

Tru64 UNIX Best Practice

Using LSM to Create Mirrored Volumes for Boot Disk Partitions

September 2001

This Best Practice describes how to use the Logical Storage Manager (LSM) to encapsulate and create mirrored volumes for each boot disk partition on a system running the Tru64 UNIX Operating System software Version 5.0 or higher.

This procedure applies only to single systems that are not part of a TruCluster cluster.

Contents

Using LSM to Create Mirrored Volumes for Boot Disk Partitions

Is This Best Practice Right for You?	1
Before You Begin	3
Applying the Best Practice	3
Step 1: Encapsulating the Boot Disk	3
Step 2: Creating Mirrors for the Boot Disk Partition and Swap Volumes	6
Verifying Success	6
Troubleshooting	8
Unencapsulating the Boot Disk	8
Manually Resetting Changes Made During the Encapsulation Procedure	10
Alternative Practices	12
Comments and Questions	13
Legal Notice	13

Using LSM to Create Mirrored Volumes for Boot Disk Partitions

Creating LSM mirrored volumes for the encapsulated boot disk partitions copies the data in the partitions onto another disk, which provides complete redundancy and recovery capability if the boot disk fails. For example, if the primary boot disk fails, the system continues running off the surviving mirror. You can also reboot the system using the surviving mirror.

Not all Best Practices apply to all configurations, so you must be sure that this Best Practice is appropriate for your system and circumstances. See *Is This Best Practice Right for You?* for more information.

See the Tru64 UNIX Best Practices Web page for more information about Best Practices documentation:

http://www.tru64unix.compaq.com/docs/best_practices/index.html

Is This Best Practice Right for You?

To use this Best Practice, you must meet the requirements described in the following table.

Requirement	Description
Operating System	Tru64 UNIX Version 5.0 or higher
System Configuration	Single systems that are not part of a TruCluster cluster
Restrictions	You must encapsulate the root file system and the primary swap partition at the same time. They do not have to be on the same disk.

Requirement	Description
Disk Space	<p>Original Boot Disk — The partition table for the boot disk (and primary swap disk, if different) must have at least one unused partition for the LSM private region, which cannot be the a or c partition. The unused partition does not have to have any space in it; if necessary, LSM takes the required space (4096 blocks by default) from the swap space and relabels the disk partitions accordingly. If there is no space (or not enough) on any unused partition for LSM to use, the encapsulation fails.</p> <p>Mirror Disk — If the primary swap space is on the boot disk, you need one separate disk to create the mirrors for the boot partition and swap space volumes. This disk must be as large as the total of the root file system and swap partitions on the primary boot disk, plus the size of the private region, which by default is 4096 sectors.</p> <p>If the primary swap space is on a separate disk, you need two separate disks to create the mirrors for the boot partition and swap space volumes.</p> <ul style="list-style-type: none"> • The disk for the boot partition volume mirror must be as large as the total of the root file system and the private region, which by default is 4096 sectors. • The disk for the swap volume mirror must be as large as the total of the swap partition and the private region, which by default is 4096 sectors. <p>In addition, the disk or disks you use to create the LSM volume mirrors must not be under LSM control and must have a disk label with all partitions marked unused. See the <code>disklabel(8)</code> reference page for your version of the operating system for more information.</p>
Impact on Availability	<p>This process takes approximately five minutes to perform and automatically reboots the system. The system will be down for three to five minutes, depending on the number of CPUs. You can perform this procedure at any time, but for the convenience of users, you can schedule this as part of your regular maintenance processes.</p>

If you do not meet the previous requirements, see *Alternative Practices* for information.

Before You Begin

Before you apply this Best Practice, ensure there are no existing LSM volumes for the boot disk partitions on the system. This procedure creates for each boot disk partition an LSM volume with the default properties, which are the best for most configurations. Specifically:

- The following LSM volume names are automatically assigned:
 - `rootvol` is assigned to the root file system.
 - `swapvol` is assigned to the swap space.
 - All other boot disk partitions are assigned an LSM volume name that matches the original partition name with the prefix `vol-`; for example, `vol-dsk0g`.

Do not change these names.

- The boot disk (and swap disk, if applicable) and the disk or disks you specify for the mirrors are initialized for use by LSM and are placed in the `rootdg` disk group.

Applying the Best Practice

Before you proceed, be sure to follow the recommendations in *Before You Begin*.

To create mirrored volumes for the boot disk partitions:

1. Encapsulate the boot disk and swap space, which creates the LSM command scripts that convert all the boot disk partitions and swap space to LSM volumes.
2. Create mirrors for the boot disk partition and swap space volumes.

