

Re: A little story of failed raid5 (3ware 8000 series)

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/stable/2007-08/msg00345.html>

- *From:* Tom Judge <tom@xxxxxxxxxxxx>
 - *Date:* Sat, 25 Aug 2007 12:05:05 +0100
-

Tom Samplonius wrote:

----- "Artem Kuchin" <matrix@xxxxxxxxxxxx> wrote: ...

But i don't understand how and why it happened. ONLY 6 hours ago (a night before) all those files were backed up fine w/o any read error. And now, right after replacing the driver and starting rebuild it said that there are bad sectors all over those file. How come?

What happened to you was an extremely common occurrence. You had a disk develop a media failure sometime ago, but the controller never detected it, because that particular bad area was not read. Your backups worked because they never touched this portion of the disk (ex. empty space, meta data, etc). And then another drive developed a electronics failure, which is instantly detected, putting the array into a degraded mode. When you did a rebuild onto a replace drive, the controller discovered that there was a second failed disk, and this is unrecoverable.

3ware controllers can recover from this situation, all you need to do is tell the controller not to verify the source data. This is a little dangerous but it has saved me in the past where 1 drive died in a raid 10 array and 2 of the 3 remaining drives had surface defects. The trick was to replace each drive 1 at a time and rebuild without data verification. After 10 painful hours the array was rebuild with out any noticeable data corruption.

RAID, of any level, isn't magic. It is important to understand how it works, and realize that drives can passively fail. BTW, if you were using RAID1 or RAID10, you would likely have had the same problem (well, RAID10 can survive some double-disk failures). RAID6 is the only RAID level that can survive failure of any two disks.

This is not all true RAID 1 can survive multiple disk failures as it has the storage capacity of 1 spindle and can tolerate the failure of N-1

Re: A little story of failed raid5 (3ware 8000 series)

spindles where N is the number of spindles in the mirror set. This also is kind of true in RAID 10, the more spindles in your mirror sets the more chance you have of being able to survive multiple failures in the array (Say use 6 disks in 2 3 disk mirror sets striped together).

The real solution is RAID scrubbing: a low level background process that reads every sector of every disk. All of the real RAID systems do this (usually scheduled weekly, or every other week). Most 3ware RAID card don't have this feature.

So rather than not using RAID5 or RAID6 again, you should just not use 3ware anymore.

If you use the 3dm2 management interface you can schedule verify and rebuild tasks to run on a regular basis. I think that 7500 series controllers can do this, 9500 and 9550's definitely can.

We have 50+ systems that are using 3ware cards (7500-9550 4 and 8 channel models) with 200+ spindles in use (no host spares unfortunately) and drives in that pool failing on average around once a month. We have only ever had trouble recovering from failed drives on 7500 series controllers that have been in production for a reasonably long time.

I don't think that you are justified in your slagging off of 3ware controllers.

Tom

freebsd-stable@xxxxxxxxxxx mailing list

<http://lists.freebsd.org/mailman/listinfo/freebsd-stable>

To unsubscribe, send any mail to "freebsd-stable-unsubscribe@xxxxxxxxxxx"