Linux−RAID FAQ

Table of Contents

Linux−RAID FAQ

Gregory Leblanc

gleblanc (at) cu−portland.edu

Revision History

Revision v0.0.9	9 October 2000	Revised by: gml
Updates to the location of the patches, and a couple of other things which I can't remember.		
Revision v0.0.8	6 September 2000	Revised by: gml
The info/welcome message on vger.kernel.org has a pointer to this FAQ. New section on recovery, and fixed a few markup things.		

This is a FAQ for the Linux−RAID mailing list, hosted on vger.kernel.org. vger.rutgers.edu is gone, so don't bother looking for it. It's intended as a supplement to the existing Linux−RAID HOWTO, to cover questions that keep occurring on the mailing list. PLEASE read this document before your post to the list.

---

1. General

1.1. Where can I find archives for the linux−raid mailing list?

My favorite archives are at http://www.geocrawler.com/lists/3/Linux/57/0/.

Other archives are available at http://marc.theaimsgroup.com/?l=linux−raid&r=1&w=2

Another archive site is http://www.mail−archive.com/linux−raid@vger.rutgers.edu/

1.2. Where can I find the latest version of this FAQ?

The latest version of this FAQ will be available from the LDP website at http://www.LinuxDoc.org/FAQ/. As soon as I get my server at home fixed I'll make it available there as well.

1.3. What sorts of things does this list cover?

Well, obviously this list covers RAID in relation to Linux. Most of the discussions are related to the raid code that's been built into the Linux kernel. There are also a few discussions on getting hardware based RAID controllers working using Linux as the operating system. Any and all of these discussions are valid for this list.

2. Kernel

2.1. I'm running *`[insert your linux distribution here]`*. Do I need to patch my kernel to make RAID work?

Well, the short answer is, it depends. Some distributions are using the RAID 0.90 patches, while others leave the kernel with the older md code. Unfortunately, I don't have a list of which distributions have which kernels. If you'd like to maintain such a list, please email me `<<gleblanc@cu−portland.edu>>` as well as the linux−raid mailing list.

If you download a 2.2.x kernel from ftp.kernel.org, then you will need to patch your kernel.

2.2. How can I tell if I need to patch my kernel?

That depends on which kernel series you're using. If you're using the 2.4.x kernels, then you've already got the latest RAID code that's available. If you're running 2.2.x, see the following instructions on how to find out.

The easiest way is to check what's in `/proc/mdstat`. Here's a sample from a 2.2.x kernel, *with* the RAID patches applied.

```
[gleblanc@grego1 gleblanc]$ cat /proc/mdstat
Personalities : [linear] [raid0] [raid1] [raid5] [translucent]
read_ahead not set
unused devices: 60;none62;
```

If the contents of `/proc/mdstat` looks like the above, then you don't need to patch your kernel.

The "Personalities" line in your kernel may not look exactly like the above, if you have RAID compiled as modules. Most distributions will have RAID compiled as modules to save space on the boot diskette. If you're not using any RAID sets, then you will probably see a blank space at the end of the "Personalities" line, don't worry, that just means that the RAID modules aren't loaded yet.

Here's a sample from a 2.2.x kernel, *without* the RAID patches applied.

```
13;[root@serek ~]# cat /proc/mdstat
Personalities : [1 linear] [2 raid0]
read_ahead not set
md0 : inactive
md1 : inactive
md2 : inactive
md3 : inactive
```

If your `/proc/mdstat` looks like this one, then you need to patch your kernel.

2.3. Where can I get the latest RAID patches for my kernel?

The patches for the 2.2.x kernels up to, and including, 2.2.13 are available from ftp.kernel.org. Use the kernel patch that most closely matches your kernel revision. For example, the 2.2.11 patch can also be used on 2.2.12 and 2.2.13.

The patches for 2.2.14 and later kernels are at http://people.redhat.com/mingo/raid−patches/. Use the right patch for your kernel, these patches haven't worked on other kernel revisions yet. Please use something like wget/curl/lftp to retrieve this patch, as it's easier on the server than using a client like Netscape. Downloading patches with Lynx has been unsuccessful for me; wget may be the easiest way.

