

June 2010 Newsletter

Speed or Security?

This month's issue is a follow-up from May's High Performance Computing theme, and I hope to explain in more detail a term that I used: "RAID Hard Disks". RAID hard disk systems are a topic that can get pretty technical, and there wasn't a lot of space to put it in last month. I hope to do it justice this month.

Missed or lost a newsletter? You can always get back issues in Adobe PDF format online at my website - www.MedOfficeSystems.com/newsletter_tips.htm. If you have any questions about computer specifications, please call me at 630-852-1736 or 630-373-7429.

RAID Hard Disks

RAID is an acronym for "Redundant Array of Inexpensive Disks" – and the basic concept is 1) prevent data loss, or 2) speed up file read/write activity by using special multi-disk systems rather than the traditional single hard disk computer. RAID systems are specified by levels from 0 to 10. The two issues that RAID systems address are:

1. Data Loss Prevention

Quick question: how many typical hard disks will fail? Answer: like basement sump pumps, they ALL will fail eventually. Given that certainty, it makes sense to have a backup. Long-time clients and readers of this newsletter know that I am a big proponent of frequent, multiple backups. But even frequent backups have one limitation: by their periodic nature, they capture data only at a point in time. After that point has passed, any new data is subject to loss until the next backup is made, typically every 24 hours. What if you saved an important file just before lunch, and your hard disk crashes in the afternoon?

2. Maximum Speed

Some applications require ultra-fast hard disk speed to function properly, such as video editing, streaming media, and some high-volume transactional data. Even the fastest single hard disks cannot keep up. In fact, one of the slowest data handlers in your computer is the hard disk. So how do movie producers edit and produce movies, and can you get the same performance for your home computer?

RAID to the Rescue

RAID systems address both of these issues using two very different methodologies, but both methodologies have a common factor: multiple disks and multi-disk controllers.

Data Loss Prevention: RAID 1 and RAID 5

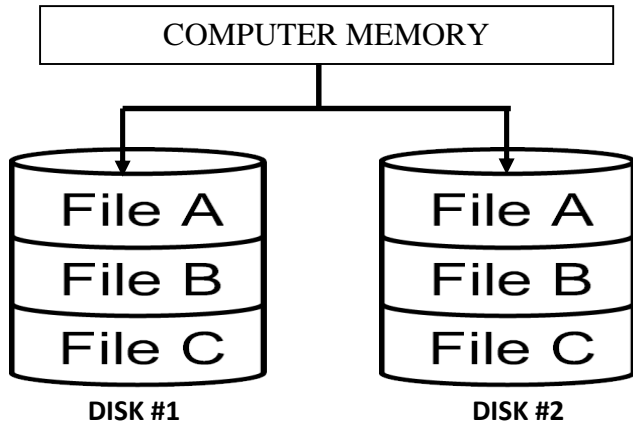
There are many levels of RAID systems, but the two that specifically are targeted to prevent data loss are RAID 1 and RAID 5. A RAID 1 system contains two disks of identical capacity (ie: both are 500GB) and the controller card is configured so that every write operation (save a file) is recorded on both disks. The probability of both disks failing at the same time is very low, thus almost completely eliminating data loss. An extra benefit of RAID systems is if one disk does fail, the computer can continue to operate until the failed unit can be replaced.

In a RAID 5 system, a minimum of three hard disks and controller card are set so that file write operations are recorded over three (or more) hard disks. The additional disk(s) give RAID 5 a speed advantage over RAID1.

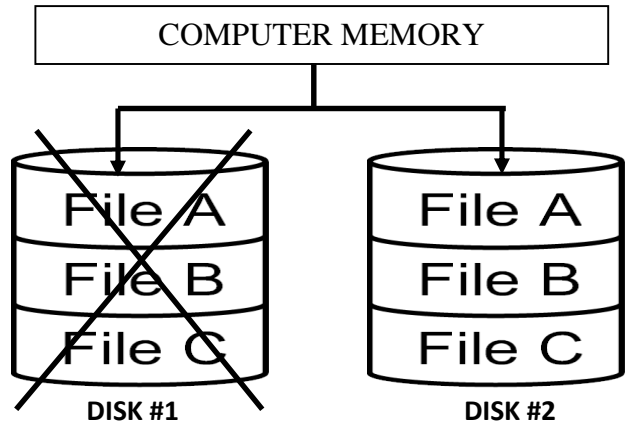
[continued]

RAID 1 Illustration:

