Dynamic Root Disk and MirrorDisk/UX on HP-UX 11iv2

5991-7582

Introduction

System administrators frequently use MirrorDisk/UX to create a redundant copy of an HP-UX system as protection against hardware failures. The Dynamic Root Disk, on the other hand, provides a means of protecting against software failures: changes can be staged on an inactive copy of the system and an unchanged copy can be retained for recovery.

Since a DRD clone or patched clone will become a booted system, it is desirable to mirror the clone as well. Depending on the number of available disks and policies about use of SAN storage, DRD users can mirror their clone when it is created, mirror it after creation but before boot, or mirror it after the clone is booted.

Some data centers always boot from internal disks on a system; others use SAN disks for mirrors or DRD clones. This document describes the various mirroring strategies a system administrator can use for combining the benefits of DRD clones and LVM mirrors.

The scenarios below describe strategies that a system administrator might use in four-disk, three-disk, and two-disk environments.

Four Disk Scenario – Mirroring the clone when it is created

In this scenario, the four disks are the boot disk, the boot mirror, the clone disk, and the clone mirror.

To apply this scenario, a system administrator needs to have the following resources available:

- *A licensed copy of MirrorDisk/UX:*

One way to determine if you have a licensed copy of MirrorDisk/UX, is to issue the command

```
lvextend -?
```

If the usage statement displayed shows the option

```
-m MirrorCopies
```

you have access to MirrorDisk/UX.

Alternatively, you can check for the LVM.LVM-MIRROR-RUN fileset with the following command:

```
swlist –l fileset LVM.LVM-MIRROR-RUN
```

If you do not have access to MirrorDisk/UX and wish to mirror a DRD clone, you will need to license MirrorDisk/UX. (The initial release of DRD does not support mirroring of VxVM volumes.) This product is available either through several of the Operating Environments (including the Mission Critical Operating Environment) as well as from the bundle B2491BA. To obtain a license and a copy of the software, contact your HP support representative or search for "MirrorDisk" at http://hp.software.com.

- *A copy of Dynamic Root Disk.*

DRD is available at no charge by searching on DRD from HP's software depot: http://hp.software.com.

- *Two disks that are currently in use as the boot disk and the boot mirror*

(If the boot mirror has not already been created, the procedure described below in "How to mirror `vg00`" may be used.).

- *Two disks available that can be used for the DRD clone and the clone mirror.*

Each disk must be large enough to contain all the logical volumes that currently in vg00. The logical volumes in the clone volume group, `drd00`, will be the same sizes as those in `vg00`. In the case of an Integrity system, the disks must also be large enough to hold EFI (Extensible Firmware Interface) and HPSP (HP Service Partition) partitions of the same size as the current root disk.

The disks must be "standalone" disk, as opposed to disk partitions, but they may actually be areas in a Storage Area Network that have been presented to the HP-UX operating system as a single disk.

When the necessary resources have been identified, issue the following command

```
drd clone –t target_disk –x mirror_disk=<mirror_disk>
```

where <`target_disk`> is the block device file of the target disk and `mirror_disk` is the block device file of the mirror.

The DRD clone disk is then mirrored from its creation, through its modification by `drd runcmd` and its activation as a boot disk by `drd activate`. For customers with sufficient internal disks or the ability to use SAN allocations, this is the ideal scenario.

Three Disk Scenario – Mirroring the clone when it is booted

In this scenario, the customer has a mirrored boot disk and wants to use DRD for either offline patching or hot recovery. (See the whitepaper "Using the Dynamic Root Disk Toolset" for more information on using DRD for offline patching and hot recovery.) However, due to the limitation of three available disks, or policies preventing use of SAN storage for boot disks, the clone cannot be mirrored when it is created. In this case, the clone will be mirrored when it is booted. The steps below describe a procedure that can be used after booting the clone.

In this case, the administrator needs to check for the following resources:

- *A licensed copy of MirrorDisk/UX.*

See The Four Disk Scenario for more information on checking whether MirrorDisk/UX is available and acquiring it if it is not currently installed.

- *A copy of Dynamic Root Disk.*

DRD is available at no charge by searching on DRD from HP's software depot: http://hp.software.com.

