queuing (queue forever till the path comes alive)	18 fail queue

Setting up Device Mapper Multipath daemon

You can set the Device Mapper Multipath daemon to start at boot time.

For RHEL hosts, complete the following steps to start the `multipathd` daemon at boot time:

1. Run the following command to check if the daemon is configured to start at boot time:
`chkconfig --list multipathd`
2. Run the following commands to start the Device Mapper Multipath daemon:
`chkconfig [--level levels] multipathd on`
`chkconfig multipathd`

For SLES hosts, complete the following steps to start the `multipathd` daemon at boot time:

1. Run the following commands to check if the daemon is configured to start at boot time:
`chkconfig --list boot.device-mapper`
`chkconfig --list boot.multipath`

```
# chkconfig --list multipathd
```

2. Run the following commands to start the Device Mapper Multipath daemons:

```
# chkconfig boot.device-mapper [levels]
```

```
# chkconfig boot.multipath [levels]
```

```
# chkconfig multipathd [levels]
```

4 Using DM Multipath

This chapter addresses the following topics:

- [Creating an alias using /etc/multipath.conf file](#)
- [Clustering Solutions with DM Multipath](#)

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/multipath.conf file.

Example 1. 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                123456789987654321
        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 RHEL5) 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               100
    rr_weight               uniform
    failback               immediate
    no_path_retry           18
    user_friendly_names     yes
}

multipaths
{
    multipath
    {
        wwid                123456789987654321
        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                "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
        failback                immediate
        hardware_handler        "0"
        no_path_retry           18
        rr_min_io               100
    }

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 8000 through two switches. In this case, if all the paths are available, the host has eight I/O paths for any LUN presented from the EVA 8000. The `multipath.conf` file enables the host to view the multipath map, as follows:

```

mpath85 (3600508b4001072a100039000277a0000) dm-2 HP,HSV210
    [size=1.0GB] [features=1 queue_if_no_path] [hwhandler=0] [rw]
        \_ round-robin 0 [prio=240] [active]
        \_ 57:0:0:19 sdb 8:16 [active] [ready]
        \_ 56:0:0:19 sda 8:0 [active] [ready]
        \_ 64:0:0:19 sdf 8:80 [active] [ready]
        \_ 65:0:0:19 sdi 8:128 [active] [ready]
        \_ 66:0:0:19 sdg 8:96 [active] [ready]
        \_ 67:0:0:19 sdh 8:112 [active] [ready]
        \_ 80:0:0:19 sdac 65:192 [active] [ready]
        \_ 81:0:0:19 sdae 65:224 [active] [ready]
mpath86 (3600508b4001072a10003900027800000) dm-3 HP,HSV210
    [size=1.0G] [features=1 queue_if_no_path] [hwhandler=0] [rw]
        \_ round-robin 0 [prio=200] [active]
        \_ 57:0:0:20 sdc 8:32 [active] [ready]
        \_ 64:0:0:20 sdk 8:160 [active] [ready]
        \_ 56:0:0:20 sdj 8:144 [active] [ready]
        \_ 65:0:0:20 sdt 65:48 [active] [ready]
        \_ round-robin 0 [prio=40] [enabled]
        \_ 66:0:0:20 sdr 65:16 [active] [ready]
        \_ 67:0:0:20 sdu 65:64 [active] [ready]
        \_ 80:0:0:20 sdv 66:0 [active] [ready]
        \_ 81:0:0:20 sdw 66:16 [active] [ready]

```

Path grouping

The information in the multipath 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 following the unique LUN identifier.

Paths are grouped based on the path grouping policy. In the previous example, for LUN `mpath85`, the paths are grouped with the policy as `multibus` and the devices `sdb`, `sda`, `sdf`, `sdi`, `sdg`, `sdh`, `sdac`, and `sdae` belong to the same path group.

For the LUN `mpath86`, the grouping is done with the policy as `group_by_prio`.

The devices `sdc`, `sdk`, `sdj`, and `sdt` belong to one path group, and the devices `sdr`, `sdu`, `sdv`, and `sdw` belong to a different path group. This path grouping is based on *one path group per path* priority value. Path priority value is determined by ALUA state of the path.

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 2. Failover and Failback

If all the paths in a path group where I/Os are active fails, then I/Os are failed over to other path groups. I/Os are failed back as soon as the failed paths become alive.

The state of the path is given as `[active]` `[ready]` if the path is up, and ready for I/O. If the path is down, this state is shown as `[failed]` `[faulty]`. 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 3 lists the basic operations supported by the `multipath` CLI utility provided with DM Multipath.

Table 3 Basic operations of DM Multipath

Command	Description
# <code>multipath -F</code>	Deletes all DM Multipath devices.
# <code>multipath -d</code>	Displays potential paths, but does not create any device.
# <code>multipath</code>	Creates DM Multipath devices.
# <code>multipath -l</code> # <code>multipath -ll</code>	Displays the list of device status. Displays the detailed list of device status.
# <code>multipath -v0</code>	Configures multipath map information

Clustering Solutions with DM Multipath

DM 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 DM Multipath devices

DM Multipath devices can be used for data storage under Serviceguard protection. To configure DM 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 [Configure DM Multipath](#) for configuring the LUN's for DM Multipath

3. Use the DM 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:

DM Multipath devices can be used with HP Serviceguard A.11.18 for Linux . The LVM volume groups must be configured on the DM Multipath devices. In HP Service Guard package configuration, the disk monitor must be configured using the DM 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 DM Multipath devices

DM Multipath devices in the clustered environment can be used for file systems having Lifekeeper as the clustering solution. To configure DM 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 DM Multipath device.
3. See [Configure DM Multipath](#) for configuring the LUN's for DM Multipath
4. Install and start the LifeKeeper for Linux with DM Multipath Recovery kit installed on all the hosts.
5. Create a Resource Hierarchy on LifeKeeper using the DM 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` 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`, execute the `# ./INSTALL` command.
3. Follow the on-screen instructions.

6 Troubleshooting

Following are the troubleshooting steps for issues that you may encounter while using DM 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>`
- 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 commands:

```
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 the 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 recovered 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 4 lists the abbreviations or terms used in this document.

Table 4 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
iSCSI	Internet SCSI
SAS	Serial Attached SCSI

