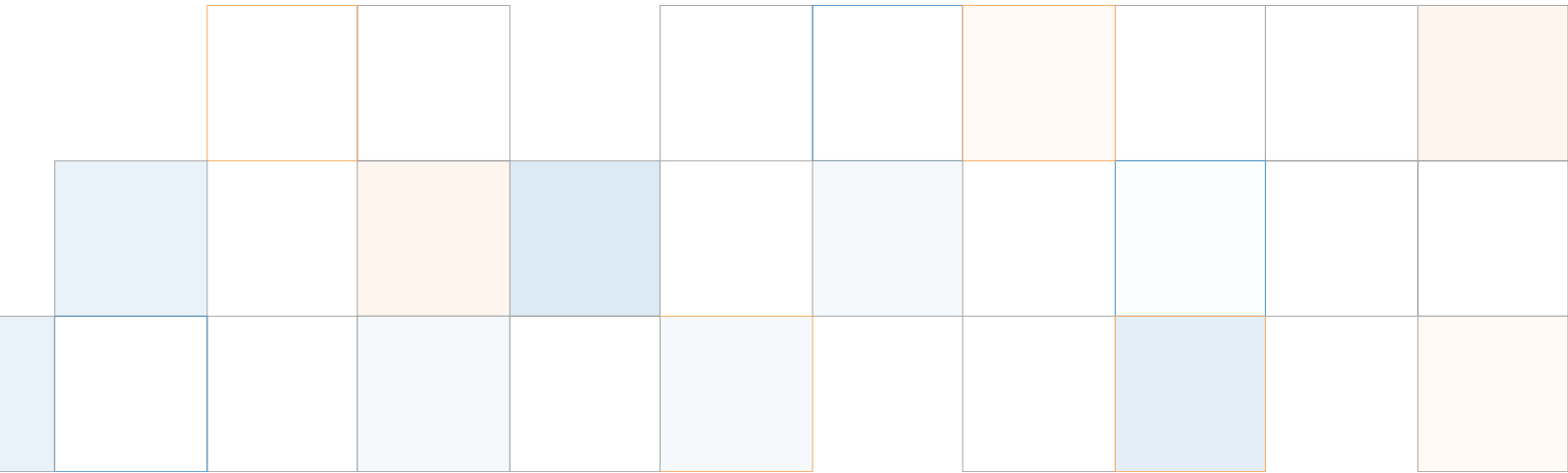




tech docs



Replacing Mirrored Disks v890

To replace a disk, use the `luxadm (1M)` command to remove it and insert the new disk. This procedure causes an update of the Solaris OS device framework, so that the new disk's DevID is inserted and the old disk's DevID is removed.

Procedure for Replacing Mirrored Disks

The following set of commands should work in all cases. Follow the exact sequence to ensure a smooth operation.

To replace a disk, which is controlled by SVM, and is part of a mirror, perform the following steps:

1. Run “`metadetach`” to detach all the submirrors on the failing disk from their respective mirrors (see the following):

```
# metadetach -f <mirror> <submirror>
```

NOTE: The “-f” option is not required if the metadevice is in an “okay” state.

2. Run `metaclear` to remove the `<submirror>` configuration from the disk:

```
# metaclear <submirror>
```

You can verify there are no existing metadevices left on the disk, by running the following:

```
# metastat -p | grep c#t#d#
```

3. If there are any replicas on this disk, note the number of replicas, and remove them using the following:

```
# metadb -i (number of replicas to be noted) .  
# metadb -d c#t#d#s#
```

Verify that there are no existing replicas left on the disk by running the following:

```
# metadb | grep c#t#d#
```

4. If there are any open filesystems on this disk not under SVM control, or non-mirrored metadevices, unmount them.
5. Run “`format`” or “`prtvtoc/fmthard`” to save the disk partition table information.

```
# prtvtoc /dev/rdisk/c#t#d#s2 > file
```

6. Run the ‘`luxadm`’ command to remove the failed disk.

```
#luxadm remove_device -F /dev/rdisk/c#t#d#s2
```

At the prompt, physically remove the disk and continue.

The `picld` daemon notifies the system that the disk has been removed.

7. Initiate `devfsadm` cleanup subroutines by entering the following command:

```
# /usr/sbin/devfsadm -C -c disk
```

The default `devfsadm` operation is, to attempt to load every driver in the system, and attach these drivers to all possible device instances. The `devfsadm` command then creates device special files, in the `/devices` directory, and logical links in `/dev`.

