Mirroring the rootvg Volume Group for AIX 4.1/4.2

Special Notices

Information in this document is correct to the best of our knowledge at the time of this writing. Please send feedback by fax to "AIXServ Information" at (512) 823-4009.

Please use this information with care. IBM will not be responsible for damages of any kind resulting from its use. The use of this information is the sole responsibility of the customer and depends on the customer's ability to evaluate and integrate this information into the customer's operational environment.

About This Document

This document is to specify the supported method for mirroring of the rootvg volume group for high availability access of the AIX operating system. Because the rootvg is special in that it contains the AIX operating system, several restrictions exist for this configuration that do not exist for the non-rootvg volume group mirroring.

This documentation supersedes previous IBM documentation and instructions on mirroring the rootvg volume group. the latest update refer to the document date on the second line to determine which is the latest set of instructions for this subject.

Function:

1. Provides continuous operation of the AIX operating system in the event that a disk that is part of the operating system experiences failure and there exists an active mirrored copy of the operating system on some other disk.
2. Provides the ability to boot more than one disk of the rootvg in the event that another boot disk has failed. In some cases, the ability to boot from alternate disks may require some user intervention.

Restrictions:

1. This is only valid for AIX 4.1 and later releases.
2. Users should contact their local IBM Branch office to determine if this function is supported on versions of AIX older than AIX 4.1.
3. This function is not supported on /usr client, diskless client, or dataless client systems.

4. APAR ix56564 must be applied on the system for mksysb to correctly restore a mirrored volume group back into a system in mirrored format. If this apar fix is not applied, then the mksysb will still restore the rootvg onto a system, however the base mirrors of the logical volumes will not be automatically created and must be recreated by the user. To determine if this fix is installed on a machine, type:

   ```
   instfix -i -k IX56564
   ```

5. In some instances, the user will not be able to capture the dump image from a crash or the dump image may be corrupted. The design of LVM prevents mirrored writes of the dump device. Depending on the boot sequence and disk availability after a crash, the dump may be in three states:

   a. not available

   b. available and not corrupted

   c. available and corrupted

   State (1) will always be a possibility. If the user prefers to prevent the risk of encountering State (3), then they can create a non-mirrored logical volume (that is not hd6) and set the dump device to that non-mirrored logical volume.

6. Parallel mirroring policy (the default) must be the policy used on the base logical volumes that constitute AIX (hd1, hd2, ...hd9var). This restriction is for performance reasons only.

7. Mirror strictness (the default) must be maintained on all the base logical volumes that constitute AIX (hd1, hd2, ...hd9var). If two copies of a logical volume reside on the same disk, this would defeat and invalidate the original goal of mirroring a rootvg volume group against possible disk failure.

8. The base logical volumes that constitute AIX:

   ```
   hd1, hd2, hd3, hd4, hd5, hd6, hd8, and hd9var
   ```

   must all exist on the same disk. Their equivalent mirrors must also reside on a common disk (another disk). These logical volumes and their mirrors cannot span multiple disks.

9. hd6 must exist in rootvg and must be an active paging device. Other paging devices on other volume groups are permitted, but hd6 (and it's mirror) must exist on the disks that make up rootvg.

10. The disks that constitute the boot drives for AIX, must be disks that are supported to boot on RS/6000 platforms. A machine and disk configuration may be queried to determine if that combination of machine and disk supports booting from disk:

    ```
    bootinfo -B hdiskX
    ```

    If the command returns a "1" then it will be bootable by AIX. Any other value indicates that this disk is not a candidate for rootvg mirroring.

11. In case of disk failure and then replacement, the user should consult the AIX System Management Guide appropriate to their level of AIX.

    When replacing a mirrored volume group, the steps are identical (with the exception of Restriction [13]) for rootvg and non-rootvg volume groups:

    a. Mirror copies referenced on the bad disk are removed.

    b. The bad disk reference is removed (via reducevg) from the volume group.

    c. Follow the procedures to re-mirror a volume group.

12. Physical and design limitations (see Procedure 1) may exist that prevent the mirroring of the rootvg.

13. When the user runs the commands:

    ```
    migratepv, reorgvg, or reducevg
    ```

    And the logical volume that constitutes the dump device (typically hd6) OR its mirror resides on the disk, the user must first set the dump device to /dev/sysdumpnull before executing the command.

    After the command is complete, the user may reset the dump device to the original dump logical volume.

14. Some models of RS/6000 do not support the bootlist command (see the section titled "Procedure). The rootvg can still be mirrored on these systems but the possible reboot following a failure of the original boot disk will require user intervention to select alternate boot disks for subsequent boots.

15. The syncvg command is not available to rootvg at the time of the varyon of the rootvg volume group. Thus if after a mirrored rootvg has booted and there exists stale logical volumes in rootvg, then the user must run syncvg on rootvg to synchronize any stale logical volumes.

