

Solaris Volume Manager – RAID-1

For dataless workstations, disk mirroring (RAID-1) is seldom done. However, in order to protect data on this standalone machine mirroring was required.

The first step

Recreating on the so far unused second internal disk the same partition (slice) layout as on the first one.

```
# prtvtoc -h /dev/dsk/c0t0d0s2 | fmthard -s - /dev/rdisk/c0t1d0s2
fmthard: New volume table of contents now in place.
#
```

The next is to create Solaris Volume Manager metadvice state database. The SVM software for its proper operations keeps multiple replicas and requires that more than half of them is available and not corrupted.

I had chosen 3 disks to keep 3 replicas each. They are on the both internal drives and on the first one (one with the lowest SCSI target number) of the external disks.

If the Solaris Volume Manager metadvice state database were to be kept on (any) two hard drives, then even a single drive failure could not be tolerated.

Disk drive names depend on hardware configuration and likely would need to be adjusted for any other system.

```
# umount /SVM
# metadb -a -f -c 3 c0t0d0s7 c0t1d0s7
# metadb -i
      flags          first blk      block count
a      u             16             8192      /dev/dsk/c0t0d0s7
a      u            8208             8192      /dev/dsk/c0t0d0s7
a      u           16400             8192      /dev/dsk/c0t0d0s7
a      u             16             8192      /dev/dsk/c0t1d0s7
a      u            8208             8192      /dev/dsk/c0t1d0s7
a      u           16400             8192      /dev/dsk/c0t1d0s7
r - replica does not have device relocation information
o - replica active prior to last mddb configuration change
u - replica is up to date
l - locator for this replica was read successfully
c - replica's location was in /etc/lvm/mddb.cf
p - replica's location was patched in kernel
m - replica is master, this is replica selected as input
W - replica has device write errors
a - replica is active, commits are occurring to this replica
M - replica had problem with master blocks
D - replica had problem with data blocks
F - replica had format problems
S - replica is too small to hold current data base
R - replica had device read errors
#
```

As this is a software RAID-1, the partitions will be mirrored and not the whole drives (like with the hardware RAID-1).

There are 4 partitions on the internal drive(s): /, swap, /var and /localhost. They correspond to SVM devices d10, d20, d30 and d40. These names and corresponding names of submirrors are the ones that would be created automatically if Solaris were to be installed by means of jumpstart.

It is not mandatory to use /etc/lvm/md.tab, however its use is preferred over typing all configuration on the command line.

```
# cat >> /etc/lvm/md.tab
d10 -m d1
d1 1 1 c0t0d0s0
d2 1 1 c0t1d0s0
d20 -m d11
d11 1 1 c0t0d0s1
d12 1 1 c0t1d0s1
d30 -m d21
d21 1 1 c0t0d0s4
d22 1 1 c0t1d0s4
d40 -m d31
d31 1 1 c0t0d0s6
d32 1 1 c0t1d0s6

# for f in d1 d2 d11 d12 d21 d22 d31 d32; do metainit -f $f; done
d1: Concat/Stripe is setup
d2: Concat/Stripe is setup
d11: Concat/Stripe is setup
d12: Concat/Stripe is setup
d21: Concat/Stripe is setup
d22: Concat/Stripe is setup
d31: Concat/Stripe is setup
d32: Concat/Stripe is setup
# for f in d10 d20 d30 d40; do metainit $f; done
d10: Mirror is setup
d20: Mirror is setup
d30: Mirror is setup
d40: Mirror is setup
[ These 4 mirrors at this moment have only one component each. ]
# metaroot d10
[ The command above places / filesystem on metadvice d10. ]
[ It modifies /etc/system and /etc/vfstab. ]
#
# cat /etc/vfstab
#device          device          mount          FS          fsck          mount  mount
#to mount        to fsck          point          type         pass         at boot options
#
fd      -          /dev/fd fd      -          no          -          -          -
/proc   -          /proc  proc   -          no          -          -          -
/dev/dsk/c0t0d0s1 -          -          swap   -          no          -          -          -
/dev/md/dsk/d10 /dev/md/rdisk/d10 /          ufs     1          no          -          -          -
/dev/dsk/c0t0d0s4 /dev/rdisk/c0t0d0s4 /var      ufs     1          no          -          -          -
/dev/dsk/c0t0d0s6 /dev/rdisk/c0t0d0s6 /localhost ufs     2          yes         -          -          -
#
# These two slices keep replicas of the Solaris Volume Manager
# metadvice state database.
# Under no circumstances these partitions should be resized
# or used for any other purpose.
#
#/dev/dsk/c0t0d0s7 /dev/rdisk/c0t0d0s7 /SVM      ufs     2          yes         -          -          -
#/dev/dsk/c0t1d0s7 /dev/rdisk/c0t1d0s7 /SVM      ufs     2          yes         -          -          -
#
/devices          -          /devices          devfs   -          no          -          -          -
ctfs              -          /system/contract  ctfs    -          no          -          -          -
objfs             -          /system/object    objfs   -          no          -          -          -
swap              -          /tmp              tmpfs   -          yes         -          -          -
#
# tail /etc/system
*          set nautopush=32
*          set maxusers=40
*
*          To set a variable named 'debug' in the module named 'test_module'
*
*          set test_module:debug = 0x13
```

```
* Begin MDD root info (do not edit)
rootdev:/pseudo/md@0:0,10,blk
* End MDD root info (do not edit)
#
# init 6
```

Final steps in mirroring the root disk

First /etc/vfstab needs to be adjusted, so also swap, /var and /localhost filesystems point to SVM metadevices instead of disk partitions.

```
# cat /etc/vfstab
#device      device      mount      FS      fsck      mount      mount
#to mount    to fsck      point      type     pass      at boot    options
#
fd          -          /dev/fd fd    -          no         -
/proc      -          /proc      proc   -          no         -
/dev/md/dsk/d10 /dev/md/rdisk/d10    /                  ufs        1          no         -
/dev/md/dsk/d20 -          -          -                  swap        -          no         -
/dev/md/dsk/d30 /dev/md/rdisk/d30    /var             ufs        1          no         -
/dev/md/dsk/d40 /dev/md/rdisk/d40    /localhost       ufs        2          yes        -
#
# These two slices keep replicas of the Solaris Volume Manager
# metadevice state database.
# Under no circumstances these partitions should be resized
# or used for any other purpose.
#
#/dev/dsk/c0t0d0s7      /dev/rdisk/c0t0d0s7      /SVM      ufs        2          yes        -
#/dev/dsk/c0t1d0s7      /dev/rdisk/c0t1d0s7      /SVM      ufs        2          yes        -
#
/devices          -          /devices      devfs      -          no         -
ctfs              -          /system/contract      ctfs      -          no         -
objfs             -          /system/object      objfs      -          no         -
swap              -          /tmp            tmpfs      -          yes        -
```

And finally SVM metadevices become two-way mirrors, when components from the other disk are added.

```
# metattach d10 d2
d10: submirror d2 is attached
# metastat | fgrep -i sync
      State: Resyncing
      Resync in progress: 0 % done
      State: Resyncing
# [ Waiting for approximately 20 minutes, until resyncing is completed. ]
#
# metattach d20 d12
d20: submirror d12 is attached
#
# [ Waiting... ]
#
# metattach d30 d22
d30: submirror d22 is attached
#
# [ Waiting... ]
#
# metattach d40 d32
d40: submirror d32 is attached
#
# [ Waiting... ]
#
# metastat | fgrep -i sync
# init 6
```