

Mirroring a Boot Disk Using Solstice DiskSuite

Mr. Gautam Das
Senior Systems Programmer
Northeast Regional Data Center
University of Florida

PO Box 112050
Gainesville
Florida
32611-2050

Table of Contents

1. Installation of Solstice DiskSuite 4.2.1 on Solaris 8	1
2. Prepare the disks for Solstice DiskSuite	2
3. Solstice DiskSuite Configuration	3

Solstice DiskSuite is a free RAID volume manager from Sun Microsystems. This document describes the method of mirroring a boot disk with Solstice DiskSuite using shell commands. The purpose of this document is to provide an example and quick reference. This document is not an exhaustive treatise on Solstice DiskSuite. For more details, please refer to documentation from Sun Microsystems on Solstice DiskSuite 4.2.1.

Solstice DiskSuite comes bundled with the Solaris operating system.

How to recover from a mirrored boot disk failure is covered in a separate document titled Recovering from a boot disk failure using Solstice DiskSuite [http://open-systems.ufl.edu/docs/solaris/bootdisk_recover.html].

The process involves three steps:

-
-
-

1. Installation of Solstice DiskSuite 4.2.1 on Solaris 8

Insert Solaris 8 Software CD 2 of 2 and change to the directory as shown below:

```
# cd /cdrom/sol_8_401_sparc_2/Solaris_8/EA/products/DiskSuite_4.2.1/sparc/Packages
```

Install Solstice DiskSuite using the **pkgadd** command as follows:

```
# pkgadd -d .
The following packages are available:
 1 SUNWmdg Solstice DiskSuite Tool
   (sparc) 4.2.1,REV=1999.11.04.18.29
 2 SUNWmdja Solstice DiskSuite Japanese localization
   (sparc) 4.2.1,REV=1999.12.09.15.37
 3 SUNWmdnr Solstice DiskSuite Log Daemon Configuration Files
   (sparc) 4.2.1,REV=1999.11.04.18.29
 4 SUNWmdnu Solstice DiskSuite Log Daemon
   (sparc) 4.2.1,REV=1999.11.04.18.29
 5 SUNWmdr Solstice DiskSuite Drivers
   (sparc) 4.2.1,REV=1999.12.03.10.00
 6 SUNWmdu Solstice DiskSuite Commands
   (sparc) 4.2.1,REV=1999.11.04.18.29
 7 SUNWmdx Solstice DiskSuite Drivers(64-bit)
   (sparc) 4.2.1,REV=1999.11.04.18.29
```

Select 1,3,4,5,6, and 7 and answer "y" to continue with the installation for all questions.

After successful installation reboot the machine.

```
# init 6
```

2. Prepare the disks for Solstice DiskSuite

Solstice DiskSuite uses metadvice state databases to store information on disk about the state of the DiskSuite configuration. The metadvice state database records and tracks changes made to the configuration.

These databases must reside on a dedicated slice of a boot disk. Two slices of approximately 50 megabytes size may be created for this purpose. If you are working with a boot disk that has the operating system already installed on it and you do not wish to reinstall the OS, and there is no unused space on the disk, then you can steal space from the swap partition. Refer to documentation [Ab2Enc=iso-\[http://docs.sun.com/ab2/coll.260.2/DISKSUITEUG/@Ab2PageView/2360?Ab2Lang=C&8859-1\]](http://docs.sun.com/ab2/coll.260.2/DISKSUITEUG/@Ab2PageView/2360?Ab2Lang=C&8859-1) from Sun if you need to do that.

In this example we will be using two Sun 18gigabyte disks. We will mirror the boot disk c0t10d0 onto an identical second disk c0t11d0.

Use the **format** command to select the boot disk and create the slices that will hold the state database.

In this example we have 2 slices of approximately 50 megabytes size for this purpose - they are slices 3 and 4.

The partition table of our example 18Gig Sun disk looks like the following:

```
partition> pr
Current partition table (original):
```

Total disk cylinders available: 7506 + 2 (reserved cylinders)

Part	Tag	Flag	Cylinders	Size	Blocks
0	root	wm	891 - 1483	1.33GB	(593/0/0) 2794216
1	swap	wu	0 - 890	2.00GB	(891/0/0) 4198392
2	backup	wm	0 - 7505	16.86GB	(7506/0/0) 35368272
3	unassigned	wm	7463 - 7484	50.62MB	(22/0/0) 103664
4	unassigned	wm	7485 - 7505	48.32MB	(21/0/0) 98952
5	var	wm	1484 - 1706	513.07MB	(223/0/0) 1050776
6	unassigned	wm	1707 - 1929	513.07MB	(223/0/0) 1050776
7	home	wm	1930 - 7462	12.43GB	(5533/0/0) 26071496

Note: The metadb database slices can be as small as 6 MB.

The partition table of the mirrored disk must be identical to the boot disk. Copy the partition table of the boot disk to the mirror disk.

```
# prtvtoc /dev/rdisk/c0t10d0s2 | fmthard -s - /dev/rdisk/c0t11d0s2
fmthard: New volume table of contents now in place
```

Here c0t10d0s2 is the boot disk and c0t11d0s2 is the mirror disk. They are both on the same controller. It is better to mirror disks across different controllers if possible. The **fmthard** command reads the partition table of the boot disk from standard in and writes it to the target disk specified. Use the **format** command to verify that the partition tables on both disks are identical.

