Quick Search:

RAID1+boot+root+grub+mdadm without raidtab,etc.

newtopic   postreply   **Gentoo Forums Forum Index -> Documentation, Tips & Tricks**

View previous topic :: View next topic

Author	Message

mkli
n00b

Joined: 10 Dec 2002
Posts: 6

Posted: Sat Aug 02, 2003 4:45 am   Post subject: RAID1+boot+root+grub+mdadm without raidtab,etc.   quote

```
===================================================
In the Software-RAID HOWTO it is mentioned that it is not known how
to set up GRUB to boot off RAID. Here is how I did it:
**Follow at your own risk. If you break something it's your fault.**
================================================================
Configuration:
- /dev/hda (Pri. Master) 60 GB Seagate HDD (blank)
- /dev/hdc (Sec. Master) 60 GB Seagate HDD (blank)
- /dev/hdd (Sec. Slave) CDROM Drive

Setup Goals:
- /boot as /dev/md0: RAID1 of /dev/hda1 & /dev/hdc1 for redundancy
- / as /dev/md1: RAID1 of /dev/hda2 & /dev/hdc2 for redundancy
- swap*2 with equal priority: /dev/hda3 & /dev/hdc3 for more speed
- GRUB installed in boot records of /dev/hda and /dev/hdc so either
drive can fail but system still boot.
```

Tools:

- mdadm (http://www.cse.unsw.edu.au/~neilb/source/mdadm/)

(I used 1.2.0, but notice that as of 20030729 1.3.0 is available)

1. Boot up off rescue/installation CD/disk/HDD/whatever with mdadm

tools installed.

2. Partitioning of hard drives:

(I won't show you how to do this. See: # man fdisk ; man sfdisk)

But here's how stuff was arranged:

Code:

```
# sfdisk -l /dev/hda

Disk /dev/hda: 7297 cylinders, 255 heads, 63 sectors/track
Units = cylinders of 8225280 bytes, blocks of 1024 bytes, counting
from 0

  Device Boot Start    End  #cyls   #blocks  Id System
/dev/hda1  *       0+   16     17-   136521  fd Linux raid autodetect
/dev/hda2         17  7219   7203 57858097+ fd Linux raid autodetect
/dev/hda3       7220  7296     77   618502+ 82 Linux swap
/dev/hda4          0    -       0        0   0 Empty
```

To make /dev/hdc the same:

Code:

```
# sfdisk -d /dev/hda | sfdisk /dev/hdc
```

/dev/hd[ac]1 for /dev/md0 for /boot

/dev/hd[ac]2 for /dev/md1 for /

/dev/hd[ac]3 for 2*swap

It is important to make md-to-be partitions with ID 0xFD, not 0x83.

3. Set up md devices: (both are **RAID1** [mirrors])

Code:

```
# mdadm --create /dev/md0 --level=1 \
    --raid-devices=2 /dev/hda1 /dev/hdc1
# mdadm --create /dev/md1 --level=1 \
    --raid-devices=2 /dev/hda2 /dev/hdc2
```

----

4. Make filesystems:

Code:

```
# mke2fs /dev/md0
# mkreiserfs /dev/md1
# mkswap /dev/hda3
# mkswap /dev/hdc3
```

----

5. Install Your distribution:
Simply treat /dev/md0 and /dev/md1 as the partitions to install on,
and install the way your normally do. Eg, for Gentoo:

Code:

```
# mkdir newinst
# mount -t reiserfs /dev/md1 ./newinst
# cd newinst
# mkdir boot
# mount -t ext2 /dev/md0 ./boot
# tar -xvjpf ../stage1-x86-1.4_rc2.tbz2
# mount -o bind /proc ./proc
# chroot ./
...
```

Here're the relevant entries /etc/fstab for the newly created
partitions:

Code:

```
/dev/md0      /boot       ext2      noauto,noatime          1 1
/dev/md1      /       reiserfs      noatime                 1 1
/dev/hda3     none       swap       sw,pri=1                0 0
/dev/hdc3     none       swap       sw,pri=1                0 0
```

The "pri=1" for each of the swap partitions makes them the same
priority so the kernel does striping and that speeds up vm. Of
course, this means that if a disk dies then the system may crash,
needing a reboot. Perhaps it would be wiser to make hd[ac]3 a **RAID1** as /dev/md2 array too, and just use that as swap. That
way, swap will be a little slower because it's raid'd, but in the case of a HDD failing while the system is running you won't have a
segfault and need to reboot.