This procedure uses command-line commands.

Step 1: Encapsulating the Boot Disk

The steps to encapsulate the boot disk are the same whether you are using AdvFS or UFS.

The encapsulation process changes the following files:

- For AdvFS, the `/etc/fdmns/*` directory is changed to use LSM volumes instead of disk partitions for domains associated with the root disk.

- For UFS, the `/etc/fstab` file is changed to use LSM volumes instead of disk partitions.
- For swap space, the `/etc/sysconfigtab` file is changed to update the `swapdevice` entry to use LSM volumes and to set the `lsm_rootdev_is_volume` entry to 1.

To encapsulate the boot disk and swap space:

1. Log in as root.
2. Determine the identities of your boot disk and primary (default) swap disk:

- To identify the boot disk, enter:

```
# consvar -l
```

Output similar to the following is displayed:

```
auto_action = HALT
boot_dev = dsk0
bootdef_dev = dsk0
booted_dev = dsk0
:
```

- To identify the swap disk, enter:

```
# swapon -s
```

Output similar to the following is displayed:

```
Swap partition /dev/disk/dsk0b (default swap):
  Allocated space:      25600 pages (200MB)
  In-use space:         1018 pages ( 3%)
  Free space:           24582 pages ( 96%)
```

3. Enter the `disklabel` command to verify that there is at least one unused partition on the boot disk other than a or c:

```
# disklabel dsk0
```

Output similar to the following is displayed:

```
:
:
8 partitions:
#          size      offset  fstype  [fsize bsize  cpg]  # NOTE: values not exact
a:    262144         0   AdvFS   # (Cyl.   0 - 115*)
b:    262144    262144    swap   # (Cyl. 115*- 231*)
c:    8380080         0  unused    0    0   # (Cyl.   0 - 3707)
d:         4096    8375984  unused    0    0   # (Cyl. 3706*- 3707)
e:    2618597    3142885  unused    0    0   # (Cyl. 1390*- 2549*)
f:    2614502    5761482  unused    0    0   # (Cyl. 2549*- 3706*)
g:    1433600    524288   AdvFS   # (Cyl.  231*- 866*)
h:    6418096    1957888  unused    0    0   # (Cyl.  866*- 3706*)
```

4. If the swap partition is on a separate disk, repeat step 3, specifying the swap disk name.
5. Enter the `volencap` command, specifying the boot disk and the swap disk, if applicable; for example:

```
# volencap dsk0 dsk4
```

Output similar to the following is displayed:

```
Setting up encapsulation for dsk0.
```

- Creating simple disk dsk0d for config area (privlen=4096).
- Creating nopriv disk dsk0a for rootvol.
- Creating nopriv disk dsk0b.
- Creating nopriv disk dsk0g.

```
Setting up encapsulation for dsk4.
```

- Creating simple disk dsk4h for config area (privlen=4096).
- Creating nopriv disk dsk4b for swapvol.

```
The following disks are queued up for encapsulation or use by LSM:  
dsk0d dsk0a dsk0b dsk0g dsk4h dsk4b
```

You must now run `/sbin/volreconfig` to perform actual encapsulations.

6. If appropriate, send a warning to your user community to alert them of the impending shutdown of the system.

When there are no users on the system, proceed with step 7.

7. Enter the `volreconfig` command to complete the encapsulation process. Enter **now** when prompted to shut down the system:

```
# volreconfig
```

Output similar to the following is displayed:

```
The system will need to be rebooted in order to continue with  
LSM volume encapsulation of:  
dsk0d dsk0a dsk0b dsk0g  
dsk4h dsk4b
```

```
Would you like to either quit and defer encapsulation until later  
or commence system shutdown now? Enter either 'quit' or time to be  
used with the shutdown(8) command (e.g., quit,  
now, 1, 5): [quit] now
```

The system shuts down, performs the encapsulation, and automatically reboots.

Step 2: Creating Mirrors for the Boot Disk Partition and Swap Volumes

After you encapsulate the boot disk partitions and swap space into LSM volumes, mirror the volumes to provide redundancy. This process might take a few minutes, but the root file system and swap space are available during this time.

- If the swap space is on the boot disk, enter:

```
# volrootmir -a target_disk
```

- If the swap space is on a separate disk, enter:

```
# volrootmir -a swap=swap_target_disk boot_target_disk
```

See the `volrootmir(8)` reference page for your version of the operating system for more information.

