Why RAID 5 stops working in 2009

By [Robin Harris](#) for [Storage Bits](#) | July 18, 2007 -- 06:18 GMT (23:18 PDT)

The storage version of Y2k? No, it's a function of capacity growth and RAID 5's limitations. If you are thinking about SATA RAID for home or business use, or using RAID today, you need to know why.

RAID 5 protects against a single disk failure. You can recover all your data if a single disk breaks. The problem: once a disk breaks, there is another increasingly common failure lurking. And in 2009 it is highly certain it will find you.

Disks fail While disks are incredibly reliable devices, they do fail. Our best data - from CMU and Google - finds that over 3% of drives fail each year in the first three years of drive life, and then failure rates start rising fast.

With 7 brand new disks, you have ~20% chance of seeing a disk failure each year. Factor in the rising failure rate with age and over 4 years you are almost certain to see a disk failure during the life of those disks.

But you're protected by RAID 5, right? Not in 2009.

Reads fail SATA drives are commonly specified with an unrecoverable read error rate (URE) of 10^{14}. Which means that once every 100,000,000,000,000 bits, the disk will very politely tell you that, so sorry, but I really, truly can't read that sector back to you.

One hundred trillion bits is about 12 terabytes. Sound like a lot? Not in 2009.

Disk capacities double Disk drive capacities double every 18-24 months. We have 1 TB drives now, and in 2009 we'll have 2 TB drives.

With a 7 drive RAID 5 disk failure, you'll have 6 remaining 2 TB drives. As the RAID controller is busily reading through those 6 disks to reconstruct the data from the failed drive, it is almost certain it will see an URE.

So the read fails. And when *that* happens, you are one unhappy camper. The message "we can't read this RAID volume" travels up the chain of command until an error message is presented on the screen. 12 TB of your carefully protected - you thought! - data is gone. Oh, you didn't back it up to tape? Bummer!

So now what? The obvious answer, and the one that storage marketers have begun trumpeting, is RAID 6, which protects your data against 2 failures. Which is all well and good, until you consider this: as drives increase in size, any drive failure will *always* be accompanied by a read error. So RAID 6 will give you no more protection than RAID 5 does now, *but you'll pay more anyway* for extra disk capacity and slower write performance.

Gee, paying more for less! I can hardly wait!

The Storage Bits take Users of enterprise storage arrays have less to worry about: your tiny costly disks have less capacity and thus a smaller chance of encountering an URE. And your spec'd URE rate of 10^{15} also helps.

There are some other fixes out there as well, some fairly obvious and some, I'm certain, waiting for someone much brighter than me to invent. But even today a 7 drive RAID 5 with 1 TB disks has a 50% chance of a rebuild failure. RAID 5 is reaching the end of its useful life.

Update: I've clearly tapped into a rich vein of RAID folklore. Just to be clear I'm talking about a failed drive (i.e. all sectors are gone) plus an URE on another sector during a rebuild. With 12 TB of capacity in the remaining RAID 5 stripe and an URE rate of 10^{14}, you are highly likely to encounter a URE. Almost certain, if the drive vendors are right.

As well-informed commenter **Liam Newcombe** notes:

> "*The key point that seems to be missed in many of the comments is that when a disk fails in a RAID 5 array and it has to rebuild there is a significant chance of a non-recoverable read error during the rebuild (BER / UER). As there is no longer any redundancy the RAID array cannot rebuild, this is not dependent on whether you are running Windows or Linux, hardware or software RAID 5, it is simple mathematics. An honest RAID controller will log this and generally abort, allowing you to restore undamaged data from backup onto a fresh array.*"

Thus my comment about hoping you have a backup.

Mr. Newcombe, just as I was beginning to like him, then took me to task for stating that "RAID 6 will give you no more protection than RAID 5 does now". What I had hoped to communicate is this: in a few years - if not 2009 then not long after - all SATA RAID failures will consist of a disk failure + URE.

RAID 6 will protect you against this quite nicely, just as RAID 5 protects against a single disk failure today. In the future, though, you will *require* RAID 6 to protect against single disk failures + the inevitable URE and so, effectively, RAID 6 in a few years will give you no more protection than RAID 5 does today. This isn't RAID 6's fault. Instead it is due to the increasing capacity of disks and their steady URE rate. RAID 5 won't work at all, and, instead, RAID 6 will replace RAID 5.

Originally the developers of RAID suggested RAID 6 as a means of protecting against 2 disk failures. As we now know, a single disk failure means a second disk failure is much more likely - see the CMU pdf [Disk Failures in the Real World: What Does an MTTF of 1,000,000 Hours Mean to You?](#) for details - or check out my synopsis in [Everything You Know About Disks Is Wrong](#). RAID 5 protection is a little dodgy today due to this effect and RAID 6 - in a few years - won't be able to help.

Finally, I recalculated the AFR for 7 drives using the 3.1% AFR from the CMU paper, using the formula suggested by a couple of readers - 1-96.9 ^# of disks - and got 19.8%. So I changed the ~23% number to ~20%.

Comments welcome, of course. And I got home despite a blow out on the Scottsdale's 101N in 110 degree heat. I thought of it as a Bikram Tire Changing Asana.

Topic: [Hardware](#)

About [Robin Harris](#)

Robin Harris has been messing with computers for over 30 years and selling and marketing data storage for over 20 in companies large and small.

[Google+](#) [Contact](#) [Disclosure](#)

Kick off your day with ZDNet's [daily email newsletter](#). It's the freshest tech news and opinion, served hot. [Get it.](#)