

# RAID

## Disk arrays

The Redundant Array of Independent Disk Drives technology, also known as RAID, is an increasingly attractive technology for CCTV. It uses several hard drives in an array so that their data transfer rates can be bundled and data security increased. Depending on the number of drives present, you can choose different modes to fulfil increase data security (in case a drive crashes), increase performance, or both. There are several RAID standards explained in this article.

The cry for more performance is as old as the computer itself. Especially when speedy action is called for, the hard drive has more influence than many users realize.

Various operating systems, such as Windows, Unix, Mac, Linux, etc, offer ways to optimise performance, but there are limits to the physical potential of any hard drive. Using several hard drives simultaneously helps to alleviate this problem.

The technology required for this is known as RAID (Redundant Array of Independent Disk Drives). To put it simply, several hard drives are used jointly so that their data transfer rates can be bundled and data security increased. Depending on the number of drives present, you can choose different modes to fulfil the following purposes: increase data security in case a drive



crashes, increase performance, or both. For a functioning drive array, you'll need the right RAID controller that will support the selected mode, in addition to the hard drives. These can be housed in nearly any PC in a free PCI slot, or they may come integrated on the motherboard.

Two things need to be considered in selecting the right hard drives: drive capacity and rotational speed. Today's interfaces correspond exclusively to UltraATA/100 or even UltraATA/133, so they are always fast enough. High rotational speeds allow maximum data transfer rates and minimal access times, but they are accompanied by an increase in both heat and operating noise.



RAID can principally be used with any hard drive; this article will focus on the finer points of this technology and discuss all of the important details.

## Elementary: Access Time and Data Transfer Rate

The next drive generation will offer capacities of over 200 GB per drive for the first time ever - a size that would have been flabbergasting two years ago.

Apart from this, there are two primary factors of interest in terms of a hard drive's performance, the access time and the data transfer rate; both are really quite self-explanatory. Access time specifies how long it takes from the seek request to the hard drive to the actual read process of the desired data. Because the hard drive has to address new sectors on the disk surface continuously in day-to-day operations, this factor becomes even more important, especially if small amounts of data are read or written. The shorter, the better. Seek time, by the way, is also a term that is commonly used at computer stores - it does not include the access time per se and is therefore usually noticeably shorter. Thus, unless you want to deal with an apples-to-oranges comparison, you should really take a good look.

The second factor is the data transfer rate. It mostly depends on the data density (expressed in memory space per surface unit, or in GB per disk) and the rotational speed of the medium. The shorter the distance between the data and the faster they pass the read-write heads, the more data can be read or written per time unit.

In reality, the rotational speed does also have a high impact on the access times, as latency times before the actual access are slashed significantly. Having as much buffer storage as possible (known as the "cache") is yet another performance booster. Cache today generally has a size of 2 MB, but you'll also find models in the IDE sector with up to 8 MB. Hard drive electronics plays a significant role too. The good quality and fast electronics are responsible for executing all

accesses, and their strengths and weaknesses become particularly obvious when several accesses must be responded to simultaneously.

Last but not least, as discussed in the previous article, as well as in the previous issue of *CCTV focus*, the operating system plays a very important role.

Unlike many other system components, a hard drive's performance cannot be increased by tuning or "overclocking" it. You can alter the access speed with small software tools, but this method is primarily a means to reduce the operating noise or to defragment the data and thereby increase the efficiency.



## Wanted: More Capacity

The single most efficient way to effectively expand system storage memory is to use several hard drives. Often it's enough to install an additional hard drive in order to give performance a bit of a boost. Then you can move the bandwidth-hogging Windows swap file (the so-called virtual memory) to the second drive, so that system files can be accessed independently. Furthermore, using applications like Photoshop is really only comfortable if the swap file is not located in the system partition.

Of course, the more elegant and truly performance-boosting method for all application areas is to set up a hard drive array in RAID mode. Here, the existing hard drives are not run independently, but are managed by the RAID controller and described according to a predefined schema.



## So How Does a RAID Array Work?

The array itself is set up in the BIOS of the RAID controller by adding all of the installed hard drives to be part of the array. Depending on the controller, the block size and the total capacity, initialization may

take up to several hours. The selection of the block size deserves special attention. Large blocks ensure maximum data transfer performance, but for mostly small files (smaller than the size of a block), an annoying amount of memory is wasted (this is at least the case for Windows OS's). Hence a block with 64 kB will always need at least those 64 kB - even if the information to be written is clearly of a smaller size.

After the operating system has booted, the new array has to be formatted. Performing a Quick Format under Windows takes no time at all, while the full format checks every single sector and in turn may take quite some time. After the system has re-booted, the RAID array is available as a new drive. Its use is no different from that of a single hard drive.