Verifying Success

If you followed the steps in this Best Practice and did not receive error messages, you can assume that the Best Practice was successful. However, you can also display the results using the following commands:

- To display simplified volume information, enter:

```
# volprint -pt
```

Output similar to the following is displayed:

PL NAME	VOLUME	KSTATE	STATE	LENGTH	LAYOUT	NCOL/WID	MODE
pl rootvol-01	rootvol	ENABLED	ACTIVE	262144	CONCAT	-	RW
pl rootvol-02	rootvol	ENABLED	ACTIVE	262144	CONCAT	-	RW
pl swapvol-01	swapvol	ENABLED	ACTIVE	333824	CONCAT	-	RW
pl swapvol-02	swapvol	ENABLED	ACTIVE	333824	CONCAT	-	RW
pl vol-dsk0g-01	vol-dsk0g	ENABLED	ACTIVE	1450796	CONCAT	-	RW
pl vol-dsk0g-02	vol-dsk0g	ENABLED	ACTIVE	1450796	CONCAT	-	RW

In this example, there are three volumes, called `rootvol`, `swapvol`, and `vol-dsk0g`, which contains the `/usr` and `/var` file systems. Each volume contains two plexes (listed in the column labeled `PL NAME`), indicating that the volumes were successfully mirrored.

- Confirm that the system recognizes both boot devices:

```
# consvar -l
```

Output similar to the following is displayed:

```
auto_action = HALT
boot_dev = dsk0,dsk4
bootdef_dev = dsk0,dsk4
```

```
booted_dev = dsk0
:
```

- If you are using AdvFS, confirm that the domains point to the LSM volumes instead of the original partitions:

```
# ls -R /etc/fdmns/*
```

Information similar to the following is displayed, showing the volume name (*rootdg.volume*) instead of the original disk partition:

```
/etc/fdmns/root_domain:
rootdg.rootvol
```

```
/etc/fdmns/usr_domain:
rootdg.vol-dsk0g
```

- If you are using UFS, enter:

```
# mount
```

Information similar to the following is displayed:

```
/dev/vol/rootdg/rootvol on / type ufs (rw)
/dev/vol/rootdg/vol-dsk0g on /usr type ufs (rw)
/proc on /proc type procfs (rw)
```

- To display the swap volume information, enter:

```
# swapon -s
```

Information similar to the following is displayed, showing the volume name (*/dev/vol/rootdg/volume_name*) instead of the original disk partition:

```
Swap partition /dev/vol/rootdg/swapvol (default swap):
  Allocated space:      25600 pages (200MB)
  In-use space:         426 pages ( 1%)
  Free space:           25174 pages ( 98%)
```

```
Total swap allocation:
  Allocated space:      25600 pages (200.00MB)
  Reserved space:       9015 pages ( 35%)
  In-use space:         426 pages ( 1%)
  Available space:      16585 pages ( 64%)
```

If the Best Practice was not successful, see *Troubleshooting* for information about identifying and solving problems.

Troubleshooting

If you determine that the Best Practice was not successful, as described in *Verifying Success*, use the following table to identify and solve problems.

Problem	Possible Solutions
Encapsulation command failed to execute, or you cannot boot into multiuser mode	Unencapsulate the boot disk. If this fails, manually reset the encapsulation procedure changes.

Unencapsulating the Boot Disk

You can unencapsulate the boot disk to convert volumes on the boot disk to partitions.

To unencapsulate the boot disk, you must reboot the system using the disk that was last used for the `rootvol` and `swapvol` volumes (which might require that you first change the default boot device on the system console). The unencapsulation process changes the following files:

- For AdvFS, the `/etc/fdmns/*` directory is updated to use disk partitions instead of LSM volumes for domains associated with the root disk.
- For UFS, the `/etc/fstab` file is changed to use disk partitions instead of LSM volumes.
- For swap space, the `/etc/sysconfigtab` file is changed to update the `swapdevice` entry to not use LSM volumes and to set the `lsm_rootdev_is_volume` entry to 0.

This procedure dismantles LSM volumes and requires a reboot. Ensure there are no users on the system.