6. Setting up GRUB: (assuming you've already installed it)

Code:

```
# grub
grub> root (hd0,0)
 Filesystem type is ext2fs, partition type 0xfd

grub> setup (hd0)
 Checking if "/boot/grub/stage1" exists... yes
 Checking if "/boot/grub/stage2" exists... yes
 Checking if "/boot/grub/e2fs_stage1_5" exists... yes
 Running "embed /boot/grub/e2fs_stage1_5 (hd0)"...  16 sectors are
embedded.
succeeded
 Running "install /boot/grub/stage1 (hd0) (hd0)1+16 p
(hd0,0)/boot/grub/stage2 /boot/grub/grub.conf"... succeeded
Done.
```

Ok, now that you've installed grub into hda's MBR, you'll need to do the same for hdc's MBR, in case you're booting up and hda is
dead:

Code:

```
grub> root (hd1,0)
 Filesystem type is ext2fs, partition type 0xfd
```

```
grub> setup (hd1)
 Checking if "/boot/grub/stage1" exists... yes
 Checking if "/boot/grub/stage2" exists... yes
 Checking if "/boot/grub/e2fs_stage1_5" exists... yes
 Running "embed /boot/grub/e2fs_stage1_5 (hd1)"...  16 sectors are
embedded.
succeeded
 Running "install /boot/grub/stage1 (hd1) (hd1)1+16 p
(hd1,0)/boot/grub/stage2 /boot/grub/grub.conf"... succeeded
Done.

grub> quit
```

Here is how /boot/grub/grub.conf is: (/dev/md0 mounted as /boot)
(Assuming kernel is installed as /boot/bzImage, and **RAID1** support
compiled into the kernel).

Code:

```
# Boot automatically after 30 secs.
timeout 30

# By default, boot the first entry.
default 0

# Fallback to the second entry.
fallback 1

# For booting with disc 0 kernel
title  GNU/Linux (hd0,0)
kernel (hd0,0)/bzImage root=/dev/md1

# For booting with disc 1 kernel, if (hd0,0)/bzImage is unreadable
title  GNU/Linux (hd1,0)
kernel (hd1,0)/bzImage root=/dev/md1
```
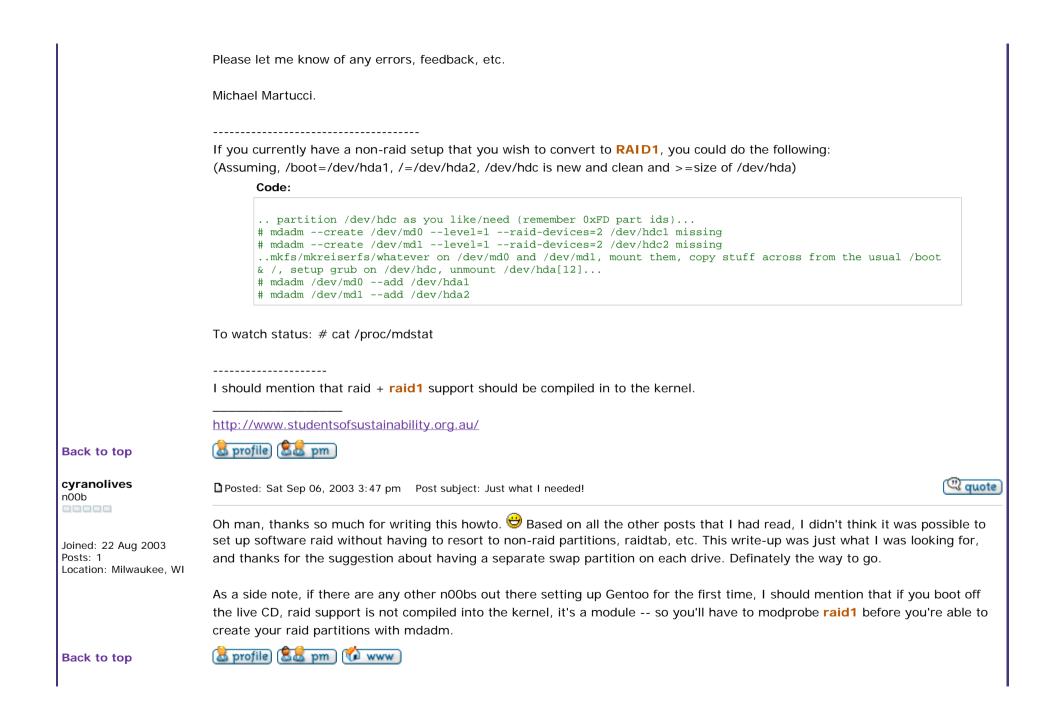

Now you should be able to reboot your system and play!
==

Please let me know of any errors, feedback, etc.

Michael Martucci.

If you currently have a non-raid setup that you wish to convert to **RAID1**, you could do the following:
(Assuming, /boot=/dev/hda1, /=/dev/hda2, /dev/hdc is new and clean and >=size of /dev/hda)

Code:

```
.. partition /dev/hdc as you like/need (remember 0xFD part ids)...
# mdadm --create /dev/md0 --level=1 --raid-devices=2 /dev/hdc1 missing
# mdadm --create /dev/md1 --level=1 --raid-devices=2 /dev/hdc2 missing
..mkfs/mkreiserfs/whatever on /dev/md0 and /dev/md1, mount them, copy stuff across from the usual /boot
& /, setup grub on /dev/hdc, unmount /dev/hda[12]...
# mdadm /dev/md0 --add /dev/hda1
# mdadm /dev/md1 --add /dev/hda2
```

To watch status: # cat /proc/mdstat

I should mention that raid + **raid1** support should be compiled in to the kernel.
_____
http://www.studentsofsustainability.org.au/

Back to top   [profile] [pm]

---

cyranolives
n00b

Joined: 22 Aug 2003
Posts: 1
Location: Milwaukee, WI

[ ] Posted: Sat Sep 06, 2003 3:47 pm   Post subject: Just what I needed!   [quote]

Oh man, thanks so much for writing this howto. 😃 Based on all the other posts that I had read, I didn't think it was possible to set up software raid without having to resort to non-raid partitions, raidtab, etc. This write-up was just what I was looking for, and thanks for the suggestion about having a separate swap partition on each drive. Definately the way to go.

