

Mdstat

From Linux Raid Wiki

Contents

- 1 /proc/mdstat
 - 1.1 querying the status
 - 1.2 Personalities line
 - 1.3 md device line
 - 1.4 md config/status line
 - 1.5 bitmap line
 - 1.6 recovery

/proc/mdstat

The **/proc/mdstat** file shows a snapshot of the kernel's RAID/md state.

querying the status

The kernel md state is easily viewed by running:

```
cat /proc/mdstat
```

It won't hurt. Let's learn how to read the file. Here are some examples:

example 1:

```
Personalities : [raid1] [raid6] [raid5] [raid4]
md_d0 : active raid5 sde1[0] sdf1[4] sdb1[5] sdd1[2] sdc1[1]
        1250241792 blocks super 1.2 level 5, 64k chunk, algorithm 2 [5/5] [UUUUU]
        bitmap: 0/10 pages [0KB], 16384KB chunk

unused devices: <none>
```

example 2:

```
Personalities : [raid6] [raid5] [raid4]
md0 : active raid5 sda1[0] sdd1[2] sdb1[1]
        1465151808 blocks level 5, 64k chunk, algorithm 2 [4/3] [UUU_]
        unused devices: <none>
```

example 3:

```
Personalities : [raid1] [raid6] [raid5] [raid4]
md1 : active raid1 sdb2[1] sda2[0]
        136448 blocks [2/2] [UU]
```

```
md2 : active raid1 sdb3[1] sda3[0]
      129596288 blocks [2/2] [UU]

md3 : active raid5 sdl1[9] sdk1[8] sdj1[7] sdil[6] sdh1[5] sdg1[4] sdf1[3] sde1[2] sdd1[1] sdc1[0]
      1318680576 blocks level 5, 1024k chunk, algorithm 2 [10/10] [UUUUUUUUUU]

md0 : active raid1 sdb1[1] sda1[0]
      16787776 blocks [2/2] [UU]

unused devices: <none>
```

example 4:

```
Personalities : [raid1] [raid6] [raid5] [raid4]
md127 : active raid5 sdh1[6] sdg1[4] sdf1[3] sde1[2] sdd1[1] sdc1[0]
      1464725760 blocks level 5, 64k chunk, algorithm 2 [6/5] [UUUUU_]
[==>.....] recovery = 12.6% (37043392/292945152) finish=127.5min speed=33440K/sec

unused devices: <none>
```

example 5:

```
Personalities : [linear] [raid0] [raid1] [raid5] [raid4] [raid6]
md0 : active raid6 sdf1[0] sde1[1] sdd1[2] sdc1[3] sdb1[4] sda1[5] hdb1[6]
      1225557760 blocks level 6, 256k chunk, algorithm 2 [7/7] [UUUUUUUU]
      bitmap: 0/234 pages [0KB], 512KB chunk

unused devices: <none>
```

example 6:

```
Personalities : [raid1]
md1 : active raid1 sde1[6](F) sdg1[1] sdb1[4] sdd1[3] sdc1[2]
      488383936 blocks [6/4] [_UUUU_]

unused devices: <none>
```

Personalities line

The "Personalities" line tells you what RAID level the kernel currently supports. This can be changed by either changing the raid modules or recompiling the kernel. Possible personalities include: [raid0] [raid1] [raid4] [raid5] [raid6] [linear] [multipath] [faulty]

Note: [faulty] is a diagnostic personality; it does not mean there is a problem with the array.

md device line

Each array is then described. from example 1:

```
md_d0 : active raid5 sde1[0] sdf1[4] sdb1[5] sdd1[2] sdc1[1]
```

This means we're looking at the device /dev/md_d0.

It is active or 'started'. An inactive array is usually faulty. Stopped arrays aren't visible here.

It is a raid5 array and the component devices are:

```
/dev/sde1 is device 0
/dev/sdf1 is device 4
/dev/sdb1 is device 5
/dev/sdd1 is device 2
/dev/sdc1 is device 1
```

The order in which the devices appear in this line means nothing.

The raid role numbers [#] following each device indicate its role, or function, within the raid set. Any device with "n" or higher are spare disks. 0,1,...,n-1 are for the working array. Notice that there is no device 3. At some point that device will have failed and been replaced by device 5. (check if this is true!)

To identify a spare devices, first look for the [#/#] value on a line. The first number is the number of a complete raid device as defined. Lets say it is "n".

Also, if you have a failure, the failed device will be marked with (F) after the [#] (see example 6 where sde1 has failed. The spare that replaces this device will be the device with the lowest role number n or higher that is not marked (F). Once the resync operation is complete, the device's role numbers are swapped.

To identify the spare devices, first determine the number of device the array needs to be fully operational (see below for the [#/#] value). Lets say it is "m". The raid role numbers [#] following each device indicate its role, or function, within the raid set. Any device with "m" or higher are spare disks. 0,1,...,m-1 are for the working array.

md config/status line

The next line continues the decription of the array; in example 1 it is:

```
1250241792 blocks super 1.2 level 5, 64k chunk, algorithm 2 [5/5] [UUUUU]
```

This line provides some basic data about the fixed size and layout: it indicates the useable size of the array in blocks is 1250241792; the array uses a 1.2 superblock and confirms a level 5 (this is redundant!) array with a chunk size of 64k using algorithm 2. (See RAID Creation for more details).

The final 2 entries on this line

```
[5/5]
[UUUUU]
```

are more dynamic.

[n/m] means that ideally the array would have n devices however, currently, m devices are in use. Obviously when m >= n then things are good.

The

```
[UUUUU]
```

represents the status of each device, either U for up or _ for down. So examples 2 and 6 show 'degraded' arrays with some devices 'down'.

bitmap line

If an array has a bitmap then this line describes the state. Example 1 shows

```
bitmap: 0/10 pages [0KB], 16384KB chunk
```

What would it mean when it's, eg: 23/234

This refers to the in-memory bitmap (basically a cache of what's in the on-disk bitmap – it allows bitmap operations to be more efficient).

If it's 23/234 that means there are 23 of 234 pages allocated in the in-memory bitmap. The pages are allocated on demand, and get freed when they're empty (all zeroes). The in-memory bitmap uses 16 bits for each bitmap chunk to count all ongoing writes to the chunk, so it's actually up to 16 times larger than the on-disk bitmap.

External bitmap references look like this:

```
bitmap: 5/113 pages [20KB], 8192KB chunk, file: /WIBS/<node>:md0/WIB_<node>:md0
```

recovery

example 4 is clearly showing some recovery activity:

```
[==>.....] recovery = 12.6% (37043392/292945152) finish=127.5min speed=33440K/sec
```

The first part is simply a graphical representation of the progress. The rest of the line is fairly self explanatory. The finish time is only an approximation since the resync speed will vary according to other I/O demands. See the resync page for more details.

Retrieved from "<https://raid.wiki.kernel.org/index.php?title=Mdstat&oldid=4731>"

-
- This page was last modified on 5 October 2013, at 09:39.