When Good Disks Go Bad:
Dealing with Disk Failures under LVM

Abstract

This white paper discusses how to deal with disk failures under the HP-UX Logical Volume Manager (LVM). It is intended for system administrators or operators who have experience with LVM. It includes strategies for preparing for disk failure, means for recognizing that a disk has failed, and steps for removing or replacing a failed disk.

Background

Whether managing a workstation or server, your goals include minimizing system downtime and maximizing data availability. Hardware problems such as disk failures can disrupt those goals. Replacing disks can be a daunting task, given the variety of hardware features such as hot-swappable disks, and software features such as mirroring or online disk replacement you can encounter.

LVM provides features to maximize data availability and improve system uptime. This paper explains how you can use LVM to minimize the impact of disk failures to your system and your data. It addresses the following topics:

- Preparing for Disk Recovery: what you can do *before* a disk goes bad. This includes guidelines on logical volume and volume group organization, software features to install, and other best practices.
- Recognizing a Failing Disk: how you can tell that a disk is having problems. This covers some of the error messages related to disk failure you might encounter in the system's error log, in your electronic mail, or from LVM commands.
- Confirming Disk Failure: what you should check to make sure the disk is failing. This includes a simple three-step approach to validating a disk failure if you do not have online diagnostics.
- Choosing a Course of Action: what you must know before you either remove or replace the disk. This includes whether the disk is hot-swappable, what logical volumes are located on the disk, and what recovery options are available for the data.
- Removing the Disk: how to remove the disk permanently from your LVM configuration, rather than replace it.
- Replacing the Disk: how to replace a failing disk, while minimizing system downtime and data loss. This section gives a high-level overview of the process and the specifics of each step. The exact procedure varies, depending on your LVM configuration and what hardware and software features you have installed, so several disk replacement scenarios are included. The section concludes with a flowchart of the disk replacement process.

In addition, there are two appendices. The first appendix contains step-by-step procedures for creating a mirror of your root disk. The second appendix lists some common LVM error messages, what triggers them, and what you should do to recover from them.

You do not have to wait for a disk failure to begin preparing for failure recovery. This paper can help you be ready when a failure does occur.

1. Preparing for Disk Recovery

Forewarned is forearmed. Knowing that hard disks will fail eventually, you can take some precautionary measures to minimize your downtime, maximize your data availability, and simplify the recovery process. Consider the following guidelines before you experience a disk failure.

Defining a Recovery Strategy

As you create logical volumes, choose one of the following recovery strategies. Each choice strikes a balance between cost, data availability, and speed of data recovery.

1. **Mirroring**: If you mirror a logical volume on a separate disk, the mirror copy is online and available while recovering from a disk failure. With hot-swappable disks, users will have no indication that a disk was lost.
2. **Restoring from backup**: If you choose not to mirror, make sure you have a consistent backup plan for any important logical volumes. The tradeoff is that you will need fewer disks, but you will lose time while you restore data from backup media, and you will lose any data changed since your last backup.
3. **Initializing from scratch**: If you do not mirror or back up a logical volume, be aware that you will lose data if the underlying hard disk fails. This can be acceptable in some cases, such as a temporary or scratch volume.

Using Hot-Swappable Disks

The hot-swap feature implies the ability to remove or add an inactive hard disk drive module to a system while power is still on and the SCSI bus is still active. In other words, you can replace or remove a hot-swappable disk from a system without turning off the power to the entire system.

Consult your system hardware manuals for information about which disks in your system are hot-swappable. Specifications for other hard disks are available in their installation manuals at http://docs.hp.com.

Installing the Patches that Enable LVM Online Disk Replacement

LVM online disk replacement (LVM OLR) simplifies the replacement of disks under LVM. With LVM OLR, you can temporarily disable LVM use of a disk in an active volume group. Without it, you cannot keep LVM from accessing a disk unless you deactivate the volume group or remove the logical volumes on the disk.

LVM OLR is delivered in two patches, one patch for the kernel and one patch for the `pvchange` command. The patches introduce a new option, `-a`, to `pvchange`. The `-a` option disables or re-enables a specified path to an LVM disk. For more information on LVM OLR, refer to the white paper *LVM Online Disk Replacement (LVM OLR)*.

Both command and kernel components are required to enable LVM OLR:

*   For HP-UX 11i version 1, install patches PHKL_31216 and PHCO_30698 or their superseding patches.
*   For HP-UX 11i version 2, install patches PHKL_32095 and PHCO_31709 or their superseding patches.

Mirroring Critical Information, Especially the Root Volume Group

By using mirror copies of the root, boot, and primary swap logical volumes on another disk, you can use the copies to keep your system in operation if any of these logical volumes fail.

Mirroring requires the add-on product HP MirrorDisk/UX (B5403BA). This is an extra-cost product available on the HP-UX 11i application release media. Use the `swlist` command to confirm that you have HP MirrorDisk/UX installed. For example:

```
# swlist -l fileset | grep -i mirror
LVM.LVM-MIRROR-RUN B.11.23 LVM Mirror
```

The process of mirroring is usually straightforward, and can be easily accomplished using the system administration manager SAM, or with a single `lvextend` command. These processes are documented in [Managing Systems and Workgroups](#). The only mirroring setup task that takes several steps is mirroring the root disk; for the recommended procedure for adding a root disk mirror, refer to *[Appendix A: Procedures](#)*.

Three corollaries to the mirroring recommendation are:

1. Use the strict allocation policy for all mirrored logical volumes. Strict allocation forces mirrors to occupy different disks. Without strict allocation, you can have multiple mirror copies on the same disk; if that disk fails, you will lose all your copies. The allocation policy is controlled with the `-s` option of the `lvcreate` and `lvchange` commands. By default, strict allocation is enabled.

2. To improve the availability of your system, keep mirror copies of logical volumes on separate I/O busses if possible. With multiple mirror copies on the same bus, the bus controller becomes a single point of failure—if the controller fails, you lose access to all the disks on that bus, and thus access to your data. If you create physical volume groups and set the allocation policy to *PVG-strict*, LVM helps you avoid inadvertently creating multiple mirror copies on a single bus. For more information about physical volume groups, refer to *lvmpvg(4)*.

3. Consider using one or more free disks within each volume group as spares. If you configure a disk as a spare, then a disk failure causes LVM to reconfigure the volume group so that the spare disk takes the place of the failed one. That is, all the logical volumes that were mirrored on the failed disk are automatically mirrored and resynchronized on the spare, while the logical volume remains available to users. You can then schedule the replacement of the failed disk at a time of minimal inconvenience to you and your users. Sparing is particularly useful for maintaining data redundancy when your disks are not hot-swappable, since the replacement process may have to wait until your next scheduled maintenance interval. Disk sparing is discussed in [Managing Systems and Workgroups](#).

Creating Recovery Media

Ignite/UX enables you to create a consistent, reliable recovery mechanism in the event of a catastrophic failure of a system disk or root volume group. You can back up essential system data to a tape device, CD, DVD, or a network repository, and recover the system configuration quickly. While Ignite/UX is not intended to be used to back up all system data, you can use it with other data recovery applications to create a means of total system recovery.

Ignite/UX is a free add-on product, available from www.hp.com/go/softwaredepot. Documentation is available at the Ignite/UX [website](#).

Other Guidelines for Optimal System Recovery

Here are some other recommendations, summarized from [Managing Systems and Workgroups](#), that simplify recoveries after catastrophic system failures:

*   Keep the number of disks in the root volume group to a minimum; no more than three, even if the root volume group is mirrored. The benefits of a small root volume group are threefold. First, fewer disks in the root volume group means less opportunities for disk failure in that group. Second, more disks in any volume group leads to a more complex LVM configuration, which will be more difficult to recreate after a catastrophic failure. Finally, a small root volume group is quickly recovered. In some cases, you can

reinstall a minimal system, restore a backup, and be back online within three hours of diagnosis and replacement of hardware.

Three disks in the root volume group are better than two, because of quorum restrictions. With a two-disk root volume group, a loss of one disk can require you to override quorum to activate the volume group; if you must reboot to replace the disk, you must interrupt the boot process and use the `-lq` boot option. If you have three disks in the volume group, and they are isolated from each other such that a hardware failure only affects one of them, then failure of only one disk enables the system to maintain quorum.

- Keep your other volume groups small, if possible. Many small volume groups are preferable to a few large volume groups, for most of the same reasons mentioned previously. In addition, with a very large volume group, the impact of a single disk failure can be widespread, especially if you must deactivate the volume group. With a smaller volume group, the amount of data that is unavailable during recovery is much smaller, and you will spend less time reloading from backup. If you are moving disks between systems, it is easier to track, export, and import smaller volume groups. Several small volume groups often have better performance than a single large one. Finally, if you ever have to recreate all the disk layouts, a smaller volume group is easier to map.