To unencapsulate the boot disk:

1. Display all volume information by entering the following command:

```
# volprint -htq
```

In the output, note the names of the secondary plexes for the system partition volumes (usually those with the `-02` suffix):

```
⋮
v rootvol      root          ENABLED ACTIVE 262145 ROUND -
pl rootvol-01 rootvol      ENABLED ACTIVE 262145 CONCAT - RW
sd root01-01p rootvol-01   root01     0      16      0      dsk0a  ENA
sd root01-01  rootvol-01   root01     16     262129 16      dsk0a  ENA
pl rootvol-02 rootvol      ENABLED ACTIVE 262145 CONCAT - RW
sd root02-02p rootvol-02   root02     0      16      0      dsk4a  ENA
```

```

sd root02-02    rootvol-02    root02    16        262129    16        dsk4a    ENA

v  swapvol      swap          ENABLED    ACTIVE    409600    ROUND    -
pl swapvol-01  swapvol      ENABLED    ACTIVE    409600    CONCAT   -        RW
sd swap01-01  swapvol-01  swap01    0         409600    0        dsk0b    ENA
pl swapvol-02  swapvol      ENABLED    ACTIVE    409600    CONCAT   -        RW
sd swap02-02  swapvol-02  swap02    0         409600    0        dsk4b    ENA

v  vol-dsk0g    fsgen        ENABLED    ACTIVE    1375019   SELECT   -
pl vol-dsk0g-01 vol-dsk0g    ENABLED    ACTIVE    1375019   CONCAT   -        RW
sd dsk0g-01    vol-dsk0g-01 dsk0g-AdvFS 0        1375019   0        dsk0g    ENA
pl vol-dsk0g-02 vol-dsk0g    ENABLED    ACTIVE    1375019   CONCAT   -        RW
sd dsk4g-01    vol-dsk0g-02 dsk4g-AdvFS 0        1375019   0        dsk4g    ENA

```

2. Remove the secondary plexes (break the mirror) for the system partition volumes. Ensure that you remove the plexes on the disk or disks you used for the mirrors, not those on the original system disk.

For example, to remove the secondary plexes for the `rootvol`, `swapvol` and `vol-dsk0g` volumes (which are all on `dsk4`), enter:

```

# volplex -o rm dis rootvol-02
# volplex -o rm dis swapvol-02
# volplex -o rm dis vol-dsk0g-02

```

This leaves the disks under LSM control, as members of the `rootdg` disk group. If you want to use these disks for another LSM volume after you reboot, you must reinitialize them. See the *Logical Storage Manager* manual for your version of the operating system for more information.

3. Optionally, verify that all the system volumes have only one plex.

For `volunroot` to work properly, only one plex must exist for each of the affected volumes. Each plex must be the one created by `volrootmir` or the original plex created when the root disk was encapsulated.

```

# volprint -htq

:
:
v  rootvol      root          ENABLED    ACTIVE    262145    ROUND    -
pl rootvol-01  rootvol      ENABLED    ACTIVE    262145    CONCAT   -        RW
sd root01-01p  rootvol-01  root01    0         16        0        dsk0a    ENA
sd root01-01  rootvol-01  root01    16        262129    16        dsk0a    ENA

v  swapvol      swap          ENABLED    ACTIVE    409600    ROUND    -
pl swapvol-01  swapvol      ENABLED    ACTIVE    409600    CONCAT   -        RW
sd swap01-01  swapvol-01  swap01    0         409600    0        dsk0b    ENA

v  vol-dsk0g    fsgen        ENABLED    ACTIVE    1375019   SELECT   -
pl vol-dsk0g-01 vol-dsk0g    ENABLED    ACTIVE    1375019   CONCAT   -        RW
sd dsk0g-01    vol-dsk0g-01 dsk0g-AdvFS 0        1375019   0        dsk0g    ENA

```

4. Enter the `volunroot` command to convert the system partition volumes to boot disk partitions. Enter `now` when prompted to shut down the system:

```
# volunroot -a
```

```
This operation will convert the following file systems on the
system/swap disk dsk0 from LSM volumes to regular disk partitions:
```

```
Replace volume rootvol with dsk0a.
Replace volume swapvol with dsk0b.
Replace volume vol-dsk0g with dsk0g.
```

```
This operation will require a system reboot. If you choose to
continue with this operation, your system files will be updated
to discontinue the use of the above listed LSM volumes.
/sbin/volreconfig should be present in /etc/inittab to remove
the named volumes during system reboot.
```

```
Would you like to either quit and defer volunroot until later
or commence system shutdown now? Enter either 'quit' or time to be
used with the shutdown(8) command (e.g., quit, now, 1, 5): [quit] now
```

If the `volunroot` command fails to restore the system to its original state, you can manually reset the changes.

Manually Resetting Changes Made During the Encapsulation Procedure

If problems occur during the encapsulation procedure, LSM tries to undo all the changes and restore the use of partitions for the root file system. Under some circumstances, such as if you cannot boot the system into multiuser mode, you might need to manually reset the changes made as a result of encapsulating the boot disk.

