

Information Security for the Digital Photographer

When you come across the term "information security," what springs to mind? If you're like most people, probably images of industrial espionage, Matrix-style hackers breaking into high-security systems, and government snooping. In other words, stuff that doesn't apply to you. In actual fact, it's a lot simpler than that -- and information security applies to everyone with information they care about enough to want to protect it against loss or falling into the wrong hands. If you care about your digital photos, you need to think about information security.

Information leaks get all the press. Whether it's Paris Hilton's home video or the Abu Ghraib pictures, the digital media's potential to spread hot information quickly is spectacular. However, in actual fact, this type of information security breach is far rarer, and the consequences almost always far less severe, than the other type -- plain ol' data loss. Most of us don't have Abu Ghraib or Paris Hilton-grade stuff on our computers anyway, and the damage caused by it leaking would be commensurately less; amounting to embarrassment or at most fairly minor economic consequences (such as having to re-buy software licenses if the vendor revoked yours because it had leaked to use in piracy). Therefore, your main information security concern is almost certainly to secure your information against loss, with the other kind of information security coming in second. In practice, this means a solid backup regimen -- and this is what this article will mostly discuss.

Mishaps, Malice, and Disasters

We humans deal with risks in rather irrational ways. We tend to greatly overemphasize risks over which we have little or no control, and play down risks in which we or people close to us play a part. For example, most people are more nervous about flying than about driving, even though the former is statistically safer. Similarly, when it comes to information security, we tend to overestimate risks like hardware failure, viruses, burglars, and hackers, and underestimate risks like mishaps or stuff done in the heat of the moment. Yet, statistically, the latter are far more common than the former. A software designer colleague of mine recently mused that there are two kinds of people: those who have accidentally nuked significant data, and those who will do it. (His personal history includes nuking the entire source code of a medium-sized project by typing in `rm -rf *.java` instead of what he intended, `rm -rf *.class`. Needless to say, he's extremely information-security conscious now, and our backup regimen is state of the art.)

The causes of information loss can be grouped into three categories: *malice*, *mishaps*, and *disasters*. Malice means someone doing something intentionally to cause damage. Mishaps are accidents or failures that happen unintentionally, whether through human error or equipment failure. And disasters such as fires or floods are major events that impact entire locations. It is impossible to protect your information against every possible eventuality, and once past a certain level of protection, increasing it becomes geometrically more expensive, inconvenient, or both. However, it is comparatively simple to work out a backup regimen that gives a quite a high level of protection against all three types of event.

Mishaps -- shit happens

Hard drives fail. Coffee gets spilled. People screw up. If the only thing between you and permanent data

loss is one moderately fragile electronic device, losing your data is a matter of when, not if. Mishaps are common, but also fairly easy to protect against -- it's usually enough to have a single backup copy on a separate device. I can't think of too many mishaps that would destroy two computers at once, and even mishaps that would destroy two hard disks inside one computer at once are a lot rarer than mishaps that would destroy just one disk.

Malice -- doesn't apply to me, right?

Information loss through malicious acts may sound far-fetched. However, think about it a bit more, and you'll probably realize it's not that far-fetched after all. Tempers flare, and people do things they regret later. People close to you can really get under your skin. And, sadly, worse things than flared-up tempers happen among families and friends. I know of cases where the object of someone's wrath has been precious data. In such a case, having a backup of the stuff can not only prevent a very painful loss, it can even help save interpersonal relationships: if no permanent damage was done, it's a lot easier to forgive, forget, and resolve not to do it again.

Malicious acts can range from the quick burst of wrath that smashes a computer to bits or formats a hard drive, to the genuinely evil, systematic destruction of data across multiple systems. And, of course, the old classic of posting naughty pictures of the ex on the Internet. It's very difficult to come up with a system that's really proof against the coldly calculated systematic kind of malice, but luckily that sort of thing is pretty rare. However, it is quite easy to come up with something that gives adequate protection against most malicious acts.