  Consider organizing your volume groups such that the data in each volume group is dedicated to a particular task. If a disk failure makes a volume group unavailable, then only its associated task is affected during the recovery process.

- Maintain adequate documentation of your I/O and LVM configuration, specifically the outputs from the following commands:

Command	Scope	Purpose
`ioscan -f`		Print I/O configuration
`lvlnboot -v`		Print information on root, boot, swap, and dump logical volumes
`vgcfgrestore -l`	for all volume groups	Print volume group configuration from backup file
`vgdisplay -v`	for all volume groups	Print volume group information, including status of logical volumes and physical volumes
`lvdisplay -v`	for all logical volumes	Print logical volume information, including mapping and status of logical extents
`pvdisplay -v`	for all physical volumes	Print physical volume information, including status of physical extents

  With this information in hand, you or your HP support representative may be able to reconstruct a lost configuration, even if the LVM disks have corrupted headers. A hard copy is not required or even necessarily practical, but accessibility during recovery is important and you should plan for this.

- Make sure that your LVM configuration backups stay up-to-date. Make an explicit configuration backup using the `vgcfgbackup` command immediately after importing any volume group or activating any shared volume group for the first time. Normally, LVM backs up a volume group configuration whenever you run a command to change that configuration; if an LVM command prints a warning that `vgcfgbackup` failed, be sure to investigate it.

While this list of preparatory actions does not keep a disk from failing, it makes it easier for you to deal with failures when they occur.

2. Recognizing a Failing Disk

The guidelines in the previous section will not prevent disk failure on your system. Assuming you follow all the recommendations, how can you tell when a disk has failed? This section explains how to look for signs that one of your disks is having problems, and how to determine which disk it is.

I/O Errors in the System Log

Often an error message in the system log file is your first indication of a disk problem. In `/var/adm/syslog/syslog.log`, you might see the following error:

```
SCSI: Request Timeout -- lbolt: 329741615, dev: 1f022000
```

To map this error message to a specific disk, look under the `/dev` directory for a device file with a device number that matches the printed value. More specifically, search for a file whose minor number matches the lower six digits of the number following `dev:`. The device number in this example is 1f022000; its lower six digits are 022000, so search for that value using the following command:

```
# ll /dev/*dsk | grep 022000
brw-r-----   1 bin          sys          31 0x022000 Sep 22  2002 c2t2d0
crw-r-----   1 bin          sys         188 0x022000 Sep 25  2002 c2t2d0
```

This gives you a device file to use for further investigation.

Disk Failure Notification Messages from Diagnostics

If you have EMS hardware monitors installed on your system, and you enabled the disk monitor `disk_em`, a failing disk can trigger an event to the Event Monitoring Service (EMS). Depending on how you configured EMS, you might get an email message, information in `/var/adm/syslog/syslog.log`, or messages in another log file. EMS error messages identify a hardware problem, what caused it, and what must be done to correct it. The following is part of an example message:

```
Event Time..........: Tue Oct 26 14:06:00 2004
Severity............: CRITICAL
Monitor.............: disk_em
Event #.............: 18
System..............: myhost

Summary:
    Disk at hardware path 0/2/1/0.2.0 : Drive is not responding.

Description of Error:

    The hardware did not respond to the request by the driver. The I/O request
    was not completed.

Probable Cause / Recommended Action:

    The I/O request that the monitor made to this device failed because the
    device timed-out. Check cables, power supply, ensure the drive is powered
    ON, and if needed contact your HP support representative to check the
    drive.
```

For more information on EMS, see the [diagnostics section](http://docs.hp.com) of http://docs.hp.com.

LVM Command Errors

Sometimes LVM commands, such as `vgdisplay`, return an error suggesting that a disk has problems. For example:

```
# vgdisplay –v | more
…
   --- Physical volumes ---
   PV Name                      /dev/dsk/c0t3d0
   PV Status                    unavailable
   Total PE                     1023
   Free PE                      173
…
```

The physical volume status of `unavailable` indicates that LVM is having problems with the disk. You can get the same status information from `pvdisplay`.

The next two examples are warnings from `vgdisplay` and `vgchange` indicating that LVM has no contact with a disk:

```
# vgdisplay -v vg
vgdisplay: Warning: couldn't query physical volume "/dev/dsk/c0t3d0":
The specified path does not correspond to physical volume attached to this volume
group
vgdisplay: Warning: couldn't query all of the physical volumes.
```

```
# vgchange -a y /dev/vg01
vgchange: Warning: Couldn't attach to the volume group physical volume
"/dev/dsk/c0t3d0":
A component of the path of the physical volume does not exist.
Volume group "/dev/vg01" has been successfully changed.
```

Another sign that you might have a disk problem is seeing stale extents in the output from `lvdisplay`. If you have stale extents on a logical volume even after running the `vgsync` or `lvsync` commands, you might have an issue with an I/O path or one of the disks used by the logical volume, but not necessarily the disk showing stale extents. For example:

```
# lvdisplay –v /dev/vg01/lvol3 | more
…
LV Status                      available/stale
…
   --- Logical extents ---
   LE    PV1                PE1  Status 1  PV2               PE2  Status 2
   0000 /dev/dsk/c0t3d0   0000 current   /dev/dsk/c1t3d0   0100 current
   0001 /dev/dsk/c0t3d0   0001 current   /dev/dsk/c1t3d0   0101 current
   0002 /dev/dsk/c0t3d0   0002 current   /dev/dsk/c1t3d0   0102 stale
   0003 /dev/dsk/c0t3d0   0003 current   /dev/dsk/c1t3d0   0103 stale
…
```

All LVM error messages tell you which device file is associated with the problematic disk. This is useful for the next step, confirming disk failure.

3. Confirming Disk Failure

Once you suspect a disk has failed or is failing, make certain that the suspect disk is indeed failing. Replacing or removing the incorrect disk makes the recovery process take longer. It can even cause data loss. For example, in a mirrored configuration, if you were to replace the wrong disk—the one holding the current good copy rather than the failing disk—the mirrored data on the good disk is lost.

It is also possible that the suspect disk is not failing. What seems to be a disk failure might be a hardware path failure; that is, the I/O card or cable might have failed. If a disk has multiple hardware paths, also known as *pvlinks*, one path can fail while an alternate path continues to work. For such disks, try the following steps on all paths to the disk.

If you have isolated a suspect disk, you can use hardware diagnostic tools like Support Tools Manager to get detailed information about it. Use these tools as your first approach to confirm disk failure. They are documented on http://docs.hp.com in the diagnostics area. If you do not have diagnostic tools available, use the following three-step approach to confirm that a disk has failed or is failing:

1. Use the `ioscan` command to check the S/W state of the disk. Only disks in state CLAIMED are currently accessible by the system. Disks in other states such as NO_HW or disks that are completely missing from the `ioscan` output are suspicious. If the disk is marked as CLAIMED, then its controller is responding. For example:

    ```
    # ioscan -fCdisk
    Class  I  H/W Path     Driver  S/W State H/W Type  Description
    ================================================================
    disk   0  8/4.5.0      sdisk   CLAIMED   DEVICE    SEAGATE ST34572WC
    disk   1  8/4.8.0      sdisk   UNCLAIMED UNKNOWN   SEAGATE ST34572WC
    disk   2  8/16/5.2.0   sdisk   CLAIMED   DEVICE    TOSHIBA CD-ROM XM-5401TA
    ```

    In this example the disk at hardware path 8/4.8.0 is not accessible.

    If the disk has multiple hardware paths, be sure to check all the paths.

2. If the disk responds to the `ioscan` command, test it with the `diskinfo` command. The reported size must be nonzero, otherwise the device is not ready. For example:

    ```
    # diskinfo /dev/rdsk/c0t5d0
    SCSI describe of /dev/rdsk/c0t5d0:
                vendor: SEAGATE
            product id: ST34572WC
                  type: direct access
                  size: 0 Kbytes
       bytes per sector: 512
    ```

    In this example the size is 0, so the disk is malfunctioning.

3. If both `ioscan` and `diskinfo` succeed, the disk might still be failing. As a final test, try to read from the disk using the `dd` command. Depending on the size of the disk, a comprehensive read can be time-consuming, so you might want to read only a portion of the disk. If the disk is functioning properly, no I/O errors are reported.

For example, the following shows a successful read of the first 64 megabytes of the disk:

```
# dd if=/dev/rdsk/c0t5d0 of=/dev/null bs=1024k count=64
64+0 records in
64+0 records out
```

The following shows an unsuccessful read of the whole disk:

```
# dd if=/dev/rdsk/c1t3d0 of=/dev/null bs=1024k
dd read error: I/O error
0+0 records in
0+0 records out
```

4. Gathering Information about a Failing Disk

Once you know which disk is failing, you can decide how to deal with it. You can choose to remove the disk if your system does not need it, or you can choose to replace it. Before deciding on your course of action, you must gather some information to help guide you through the recovery process.