To manually reset the changes made during the encapsulation procedure:

1. Shut down the machine:

```
# shutdown -h now
```

2. At the console prompt, boot the machine into interactive mode:

```
>>> boot -fl i
```

3. When prompted for the kernel file name, enter:

```
/vmunix lsm_rootdev_is_volume=2
```

4. When the system enters single-user mode, mount the root file system partition. For example:

```
# mount -u /
```

5. Restore the root file system files:

- For AdvFS, do the following:

- For the root (/) file system, enter:

```
# cd /etc/fdmns/root_domain
# rm rootvol
# ln -s /dev/disk/dskNa dskNa
```

- For the /usr file system, enter:

```
# cd /etc/fdmns/usr_domain
# rm vol-dsk0g
# ln -s /dev/disk/dskNg dskNg
```

- If the /var file system is separate from the /usr file system, enter:

```
# cd /etc/fdmns/var_domain
# rm vol-dsk0h
# ln -s /dev/disk/dskNh dskNh
```

- For UFS, make a backup copy of the /etc/fstab file and edit the original file as follows:
 - For the root (/) file system, change the device special file from /dev/vol/rootdg/rootvol to the a partition of the boot disk.
 - For the /usr file system, change the device special file from /dev/vol/rootdg/vol-dsk0g to the g partition of the boot disk.
 - If the /var file system is on a separate partition, change the device special file from /dev/vol/rootdg/vol-dsk0h to the h partition of the boot disk.

6. Make a backup copy of the /etc/sysconfigtab file and edit the original file as follows:

- In the LSM entry for `lsm_rootdev_is_volume = n`, change the value of `n` from 1 to 0.
- Edit the `swapdevice` entry under the `vm: tag` to use the original swap partition instead of the LSM volume.

See the *System Administration* manual and the `swapon(8)` reference page for your version of the operating system for more information.

7. Remove files related to the conversion. For example:

```
# rm -rf /etc/vol/reconfig.d/disk.d/*
```

8. Edit the disk label to remove the LSMnopr_v file system type (fstype) that was set by the volencap command for the swap, /usr, and /var partitions, and reset the fstype to its original value.

For the /usr and /var partitions, enter AdvFS or ufs depending on the file system you are using.

For example:

- If the swap space is on the b partition, enter:

```
# disklabel -sF dskMb swap
```
- If the /usr file system is on the g partition, enter:

```
# disklabel -sF dskNg AdvFS
```
- If the /var file system is on the h partition, enter:

```
# disklabel -sF dskNh AdvFS
```

9. Reboot the system on the same boot disk. When the system restarts, the root file system uses disk partitions.

If you successfully restored the system to its initial working state, you can try the encapsulation and mirroring procedures again. Be sure to observe the prerequisites described in *Is This Best Practice Right for You?*. If the encapsulation procedure fails again, contact Customer Support.

Alternative Practices

- If you are running Tru64 UNIX Version 5.0 through Version 5.1 in a TruCluster environment, use RAID controllers (for example, an HSZ70) to provide high availability for the root file system, member boot partitions, the quorum disk, and swap partitions in a cluster. See the following URL for user documentation for hardware RAID controllers:

```
http://www.compaq.com/products/storageworks/ \ array-and-scsi-controllers/Hsxuserdocs.html
```

- If you are running Tru64 UNIX Version 5.1A in a TruCluster environment, you can use the volmigrate command to provide high availability for the cluster root file system and the volencap swap command for member swap partitions.

You cannot use LSM on member boot partitions or the quorum disk in a cluster; to provide high availability for these, use RAID controllers. See the following URL for user documentation for hardware RAID controllers:

```
http://www.compaq.com/products/storageworks/ \ array-and-scsi-controllers/Hsxuserdocs.html
```

See the *TruCluster Server Cluster Administration* manual for your version of the product for more information on configuring LSM in a TruCluster environment.

Comments and Questions

We value your comments and questions on the information in this document. Please mail your comments to us at this address:

`best_practices@zk3.dec.com`

Legal Notice

Compaq, the Compaq logo, TruCluster, and Tru64 are trademarks of Compaq Information Technologies Group, L.P. UNIX is a trademark of The Open Group. All other product names mentioned herein may be trademarks of their respective companies.

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

Compaq shall not be liable for technical or editorial errors or omissions contained herein. The information in this document is provided "as is" without warranty of any kind and is subject to change without notice. The warranties for Compaq products are set forth in the express limited warranty statements accompanying such products. Nothing herein should be construed as constituting an additional warranty.