Disasters

Disasters like fires, floods, burglaries, and acts of terror are mercifully rare -- but not so rare that most people wouldn't get home insurance. If you care about your home enough to insure it, shouldn't you have a plan for recovering at least a part of your most precious data if that kind of disaster strikes? To completely protect your data against this type of event is very expensive or inconvenient, but it is possible to have a reasonably light system that will ensure that you'll at least get most of it back. In practice, this means having an off-site backup of at least your most important data. This is a drag in many ways, and the higher the level of protection you want the bigger the annoyance becomes. However, a minimal system is not too onerous, and will make the difference between a personal tragedy and a minor loss.

On-line and off-line backups

There are two types of backups to consider: on-line and off-line. Each has its strengths and its weaknesses, and ideally a backup system should combine elements of both.

An on-line backup is, well, on-line: it is continuously accessible and should be frequently updated. The advantages are that because of the high refresh frequency it will minimize the damage if something does happen, it's generally very little trouble to maintain, and you will notice immediately if something goes wrong with it. The downside is that at this time it's really expensive to have an off-site on-line backup of your data, which means that it offers limited protection against disasters and only partial protection against malice. Moreover, on-line backups offer limited protection against "bit rot" -- data corruption due to device or software failure or accidental screw-ups that affect individual files: unless the backup system maintains a longish history, you will simply overwrite your "unspoiled" backups with "rotted" data.

In the near future, it will become possible to automatically mirror your data on a service accessible through the Internet. This will provide a high degree of convenience and security against data loss, but of course carries its own risks -- in particular, that of data compromise (once out of your hands, you no longer have full control over who gets the data and under which circumstances). It'll be a while yet, though. For example, with my current broadband connection's upstream speed, it would take about 40 days, non-stop, to upload my photo library to the Internet. And, of course, 20 days to download it, should I have to recover from the total loss of my library. In other words, this won't really be feasible until the Internet's bandwidth has gone up by at least two orders of

magnitude: I could deal with 0.4 days, but not 4 let alone 40. And of course by the time bandwidth has gone up by a factor of 100, who knows how big my library will be -- if the past is any guide to the future, it will have expanded by a couple of orders of magnitude too.

Off-line backups are not running except when they are being refreshed or actually used. Typically this means copies of your photo library burned onto DVD's or perhaps an external hard disk that you only plug in when needed. The advantage with off-line backups is that they're easy to move off-site, which means that they're the only thing that can give you good protection against disasters. The downside is that they involve a lot more housekeeping than on-line backups: you need to physically carry the media around, if you don't periodically verify and refresh them the risk is that when something bad does happen you'll find that the media have decayed or become obsoleted and the data is lost anyway.

In other words, on-line backups fail to protect against the worst events, and off-line backups fail through lack of a disciplined process to keep them going.

Recovery: the Chernobyl effect

The Chernobyl nuclear disaster was one of the worst accidents in human history. While the technology used in the nuclear power plant certainly played a part in it (that exact type of disaster cannot happen in reactors of different types), the main cause of the disaster was a chain of mishaps and bad decisions. The power plant had numerous fail-safe mechanisms, alarms, and other safeguards active and working, and they did not all mysteriously fail at once. Instead, someone made a bad decision, then something bad but recoverable happened, then someone else decided to cut a corner when recovering from it, which made things worse, and eventually things spun out of control. This pattern is very common in disasters, whether it's about nuclear power plants, airliners, or losing your photos. The fact is that if you have a recovery system in place, *the most dangerous moment by far is the moment of recovery*. Recovery is usually tedious, and there are many possibilities for cutting corners. For example, if you do lose your main storage medium, and you only have one backup, there will be a period during which you will only have one copy of your data. Now, if your data is on a rewritable medium, and given the fact that you're doing stuff to your disks and files, there's a very real possibility that you'll accidentally nuke your backup as well (all it takes is a slip and fall when you're carrying it home!), in which case you're really screwed.