Is the Questionable Disk Hot-Swappable?

This determines whether you must power down your system to replace the disk. If you do not want to power down your system and the failing disk is not hot-swappable, the best you can do is disable LVM access to the disk.

Is it the Root Disk or Part of the Root Volume Group?

If the root disk is failing, the replacement process has a few extra steps to set up the boot area; in addition, you might have to boot from the mirror of the root disk if the primary root disk has failed. If a failing root disk is not mirrored, you must reinstall to the replacement disk, or recover it from an Ignite-UX backup.

To determine whether the disk is in the root volume group, use the `lvlnboot` command with the `-v` option. It lists the disks in the root volume group, and any special volumes configured on them. For example:

```
# lvlnboot -v
Boot Definitions for Volume Group /dev/vg00:
Physical Volumes belonging in Root Volume Group:
        /dev/dsk/c0t5d0 (0/0/0/3/0.5.0) -- Boot Disk
Boot: lvol1     on:     /dev/dsk/c0t5d0
Root: lvol3     on:     /dev/dsk/c0t5d0
Swap: lvol2     on:     /dev/dsk/c0t5d0
Dump: lvol2     on:     /dev/dsk/c0t5d0, 0
```

What Recovery Strategy Do You Have for the Logical Volumes on this Disk?

Part of the disk removal or replacement process is based on what recovery strategy you have for the data on that disk. You can have different strategies (mirroring, restoring from backup, reinitializing from scratch) for each logical volume.

You can find the list of logical volumes using the disk with the `pvdisplay` command. For example:

```
# pvdisplay -v /dev/dsk/c0t5d0 | more
…
    --- Distribution of physical volume ---
    LV Name            LE of LV  PE for LV
    /dev/vg00/lvol1    75        75
    /dev/vg00/lvol2    512       512
    /dev/vg00/lvol3    50        50
    /dev/vg00/lvol4    50        50
    /dev/vg00/lvol5    250       250
    /dev/vg00/lvol6    450       450
    /dev/vg00/lvol7    350       350
    /dev/vg00/lvol8    1000      1000
    /dev/vg00/lvol9    1000      1000
    /dev/vg00/lvol10   3         3
…
```

If `pvdisplay` fails, you have several options. You can refer to any configuration documentation you created in advance. Alternately, you can run `lvdisplay -v` on all the logical volumes in the volume group and see if any

extents are mapped to an unavailable physical volume. The `lvdisplay` command shows '???' for the physical volume if it is unavailable.

The problem with this approach is that it is not precise if more than one disk is unavailable; to ensure that multiple simultaneous disk failures have not occurred, run `vgdisplay` to see if the active and current number of physical volumes differs by exactly one.

A third option for determining which logical volumes are on the disk is to use the `vgcfgdisplay` command. This command is available from your HP support representative.

If you have mirrored any logical volume onto a separate disk, confirm that the mirror copies are current. For each of the logical volumes affected, use `lvdisplay` to determine if the number of mirror copies is greater than zero. This verifies that the logical volume is mirrored. Then use `lvdisplay` again to determine which logical extents are mapped onto the suspect disk, and whether there is a current copy of that data on another disk. For example:

```
 # lvdisplay -v /dev/vg00/lvol1
--- Logical volumes ---
LV Name                 /dev/vg00/lvol1
VG Name                 /dev/vg00
LV Permission           read/write
LV Status               available/syncd
Mirror copies           1
Consistency Recovery    MWC
Schedule                parallel
LV Size (Mbytes)        300
Current LE              75
Allocated PE            150
Stripes                 0
Stripe Size (Kbytes)    0
Bad block               off
Allocation              strict/contiguous
IO Timeout (Seconds)    default

# lvdisplay -v /dev/vg00/lvol1 | grep -e /dev/dsk/c0t5d0 -e '???'
   00000 /dev/dsk/c0t5d0  00000 current  /dev/dsk/c2t6d0  00000 current
   00001 /dev/dsk/c0t5d0  00001 current  /dev/dsk/c2t6d0  00001 current
   00002 /dev/dsk/c0t5d0  00002 current  /dev/dsk/c2t6d0  00002 current
   00003 /dev/dsk/c0t5d0  00003 current  /dev/dsk/c2t6d0  00003 current
   00004 /dev/dsk/c0t5d0  00004 current  /dev/dsk/c2t6d0  00004 current
   00005 /dev/dsk/c0t5d0  00005 current  /dev/dsk/c2t6d0  00005 current
…
```

The first `lvdisplay` command output shows that `lvol1` is mirrored. In the second `lvdisplay` command output, you can see that all extents of the failing disk have a current copy elsewhere on the system, specifically on `/dev/dsk/c2t6d0`. If the disk is unavailable when the volume group is activated, its column contains a '???' instead of the disk name.

With this information in hand, you can now decide how best to resolve the disk failure.

5. Removing the Disk

If you have a copy of the data on the failing disk, or you can move the data to another disk, you can choose to remove the disk from the system instead of replacing it.

Removing a Mirror Copy from a Disk

If you have a mirror copy of the data already, you can stop LVM from using the copy on the failing disk by reducing the number of mirrors. To remove the mirror copy from a specific disk, use `lvreduce`, and specify the disk from which to remove the mirror copy. For example:

```
# lvreduce -m 0 -A n /dev/vgname/lvname pvname    (if you have a single mirror copy)
or
# lvreduce -m 1 -A n /dev/vgname/lvname pvname    (if you have two mirror copies)
```

The `-A n` option is used to prevent the `lvreduce` command from performing an automatic `vgcfgbackup` operation, which could hang while accessing a defective disk.

If you have only a single mirror copy and want to maintain redundancy, create a second mirror of the data on a different, functional disk, subject to the mirroring guidelines mentioned in [Preparing for Disk Recovery](#), *before* you run `lvreduce`.

If the disk was not available at boot time (`pvdisplay` failed) then the `lvreduce` command fails with an error that it could not query the physical volume. You can still remove the mirror copy, but you must specify the physical volume ***key*** rather than the name. You can get the key using `lvdisplay` with the `-k` option as follows:

```
# lvdisplay -v -k /dev/vg00/lvol1
…
--- Logical extents ---
   LE    PV1              PE1    Status 1 PV2              PE2    Status 2
   00000       0         00000 stale          1         00000 current
   00001       0         00001 stale          1         00001 current
   00002       0         00002 stale          1         00002 current
   00003       0         00003 stale          1         00003 current
   00004       0         00004 stale          1         00004 current
   00005       0         00005 stale          1         00005 current
…
```

Compare this output with the output of `lvdisplay` without `-k`, which you did to check the mirror status. The column that contained the failing disk (or '???') now holds the key. For this example, the key is 0. Use this key with `lvreduce` as follows:

```
# lvreduce -m 0 -A n -k /dev/vgname/lvname key    (if you have a single mirror copy)
or
# lvreduce -m 1 -A n -k /dev/vgname/lvname key    (if you have two mirror copies)
```

Moving the Physical Extents to Another Disk

If the disk is marginal and you can still read from it, you can move the data onto another disk by moving the physical extents onto another disk.

The `pvmove` command moves logical volumes or certain extents of a logical volume from one physical volume to another. It is typically used to free up a disk—that is, to move all data from that physical volume so it can be removed from the volume group. In its simplest invocation, you specify the disk to free up, and LVM moves all the

physical extents on that disk to any other disks in the volume group, subject to any mirroring allocation policies. For example:

```
# pvmove pvname
```

You can select a particular target disk or disks, if desired. For example, to move all the physical extents from c0t5d0 to the physical volume c0t2d0, enter the following command:

```
# pvmove /dev/dsk/c0t5d0 /dev/dsk/c0t2d0
```

You can choose to move only the extents belonging to a particular logical volume. Use this option if only certain sectors on the disk are readable, or if you want to move only unmirrored logical volumes. For example, to move all physical extents of lvol4 that are located on physical volume c0t5d0 to c1t2d0, enter the following command:

```
# pvmove -n /dev/vg01/lvol4 /dev/dsk/c0t5d0 /dev/dsk/c1t2d0
```

Note that pvmove is not an atomic operation, and moves data extent by extent. If pvmove is abnormally terminated by a system crash or kill -9, the volume group can be left in an inconsistent configuration showing an additional pseudo mirror copy for the extents being moved. You can remove the extra mirror copy using the lvreduce command with the -m option on each of the affected logical volumes; there is no need to specify a disk.