> **Note:** These patches should also be available from ftp://ftp.kernel.org/pub/linux/kernel/people/mingo/raid−patches/ I could not find them on my local mirror, but please check yours before using the main kernel.org site. You can find a list of the local mirrors at http://www.kernel.org/mirrors/.

2.4. How do I apply the patch to a kernel that I just downloaded from ftp.kernel.org?

First, unpack the kernel into some directory, generally people use `/usr/src/linux`. Change to this directory, and type **patch −p1 < /path/to/raid−version.patch**.

On my RedHat 6.2 system, I decompressed the 2.2.16 kernel into `/usr/src/linux−2.2.16`. From `/usr/src/linux−2.2.16`, I type in **patch −p1 < *`/home/gleblanc/raid−2.2.16−A0`***. Then I rebuild the kernel using **make menuconfig** and related builds.

2.5. What kind of drives can I use RAID with? Do only SCSI or IDE drives work? Do I need different patches for different kinds of drives?

Software RAID works with any block device in the Linux kernel. This includes IDE and SCSI drives, as well as most harware RAID controllers. There are no different patches for IDE drives vs. SCSI drives.

3. RAIDtools

3.1. Why are the RAIDtools at http://people.redhat.com/mingo/raid−patches/ labeled *dangerous*, and if they're dangerous, should I use them?

The tools are labeled *dangerous* because the RAID code isn't part of the "stable" Linux kernel.

The tools found at the above URL are the latest and greatest. You *should* use these tools with the kernel patches from the same location.

3.2. Are there any tools other than the *dangerous* ones available?

No, the *dangerous* tools available from http://people.redhat.com/mingo/raid−patches/ are the most current tools to use. *Everyone using RAID with the patches at the above location should be using these dangerous tools.*

4. Disk Failures and Recovery

4.1. How can I tell if one of the disks in my RAID array has failed?

A couple of things should indicate when a disk has failed. There should be quite a few messages in `/var/log/messages` indicating errors accessing that device, which should be a good indication that something is wrong.

You should also notice that your `/proc/mdstat` looks different. Here's a snip from a good /proc/mdstat

```
[gleblanc@grego1 gleblanc]$ cat /proc/mdstat
Personalities : [linear] [raid0] [raid1] [raid5] [translucent]
read_ahead not set
md0 : active raid1 sdb5[0] sda5[1] 32000 blocks [2/2] [UU]
unused devices: 60;none62;
```

And here's one from a `/proc/mdstat` where one of the RAID sets has a missing disk.

```
[gleblanc@grego1 gleblanc]$ cat /proc/mdstat
Personalities : [linear] [raid0] [raid1] [raid5] [translucent]
read_ahead not set
md0 : active raid1 sdb5[0] sda5[1] 32000 blocks [2/1] [U_]
unused devices: 60;none62;
```

I don't know if `/proc/mdstat` will reflect the status of a HOT SPARE. If you have set one up, you should be watching `/var/log/messages` for any disk failures. I'd like to get some logs of a disk failure, and `/proc/mdstat` from a system with a hot spare.

4.2. So my RAID set is missing a disk, what do I do now?

RAID generally doesn't mark a disk as bad unless it is, so you probably need a new disk. Most disks have a 3 year warranty, but some good SCSI hard drives may have a 5 year warranty. See if you can get the manufacturer to replace the failed disk for you.

When you get the new disk, power down the system, and install it, then partition the drive so that it has partitions the size of your missing RAID partitions. After you're finished partitioning the disk, use the command **raidhotadd** to put the new disk into the array and begin reconstruction. See Chapter 6 of the Software RAID HOWTO for more information.

4.3. dmesg shows "md: serializing resync, md4 has overlapping physical units with md5". What does this mean?

In that message "physical units" refers to disks, and not to blocks on the disks. Since there is more than 1 RAID array that needs resyncing on a disk, the RAID code is going to sync md4 first, and md5 second, to avoid excessive seeks (also called thrashing), which would drastically slow the resync process.