… Created by Mike Bloomfield

'Backup Theory & Practice'

August 2004

Sorry Miss – The dog ate my hard disk!

By Plamen Pazov

Xyber Data Recovery
Xyber Computer Service Centre Pty Ltd
28 Chandos St – St Leonards – NSW 2065
Tel (02) 9906 7976, Fax (02) 9906 3620
http://www.xyber.com.au/Contact.html

Data Recovery is a lifesaver - But don't count on it.

We are constantly amazed at the small miracles of extracting years' worth of hard work from the surface of a small silvery disk. Unfortunately there are limits to everything.

- Older disk drives were manufactured in such a way that it was impossible to replace a failed disk head, as there were no reference points to aligning the replacement (Thankfully, most current devices contain servo information on every surface, so a big barrier has been lifted).
- The information is stored in very high density (up to 50 000 million bits of information per square inch) on the metallised surface of a few disks (made of glass or aluminium). What happens if this surface gets scratched? That's exactly what a head crash does!
- What happens if the computer (presumably laptop) gets dropped and a disk slips a little? It becomes excentric and basically unreadable. Although we've had some success with these, don't count on it.
- What happens if the computer simply gets stolen?
  Somehow, I just can't imagine receiving anonymously a bunch of CDs in the mail …

You mean I was supposed to back-up?

Yup!
And don't you forget it!

In fact, the best way to make sure it happens is to automate it. Have a system in place that comes in at night, backs up your computer and then shuts it down. Yes, they do exist and they are not that expensive.
You wouldn't believe how many people have come to us for a data recovery saying, "I just bought an external disk drive last week and was going to backup everything but…"
And while you are at it, consider installing a backup on the power supply of an important computer – an Uninterruptible Power Supply (or UPS for short). Again, they are quite inexpensive these days. Certainly inexpensive compared to losing lots of days, or even months' worth of work.

They also totally preserve your hardware and software health beautifully in the long-term as well as protecting against lightning and electrical surge damage if you select a good one ***- Mike***

Where do things go wrong?

Hardware failure.

This is the scariest one, because there is a thing that we can point at and say: this is the baddie. Also it looks so forbidding. It's completely sealed, with warning stickers all over threatening dire consequences if you break the seals. It certainly should not be underestimated and touched by the faint hearted. At the same time, data recovery from a failed hard disk is quite common and nowhere near as expensive as most of us are led to believe.
Actually, hardware failure only accounts for a small percentage of all problems. The biggest culprits are:

Directory damage.

This is by far the biggest culprit. The directory keeps track of what files are on your hard disk and which sector belongs to which file.

You.

Unfortunately, we've all been guilty of erasing an important document by mistake, erasing the document instead of the shortcut/alias, "cleaning-up" the wrong folder, etc. Generally recovering erased data works well but make sure that you talk to a professional immediately! Any work that you do subsequently to the erasure (even rebooting) may overwrite the data irretrievably.

Viruses and other problems.

Viruses - Yes, there have been some particularly nasty viruses out there, but it's surprising that the vast majority of them are completely innocuous. The biggest problems that viruses seem to create are consequential (wasting bandwidth and clogging up email boxes – rather than doing actual harm themselves). We hope that this doesn't change because it's actually very simple to issue a Format command and there is not much that you can recover after that…
Failed system or application upgrades. Usually this is a consequence of latent directory or disk surface problems that come to the surface when a large amount of files are being shuffled around. If some of these are important system files, the consequence can be a complete lock-up.
Badly setup directories. In particular, there were a number of limitations to the size of a volume / partition and the number of possible clusters / allocation blocks in a volume. So, some started trying to cheat, manually

making the size of clusters larger than permitted. In addition to the resultant inefficiency of use of the disk space, these forms of cheating inevitably tend to come and haunt us. Beware – the gain isn't worth the pain.
Unnecessary partitioning. Again in order to limit the impact of the limit of allocation blocks / clusters in earlier operating systems.

What kind of backup options are there?

The basic concept is to keep a copy of everything important somewhere else. That basically means on a different storage medium.
Sounds simple, but (for example) are you sure those two network drives that you are using aren't sitting on the same RAID array somewhere in the IT department?

Standard backup systems

There are a number of standard backup solutions that will collect data across a network and store it on some form of removable medium. Usually, on magnetic tape (DDS, Travan, AIT, DLT, etc.)
Generally, these can be set up to backup everything every time (Full backups) or only whatever has changed since the last backup (Incremental backups).
These systems work well and any medium to large organization would be dead without them.
These are mainly aimed at networks and need some maintenance, at least to manage the large amount of backup media that they generate constantly.
It is important to restart the backup cycle every so often and ideally maintain a current set of media off-site, in case of fire or other calamity.