Removing the Disk from the Volume Group

After the disk no longer holds any physical extents, you can use the vgreduce command to remove it from the volume group so it is not inadvertently used again. Check for alternate links before removing the disk, since you must remove all the paths to a multipathed disk. Use the pvdisplay command as follows:

```
# pvdisplay /dev/dsk/c0t5d0
--- Physical volumes ---
PV Name                /dev/dsk/c0t5d0
PV Name                /dev/dsk/c1t6d0  Alternate Link
VG Name                /dev/vg01
PV Status              available
Allocatable            yes
VGDA                   2
Cur LV                 0
PE Size (Mbytes)       4
Total PE               1023
Free PE                1023
Allocated PE           0
Stale PE               0
IO Timeout (Seconds)   default
Autoswitch             On
```

In this example, there are two entries for the PV Name. Use the vgreduce command to reduce each path as follows:

```
# vgreduce vgname /dev/dsk/c0t5d0
# vgreduce vgname /dev/dsk/c1t6d0
```

If the disk is unavailable, the vgreduce command fails. You can still forcibly reduce it, but you must then rebuild the lvmtab, which has two side effects. First, any deactivated volume groups are left out of the lvmtab, so you must manually vgimport them later. Second, any multipathed disks have their link order reset; if you arranged your pvlinks to implement load-balancing, you might have to arrange them again. The procedure to remove the disk and rebuild lvmtab is as follows:

```
# vgreduce -f vgname
# mv /etc/lvmtab /etc/lvmtab.save
# vgscan -v
```

This completes the procedure for removing the disk from your LVM configuration. If the disk hardware allows it, you can remove it physically from the system. Otherwise, physically remove it at the next scheduled system reboot.

6. Replacing the Disk

If you decide to replace the disk, you must perform a five-step procedure. How you perform each step depends on the information you gathered earlier (hot-swap information, logical volume names, and recovery strategy), so this procedure varies.

This section also includes several common scenarios for disk replacement, and a flowchart summarizing the disk replacement procedure.

The five steps are:

1. Temporarily halt LVM attempts to access the disk.
2. Physically replace the faulty disk.
3. Configure LVM information on the disk.
4. Re-enable LVM access to the disk.
5. Restore any lost data onto the disk.

In the following steps, *pvname* is the character device special file for the physical volume. This name might be `/dev/rdsk/c2t15d0` or `/dev/rdsk/c2t1d0s2`.

Step 1: Halting LVM Access to the Disk

This is known as **detaching** the disk. The actions you take to detach the disk depend on whether the data is mirrored, if the LVM online disk replacement functionality is available, and what applications are using the disk. In some cases (for example, if an unmirrored file system cannot be unmounted), you must shut down the system.

- If the disk is not hot-swappable, you must power down the system to replace it. By shutting down the system, you halt LVM access to the disk, so you can skip this step.

- If the disk contains any unmirrored logical volumes or any mirrored logical volumes without an available and current mirror copy, halt any applications and unmount any file systems using these logical volumes. This prevents the applications or file systems from writing inconsistent data over the newly restored replacement disk. For each logical volume on the disk:

    o If the logical volume is mounted as a file system, try to unmount the file system.

       # **umount /dev/*vgname*/*lvname***

       Attempting to unmount a file system that has open files (or that contains a user's current working directory) causes the command to fail with a `Device busy` message. You can use the following procedure to determine what users and applications are causing the unmount operation to fail:

       1. Use the `fuser` command to find out what applications are using the file system as follows:

          # **fuser –u /dev/*vgname*/*lvname***

          This command displays process IDs and users with open files mounted on that logical volume, and whether it is a user's working directory.

       2. Use the `ps` command to map the list of process IDs to processes, and then determine whether you can halt those processes.

3. To kill processes using the logical volume, enter the following command:

    # **fuser –ku /dev/*vgname*/*lvname***

4. Then try to unmount the file system again as follows:

    # **umount /dev/*vgname*/*lvname***

o If the logical volume is being accessed as a raw device, you can use `fuser` to find out which applications are using it. Then you can halt those applications.

If for some reason you cannot disable access to the logical volume—for example, you cannot halt an application or you cannot unmount the file system—you must shut down the system.

- If you have LVM online replacement (OLR) functionality available, detach the device using the `–a` option of the `pvchange` command:

  # **pvchange –a N *pvname***

  If `pvchange` fails with a message that the `–a` option is not recognized, then the LVM OLR feature is not installed.

- If you do not have LVM OLR functionality, LVM continues to try to access the disk as long as it is in the volume group and has ever been available. You can make LVM stop accessing the disk in one of three ways:

  1. Remove the disk from the volume group. This means reducing any logical volumes that have mirror copies on the faulty disk so that they no longer mirror onto that disk, and reducing the disk from the disk group, as described in [Removing the Disk](#). This maximizes access to the rest of the volume group, but requires more LVM commands to modify the configuration and then re-create it on a replacement disk.

  2. Deactivate the volume group. You do not have to remove and re-create any mirrors, but all data in the volume group is inaccessible during the replacement procedure.

  3. Shut down the system. This halts LVM access to the disk, but it makes the entire system inaccessible. Use this option only if you do not want to remove the disk from the volume group, and you cannot deactivate it.

The following recommendations are intended to maximize system uptime and access to the volume group, but you can use a stronger approach if your data and system availability requirements allow.

- If `pvdisplay` shows `PV status` as available, halt LVM access to the disk by removing it from the volume group.

- If `pvdisplay` shows `PV status` as unavailable, or if `pvdisplay` fails to print the status, use `ioscan` to determine if the disk can be accessed at all. If `ioscan` reports the disk status as NO_HW on all its hardware paths, then you can simply remove the disk. If `ioscan` shows any other status, halt LVM access to the disk by deactivating the volume group.

Step 2: Replacing the Faulty Disk

If the disk is hot-swappable, you can replace it without powering down the system. Otherwise, power down the system before replacing the disk. For the hardware details on how to replace the disk, refer to the hardware administrator's guide for the system or disk array.

If you powered down the system, reboot it normally. The only exception is if you replaced a disk in the root volume group.

- If you replaced the disk that you normally boot from, the replacement disk will not contain the information needed by the boot loader. If your root disk is mirrored, boot from it by using the alternate boot path. If the root disk was not mirrored, you have no recourse but to reinstall or recover your system.

- If there are only two disks in the root volume group, the system will probably fail its quorum check, and might panic early in the boot process with the message "panic: LVM: Configuration failure". In this situation, you must override quorum to boot successfully. Do this by interrupting the boot process and adding the option -lq to the boot command normally used by the system. Boot process and options are discussed in Chapter 5 of Managing Systems and Workgroups.

Step 3: Initializing the Disk for LVM

This step copies LVM configuration information onto the disk, and marks it as owned by LVM so it can subsequently be attached to the volume group.

If you replaced a mirror of the root disk on an Integrity server, run the `idisk` command as described in step 1 of *Appendix A: Mirroring the Root Volume on Integrity Servers*. For PA-RISC servers or non-root disks, this step is unnecessary.

For any replaced disk, restore LVM configuration information to the disk using the `vgcfgrestore` command as follows:

vgcfgrestore –n *vgname pvname*

Step 4: Re-enabling LVM Access to the Disk

This is known as **attaching** the disk. The action you take here depends on whether LVM OLR is available.

- If you have LVM OLR, attach the device using the -a option with the `pvchange` command as follows:

  # **pvchange -a y *pvname***

  After LVM processes the `pvchange` command, it resumes using the device if possible.

- If you do not have LVM OLR, or you want to ensure that any alternate links are attached, use the `vgchange` command to activate the volume group and bring any detached devices online:

  # **vgchange -a y *vgname***

  The `vgchange` command attaches all paths for all disks in the volume group, and resumes automatically recovering any unattached failed disks in the volume group. Therefore, run `vgchange` only after all work has been completed on all disks and paths in the volume group, and it is desirable to attach them all.

Step 5: Restoring Lost Data to the Disk

This final step can be a straightforward resynchronization for mirrored configurations, or a recovery of data from backup media.

- If a mirror of the root disk was replaced, initialize its boot information:
  - For an Integrity server, follow steps 5, 6, and 8 in *Appendix A: Mirroring the Root Volume on Integrity Servers*.
  - For a PA-RISC server, follow steps 4, 5, and 7 in *Appendix A: Mirroring the Root Volume on PA-RISC Servers*.

- If all the data on the replaced disk was mirrored, you do not have to do anything; LVM automatically synchronizes the data on the disk with the other mirror copies of the data.

- If the disk contained any unmirrored logical volumes (or mirrored logical volumes that did not have a current copy on the system), restore the data from backup, mount the file systems, and restart any applications you halted in step 1.