3. Solstice DiskSuite Configuration

The state database replicas must be created first. The minimum requirement is 2 database replicas. Earlier we have created 2 disk slices to hold the replicas. We will create 2 replicas on each slice, thus totaling 4 replicas. A state database replica stores DiskSuite configuration and state information. We use the **metadb** command to create the database replicas.

```
# metadb -a -f -c2 /dev/dsk/c0t10d0s3 /dev/dsk/c0t11d0s3
# metadb -a -f -c2 /dev/dsk/c0t10d0s4 /dev/dsk/c0t11d0s4
```

The metadb command options used above are explained below:

- *number*
-

3.1. Create a mirror for the / filesystem

To start we create a one-way mirror which is composed of a single disk. Later we attach the second disk to the mirror. The metainit command defines the metadevices that the mirror will use. The device numbers (d##) are arbitrary. The convention we have adopted here is:

```
d10 - 1st set of submirrors
d20 - 2nd set of submirrors
d30 - mirror containing submirrors
```

The following **metainit** commands are used to create the mirror for disk slice 0:

```
# metainit -f d10 1 1 c0t10d0s0
d10: Concat/Stripe is setup

# metainit -f d20 1 1 c0t11d0s0
d20: Concat/Stripe is setup

# metainit d30 -m d10
d30: Mirror is setup
```

The **-f** means force the creation. **1 1** means create a 1 way 1 slice metadvice. The last command creates the metadvice mirror and attaches the submirror d10 to it.

Next update the `/etc/vfstab` file for the `/` filesystem, and the `/etc/system` file. Do not edit `/etc/vfstab` or `/etc/system` manually - Use the **metaroot** command as below:

```
# metaroot d30
```

Take a look at `/etc/vfstab` and note that the `/` filesystem will be mounted on `/dev/md/dsk` rather than on `/dev/dsk`.

3.2. Create mirrors for the other filesystems

The swap filesystem:

```
# metainit -f d11 1 1 c0t10d0s1
# metainit -f d21 1 1 c0t11d0s1
# metainit d31 -m d11
```

The `/var` filesystem:

```
# metainit -f d12 1 1 c0t10d0s5
# metainit -f d22 1 1 c0t11d0s5
# metainit d32 -m d12
```

The `/opt` filesystem:

```
# metainit -f d13 1 1 c0t10d0s6
# metainit -f d23 1 1 c0t11d0s6
# metainit d33 -m d13
```

The `/export/home` filesystem:

```
# metainit -f d14 1 1 c0t10d0s7
# metainit -f d24 1 1 c0t11d0s7
# metainit d34 -m d14
```

Edit the `/etc/vfstab` file to mount the new mirrors on boot up.

The new `/etc/vfstab` file should look like this:

#device	device	mount	FS	fsck	mount	mount
#to mount	to fsck	point	type	pass	at boot	options
#						
#/dev/dsk/cl1d0s2	/dev/rdisk/cl1d0s2	/usr	ufs	1	yes	-
FD	-	/dev/fd	fd	-	no	-
/proc	-	/proc	proc	-	no	-
/dev/md/dsk/d31	-	-	swap	-	no	-
/dev/md/dsk/d30	/dev/md/rdsk/d30	/	ufs	1	no	-
/dev/md/dsk/d32	/dev/md/rdsk/d32	/var	ufs	1	no	-
/dev/md/dsk/d34	/dev/md/rdsk/d34	/export/home	ufs	2	yes	-
/dev/md/dsk/d33	/dev/md/rdsk/d33	/opt	ufs	2	yes	-
swap	-	/tmp	tmpfs	-	yes	-

Needless to say extra care must be taken when editing the `/etc/vfstab` file. A single mistake may result in not being able to boot the system.

3.2.1. Suppress harmless warning messages (optional)

Typically, after a Solstice DiskSuite install, you will receive harmless but annoying messages on boot-up like `"WARNING: forceload of misc/md_hotspares failed"`. These messages can be suppressed by creating an empty hot spare pool. The following `metainit` command does just that:

```
# metainit hsp001
```

3.3. Attach submirrors to the mirror

Reboot and allow the system to mount the mirrors.

```
# lockfs -fa
# init 6
```

The following warning messages are harmless, and may be safely ignored. They are an artifact of the way drivers are loaded during the boot process when you have a mirrored root or `/usr` file system.:

```
WARNING: forceload of misc/md_trans failed
WARNING: forceload of misc/md_raid failed
WARNING: forceload of misc/md_hotspares failed
```

Attach the second submirror to the mirror.

This will cause the data from the boot disk to be synchronized with the mirrored disk.

```
# metattach d30 d20
```

Run `metastat` to check the progress of the synchronization.

```
# metastat d30
```

Now attach the rest of the submirrors.

```
# metattach d31 d21
# metattach d32 d22
```

```
# metattach d33 d23
# metattach d34 d24
```

Run `metastat` to check the progress.

3.4. Make the mirror disk bootable.

```
# installboot /usr/platform/`uname -i`/lib/fs/ufs/bootblk /dev/rdisk/c0t11d0s0
```

Create an nvram alias to allow booting from the backup disk.

```
# ls -l /dev/rdisk/c0t11d0s0
```

```
lrwxrwxrwx  1 root    root          54 Aug  5 20:35 /dev/rdisk/c0t11d0s0 ->
  ../../devices/sbus@3,0/SUNW,fas@3,8800000/sd@b,0:a,raw
```

The device path to be used for the alternate boot device is `/sbus@3,0/SUNW,fas@3,8800000/sd@b,0:a`

Shut down the machine and at the ok prompt type

```
ok nvalias backup /sbus@3,0/SUNW,fas@3,8800000/sd@b,0:a
```

Alternatively this can be done using the `eeprom` command as follows:

```
# eeprom nvramrc="devalias backup /sbus@3,0/SUNW,fas@3,8800000/sd@b,0:a"
```

In case of a primary boot disk failure, boot from the backup mirror disk.

```
ok boot backup
```