Probability of a RAID

by M. E. Kabay, PhD, CISSP
Associate Professor, Information Assurance
Norwich University, Northfield VT

In the past months, I have been having a terrible time with my Compaq Presario 8000T computer. It was built for me by HP-Compaq in January 2004 with quite a nice setup: dual 3.2 GHz Pentium IV processors (ordered by mistake because I didn’t notice that the base unit was a dual-processor machine – duhhhh), 2 GB RAM (definitely ordered on purpose), very nice Nvidia graphics board for my two 19-inch monitors, two DVD drives (a read-only and a burner), and a classic SoundBlaster card. The significant feature for the purpose of this article is that I also ordered two 160 GB drives in a RAID 1 array.

The system worked fine until October 2004, when it began crashing with blue screens of death; the frequency of crashes increased to several per day by January. At one point the system crashed so hard it wouldn’t boot at all any more. HP took the system back via FedEx and sent it back a week later with assurances that it was fine.

It wasn’t.

It crashed immediately after being taken out of the shipping box; I sent it back an hour after receiving it.

It returned after another week at the HP repair center. It crashed again. The techs finally realized that all my problems were due to an old Nvidia driver; after I downloaded the current version the system seemed to settle down considerably. Too bad they didn’t think of that a long time ago, n’est-ce pas?

I spent from 08:00 to 16:00 today (as I write this) loading files back onto the hard drive. By accident, I noticed that the free space on the C: drive was now 289 GB – and my heart sank.

Sure enough, the technicians at HP had, for completely unknown reasons, converted my RAID 1 array to a RAID 0 array. Now, some of you may be wondering (1) what is a RAID array; (2) what’s the difference between RAID 0 and RAID 1; (3) who cares?

- RAID stands for “Redundant Array of Independent Disks.” These arrays can be set up in a variety of ways.

- RAID 0 improves performance by “striping,” in which data are written alternately to cylinders of two or more disk drives. With multiple disk heads reading and writing data concurrently, input/output (I/O) performance improves noticeably.

- RAID 1 improves resistance to disk failure (i.e., provides fault tolerance) by making bit-for-bit copies of data from a main drive to a mirror drive. If the main drive fails, the mirror drive(s) continue(s) to provide for I/O while the defective drive is replaced. Once the new, blank drive is in place, array management software can rebuild the image on the
new drive. The frequency of mirroring updates can be defined through the management software to minimize performance degradation.

- Other RAID modes are available for increased performance, fault tolerance and both at once.

But why would I care about my drives being converted into RAID 0 from RAID 1?

The first problem is that it was not possible for me to convert the RAID 0 back to RAID 1 at all; that has to be done at the factory, and it necessitates losing all the data I had laboriously copied back from my DVDs to the hard disks.

More important, though, is that I chose RAID 1 for safety rather than RAID 0 for performance. If either of the disks fails in a RAID 0 array, then the entire array fails. That means that the likelihood of failure increases rapidly as the number of disks in the array rises. To be precise, the failure rate is calculated as follows:

Let the expected failure rate of a single disk drive be “p” (considering only drive-specific problems, not things like power failures).

Then the probability that a single drive will _not_ fail is \((1 − p)\).

So the probability that all “n” drives in a RAID 0 array will not fail at the same time due to individual drive problems is \((1 − p)^n\).

So the probability that at least one of the n drives in a RAID 0 array will indeed fail is \([1 − (1 − p)^n]\) and this number rises rapidly as a function of n.

And that’s why my Compaq computer is back on its way to the HP Repair Center as I write this article.

I’m putting a copy of this into the box and hope the technicians read it. Just think what I’ll write if they screw up again.

* * *

A Master’s degree in the management of information assurance in 18 months of online study from Norwich University – see <http://www3.norwich.edu/msia> for details.

M. E. Kabay, PhD, CISSP is Associate Professor in the Division of Business and Management at Norwich University in Northfield, VT. Mich can be reached by e-mail at <mailto:mkabay@norwich.edu>; Web site at <http://www.mekabay.com/index.htm>.

Copyright © 2005 M. E. Kabay. All rights reserved.

Permission is hereby granted to Network World to distribute this article at will, to post it without limit on any Web site, and to republish it in any way they see fit.