Replacing a LVM Disk in an HP Serviceguard Cluster Volume Group

Replacing LVM disks in an HP Serviceguard cluster follows the same procedure described in steps 1-5, unless the volume group is shared. If the volume group is shared, make these two changes:

- When disabling LVM access to the disk, perform any online disk replacement steps individually on each cluster node sharing the volume group. If you do not have LVM OLR, and you detach the disk, you will probably need to make configuration changes that require you to deactivate the volume group on all cluster nodes. However, if you have Shared LVM Single Node Online Volume Reconfiguration (SNOR) installed, you can leave the volume group activated on one of the cluster nodes.
- When re-enabling LVM access, activate the physical volume on each cluster node sharing the volume group.

Special care is required when performing a Serviceguard rolling upgrade. For details, refer to the LVM Online Disk Replacement (LVM OLR) white paper.

Disk Replacement Scenarios

Following are several scenarios that illustrate LVM disk replacement.

Scenario 1: Best Case

For this example, you have followed all the guidelines in section 1—all disks are hot-swappable, all logical volumes are mirrored, and LVM OLR patches are installed. In this case, you can detach the disk with `pvchange`, replace it, reattach it, and let LVM mirroring synchronize the logical volumes—all while the system remains up.

For this example, you assume the bad disk is at hardware path 2/0/7.15.0 and has device special files named `/dev/rdsk/c2t15d0` and `/dev/dsk/c2t15d0`.

Check that the disk is not in the root volume group, and that all logical volumes on the bad disk are mirrored with a current copy available. Run the following commands:

```
# lvlnboot -v
Boot Definitions for Volume Group /dev/vg00:
Physical Volumes belonging in Root Volume Group:
        /dev/dsk/c0t5d0 (0/0/0/3/0.5.0) -- Boot Disk
Boot: lvol1     on:     /dev/dsk/c0t5d0
```

```
Root: lvol3     on:      /dev/dsk/c0t5d0
Swap: lvol2     on:      /dev/dsk/c0t5d0
Dump: lvol2     on:      /dev/dsk/c0t5d0, 0
# pvdisplay -v /dev/dsk/c2t15d0 | more
…
    --- Distribution of physical volume ---
    LV Name            LE of LV  PE for LV
    /dev/vg01/lvol1    4340        4340
…
# lvdisplay -v /dev/vg01/lvol1 | grep "Mirror copies"
Mirror copies                    1
# lvdisplay -v /dev/vg01/lvol1 | grep -e /dev/dsk/c2t15d0 -e '???' | more
    00000 /dev/dsk/c2t15d0  00000 current  /dev/dsk/c5t15d0  00000 current
    00001 /dev/dsk/c2t15d0  00001 current  /dev/dsk/c5t15d0  00001 current
    00002 /dev/dsk/c2t15d0  00002 current  /dev/dsk/c5t15d0  00002 current
    00003 /dev/dsk/c2t15d0  00003 current  /dev/dsk/c5t15d0  00003 current
…
```

The `lvlnboot` command confirms that the disk is not in the root volume group. The `pvdisplay` command shows which logical volumes are on the disk. The `lvdisplay` command shows that all data in the logical volume has a current mirror copy on another disk. Run the following commands to continue with the disk replacement:

```
# pvchange -a N /dev/dsk/c2t15d0
# <replace the hot-swappable disk>
# vgcfgrestore -n vg01 /dev/rdsk/c2t15d0
# vgchange -a y vg01
```

Scenario 2: No Mirroring and No LVM Online Replacement

In this example, the disk is still hot-swappable, but there are unmirrored logical volumes and the LVM OLR patches are not installed. Disabling LVM access to the logical volumes is more complicated, since you must find out what processes are using them.

The bad disk is represented by device special file `/dev/dsk/c2t2d0`. Run the following commands:

```
# lvlnboot -v
Boot Definitions for Volume Group /dev/vg00:
Physical Volumes belonging in Root Volume Group:
        /dev/dsk/c0t5d0 (0/0/0/3/0.5.0) -- Boot Disk
Boot: lvol1     on:      /dev/dsk/c0t5d0
Root: lvol3     on:      /dev/dsk/c0t5d0
Swap: lvol2     on:      /dev/dsk/c0t5d0
Dump: lvol2     on:      /dev/dsk/c0t5d0, 0
# pvdisplay -v /dev/dsk/c2t2d0 | more
…
    --- Distribution of physical volume ---
    LV Name            LE of LV  PE for LV
    /dev/vg01/lvol1    4340        4340
…
# lvdisplay -v /dev/vg01/lvol1 | grep "Mirror copies"
Mirror copies                    0
```

This confirms that the logical volume is not mirrored, and it is not in the root volume group. As system administrator, you know that the logical volume is a mounted file system. To disable access to the logical volume, try to unmount it. Use `fuser` to isolate and terminate processes using the file system, if necessary. Run the following commands:

```
# umount /dev/vg01/lvol1
umount: cannot unmount /dump : Device busy
```

```
# fuser -u /dev/vg01/lvol1
/dev/vg01/lvol1:    27815c(root)    27184c(root)
# ps -fp27815 -p27184
    UID   PID  PPID  C    STIME TTY        TIME COMMAND
   root 27815 27184  0 09:04:05 pts/0      0:00 vi test.c
   root 27184 27182  0 08:26:24 pts/0      0:00 -sh
# fuser -ku /dev/vg01/lvol1
/dev/vg01/lvol1:    27815c(root)    27184c(root)
# umount /dev/vg01/lvol1
```

For this example, it is assumed that you are permitted to halt access to the entire volume group while you recover the disk. Use vgchange to deactivate the volume group and stop LVM from accessing the disk:

```
# vgchange –a n vg01
```

 Proceed with the disk replacement and recover data from backup.

```
# <replace the hot-swappable disk>
# vgcfgrestore –n vg01 /dev/rdsk/c2t2d0
# vgchange –a y vg01
# newfs [options] /dev/vg01/rlvol1
# mount /dev/vg01/lvol1 /dump
# <restore the file system from backup>
```

Scenario 3: No Hot-Swappable Disk

In this example, the disk is not hot-swappable, so you must reboot the system to replace it. Once again, the bad disk is represented by device special file /dev/dsk/c2t2d0. Run the following commands:

```
# lvlnboot –v
Boot Definitions for Volume Group /dev/vg00:
Physical Volumes belonging in Root Volume Group:
       /dev/dsk/c0t5d0 (0/0/0/3/0.5.0) -- Boot Disk
Boot: lvol1     on:     /dev/dsk/c0t5d0
Root: lvol3     on:     /dev/dsk/c0t5d0
Swap: lvol2     on:     /dev/dsk/c0t5d0
Dump: lvol2     on:     /dev/dsk/c0t5d0, 0
# pvdisplay –v /dev/dsk/c2t2d0 | more
…
   --- Distribution of physical volume ---
   LV Name            LE of LV  PE for LV
   /dev/vg01/lvol1    4340      4340
…
# lvdisplay –v /dev/vg01/lvol1 | grep "Mirror copies"
Mirror copies              0
```

