Storing data on your PC's hard drive is by no means permanent or reliable. Without proper care and maintenance the hard drive could lose valuable data over time. Prevention is the best strategy for keeping your data.

Magnetic Data Loss in Hard Disk Drives

All hard disk drives use a magnetic signal to store data. The binary 1's and 0's are translated into an analogue signal of peaks and dips and stored on the magnetic surface of the hard drive in much the same way as analogue music is stored on the magnetic surface of a video tape or music cassette. The problem is that this magnetic signal fades over time. In the case of music, this can be heard by an increase in hiss, a loss of fidelity, and a loss in volume. The main problem with digital data is the loss of volume. This is the main method the drive uses to separate 1's from 0's.

The rate of signal decay over time can be increased by heat, proximity to magnetic fields, and chemical breakdown of the magnetic medium through poor quality control or plain old age. In hard disk drives all three factors are present. The hard drive runs hot when it is busy, and there are strong static magnets built into the drive controller mechanism. Over time these can weaken the strength of the magnetic signal that contains the data. Also the signal density is pushed to its limit on a surface where any slight imperfection in the magnetic medium will show up as an unreliable area in which to store data. A drive that runs hot (over 55°C) for any length of time will lose all its data.

The decreasing signal strength first shows up in increased error rates, and later in total loss of data as the error correction mechanism starts to fail. Hard drives have built-in mechanisms to deal with this, but they are by no means fool-proof, and they rely on the hard drive controller being able to detect the flagging magnetic signal before the data is lost altogether.

If the hard drive detects an area of the drive where the signal is too weak, it can mark the area as a "bad block" and re-write the signal in a "spare block". All drives have space reserved for such an event, and once these have all been used up the drive needs to be replaced. Problems occur when the drive encounters a section of the disk that it has not been asked to read for a long time. By then the signal may be too weak to read correctly, and the data is usually lost.

Some specialised software can be employed to make sense of the weak signal and attempt to recover all or part of the data by attempting to read and re-read the data and work out from the analogue signal what the most likely digital bit pattern was.

If a drive is dropped or otherwise damaged, the magnetic surface of the drive can also be affected, giving a weak or partial signal. Data recovery services employ a wide variety of techniques to recover data from such damaged drives. But prevention is always better (and cheaper) than cure.

Another problem is that computers do not typically read *all* the files on your hard drive, only the ones you ask them to read. Some files can stay there for years without ever being read. In the case of these files the signal could fade away completely without the drive controller ever having a chance to detect it.

Preventing Data Loss

The most obvious method of avoiding data loss is to make regular backups of your data. When you make a complete backup of your data you are forcing the drive to read all the areas of the drive where the data is stored.

Hopefully this will alert the drive controller to any weak or dodgy data signals, even for files that have not been accessed for a long time. Even though you now have a backup copy, this will not stop the hard drive signal from deteriorating over time.

For various reasons most people do not make regular (full) backups. Part of the reason is that it takes a long time and can be a real hassle. Another reason is that backup storage solutions are often inadequate to store all the data from a big modern hard drive. It requires 445 DVD disks (or nearly 3000 data CDs) to make a complete backup of a modern 2TB hard drive. Backing up 2TB to “the cloud” can take many months or even years, depending on the speed of your internet connection. Restoring it all will take just as long.

In order to keep the data signal from fading, you need to re-write the data. This is often known as “hard disk maintenance”, and should be done 3 or 4 times a year.

While it does not prevent data from being corrupted or deleted, it can go a long way towards ensuring that the magnetic signal does not fade away completely. The way it works is to read every sector of the drive, and then re-write the data found there, provided the drive reported no errors. If this is done on a regular basis, the magnetic signal of every part of the drive will be refreshed long before the signal fades or becomes ambiguous.

This technique also gives the drive controller the opportunity to decide whether to retire any sectors that are becoming too unreliable, *before* any important data is lost. *DiskFresh* is not the first program to do this, but it is the first program for the PC that allows you to continue to use the drive while the refresh is being performed.

The other programs that can refresh all the data require *complete and exclusive control* of the drive, which means that a server doing disk maintenance is effectively offline for several hours, or even days. A typical refresh time for a 2TB hard drive is 21 hours. Most servers cannot be put offline for that length of time, especially as recommended once every 3-4 months. This would result in only 99.04% uptime, when most servers are expected to aim for “five nines”, i.e. 99.999%, or less than 5 *minutes* of downtime per year.

Other data storage: USB Flash Drives and Solid State Drives

Not every drive in a modern PC is a hard drive. Solid State Drives (SSDs) and small portable “thumb” drives don’t use a magnetic medium to store data. They use “flash memory” instead. This kind of memory can be rewritten only a finite number of times, so rewriting the data on a regular basis is *not* a good idea. But a regular “read-only” refresh is advisable, because the drive controller is programmed to look for bad signals and take action to fix the problem in much the same way as hard drive controllers do. Reading the data on a flash drive does not shorten the lifespan of the drive, and can certainly help to detect and avoid total drive failure.

DiskFresh has a “read-only” mode that should always be used when dealing with flash drives and SSDs. Also, you should not try to defragment an SSD or flash drive. Defrag hard disk drives only.

Conclusion:

A regular (quarterly) refresh of all hard disk drives will help the drive detect and fix errors before they turn into problems, and keep the data integrity intact. For flash drives and SSDs, use the “read-only” mode to find and fix errors before they become major problems. As a precaution against accidental deletion or other disasters, make backups as well.

- Donn Edwards (http://donnedwards.openaccess.co.za)