With the “`-c disk`” option, `devfsadm` will only update disk device files. This saves time, and is important on systems that have tape devices attached. Rebuilding these tape devices could cause undesirable results on non-Sun hardware.

The `-C` option cleans up the `/dev` directory, and removes any lingering logical links to the device link names.

This should remove all the device paths for this particular disk. This can be verified with:

```
# ls -ld /dev/dsk/cxtxd*
```

This should return no devices.

8. It is now safe to physically replace the disk. Insert a new disk, and configure it. Create the necessary entries in the Solaris OS device tree, with one of the following commands:

```
# devfsadm
```

or

```
# /usr/sbin/luxadm insert_device <enclosure_name,>sx>
```

where `sx` is the slot number

or

```
# /usr/sbin/luxadm insert_device (if enclosure name is not known)
```

Note: In many cases, `luxadm insert_device` does not require the enclosure name and slot number.

Use the following to find the slot number:

```
# luxadm display <enclosure_name>
```

To find the `<enclosure_name>` use:

```
# luxadm probe
```

Run “`ls -ld /dev/dsk/c1t1d*`” to verify that the new device paths have been created.

Procedure for Replacing a Disk In a Raid-5 Volume

Note: If a disk is used in BOTH a mirror and a RAID5, do not use the following procedure; instead, follow the instructions for the MIRRORED devices (above). This is because the RAID5 array just healed, is treated as a single disk for mirroring purposes.

To replace an SVM-controlled disk, which is part of a RAID5 metadvice, the following steps must be followed:

1. If there are any open filesystems on this disk not under SVM control, or non-mirrored metadvice, unmount them.

2. If there are any replicas on this disk, remove them using:

```
# metadb -d c#t#d#s#
```

Verify there are no existing replicas left on the disk by running:

```
# metadb | grep c#t#d#
```

3. Run “**format**” or “**prtvtoc/fmthard**” to save the disk partition table information.

```
# prtvtoc /dev/rdisk/c#t#d#s2 > file
```

4. Run the ‘**luxadm**’ command to remove the failed disk.

```
# luxadm remove_device -F /dev/rdisk/c#t#d#s2
```

At the prompt, physically remove the disk and continue.

The picld daemon notifies the system that the disk has been removed.

5. Initiate devfsadm cleanup subroutines by entering the following command:

```
# /usr/sbin/devfsadm -C -c disk
```

The default **devfsadm** operation, is to attempt to load every driver in the system, and attach these drivers to all possible device instances. The **devfsadm** command then creates device special files in the /devices directory, and logical links in /dev.

With the “**-c disk**” option, devfsadm will only update disk device files. This saves time and is important on systems that have tape devices attached. Rebuilding these tape devices could cause undesirable results on non-Sun hardware.

The “**-C**” option cleans up the /dev directory, and removes any lingering logical links to the device link names.

This should remove all the device paths for this particular disk.

This can be verified with:

```
# ls -ld /dev/dsk/c#t#d#s*
```

This should return no devices.

6. It is now safe to physically replace the disk. Insert a new disk, and configure it. Create

the necessary entries in the Solaris OS device tree with one of the following commands:

```
# devfsadm
```

or

```
# /usr/sbin/luxadm insert_device <enclosure_name,sx>
```

where sx is the slot number

or

```
# /usr/sbin/luxadm insert_device (if enclosure name is not known)
```

Note: In many cases, `luxadm insert_device` does not require the enclosure name and slot number.

Use the following to find the slot number:

```
# luxadm display <enclosure_name>
```

To find the `<enclosure_name>` you can use:

```
# luxadm probe
```

Run `ls -ld /dev/dsk/c1t1d*` to verify that the new device paths have been created.

CAUTION:

After inserting a new disk and running `devfsadm`(or `luxadm`), the old `ssd` instance number changes to a new `ssd` instance number. This change is expected, so ignore it.