This confirms that the logical volume is not mirrored, and it is not in the root volume group. Shutting down the system disables access to the disk, so you do not need to determine who is using the logical volume.

```
# shutdown –h
# <replace the disk>
# <reboot normally>
# vgcfgrestore –n vg01 /dev/rdsk/c2t2d0
# vgchange –a y vg01
# newfs [options] /dev/vg01/rlvol1
# mount /dev/vg01/lvol1 /dump
# <restore the file system from backup>
```
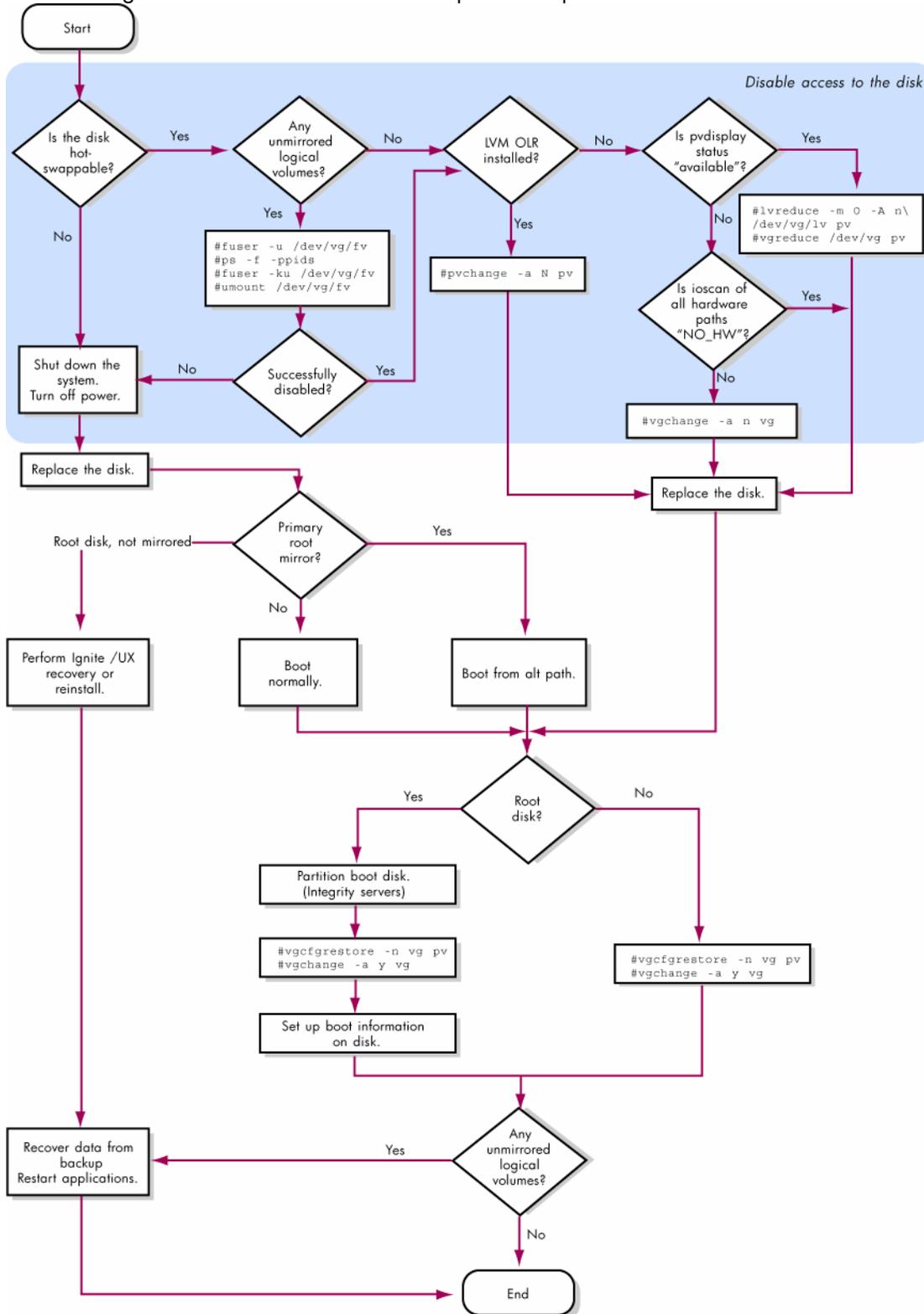
Disk Replacement Flowchart

The following flowchart summarizes the disk replacement process.

Conclusion

In your role as system manager, you will encounter disk failures. LVM can lessen the impact of those disk failures, enabling you to configure your data storage to make a disk failure transparent to users, and to keep your system and data available during the recovery process. By making use of hardware features such as hot-swappable disks and software features such as mirroring and online disk replacement, you can maximize your system availability and minimize data loss due to disk failure.

Appendix A: Procedures

This section contains details on some of the procedures described earlier.

Mirroring the Root Volume on PA-RISC Servers

To set up a mirrored root configuration, you must add a disk to the root volume group, mirror all the root logical volumes onto it, and make it bootable. For this example, the disk is at path 2/0/7.15.0 and has device special files named `/dev/rdsk/c2t15d0` and `/dev/dsk/c2t15d0`.

1. Use the `insf` command with the `-e` option to make sure the device files are in place. For example:

    ```
    # insf -e -H 2/0/7.15.0
    ```

    You should now have the following device files for this disk:

`/dev/dsk/c2t15d0`	The entire disk (block access)
`/dev/rdsk/c2t15d0`	The entire disk (character access)

2. Create a physical volume using `pvcreate` with the `-B` option:

    ```
    # pvcreate -B /dev/rdsk/c2t15d0
    ```

3. Add the physical volume to your existing root volume group using `vgextend`:

    ```
    # vgextend /dev/vg00 /dev/dsk/c2t15d0
    ```

4. Use the `mkboot` command to set up the boot area:

    ```
    # mkboot /dev/rdsk/c2t15d0
    ```

5. Use the `mkboot` command to add an autoboot file to the disk boot area. If you expect to boot from this disk only when you lose quorum, you can use the alternate string "`hpux -lq`" to disable quorum checking:

    ```
    # mkboot -a "hpux" /dev/rdsk/c2t15d0
    ```

1. Use the `lvextend` command to mirror each logical volume in `vg00` (the root volume group) onto the specified physical volume. The logical volumes must be extended in the same order that they are configured on the original boot disk. Use the `pvdisplay` command with the `-v` option to determine the list of logical volumes and their order. For example:

    ```
    # pvdisplay -v /dev/dsk/c0t6d0 | grep 'current.*0000 $'
       00000 current   /dev/vg00/lvol1     00000
       00038 current   /dev/vg00/lvol2     00000
       00550 current   /dev/vg00/lvol3     00000
       00583 current   /dev/vg00/lvol4     00000
       00608 current   /dev/vg00/lvol5     00000
       00611 current   /dev/vg00/lvol6     00000
       00923 current   /dev/vg00/lvol7     00000
       01252 current   /dev/vg00/lvol8     00000
    ```

    In this example, mirror the logical volumes as follows:
    ```
    # lvextend -m 1 /dev/vg00/lvol1 /dev/dsk/c2t15d0
    # lvextend -m 1 /dev/vg00/lvol2 /dev/dsk/c2t15d0
    # lvextend -m 1 /dev/vg00/lvol3 /dev/dsk/c2t15d0
    # lvextend -m 1 /dev/vg00/lvol4 /dev/dsk/c2t15d0
    ```

```
# lvextend -m 1 /dev/vg00/lvol5 /dev/dsk/c2t15d0
# lvextend -m 1 /dev/vg00/lvol6 /dev/dsk/c2t15d0
# lvextend -m 1 /dev/vg00/lvol7 /dev/dsk/c2t15d0
# lvextend -m 1 /dev/vg00/lvol8 /dev/dsk/c2t15d0
```

6. Update the root volume group information:

```
# lvlnboot -R /dev/vg00
```

7. Verify that the mirrored disk is displayed as a boot disk and that the boot, root, and swap logical volumes appear to be on both disks:

```
# lvlnboot -v
```

8. Specify the mirror disk as the alternate boot path in nonvolatile memory:

```
# setboot -a 2/0/7.15.0
```

9. Add a line to /stand/bootconf for the new boot disk using vi or another text editor:

```
# vi /stand/bootconf
l /dev/dsk/c2t15d0
```

where l denotes LVM.

Mirroring the Root Volume on Integrity Servers

The procedure to mirror the root disk on Integrity servers is similar to the procedure for PA-RISC servers. The difference is that Integrity server boot disks are partitioned; you must set up the partitions, copy utilities to the EFI partition, and use the HP-UX partition device files for LVM commands.

For this example, the disk is at hardware path 0/1/1/0.1.0, with a device special file named `/dev/rdsk/c2t1d0`.

1. Partition the disk using the `idisk` command and a partition description file.

   a. Create a partition description file. For example:

      # **vi /tmp/pdf**

      In this example the partition description file contains:
      ```
      3
      EFI 500MB
      HPUX 100%
      HPSP 400MB
      ```

   b. Partition the disk using `idisk` and the partition description file created in step 1a:

      # **idisk -f /tmp/pdf -w /dev/rdsk/c2t1d0**

      To verify that your partitions are correctly laid out, run the following command:

      # **idisk /dev/rdsk/c2t1d0**

2. Use the `insf` command with the `-e` option to create the device files for all the partitions. For example:

   # **insf -e -H 0/1/1/0.1.0**

   You should now have the following device files for this disk:

`/dev/dsk/c2t1d0`	The entire disk (block access)
`/dev/rdsk/c2t1d0`	The entire disk (character access)
`/dev/dsk/c2t1d0`**s1**	The EFI partition (block access)
`/dev/rdsk/c2t1d0`**s1**	The EFI partition (character access)
`/dev/dsk/c2t1d0`**s2**	The HP-UX partition (block access)
`/dev/rdsk/c2t1d0`**s2**	The HP-UX partition (character access)
`/dev/dsk/c2t1d0`**s3**	The Service partition (block access)
`/dev/rdsk/c2t1d0`**s3**	The Service partition (character access)

3. Create a physical volume using `pvcreate` with the `-B` option. Be sure to use the device file denoting the HP-UX partition.