- *Two disks that are currently in use as the boot disk and the boot mirror*

(If the boot mirror has not already been created, the procedure described below in "How to mirror vg00" may be used.).

- *One disk available that can be used for the DRD clone.*

The disk must be large enough to contain all the logical volumes that currently in vg00. The logical volumes in the clone volume group, `drd00`, will be the same sizes as those in vg00. In the case of an Integrity system, the disk must also be large enough to hold EFI (Extensible Firmware Interface) and HPSP (HP Service Partition) partitions of the same size as the current root disk.

The disk that is currently the boot mirror will be "broken off" when the clone is booted and used to mirror the clone.

When the necessary resources have been identified, issue the following command

```
drd clone –t <target_disk>
```

where <`target_disk`> is the block device file of the clone disk.

The clone will not be mirrored during the time it is modified with DRD runcmd.

When the clone is booted, the administrator probably wants to "move" the mirror from the previously booted image to the clone.

Splitting a mirror from an inactive image to provide a free disk

1. Mount the inactive (original) system image so that it is available for LVM commands.

```
drd mount
```

2. Run `vgdisplay` to determine the mirrored pair in use for `drd00`:

```
vgdisplay –v drd00 | grep "PV Name"
```

3. Tell DRD that the mirror disk is to be removed by running `drd_unregister_mirror`:

```
drd_unregister_mirror _mirror_disk
```

where `mirror_disk` is the mirror to be split off from the inactive image.

The `drd_unregister_mirror` command will find that the mirror is associated with the inactive image, and remove its record of the disk.

The un-registration is needed since an administrator may want to use `drd mount` to browse the original disk. The mount of the original system image requires information about the exact list of disks on which it resides.

Caution: ***The administrator will have the choice of using "breaking off" either disk in the mirrored pair. However, DRD is only aware of a single boot disk. This disk should not be removed! Doing so will prevent correct*** `drd activate` ***from working correctly To prevent removal of DRD's boot disk, the*** `drd_unregister_mirror` ***command will fail if the boot disk is used as an argument. If this occurs, un-register – and subsequently remove - the other disk of the mirrored pair.***

4. Prepare the inactive image to boot without its mirror.

The LVM infrastructure on the inactive system image itself will not be affected by the mirroring commands on the booted system: It will not be aware that half of its mirror pair has been broken off. However, in the event of a problem, it is desirable to be able to boot the inactive image. To facilitate this, the AUTO file on the inactive image should include the "-lq" option, which overrides the need for a quorum of the volume group to boot.

On a PA system, to change the AUTO file on the inactive system image (which is still mounted from step 1 above) issue the following commands:

```
lifrm <inactive_image_disk>:AUTO
```

```
echo “hpux –lq” > /tmp/auto
lifcp /tmp/auto <inactive_image_disk>:AUTO
```

where <inactive_image_disk> is the block device file of the (remaining) disk on the inactive system image (i.e. the disk for drd00).

On an Integrity system, to change the AUTO file on the inactive system image (which is still mounted from step 1 above) issue the following commands:

```
echo “boot vmunix –lq” > /tmp/auto
efi_cp –d <efi_part> /tmp/auto /EFI/HPUX/AUTO
```

where <efi_part> is the block device file of the EFI partition on the (remaining) disk on the inactive system image. The EFI partition is the first partition, and its device file is formed by appending “s1” to the device file of the entire disk.

5. Use lvremove to remove the mirror disk from the inactive system image:

Set the mirror environment variable to the mirror disk for PA-RISC systems, or to the second partition of the mirror disk for Integrity systems:

```
export mirror=/dev/dsk/c99t99d0  # example for PA-RISC
```
or
```
export mirror=/dev/dsk/c99t99d0s2  # example for Integrity

# for logical_volume in `vgdisplay -v drd00| grep "LV Name"
| awk '{print $3}'`
> do
> lvreduce -m 0 $logical_volume $mirror
> done
vgreduce  /dev/drd00 $mirror
pvremove  <character_device_of_mirror>
```

where <character_device_of_mirror> is the character (raw) device file of the mirror disk, or of the second partition of the mirror disk for integrity systems.