As a side note, if there are any other n00bs out there setting up Gentoo for the first time, I should mention that if you boot off the live CD, raid support is not compiled into the kernel, it's a module -- so you'll have to modprobe **raid1** before you're able to create your raid partitions with mdadm.

Back to top   [profile] [pm] [www]

Exci
Apprentice
□□□□□

Joined: 12 Jul 2002
Posts: 264
Location: The Netherlands, Zoetermeer

Back to top

Posted: Tue Dec 23, 2003 4:58 am   Post subject: quote

> **Quote:**
>
> > **Code:**
> >
> > ```
> > # For booting with disc 1 kernel, if (hd0,0)/bzImage is unreadable
> > title  GNU/Linux (hd1,0)
> > kernel (hd1,0)/bzImage root=/dev/md1
> > ```

it's hdc, so shouldn't it be hd2,0 ?

[profile] [pm] [email]

BackSeat
Apprentice
□□□□□

Joined: 12 Apr 2002
Posts: 242
Location: Reading, UK

Back to top

Posted: Tue Dec 23, 2003 11:28 am   Post subject: Re: Just what I needed!   quote

> **cyranolives wrote:**
>
> > thanks for the suggestion about having a separate swap partition on each drive. Definately the way to go.

Not "definitely". The original poster pointed this out as well, but think about why you want a RAID system. One of the key advantages of **RAID1** is that if a disk dies the system will carry on working. If, rather than RAID the swap partition, you implement two partitions, then if a disk dies and it is actively being used as swap, the system will crash. You have obviated one of the main advantages of RAID. Just make both swap partitions a single md device, and swap to that. In that way, if either disk fails you won't even notice (of course you should have monitoring software to tell you about it, but the system won't crash).

BS

[profile] [pm]

---

carpman
l33t
□□□□□

Joined: 20 Jun 2002
Posts: 676

Posted: Fri Aug 27, 2004 3:45 am   Post subject: quote

Hello, this nice howto but like all raid howtos i have read it does not address howto acces a raid system that has failed so you can attampt to repair it, see my current problem below.

The other thing that most fail to make note of is how many I/O channels you have, you can build a raid system but if it is on same channel then if one drive goes down it will take you whole system down.

Small home server built with software raid, which has died 🙄

The raid is 3 scsi drives on single scsi channel.

On boot it builds md0 (boot) in raid 1 ok but md1 (/) raid5 fails:

Code:

Reiserfs: md1: warning: sh-2006: read_super_block: bread failed (dev md1, block2, size 4096)

Reiserfs: md1: warning: sh-2006: read_super_block: bread failed (dev md1, block16, size 4096)

VFS: Cannot open root device "md1" or md1
please append a correct "root=" boot option
Kernel panic: VFS: Unable to mount root fs on md1

Now i gathered that my root partiton on md1 is not accessible due to fs error, so booted with livecd to try and sort things out.

Once booted i loaded raiddriver

Code:

modprobe md

I then downloaded a backup of raidtab:

Code:

cd /etc
wget http://www.myserver.net/raidtab

nano -w /etcc/raidtab

Raidtab checked out ok, so i checked partitions were still there:

Code:

cfdisk

Again things appeared ok.

Now i tried to assemble damaged raid array

Code:

mdadm --assemble /dev/sda5 /dev/sdb5 /dev/sdc5 /dev/md1
mdadm: /dev/sda5 does not appear to be an md device

Now i am not sure if i should go through the making raid process as i don't want to destroy data?

Code:

mkraid /dev/md1

So do i have to do the mkraid process?

If so once i have done this do i run reiserfsk on umounted array?

I have booted into the scsi controller and done verify disk, all comes back ok.

Have also booted with knoppix and it see all partitions. 🙁
_____
Gigabyte GA K8NSNXP - nforce3
Amd 2800 64bit cpu
2 x Maxtor diamond 40gb GigaRaid (ITE) raid 0 array
LeadTek 6800 128mb
Plextor 12/10/32A Burner
DVD 16x
750mb Kingston 333MHz ram

Back to top   [profile] [pm]

Display posts from previous: | All Posts ▾ | | Oldest First ▾ | | Go |

[new topic] [post reply]   **Gentoo Forums Forum Index** -> **Documentation, Tips & Tricks**   All times are GMT - 5 Hours

Page 1 of 1

| ▾ | Jump to: | Documentation, Tips & Tricks ▾ | | Go |

Links: forums.gentoo.org | www.gentoo.org | bugs.gentoo.org | forum-mods@gentoo.org