Therefore, *be aware of the Chernobyl effect, and if disaster strikes, your first action should be to prevent it from happening*. In other words, consider the risks to your backup in every step you take, and don't do anything that could conceivably nuke it. Ideally, make taking a second copy of your backup the first step in your recovery process -- if you can do that safely, that is. And if you get the feeling that you're doing something you shouldn't really be doing "but," then stop doing it.

Backup regimens

A backup regimen consists of two parts: technology and process. You can't have process without technology, but technology alone won't solve any of your problems. We'll go over some fairly simple and inexpensive backup regimens you could adopt or tailor to your needs, as well as going over a few general principles about them.

Organize your information

The key to any working backup regimen is to organize your information. The minimal level of organization isn't much: simply but everything you want to protect in one place, and everything that's expendable in another. For example, given the nature of Windows, it's usually not feasible to maintain a backup of your system disk: this is both complicated and will almost inevitably adversely affect system performance. So the key to protecting the stuff on your system disk -- programs and operating system -- is just to store your installation CD's and license numbers in a safe place, so you can rebuild your system should bad things happen. On the other hand, the really precious stuff is your personal data. So, make up a

system where you put everything that matters in one place, and consider the rest expendable. I use My Documents as my "protected zone," with everything else either recoverable or expendable.

Most importantly, **always know where your original is**, and treat that with particular respect. For example, create a photoediting routine where you never overwrite the originals. This is a first principle of data management in professional applications, and it's holds exactly as well for you. I strongly recommend that you keep your originals on your main workstation. If your library grows too big, get more disks so you can continue to keep it there. If it grows really, *really* big, you'll need to archive a part of it as an off-line copy somewhere, and make it a part of your regimen to maintain that archive -- and backups of it too, of course. In other words, your data management will get a lot more complicated, since you'll have to set up a routine to check in new originals into the archive, and then check out copies to work on them locally. Trust me, you'll need to think of it this way -- otherwise you'll get hopelessly confused and lose track of what's where and what's backed up and what isn't. Disks are cheap and data is precious, so don't complicate your life unnecessarily by archiving stuff off-line: just get another disk, or a bigger disk, instead.

Technologies

One of the first things to do is to consider the technology base of your backup regimen. Ideally, it should be a mix of on-line and off-line backups, with the off-line backup refreshed and migrated to new media at least once per year. There are lots of possible solutions for backups. Here are some common and comparatively simple and affordable ones you may want to consider:

Optical disks (CD or DVD)

Optical disks are the simplest and most common way for home users to back up their data. They're cheap (at least superficially), simple, promiscuously available, and easy to transport. However, they're far from ideal: a photo library will quickly grow larger than a single DVD (which holds only five one-gigabyte cardfuls of photos, after all) let alone CD, their longevity is somewhat suspect, they're physically somewhat fragile, and they're slow both to read and to write. In particular, backing up a multi-disk photo library takes a quite a bit of time and annoying disk-jockeying to do. Therefore, being the lazy slobs that we are, most of the time we just don't bother. What's more, optical disks aren't as cheap as they may appear: the cost per gigabyte of good-quality DVD+R is about the same or even somewhat higher than the cost of hard disk storage. (For example, at this writing a 25-disk spindle of archival-quality DVD+R's costs about 50 euros and holds about 100 GB, and 100 euros will buy you a 250 GB hard disk.). However, the worst downside of optical backups is that they stay off-line rather than on-line and the media degrade over time. Therefore, any media failures are often caught too late -- only when the backup is actually needed in a real situation.