You are now ready to mirror the booted system with LVM commands:

Since there are a number of differences in this procedure for PA-RISC and Integrity systems, the description for PA-RISC is give in its entirety, followed by the description for Integrity systems. For more information on mirroring, you may also want to consult Chapter 6 of Managing Systems and Workgroups: A Guide for HP-UX System Administrators

How to Mirror vg00 – HP 9000 system

1. Identify a disk that is not currently in use to be used as the mirror. This disk should be sufficiently large to hold all the logical volumes in the clone. It should not currently be used as a mirror of the original system image, or for any other purpose.

2. Run the `pvcreate` command on the *character* device file for the disk that will be the mirror, using the "-B" option to make the disk bootable, and the "-f" option to overwrite any pre-existing data on the disk.

   ```
   # pvcreate –f -B /dev/rdsk/<mirror_disk>
   ```

3. Extend the root volume group `vg00` to include the *block* device file of the mirror disk.

   ```
   # vgextend /dev/vg00 /dev/dsk/<mirror_disk>
   ```

4. Run the `mkboot` command to create the boot areas on the disk. (Either the character or block device may be used, but both must exist.)

   ```
   # mkboot /dev/rdsk/<mirror_disk>
   ```

5. Run the mkboot command again to update the boot string in the AUTO file. The "–lq" option ensures that the system will boot without quorum, In the event of the primary boot disk failing this will allow the mirror disk to boot.

   ```
   # mkboot -a "hpux -lq" /dev/rdsk/<mirror_disk>
   ```

6. Starting with the boot, swap, and root logical volumes, run lvextend to mirror each logical volume:

   ```
   # lvextend -m 1 /dev/vg00/lvol1 /dev/dsk/<mirror_disk>
   # lvextend -m 1 /dev/vg00/lvol2 /dev/dsk/<mirror_disk>
   # lvextend -m 1 /dev/vg00/lvol3 /dev/dsk/<mirror_disk>
   # lvextend -m 1 /dev/vg00/lvol4 /dev/dsk/<mirror_disk>
   # lvextend -m 1 /dev/vg00/lvol5 /dev/dsk/<mirror_disk>
   ...
   ```

7. Run the following lvlnboot commands to reestablish LVM's knowledge of the root and boot file systems and primary swap and dump.

   ```
   # lvlnboot -b /dev/vg00/lvol1
   # lvlnboot -r /dev/vg00/lvol3
   # lvlnboot -s /dev/vg00/lvol2
   # lvlnboot -d /dev/vg00/lvol2
   # lvlnboot -R
   ```

8. Add the *block* device file of the mirror to the file `/stand/bootconf. The lower case "l" in the following command indicates that the disk is managed by LVM or VxVM.`

   echo "l /dev/dsk/mirror_disk" >> /stand/bootconf

9. If desired, set the alternate boot path to the mirror_disk, using the hardware address determined in step 5.

```
# setboot -a  <hardware path of mirror disk>
```

How to Mirror vg00 – HP Integrity system

1. Use the following commands to determine the size in megabytes of the (1) EFI, (2) HP-UX, and (3) HPSP partitions on the current boot disk.

```
# diskinfo -b /dev/rdsk/<boot disk>s1 | awk '{print $1 / (1024)}'
# diskinfo -b /dev/rdsk/<boot disk>s2 | awk '{print $1 / (1024)}'
# diskinfo -b /dev/rdsk/<boot disk>s3 | awk '{print $1 / (1024)}'
```

2. Identify a disk that is not currently in use to be used as the mirror. This EFI partition should be as large as the EFI partition on the current disk. The HP-UX partition should be large enough to hold all the logical volumes in the clone. Although the HPSP partition is optional, if it is created, it should be the same size as the HPSP on the clone.

3. Create a temporary file, say /tmp/partitionfile, containing the number of partitions and the size of each partition to be created on the mirror disk. Use the information from in steps 1 and 2 to determine partition sizes. Here is a sample command to create the file, indicating three partitions with an EFI partition size of 500MB, an HPSP partition size of 400BM, and the remaining space allocated to the HPUX partition:

```
# print '3\nEFI 500MB\nHPUX 100%\nHPSP 400MB' > /tmp/partitionfile
```

   See idisk(1M) for more information about partitioning disks.