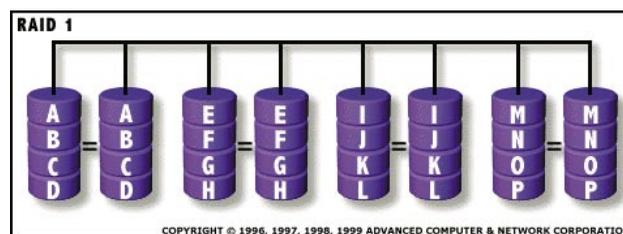
## RAID 0: Striping

Technically speaking, mode 0 doesn't adhere to the principles of a RAID, given the fact that an important factor, data redundancy, does not exist. Hence RAID 0 offers no advantages in terms of security - in fact, on the contrary. All of the data are evenly distributed to all of the existing drives; this array is called a stripe set. This process can best be described with the "zipper method." The benefits are clear: because the data stream can be allocated to all the different drives, the data transfer rate is multiplied by the number of drives. Here the upper limits are the maximum transfer rate per channel (max. 100 MB/s for UltraATA/100), or the maximum bandwidth of the controller on the PCI bus (266 MB/s at 66 MHz / 32-bit PCI). However, in reality this very drastic performance boost comes at the price of higher fault

vulnerability. Instead of one, now all of the RAID drives must work error-free. If even one of those drive crashes, all of the stored data will be lost.

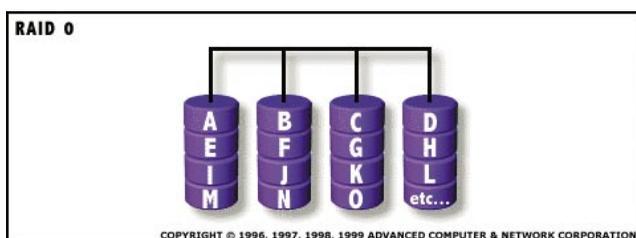
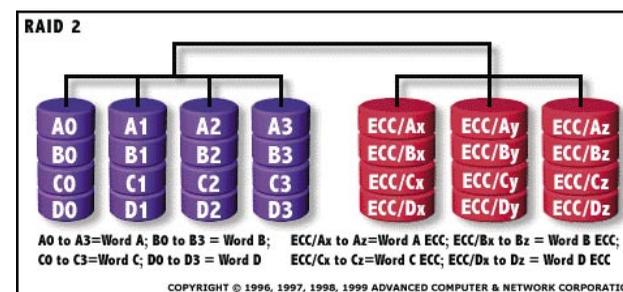
## RAID 1: Mirroring

Mode 1 is basically the complete opposite of RAID 0. The goal here is not to boost performance, but to ensure data security. When reading or writing data, all drives of the array are used simultaneously. Hence, data is written synchronously to two or more drives, which is equivalent to a perfect backup copy - perfect because the data is always 100% up-to-date.



## RAID 2: Striping

Striping is based on the same principle as RAID 0: the stripe set distributes the data to all drives, though not in block form, but, rather, on a bit level. This is necessary because an Error Correcting Code (ECC) is implemented in all transaction data. Additional hard drives are necessary to store the resulting additional volume. If you wanted to guarantee complete data security, you would have to deploy at least ten data disks and four ECC disks. The next level would entail 32 data disks and seven ECC disks. This explains why RAID 2 never caught on.

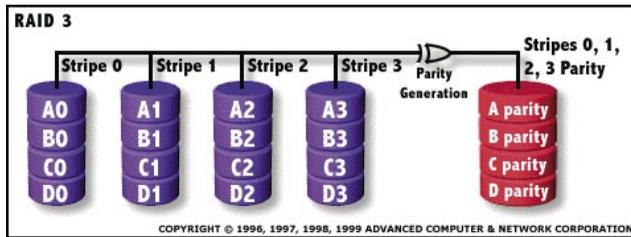


On top of that, performance is only mediocre as multiple access is not possible in bit stripe sets. The higher the number of accesses, and the shorter they are, the more lethargic RAID 2 gets.

### RAID 3: Data Striping, Dedicated Parity

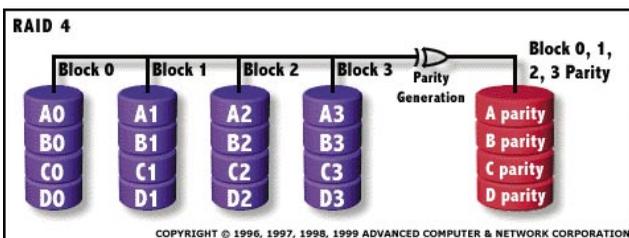
Level 3 incorporates prudent error correction. Data is allocated byte by byte to several hard drives, while the parity data is stored in a separate drive. This is exactly the disadvantage of RAID 3, as the parity drive has to be accessed with every access. So the advantage of RAID, bundling the disk performance by distributing access, is partially offset. RAID 3 needs a minimum of three drives.

