

Re: Reproducible OOPS with MD RAID-5 on 2.6.0-test11

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2003-12/1053.html>

From: Theodore Ts'o (tytso_at_mit.edu)

Date: 12/05/03

Date: Fri, 5 Dec 2003 01:55:22 -0500

To: Simon Kirby <sim@netnation.com>

On Wed, Dec 03, 2003 at 08:31:06PM -0800, Simon Kirby wrote:

>
> *Without the patches, the box gets as far as assembling the array and*
> *activating it, but dies on "mke2fs". Running mke2fs through strace shows*
> *that it stops during the early stages, before it even tries to write*
> *anything. mke2fs appears to seek through the whole device and do a bunch*
> *of small reads at various points, and as soon as it tries to read from an*
> *offset > 2 TB, it hangs.*

It sounds like mke2fs tried using BLKGETSIZE ioctl, but given that this returns the number of 512 byte sectors in a device in a 4 byte word, the BLKGETSIZE ioctl quite rightly threw up its hands and said, "sorry, I can't tell you the correct size."

The mke2fs fell back to its backup algorithm, which uses a modified binary search to find the size of the device. It started to see if the device was at least 1k, and checks to see if the device is at least 2k, 4k, 8k, 16k, 32k, 64k, 128k, etc. So it sounds like it's dieing when it tries to seek past 2TB using llseek().

It would probably be worthwhile to write a little test program which opens the disk, llseeks to 2TB+1, and then tries reading a byte. If that fiailes, then there's definitely a bug somewhere in the device driver....

– Ted

–

To unsubscribe from this list: send the line "unsubscribe linux-kernel" in the body of a message to majordomo@vger.kernel.org

More majordomo info at <http://vger.kernel.org/majordomo-info.html>

Please read the FAQ at <http://www.tux.org/lkml/>