4. Use the file created in the previous step to partition the mirror disk with idisk. (The echo will reply to the confirmation request from the idisk command.)

```
# echo yes | idisk -wf /tmp/partitionfile /dev/rdsk/<second disk>

The temporary file may then be removed if desired.
```

5. Use ioscan –fnkC disk determine the hardware path to the mirror disk.

6. Use the insf command to create device files for all the newly-created partitions:

```
 # insf –e –H <hardware path to second disk>
```

7. Run the `pvcreate` command on the *character* device file for the HPUX (second) partition of the mirror disk, using the "-B" option to make the disk bootable, and the "-f" option to overwrite any pre-existing data on the disk.

```
# pvcreate –f -B /dev/rdsk/<mirror_disk>s2
```

8. Extend the root volume group to the HPUX partition of the mirror disk, using the *block* device file of the *second* (HPUX) partition in the following command:

```
# vgextend /dev/vg00 /dev/dsk/<mirror_disk>s2
```

9. Run `mkboot` to complete the setup of the disk as a boot disk. The `-e` option tells `mkboot` to use EFI layout and the `-l` option tells `mkboot` that this volume will be used by a volume manager (even if it is not currently used by one). Note that no partition number is given to this command; it looks at the whole disk.

```
# mkboot -e -l /dev/dsk/<mirror_disk>
```

10. Create a temporary AUTO file and use the efi_cp command to copy it to the mirror, using the *block* device of the *first* (EFI) parititon. The "–lq" option ensures that the system will boot without quorum, In the event of the primary boot disk failing this will allow the mirror disk to boot.

```
print 'boot vmunix -lq' > /tmp/AUTO
efi_cp -d /dev/dsk/<mirror_disk>s1 /tmp/AUTO EFI/HPUX/AUTO
rm /tmp/AUTO
```

11. Starting with the boot, swap, and root logical volumes, run lvextend to mirror each logical volume:

```
# lvextend -m 1 /dev/vg00/lvol1 /dev/dsk/<mirror_disk>
# lvextend -m 1 /dev/vg00/lvol2 /dev/dsk/<mirror_disk>
# lvextend -m 1 /dev/vg00/lvol3 /dev/dsk/<mirror_disk>
# lvextend -m 1 /dev/vg00/lvol4 /dev/dsk/<mirror_disk>
# lvextend -m 1 /dev/vg00/lvol5 /dev/dsk/<mirror_disk>
...
```

12. Run the following lvlnboot commands to reestablish LVM's knowledge of the root and boot file systems and primary swap and dump.

```
# lvlnboot -b /dev/vg00/lvol1
# lvlnboot -r /dev/vg00/lvol3
# lvlnboot -s /dev/vg00/lvol2
# lvlnboot -d /dev/vg00/lvol2
# lvlnboot –R
```

13. Add the *block* device file of the HPUX (*second)* partition of the mirror to the file `/stand/bootconf.` The lower case "l" in the following command indicates that the disk is managed by LVM or VxVM.

    echo "l /dev/dsk/mirror_disk"s2 >> /stand/bootconf

14. If desired, set the alternate boot path to the mirror_disk, using the hardware address determined in step 5.

```
# setboot -a  <hardware path of mirror disk>
```

15. If the mirror disk includes an HPSP partition, use the efi_fsinit command on the *character* device file for the HPSP (third) partition to initialize it with an EFI file system:

    ```
    # efi_fsinit -d /dev/rdsk/<mirror_disk>s3
    ```

16. The initial release of DRD does not copy the contents of the HPSP to the clone. There are two methods you can use to populate the service partition on both the clone and the mirror. One alternative is to use the "HP Itanium Processor Family Diagnostics and Utilities CD" to install the diagnostics to one or more of the disks. If the diagnostics have already been installed to the original system disk, the dd command can be used to copy them to both the clone and its mirror:

    To see if the diagnostics are installed on the original boot disk, use the command

    ```
    # efi_ls –d /dev/rdsk/<disk>s3
    ```

    If present, copy them with the following dd commands:

    ```
    # dd if=/dev/rdsk/<disk>s3 of=/dev/rdsk/<clone_disk>s3 bs=1024k

    # dd if=/dev/rdsk/<disk>s3 of=/dev/rdsk/<mirror_disk>s3 bs=1024k
    ```

Notifying DRD that the booted system has been mirrored

After mirroring `vg00`, DRD should be notified that the booted system has been mirrored. In this way, if the inactive system image is later booted, the current system can then be mounted using `drd mount`.