   # **pvcreate -B /dev/rdsk/c2t1d0s2**

4. Add the physical volume to your existing root volume group using `vgextend`:

   # **vgextend /dev/vg00 /dev/dsk/c2t1d0s2**

5. Use the `mkboot` command to set up the boot area. Specify the `-e` and `-l` options to copy EFI utilities to the EFI partition, and use the device special file for the entire disk:

```
# mkboot -e -l /dev/rdsk/c2t1d0
```

6. Update the autoboot file in the EFI partition.

    a. Create an AUTO file in the current directory. If you expect to boot from this disk only when you have lost quorum, you can use the alternate string "`boot vmunix -lq`" to disable quorum checking:

    ```
    # echo "boot vmunix" > ./AUTO
    ```

    b. Copy the file from the current directory into the new disk EFI partition. Make sure to use the device file with the `s1` suffix:

    ```
    # efi_cp -d /dev/rdsk/c2t1d0s1 ./AUTO /efi/hpux/auto
    ```

7. Use the `lvextend` command to mirror each logical volume in `vg00` (the root volume group) onto the specified physical volume. The logical volumes must be extended in the same order that they are configured on the original boot disk. Use the `pvdisplay` command with the `-v` option to determine the list of logical volumes and their order. For example:

```
# pvdisplay -v /dev/dsk/c0t0d0s2 | grep 'current.*0000 $'
   00000 current   /dev/vg00/lvol1     00000
   00010 current   /dev/vg00/lvol2     00000
   00138 current   /dev/vg00/lvol3     00000
   00151 current   /dev/vg00/lvol4     00000
   00158 current   /dev/vg00/lvol5     00000
   00159 current   /dev/vg00/lvol6     00000
   00271 current   /dev/vg00/lvol7     00000
   00408 current   /dev/vg00/lvol8     00000
```

In this example, mirror the logical volumes as follows:
```
# lvextend -m 1 /dev/vg00/lvol1 /dev/dsk/c2t1d0s2
# lvextend -m 1 /dev/vg00/lvol2 /dev/dsk/c2t1d0s2
# lvextend -m 1 /dev/vg00/lvol3 /dev/dsk/c2t1d0s2
# lvextend -m 1 /dev/vg00/lvol4 /dev/dsk/c2t1d0s2
# lvextend -m 1 /dev/vg00/lvol5 /dev/dsk/c2t1d0s2
# lvextend -m 1 /dev/vg00/lvol6 /dev/dsk/c2t1d0s2
# lvextend -m 1 /dev/vg00/lvol7 /dev/dsk/c2t1d0s2
# lvextend -m 1 /dev/vg00/lvol8 /dev/dsk/c2t1d0s2
```

8. Update the root volume group information:

    ```
    # lvlnboot -R /dev/vg00
    ```

9. Verify that the mirrored disk is displayed as a boot disk and that the boot, root, and swap logical volumes appear to be on both disks:

    ```
    # lvlnboot -v
    ```

10. Specify the mirror disk as the alternate boot path in nonvolatile memory:

    ```
    # setboot -a 0/1/1/0.1.0
    ```

11. Add a line to `/stand/bootconf` for the new boot disk using `vi` or another text editor:

```
# vi /stand/bootconf
l /dev/dsk/c2t1d0s2
```

where `l` denotes LVM.

Appendix B: LVM Error Messages

This appendix lists some of the warning and error messages reported by LVM. For each message, the cause is listed, and an administrator action is recommended. The appendix is divided into two sections, one for LVM command errors, and one for the system log file `/var/adm/syslog/syslog.log` error messages.

LVM Command Error Messages

All LVM commands

> **vgcfgbackup: /etc/lvmtab is out of date with the running kernel:**
> **Kernel indicates # disks for "/dev/*vgname*"; /etc/lvmtab has # disks.**
> **Cannot proceed with backup.**

    Cause:

        The number of current and active physical volumes, printed by `vgdisplay` as `Cur PV` and `Act PV`, are not the same. Cur PV and Act PV must always agree for the volume group. This error also indicates that the `/etc/lvmtab` file, which is used to match physical volumes to a volume group, is out of date with the LVM data structures in memory and on disk.

    Recommended Action:

        Try to locate any missing disks. For each of the disk in the volume group, use `ioscan` and `diskinfo` to confirm that the disk is functioning properly.

lvchange

> **"m": Illegal option.**

    Cause:

        The system does not have HP MirrorDisk/UX installed.

    Recommended Action:

        Install HP MirrorDisk/UX.

lvextend

> **lvextend: Not enough physical extents available.**
> **Logical volume "/dev/*vgname*/lvname" could not be extended.**
> **Failure possibly caused by strict allocation policy**

    Cause:

        There is not enough space in the volume group to extend the logical volume to the requested size. This is typically caused by one of three situations:

        a. There are not enough free physical extents in the volume group. Run `vgdisplay` to confirm the number of available physical extents, and multiply that number by the extent size to determine the free space in the volume group. For example:

```
# vgdisplay vg00
--- Volume groups ---
VG Name                 /dev/vg00
VG Write Access         read/write
VG Status               available
Max LV                  255
Cur LV                  10
Open LV                 10
Max PV                  16
Cur PV                  1
```

```
Act PV                      1
Max PE per PV               4350
VGDA                        2
PE Size (Mbytes)            4
Total PE                    4340
Alloc PE                    3740
Free PE                     600
Total PVG                   0
Total Spare PVs             0
Total Spare PVs in use      0
```
In this example, the total free space is 600 * 4 MB, or 2400 MB.

b. The logical volume is mirrored with a strict allocation policy, and there are not enough extents on a separate disk to comply with the allocation policy. To confirm this, run `lvdisplay` to determine which disks the logical volume occupies, and then check whether there is sufficient space on the other disks in the volume group.

c. In a SAN environment, one of the disks was dynamically increased in size. LVM did not detect the asynchronous change in size.

Recommended Action:

a. Choose a smaller size for the logical volume, or add more disk space to the volume group.

b. Choose a smaller size for the logical volume, or add more disk space to the volume group. Alternatively, free up space on an available disk using `pvmove`.

c. None. LVM does not support dynamic resizing of disks.

"m": Illegal option.

Cause:
   The system does not have HP MirrorDisk/UX installed.
Recommended Action:
   Install HP MirrorDisk/UX.

lvlnboot

lvlnboot: Unable to configure swap logical volume. Swap logical volume size beyond the IODC max address.

Cause:
   The boot disk firmware cannot access the entire range of the swap logical volume. This happens with older host bus adapters when primary swap is configured past 4 GB on the disk.
Recommended Action:
   Upgrade the system firmware or use a newer host bus adapter that supports block addressing. If neither of these actions are successful, reduce the size of the primary swap logical volume so that it does not exceed 4 GB.

pvchange

"a": Illegal option.

Cause:
   LVM OLR is not installed.
Recommended Action:
   Install the patches enabling LVM OLR, or use an alternate replacement procedure.

The HP-UX kernel running on this system does not provide this feature. Install the appropriate kernel patch to enable it.

Cause:

LVM OLR is not completely installed. Both the LVM command and kernel components are required to enable LVM OLR. In this case, the command patch is installed and the kernel patch is not.

Recommended Action:

Install the appropriate kernel patch to enable LVM OLR, or use an alternate replacement procedure.

Unable to detach the path or physical volume via the pathname provided. Either use pvchange(1M) -a N to detach the PV using an attached path or detach each path to the PV individually using pvchange(1M) -a n

Cause:

The specified path is not part of any volume group, because the path has not been successfully attached to the otherwise active volume group it belongs to.

Recommended Action:

Check the specified path name to make sure it is correct. If the error occurred while detaching a physical volume, specify a different path that it was attached to before. If it is not clear whether any path was attached before, individually detach each path to the physical volume using `pvchange` with the `-a n` option.

Warning: Detaching a physical volume reduces the availability of data within the logical volumes residing on that disk. Prior to detaching a physical volume or the last available path to it, verify that there are alternate copies of the data available on other disks in the volume group. If necessary, use pvchange(1M) to reverse this operation.

Cause:

This warning is advisory only and generated whenever a path or physical volume is detached.

Recommended Action:

None.

vgcfgbackup

Invalid LVMREC on Physical Volume.

Cause:

The LVM header on the disk is incorrect. This can happen when an existing LVM disk is overwritten with a command like `dd` or `pvcreate`. If the disk is shared between two systems, it is likely that one of the systems was not aware that the disk was already in a volume group. The corruption can also be caused by running `vgchgid` incorrectly when using BCV split volumes.

Recommended Action:

Restore a known good configuration to the disk using `vgcfgrestore`. Be sure to use a valid copy dated before the first occurrence of the problem.

vgcfgrestore -n *vgname pvname*

vgcfgrestore

**Cannot restore Physical Volume *pvname*
Detach the PV or deactivate the VG, before restoring the PV.**

Cause:

The `vgcfgrestore` command was used to initialize a disk that already belongs to an active volume group.