For Example:

When the error occurs on the following disks(ssd3).

```
WARNING: /pci@8,600000/SUNW,qlc@2/fp@0,0/ssd@w21000004cfa19920,0  
(ssd3):
```

```
Error for Command: read(10)          Error Level: Retryable  
Requested Block: 15392944          Error Block: 15392958
```

After inserting a new disk, the `ssd` instance changes to `ssd10` as shown below. It is not a cause of concern as this is expected.

```
picld[287]: [ID 727222 daemon.error] Device DISK0 inserted  
qlc: [ID 686697 kern.info] NOTICE: Qlogic qlc(2): Loop ONLINE  
scsi: [ID 799468 kern.info] ssd10 at fp2: name  
w21000011c63f0c94,0, bus  
address ef
```

```
genunix: [ID 936769 kern.info] ssd10 is /pci@8,600000/SUNW,qlc@2/
fp@0,0/ssd@
w21000011c63f0c94,0
scsi: [ID 365881 kern.info] <SUN72G cyl 14087 alt 2 hd 24 sec
424>
genunix: [ID 408114 kern.info] /pci@8,600000/SUNW,qlc@2/fp@0,0/
ssd@w21000011c
63f0c94,0 (ssd10) online
```

7. Run `'format'` or `'prtvtoc'` to put the desired partition table on the new disk

```
# fmthard -s file /dev/rdisk/c#t#d#s2
```

['file' is the `prtvtoc` saved in step 3]

8. Run `'metadevadm'` on the disk, which will update the New DevID.

```
# metadevadm -u c#t#d#
```

Note: Due to BugID 4808079 a disk can show up as `"unavailable"` in the `metastat` command, after running Step 8. To resolve this, run `"metastat -i"`. After running this command, the device should show a `metastat` status of `"Okay"`. The fix for this bug has been delivered and integrated in s9u4_08,s9u5_02 and s10_35.

9. If necessary, recreate any replicas on the new disk:

```
# metadb -a c#t#d#s#
```

10. Run `metareplace` to enable and resync the new disk:

```
# metareplace -e <raid5-md> c#t#d#s#
```

Reference:

The following Infodocs were referenced:

Infodoc ID 40517:

Synopsis: Removing and Replacing the Sun Fire[™] 280R , V480 , or V880 hot-pluggable internal disk drives.

Infodoc ID 73132:

Synopsis: Solaris[™] Volume Manager software: Replacing SCSI Disks (Solaris[™] 9 Operating System and above)

Infodoc ID 40842:

Synopsis: Veritas Volume Manager - Procedure to Replace Internal FibreChannel (FC) Disks controlled by VxVM

Example

In this example, a V280R has two fc-ali internal drives. SVM is used to mirror both the root and the swap devices between c1t0d0 and c1t1d0.

Here is the 'luxadm' display before the submirror disk replacement:

```
# luxadm probe
Found Fibre Channel device(s):
Node WWN:20000004cf357186 Device Type:Disk device
Logical Path:/dev/rdisk/c1t1d0s2
Node WWN:20000004cf365024 Device Type:Disk device
Logical Path:/dev/rdisk/c1t0d0s2
```

Here is the 'format' display before the submirror disk replacement:

```
# format

AVAILABLE DISK SELECTIONS:
 0. c1t0d0 <SUN36G cyl 24620 alt 2 hd 27 sec 107>
    /pci@8,600000/SUNW,q1c@4/fp@0,0/ssd@w21000004cf365024,0
 1. c1t1d0 <SUN36G cyl 24620 alt 2 hd 27 sec 107>
    /pci@8,600000/SUNW,q1c@4/fp@0,0/ssd@w21000004cf357186,0
```

Here is the output of the 'metadb' command, showing the locations of the SVM database replicas:

```
# metadb

      flags          first blk      block count
a m p lu0          16             8192      /dev/dsk/
c1t0d0s7
a p lu0            8208            8192      /dev/dsk/
c1t0d0s7
a u                16              8192      /dev/dsk/
c1t1d0s7
a u                8208            8192      /dev/dsk/
c1t1d0s7
```