To notify DRD that the booted system has been mirrored, issue the following command:

```
drd_register_mirror <mirror_disk>
```

Updating the inactive image with the changes

With the changes described above, the current HP-UX system and the drd commands will work correctly on the booted system. However, it may be desirable in the future to boot back to the inactive system image. Two changes must be made to the inactive system image to prepare for the possibility of a boot: The DRD infrastructure must be updated, and the LVM infrastructure must be updated.

In the future, the update of the DRD infrastructure on the inactive image will be incorporated into drd commands. For the initial release, the registry file itself must be copied from the booted system to the inactive system image:

```
drd mount
cp /var/opt/drd/registry/registry.xml
/var/opt/drd/mnts/sysimage_00*/var/opt/drd/registry/registry.xml
drd umount
```

The LVM infrastructure on the inactive image cannot be updated by LVM commands run on the booted system; they will only change the LVM control information on the booted system itself. Instead, the commands must be run when the system is booted.

A way to automate this process is to create an rc script that will rename /etc/lvmtab and re-create it with "`vgscan -v`" when the inactive system image is booted, then remove itself. Here is a sample creation of such a script:

```
initfile=/sbin/init.d/lvm_update
rcfile=/sbin/rc2.d/S900lvm_update
rm -rf $rcfile $initfile
print "#!/sbin/sh" > $initfile
print "    mv –f /etc/lvmtab /etc/lvmtab.old >> $initfile
print «    vgscan –v « >> $initfile
print ”    rc = \$?
print "    /usr/bin/rm -rf $initfile $rcfile " >> $initfile
print "    exit \$rc " >> $initfile
chmod 555 $initfile
chown bin:bin $initfile
ln -s $initfile $rcfile
```

Automating the Three Disk Scenario

It is desirable to minimize the time spent with a system disk un-mirrored. On the other hand, systems that are limited to three internal disks and policies that limit use of networked storage for boot disks will need to keep the "third" disk for the mirror of the active system image. The steps above can be performed by a startup (rc) script that is copied to the clone by the administrator, runs at boot of the clone, and then removes or renames itself. (Note that this startup script is distinct from the one described in the section above. This script is created on the clone before booting. The script described in the section above is created on the original system in case there is a need to boot back to it.)

Two disk scenario – Splitting a mirror to provide a clone disk

It is not recommended that an LVM mirror be broken off to provide a DRD clone, since this leaves the operating system un-mirrored. Further, if a "broken mirror" disk is used for a DRD clone, the clone must still be created using the usual "`drd clone`" command. However, if a system administrator wants to pursue this approach (probably on a test system), the following steps can be used:

1. Break off the mirror from the currently booted disk, using the steps from the section "Splitting a mirror from an inactive image to provide a free disk "above. There is no need to notify DRD about the broken mirror in this case, since it is not yet aware of the system configuration.

2. Clone the booted system to the newly available mirror disk. If desired, use `drd runcmd` to patch the clone.
3. Boot the clone.
4. Mirror the newly booted system, using the steps described above in "How to mirror vg00" for either PA-RISC or Integrity systems. (Although DRD may be notified of the new mirror by calling `drd_register_mirror` <mirror_disk>, it is not required. There is no other disk to boot, so `drd activate` and subsequent `drd mount` commands need not be considered.)

In this case, no DRD commands are useful after the scenario is completed, since the clone has been overwritten by the mirror. However, the clone procedure may be repeated when new patches need to be installed.

Glossary

EFI Extensible Firmware Interface is the firmware interface for Itanium®-based systems. Also the name of the first partition on a HP-UX boot disk.

HPSP HP Service Partition, holds the offline diagnostics on HP Integrity systems.

LIF Logical Interchange Format.