Removable media

There was a time when 40Mbyte SyQuest cartridges looked enormous. Then there was a proliferation of incompatible models and that, coupled with their inherent sensitivity of the media to shocks made them fall from grace. Only the ZIP is still being used.
The CD-ROM on the other hand has become ubiquitous and the cost per megabyte is too good to be true.
The recent drop in prices of DVD writers is probably going to make the DVD-R take over from the CD-R.
On the other hand, be wary of erasable media CD or DVD. They are a lot more expensive and a lot less reliable. In many ways, the write-once nature of CD-R's is their greatest strength: Once it's written down, they won't be erased.
Note that we have been known to recover a lot of data from CD-R's that won't mount / read any more.

Absolutely right! I remember making this point succinctly to a client, only to then sit and watch their entire business disappear as they assured me CD-RW was as good as CD-R… not true and RW media, DVD or CD should only be used for disposable data as a rule. ***- Mike***

External storage

Back in time, this meant SCSI external drives.
Now it's FireWire or USB external disk drives. Generally, FireWire drives seem more reliable and less prone to compatibility problems than USB 2 devices. USB 1 was of course impossibly slow. Hope they improve.
Again, we recover data from all of these devices as a matter of course.

Hardened storage

These basically are some form of bulletproofed storage server. Generally will have redundant Power Supplies and some form of RAID (Redundant Array of Independent / Inexpensive Disks) structure to limit the impact of a single device failing.
The simplest is straight mirroring (2 disk drives recording exactly the same information, so that if one goes down, the other will be fine).
Beware – Unfortunately the concept of a RAID will deal with a hardware fault but won't protect you from a mushed up directory or any of the other software based nasties.
Also, a lot of RAID's are actually software RAID's (the data distribution across the multiple drives is controlled by the device driver or the operating system and if any of these get corrupted, you get distributed mush instead of centralised mush – even worse.

Network storage

Of course the best of both worlds is a stand-alone RAID server.
This would have its own processor, operating system and storage devices.
In the big end of town, there is a continuous tug of war between SAN and NAS configurations. Unless you look after hundreds of users, you can probably safely ignore this lot.
The single drive ones on the other hand can be cost effective even for a small group or individual users.

Internet Backup

With the advent of relatively inexpensive broadband connections (Cable, ADSL, etc.) it is becoming cost effective to store your data on a server across the Internet. Usually these are well protected and backed-up internally. Data is generally stored in an encrypted format and these FTP servers can generally be used with standard backup applications (such as Retrospect, Backup Exec, etc.)
There are relatively few providers of these services but they will increase in time and the prices will come down.

Pseudo-backups

There are a multitude of other ways of protecting yourself against potential threats.
Bearing in mind that a large number of problems originate from damaged / corrupted directories, it makes sense to keep a backup of just the directory data. In fact, operating systems already attempt to do that by duplicating some of the more important structures. There are a number of utilities out there that would do just that.
Mainframes have been doing similar tricks for many years under the guise of Journaling (keeping a log of all file changes, so that if things get interrupted, there is a way of "rolling back" the unfinished transactions. In fact Journaling was recently introduced to MacOS X.

Poor man's backup

This is where you have to get creative:

- Email it to yourself.
- Compress and encrypt it if you are concerned about prying eyes.

- Store it on an online server – most Internet Service Providers (ISPs) include a number of megabytes' worth of web storage in the cost of a subscription. You are paying for it – might as well use it.
- Subscribe to an on-line service. Many are free and the rest (like Apple's .Mac service) are good value.
- Even a paper printout can be a Godsend that you can scan and OCR if things get really nasty.

By

Plamen Pazov

XYBER Computer Service Centre
02 9906 7976

Plamen is a true genius in our field and a leading expert on data recovery, photo recovery and all things MacIntosh. I am refreshed as he enlightens misunderstood areas of our industry with absolute clarity and precision – and authority.
Please feel free to contact Plamen and his brilliant team at Xyber for unparalleled advice and service. There are few businesses I have found in any sphere of activity that share the integrity, honesty, ethical standards and enthusiasm that this team zealously uphold.
Little wonder our leading Television and Radio networks, I.T. professionals and the elite of Graphics Design champion his outstanding service and unique craft year in, year out!

Kind Regards,

Mike Bloomfield