Here is the SVM configuration before the submirror disk replacement.

```
# metastat -i

d7: Mirror
  Submirror 0: d5
    State: Okay
  Submirror 1: d6
    State: Okay
  Pass: 1
  Read option: roundrobin (default)
  Write option: parallel (default)
  Size: 8389656 blocks (4.0 GB)

d5: Submirror of d7
  State: Okay
  Size: 8389656 blocks (4.0 GB)
  Stripe 0:
    Device      Start Block  Dbase      State Reloc Hot
Spare
    c1t0d0s1      0           No         Okay   Yes

d6: Submirror of d7
  State: Okay
  Size: 8389656 blocks (4.0 GB)
  Stripe 0:
    Device      Start Block  Dbase      State Reloc Hot
Spare
    c1t1d0s1      0           No         Okay   Yes

d3: Mirror
  Submirror 0: d0
    State: Okay
  Submirror 1: d1
    State: Okay
  Pass: 1
  Read option: roundrobin (default)
  Write option: parallel (default)
  Size: 6292242 blocks (3.0 GB)

d0: Submirror of d3
  State: Okay
  Size: 6292242 blocks (3.0 GB)
  Stripe 0:
    Device      Start Block  Dbase      State Reloc Hot
Spare
    c1t0d0s3      0           No         Okay   Yes
```

```

d1: Submirror of d3
State: Okay
Size: 6292242 blocks (3.0 GB)
Stripe 0:
    Device      Start Block  Dbase      State Reloc Hot Spare
    c1t1d0s3    0           No         Okay   Yes

```

Device Relocation Information:

```

Device  Reloc  Device ID
c1t1d0  Yes   id1,ssd@w20000004cf357186
c1t0d0  Yes   id1,ssd@w20000004cf365024

```

Assume c1t0d0 is the drive that needs to be replaced, use 'metadetach' to detach the submirrors from that disk.

```

# metadetach d3 d0
d0: submirror d0 is detached
# metaclear d0

# metadetach d7 d5
d1: submirror d5 is detached
# metaclear d5

```

Here is the 'metastat -i' output after detaching d0 and d5:

```

d7: Mirror
Submirror 1: d6
State: Okay
Pass: 1
Read option: roundrobin (default)
Write option: parallel (default)
Size: 8389656 blocks (4.0 GB)

d6: Submirror of d7
State: Okay
Size: 8389656 blocks (4.0 GB)
Stripe 0:
    Device      Start Block  Dbase      State Reloc Hot Spare
    c1t1d0s1    0           No         Okay   Yes

d3: Mirror
Submirror 1: d1
State: Okay
Pass: 1
Read option: roundrobin (default)
Write option: parallel (default)
Size: 6292242 blocks (3.0 GB)

d1: Submirror of d3
State: Okay
Size: 6292242 blocks (3.0 GB)

```

Stripe 0:

Device	Start Block	Dbase	State	Reloc	Hot Spare
c1t1d0s3	0	No	Okay	Yes	

Device Relocation Information:

Device	Reloc	Device ID
c1t1d0	Yes	id1,ssd@w20000004cf357186

Since we still have a replica on the disk to be removed, it should be removed:

```
# metadb -d c1t0d0s7
```

Now run luxadm and devfsadm commands to physically replace the disk.

```
#luxadm remove_device -F /dev/rdisk/c1t0d0s2
```

WARNING!!! Please ensure that no filesystems are mounted on these device(s). All data on these devices should have been backed up.

The list of devices which will be removed is:

```
1: Device name: /dev/rdisk/c1t0d0s2
   Node WWN:    20000004cf3591bc
   Device Type:Disk device
   Device Paths:
     /dev/rdisk/c1t0d0s2
```

```
Please verify the above list of devices and
then enter 'c' or <CR> to Continue or 'q' to Quit. [Default: c]:
stopping: /dev/rdisk/c1t0d0s2....Done
offlining: /dev/rdisk/c1t0d0s2....Done
```

Hit <RETURN> after removing the device(s).

(At this point physically remove the disk and complete the above task).

```
#devfsadm -C
```

```
#luxadm insert_device
```

Please hit <RETURN> when you have finished adding Fibre Channel Enclosure(s)/Device(s):

(At this point physically insert the disk and complete the above task)

Once the new disk is configured, run 'format' to put the appropriate partition table onto the disk.

```
# format
```

Run 'metainit' and 'metattach' to reattach them to their respective mirrors.

```
# metainit d0 1 1 c1t0d0s3
# metattach d3 d0
```

```
d3: submirror d0 is attached
# metainit d5 1 1 c1t0d0s1
# metattach d1 d21
d7: submirror d5 is attached
```

Meanwhile run `'metadb'` to recreate the replica on the disk:

```
# metadb -a -c 2 c1t0d0s7
```

After the new disk is attached to the mirror disk, it will be resynchronized. Once resynchronization process is completed, the mirror disk will be back to fully redundant mode.

Once the mirrors are in full redundant mode, proceed to replace the next disk in a similar manner.

Be sure to correct boot-device entry, if the root disk has been replaced.