This mode requires quite a complex controller, which is why RAID 3, similar to levels 4 and 5, never caught on in the mass market.



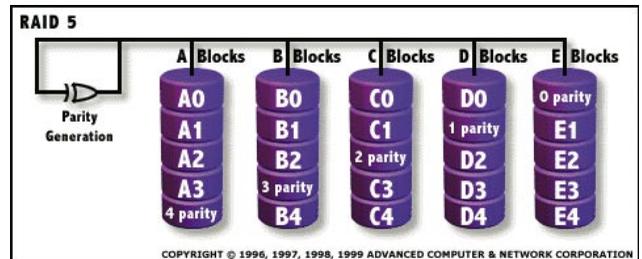
### RAID 4: Data Striping, Dedicated Parity

The technology of RAID Level 4 is similar to that of level 3, except that the individual stripes are not written in bytes, but in blocks. In theory, this should speed things up, but the parity drive still remains the bottleneck.



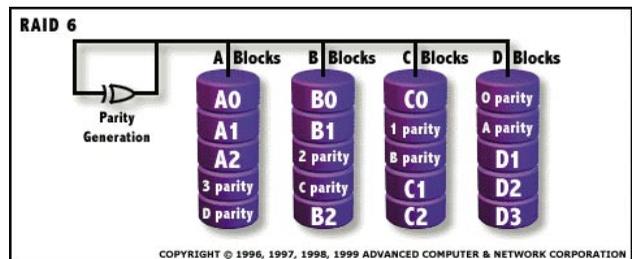
### RAID 5: Distributed Data, Distributed Parity

RAID Level 5 is generally considered the best compromise between data security and performance. Not only the data, but also the parity information, is distributed to all the existing drives. The resulting advantage is that RAID is really only a bit slower than RAID 3. However, failure safety is limited, as only one hard drive can safely crash. At least three hard drives are required in each case.



### RAID 6: Distributed Data, Distributed Parity

With RAID 6, you're really only talking about RAID 5 - except that twice the amount of parity information is stored. Though this cuts down on performance a bit, it allows up to two hard drives to crash. It does require, however, a minimum of five drives.



### RAID: Who with Whom?

So far, we've talked about several hard drives, but haven't gone into detail. You should principally set up all RAID modes with similar hard drives, as only then will you get maximum performance.

However, you can also combine different drives, with the smaller or slower drive being

the determinant drive for the entire array. For example, one 30 GB and two 40 GB hard drives in RAID 0 will give you a total capacity of 90 GB, which is three times the capacity of the smallest hard drive.

The same applies to the combination of an old 40 GB hard drive with 5,400 rpm with a new model with 7,200 rpm. If you were to use two of the slower drives, the performance level would be the same. Replacing the older disk with a second and faster one would increase the performance.

If you want to use several different hard drives together, you have the option of creating what is known as a span array. Another term is JBOD (just a bunch of drives). Here, the hard drives are simply hooked up in series, which results in a useable total capacity, but without, however, an increase in performance or data security.

Another unsettled point is the question of which drives should be hooked up to which IDE channel. If possible, every drive should be connected to its

own channel as a master. On a dual channel controller card, you'd be able to hook up a mere two hard drives. Though using four hard drives (master and slave per channel) drastically increases performance, you'd get even more out of a four channel controller with four master drives.

Another important fact is that only a few of the currently available IDE RAID controllers support the ATAPI protocol. CD-ROM or DVD-ROM drives will not necessarily work with a RAID controller (don't even bother trying in RAID mode).