16. Apar IX58121 is required to be applied on UP and SMP systems before mirroring of rootvg can be initiated. Apar IX58267 is only required on SMP systems, for performance reasons. To determine if either fix is installed on a machine, type:

    ```
    instfix -i -k <apar number>
    ```

Procedure:

The following steps assume the user has rootvg contained on hdisk0 and is attempting to mirror the rootvg to a new disk hdisk1.

1. Extend rootvg to hdisk1:

   ```
   extendvg rootvg hdisk1
   ```

   If the user encounters the error message:

```
0516-050 Not enough descriptor space left in this volume group.
Either try adding a smaller PV or use another volume group.
```

Then the user may not add hdisk1 to rootvg for mirroring. The user may attempt to mirror rootvg's logical volumes to another disk that already exists in rootvg and meets the criteria and restrictions listed in the "restrictions" section above. The user may also attempt to add a smaller disk to the rootvg. If neither option is possible, then mirroring rootvg cannot be performed on this system.

2. Disable QUORUM, by running the following:

```
chvg -Qn rootvg
```

3. Mirror the logical volumes that make up the AIX operating system:

```
mklvcopy hd1 2 hdisk1            #  /home file system
mklvcopy hd2 2 hdisk1            #  /usr file system
mklvcopy hd3 2 hdisk1            #  /tmp file system
mklvcopy hd4 2 hdisk1            #  / (root) file system
mklvcopy hd5 2 hdisk1            #  blv, boot logical volume

mklvcopy hd6 2 hdisk1            #  paging space
(if the user has other paging devices, rootvg and non-rootvg, they

must also mirror those logical volumes in addition to hd6)
mklvcopy hd8 2 hdisk1            #  file system log
mklvcopy hd9var 2 hdisk1         #  /var file system
```

If hd5 consists of more than one logical partition, then after mirroring hd5 the user must verify that the mirrored copy of hd5 resides on contiguous physical partitions. This can be verified with the command:

```
lslv -m hd5
```

If the mirrored hd5 partitions are not contiguous, then the user must delete the mirror copy of hd5 (on hdisk1) and rerun the mklvcopy for hd5 using the "-m" option. The user should consult documentation on the usage of the "-m" option for mklvcopy.

4. Synchronize the newly created mirrors:

```
syncvg -v rootvg
```

5. Bosboot to initialize all boot records and devices:

```
bosboot -a
```

6. Initialize the boot list:

```
bootlist -m normal hdisk0 hdisk1
```

Warning: Even though this command identifies the list of possible boot disks it does not guarantee that the system will boot from the alternate disk in all cases involving failures of the first disk. In such situations, it may be necessary for the user to boot from the installation/maintenance media, select maintenance, reissue the bootlist command leaving out the failing

disk, and then reboot. On some models, firmware provides a utility for selecting the boot device at boot time. This may also be used to force the system to boot from the alternate disk.

7. Shutdown and reboot the system:

```
shutdown -Fr
```

This is so that the "Quorum OFF" functionality takes effect.

Description	UP & MP	MP only	Apar Level
mksysb fix	ix56564		4.1
	------	-------	
	In 4.2.0 code		4.2
mirror poweroff crash	ix58121		4.1
	------	-------	
	ix59517		4.2
mirror read algorithm		ix58267	4.1
	------	-------	
		ix59519	4.2
mirror sync hang		ix58889	4.1
	------	-------	
		In 4.2.0 code	4.2
mirror sync crash		ix58919	4.1
	------	-------	
		ix59520	4.2
mklvcopy failure	not needed		4.1
	------	-------	
	ix59414		4.2

The fixes listed above are packaged for the clvm prpq under the cumulative apar IX58707.

10/04/96, 4fax# 4254

Reader's Comments

Please fax this form to (512) 823-4009, attention "AIXServ Information". You may also e-mail comments to: elizabet@austin.ibm.com. These comments should include the same customer information requested below.

Use this form to tell us what you think about this document. If you have found errors in it, or if you want to express your opinion about it (such as organization, subject matter, appearance) or make suggestions for improvement, this is the form to use.

If you need technical assistance, contact your local branch office, point of sale, or 1-800-CALL-AIX (for information about support offerings). These services may be billable. Faxes on a variety of subjects may be ordered free of charge from 1-800-IBM-4FAX. Outside the U.S. call 415-855-4329 using a fax machine phone.

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate without incurring any obligation to you.

Note: If you have a problem report or item number, supplying that number may help us determine why a procedure did or did not work in your specific situation.

Problem Report or Item #: Branch Office or Customer #:

Be sure to print your name and fax number below if you would like a reply:

Name: Fax Number:

END OF DOCUMENT (mirror.rvg.41.42.krn, 4FAX 4254)