Before: Each file saved simultaneously on both disks

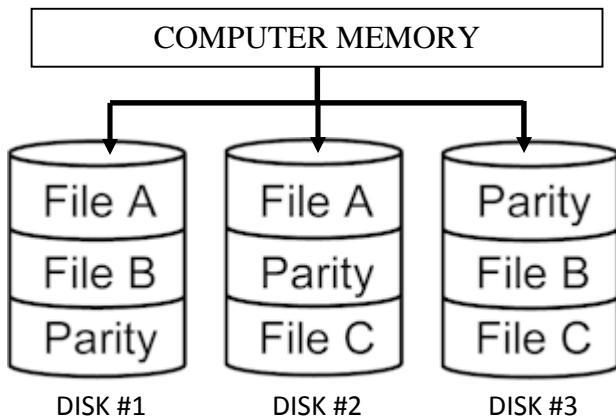


After: Disk #1 Fails, Disk #2 continues working

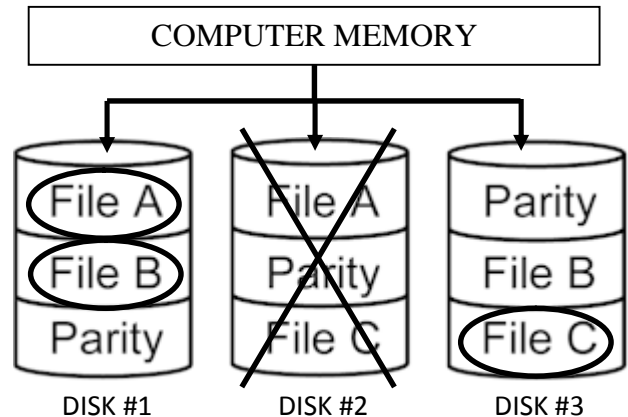


RAID 5 Illustration:

Files saved across three disks, plus parity check data



Disk #2 Fails: Disks 1 & 3 continue working, no files lost



Maximum Speed: RAID 0

Some tasks need ultimate speed, and data security is secondary. This is true for special applications such as video editing or computer gaming. In the first scenario, the data is temporary as the source and final destination of the data are usually on different media. An example is video editing: the source file is on the video camera memory, and when editing is complete, the finished product is burned to a DVD. In the gaming scenario, the game data is meaningless after the game is over, and split-second data movement during the game is more important.

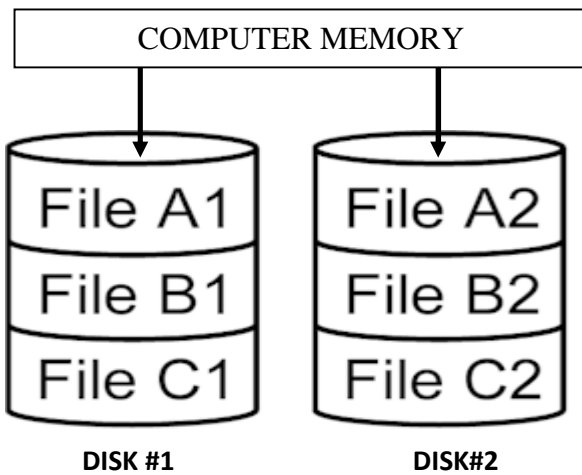
[continued]

RAID 0 (“RAID Zero”) gains its speed by paralleling the read/write process. A single file being saved is split in half, the first half saves on drive #1 while the second half saves on drive#2, cutting overall save time in half. The same is true for file retrieval: both halves of the file are read at the same time, and assembled back together in the computer’s memory. The minimum number of disks in a RAID0 is two, even further speed gains can be achieved by adding 4, 6, 8 or more disks, thus spreading the file save/open activity to more disks, in 1/8 the time.

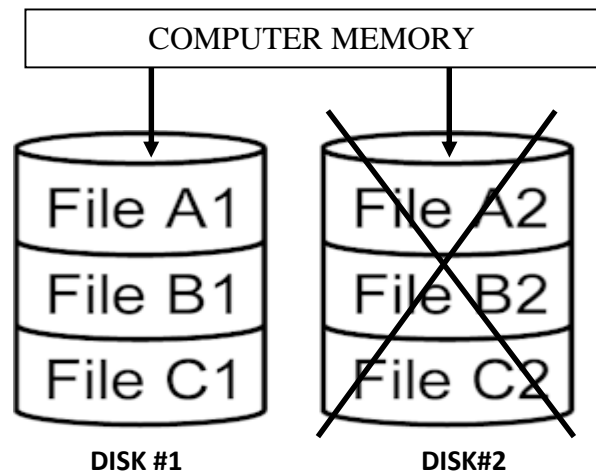
The flip side of having all this speed available is that should a hard disk fail, at least part of every file is lost, making most files completely unusable. Therefore, RAID 0 is best used for temporary, non-critical “working” storage. After the work is complete, the data should be backed up or moved to a more-permanent storage media as soon as possible.

RAID-0 Illustration:

Each file is split – half goes to each disk – very fast!



Disk #2 fails – ½ of each file is lost – Uh-Oh!



Summary:

For the best data security use RAID 1, or for slightly more speed, RAID 5. Both guard against data loss occurring between backups and allow the machine to continue working until downtime can be scheduled for disk replacement. For ultimate data save and retrieve speed, use RAID 0. There are more RAID levels available: www.acnc.com has all the RAID definitions and excellent RAID simulations.

Finally, each of the above RAID methodologies have one flaw: they do not guard against total loss from theft or physical destruction of the machine due to fire, flood or other disaster. A good, multiple copy backup plan that includes offsite storage is still essential.

Questions or comments? Call or email me, I'll be happy to discuss questions or provide more details.