If you had opted for a RAID level before a disk crash, assigning highest priority to

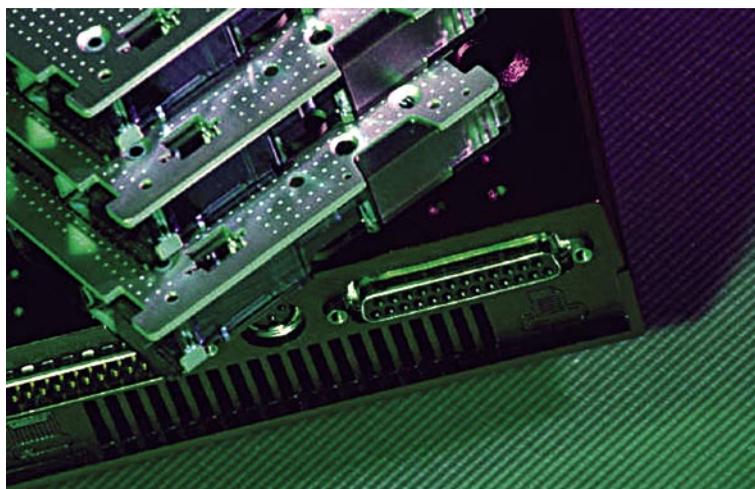
security, then you're in good shape. If you're using RAID Level 1, 3, 4, 5 or 6, the crash of a single hard drive will not affect your existing data. Depending on the controller you're using, the procedure will vary.

Most RAID controllers today notify the user of a crash with a beep and by e-mail (of course, this does not apply to the RAID levels with the system partition, which do not offer any crash protection).

Older or very simple RAID controllers require the computer to be shut down and the defective drive to be replaced. After restarting the system, the user has to go into the BIOS of the RAID controller to initiate the rebuild process.

Practically all of the RAID controllers on the market today - including the simple mod-

els - now master the exchange of defective hard drives without a need to shut down the system, a process called hot swapping. Rebuild takes place automatically, too - you really don't have to do anything yourself any more.



A clever feature is the hot-spare function. Many RAID controllers support an additional drive, which is labelled as a hot spare. If one of the array drives refuse to work, it will be removed from operation and the hot-spare drive will be connected automatically.

In the event that you were using RAID 0 or JBOD and have lost important data, you'll probably never want to use this mode again. Though there's almost always a way to restore your data, it is horrendously expensive. Companies dedicated to data recovery (such as, for example, OnTrack) are able to take hard drives apart and restore most of the data, even after a head crash, fire damage

or other catastrophic events. But be warned: restoring RAID arrays is disproportionately harder than the effort that goes into restoring one single hard drive, which is enormous enough.

### **Conclusion: Only a Backup is Truly Safe!**

It's not just aspects like performance and data security that should be considered; in many cases, enormous amounts of data must be managed and stored - the right approach to tackling this problem is a RAID array with large hard drives. Because expensive SCSI RAID adapters and SCSI hard drives were the only available options just a few years ago, high-capacity arrays were feasible only for very few individual users or companies.

Today, the prices for IDE hard drives with a capacity of 100 GB have dropped to a few hundred dollars - a downright bargain. It takes only \$500 to set up arrays with a capacity of 300 to 400 giga bytes. The introduction of new hard drives with up to 200 GB will make RAID arrays with up to 1 tera byte (5x 200GB) affordable for the first time ever.

No matter which RAID array you're using - for the operating system, it's ultimately a drive just like any other, and therefore it needs to be maintained accordingly.

You should defragment it at least a few times a year; for more heavily frequented drives, once a month. Ideally, you'll enter the defragmentation program in your task planner and have this tricky operation performed during acceptable times.

If one of your drives ever begins to snarl (louder operating noise, reduced performance or other conspicuities), don't hesitate. You should back up all of your important data, especially if you're using RAID 0. If the operating system is on the RAID array as well, you might want to try and mirror the drive in question on another computer with an identical hard drive. Otherwise, you'll have no choice but to reinstall everything.

### **RAID Without RAID**

RAID modes 2 to 6 can be implemented only if the appropriate hardware RAID controller is available. On the other hand, RAID 0 and 1 are offered directly by Windows 2000, Windows XP or Linux - as long as there are several hard drives.

Under disk administration of the computer administration console, you can, among other things, change partitions and drive letters. You can also connect two or more hard drives to form a software RAID.

### **There Are Limits**

RAID arrays are certainly an excellent approach to solving chronic performance deficits and improving your sense of security. Let us mention, though, that they are not able to perform miracles, and do not absolve the user or the administrator from backing up his or her data periodically.

For example, a RAID controller cannot withstand short circuits or lightning, meaning that, in the worst case scenario, your data could be lost. Therefore, an uninterruptible power supply (UPS) is part of the required equipment in productive or otherwise critical systems.

Furthermore, a RAID array only offers protection from technical errors - the human factor, however, should not be underestimated. Most users have had to live with lost data because they carelessly deleted or hastily clobbered them - the same holds true with the RAID.

The chapter on human-related causes also includes malicious attacks on the existing data, or acts of the "powers that be." These involve attacks on software (deleting, formatting, renaming, software bugs), as well as physical threats (theft, vandalism, arson, floods, etc.).

Don't forget - only a backup is truly safe.