Recommended Action:

Detach the physical volume or deactivate the volume group before attempting to restore the physical volume. If there is reason to believe that the data on the disk is corrupted, the disk can be detached and marked using `vgcfgrestore` then attached again without replacing the disk. This causes LVM to reinitialize the disk and synchronize any mirrored user data mapped there.

<u>vgchange</u>

Warning: couldn't query physical volume "*pvname*":
The specified path does not correspond to physical volume attached to this
volume group
Warning: couldn't query all of the physical volumes.

Cause:
This error has several possible causes:
a. The disk was missing when the volume group was activated, but was later restored. This typically occurs when a system is rebooted or the volume group is activated with a disk missing, uncabled, or powered down.
b. The disk LVM header was overwritten with the wrong volume group information. If the disk is shared between two systems, it is likely that one of the systems was not aware that the disk was already in a volume group. To confirm, check the volume group information using the `dump_lvmtab` command, available from your HP support representative, and look for inconsistencies. For example:

```
# dump_lvmtab -s | more
SYSTEM : 0x35c8cf58
TIME   : 0x3f9acc69 : Sat Oct 25 15:18:01 2003
FILE   : /etc/lvmtab
HEADER : version:0x03e8   vgnum:7
VG[00]  VGID:35c8cf58 3dd13164 (@0x00040c) pvnum:2 state:0 /dev/vg00
  (00)  VGID:35c8cf58 3dd13164 PVID:35c8cf58 3dd13164 /dev/dsk/c0t6d0
  (01)  VGID:35c8cf58 3dd13164 PVID:35c8cf58 3dda4694 /dev/dsk/c4t6d0
VG[01]  VGID:065f303f 3e63f01a (@0x001032) pvnum:92 state:0 /dev/vg01
  (00)  !VGID:35c8cf58 3f8df316 PVID:065f303f 3e63effa /dev/dsk/c40t0d0
  (01)  !VGID:35c8cf58 3f8df316 PVID:065f303f 3e63effe /dev/dsk/c40t0d4
  (02)  !VGID:35c8cf58 3f8df316 PVID:065f303f 3e63f003 /dev/dsk/c40t1d0
   ...
```

In this example, note that the volume group ids (`VGID`) for the disks in `/dev/vg01` are not consistent; inconsistencies are marked `!VGID`.

Recommended Action:
a. Use `ioscan` and `diskinfo` to confirm that the disk is functioning properly. Re-activate the volume group using the following command:
   # **vgchange −a y *vgname***
b. There are several methods of recovery from this error. If you are not familiar with the commands outlined in the following procedures, contact your HP support representative for assistance.
   1) Restore a known good configuration to the disks using `vgcfgrestore`. Be sure to use a valid copy dated before the first occurrence of the problem.
      # **vgcfgrestore −n *vgname pvname***
   2) Recreate the volume group and its logical volumes, restoring the data from the most current backup.
   3) Export and re-import the volume group using the following commands:
      # **vgexport -m *vgname*.map -v -f *vgname*.file /dev/*vgname***
      # **mkdir /dev/*vgname***
      # **mknod /dev/*vgname*/group c 64 *unique_minor_number***
      # **vgimport -m *vgname*.map -v -f *vgname*.file /dev/*vgname***

vgchange: Couldn't set the unique id for volume group "/dev/*vgname*"

Cause:

There are multiple LVM group files with the same minor number.

Recommended Action:

List the LVM group files. If there are any duplicate minor numbers, export one of the affected volume groups, create a new group file with a unique minor number, and re-import the volume group. If you are not familiar with this procedure, contact your HP support representative for assistance.

```
# ll /dev/*/group
# vgexport -m vgname.map -v -f vgname.file /dev/vgname
# mkdir /dev/vgname
# mknod /dev/vgname/group c 64 unique_minor_number
# vgimport -m vgname.map -v -f vgname.file /dev/vgname
```

vgcreate

vgcreate: Volume group "/dev/*vgname*" could not be created:
VGRA for the disk is too big for the specified parameters. Increase the extent size or decrease max_PVs/max_LVs and try again.

Cause:

The Volume Group Reserved Area at the front of each LVM disk cannot hold all the information about the disks in this volume group. This error typically occurs if you use disks larger than 100 GB.

Recommended Action:

Adjust the volume group creation parameters. Use the -s option of the vgextend command to select an extent size larger than 4 MB, or use the -p option to select a smaller number of physical volumes. Refer to *vgcreate(1M)* for information on these options.

vgdisplay

vgdisplay: Couldn't query volume group "/dev/vg00".
Possible error in the Volume Group minor number;
Please check and make sure the group minor number is unique.
vgdisplay: Cannot display volume group "/dev/vg00".

Cause:

This error has several possible causes:

a. There are multiple LVM group files with the same minor number.
b. Serviceguard was previously installed on the system, and the /dev/slvmvg device file still exists.

Recommended Action:

a. List the LVM group files. If there are any duplicate minor numbers, export one of the affected volume groups, create a new group file with a unique minor number, and re-import the volume group. If you are not familiar with this process, contact your HP support representative for assistance.

```
# ll /dev/*/group
# vgexport -m vgname.map -v -f vgname.file /dev/vgname
# mkdir /dev/vgname
# mknod /dev/vgname/group c 64 unique_minor_number
# vgimport -m vgname.map -v -f vgname.file /dev/vgname
```

b. Remove the /dev/slvmvg device file and recreate the /etc/lvmtab file using the following commands:

```
# rm /dev/slvmvg
# mv /etc/lvmtab /etc/lvmtab.old
# vgscan -v
```

```
Warning: couldn't query physical volume "pvname":
The specified path does not correspond to physical volume attached to this
volume group
Warning: couldn't query all of the physical volumes.
```

Cause:
> This error has several possible causes. They are described earlier under the `vgchange` error messages.

Recommended Action:
> Refer to the recommended actions under the `vgchange` error messages.

`vgextend`

```
vgextend: Not enough physical extents per physical volume.
Need: #, Have: #.
```

Cause:
> The disk size exceeds the volume group maximum disk size. This limitation is defined when the volume group is created, as a product of the extent size specified with the -s option of `vgcreate` and the maximum number of physical extents per disk specified with the -e option. Typically, the disk is successfully added to the volume group, but not all of the disk is accessible.

Recommended Action:
> The volume group extent size and number of physical extents per disk are not dynamic. The only way to use the entire disk is to re-create the volume group with new values for the -s and -e options. Alternatively, you can work with an HP support representative to adjust the volume group characteristics using the `vgmodify` command; note that this utility is currently unsupported and available only from your HP support representative.

`vgimport`

```
Verification of unique LVM disk id on each disk in the volume group
/dev/vgname failed.
```

Cause:
> There are two possible causes for this message:
> a. The `vgimport` command used the -s option, and two or more disks on the system have the same LVM identifier; this can happen when disks are created with BCV copy or cloned with `dd`.
> b. LVM was unable to read the disk header; this can happen when you create new logical units on a SAN array.

Recommended Action:
> a. Do not use the -s option to `vgimport`. Alternatively, use `vgchgid` to change the LVM identifiers on copied or cloned disks.
> b. Retry the `vgimport` command. For a long term solution, install patch PHKL_30510 or one of its superseding patches.

Syslog Error Messages

LVM: VG 64 0x010000: Data in one or more logical volumes on PV 188 0x072000 was lost when the disk was replaced. This occurred because the disk contained the only copy of the data. Prior to using these logical volumes, restore the data from backup.

Cause:

LVM cannot synchronize the data on a replaced disk automatically, as when LVM discovers an unmirrored logical volume residing on a disk that was just replaced. When all data on a disk is mirrored elsewhere and a copy is available, LVM automatically synchronizes the data on the replaced disk from the mirrors of the data on other disks.

Recommended Action:

Restore the contents of the logical volume from backup.

LVM: VG 64 0x010000: PVLink 188 0x072000 Detached.

Cause:

This message is advisory and generated whenever a disk path is detached.

Recommended Action:

None.

For more information

To learn more about LVM and HP-UX system administration, refer to the following documents on the HP documentation website http://docs.hp.com/:

- LVM Online Disk Replacement (LVM OLR)
  http://docs.hp.com/en/7161/LVM_OLR_whitepaper.pdf
- Managing Systems and Workgroups
  http://docs.hp.com/en/B2355-90912/B2355-90912.pdf

Call to action

HP welcomes your input. Please give us comments about this white paper, or suggestions for LVM or related documentation, through our technical documentation feedback website: http://docs.hp.com/en/feedback.html