Internal hard disk

Having a separate physical disk in your computer dedicated to backups is in many ways a better solution than periodic backups to DVD: the backup process is much simpler and faster (just drag-and-drop), and since the disk is on-line all the time, you'll notice if it starts to go bad. However, since it's inside the computer, it won't save you from any major disasters that affect the entire computer -- such as a burglary, flood, or fire. An internal backup is also still susceptible to mishaps or malice, although clearly far less so than not keeping the backup at all. In most places, though, these are far rarer eventualities than data loss through human error or disk failure, so on balance I believe that an internal backup hard disk is a more secure solution than backups on optical disk. The initial cost is somewhat higher than for DVD's since you're likely paying for storage you're not actually using, but you will recoup the difference over time since the medium is reusable and your library will expand.

Mirrored RAID pair

A mirrored RAID pair isn't really a backup, and you shouldn't think of it as one. It only protects against exactly one eventuality: disk failure. It will not protect against data loss due to you mucking with the system or the system mucking with the system, which is much more common. Therefore, if you install a RAID pair on your computer and stop making any other backups, you will have worsened your security situation rather than improved it. However, when used with other backup strategies, a RAID pair can help reduce the risks. However, if you can only afford one extra hard disk for security, don't make it a RAID pair; make it a dedicated backup disk instead: at least this will protect against a quite a few more-common-than-you'd-think human errors.

External hard drive (USB, Firewire, SATA...)

An external hard drive is a very attractive option for off-line backups. The cost per gigabyte is reasonable, they're much less fussy to deal with than DVD's, and just as portable. I would strongly recommend an external HDD over a bunch of DVD's, that you then pick up and store off-site, like at the office. The only risk related to it is that HDD's are random-access media -- meaning, it's possible to accidentally nuke one while it's more work to do that on a DVD. (On the other hand, recovering from an accidental quick format is pretty easy.)

Networked disk (NAS)

A networked disk -- actually a "plug-and-play" file server -- is a higher-cost backup platform than a simple external or internal backup disk. A NAS disk can be located in a different room from the main system and is easier to unplug and evacuate in a hurry, which provides some protection against disasters. Additionally, the higher-end ones can be set up as RAID arrays, allowing for additional redundancy or higher performance. Since they're plugged in and left running rather than spun up and down regularly they're also less likely to fail. However, they do cost a good deal more and they're not quite at the "plug and play" stage yet. When considering my own on-line backup solution, I ended up going with a fully-fledged file server instead of a NAS, because of the added flexibility. (Also, I had a suitable computer lying around.)

File server configured for backups

In my opinion, a disk-based file server configured for automatic backups is the highest-tech backup system a home user is likely to want to set up. It gets pretty close to the "ideal" in that once set up, it doesn't need any attention, if it starts to fail (or run out of space) it'll yell, and it can even maintain a history. You can also add other nifty services onto it, such as network disks and a print server. The downsides are that you need to set up a (cabled) network and the server itself, and keep it (and your computer) on long enough and often enough for it to do its work, which will use electricity and produce noise. And of course it's clearly the highest-cost backup solution of the ones discussed in this article, although perhaps not as expensive as you might think. And while it will protect you against most varieties of human error or equipment failure, being on the same network and physically in the same space, it is just as liable to getting burgled or destroyed in a fire as your main computer.

Process

To get most security for the least amount of trouble and cost, **mix a frequently-updated on-line system and a periodically-updated off-line one**. The on-line system is your first line of defense, and should that fail, you can fall back on the off-line copy, only losing the information accumulated since the last refresh. Your choice of system is a matter of preference. I have set up a networked file server that runs an automated backup program that backs up my PC once per day, and have a copy of my photo library on a 250GB external hard disk which I keep off-site, at the office.

The key to making your backup regimen work is to **create a routine** around it. For example, if your

on-line backup is an external hard disk, make it a habit to dump a copy of your My Documents folder into it last thing every day. If you can automate this, great. Same thing with your off-site backups: if you're lazy (like me) so that you only refresh them once or twice a year, tie that to some day you'll remember, such as the day you submit your tax declaration, or something like that. If you always do it the same way, you won't need to think about what's backed-up and what's not. If you don't, you will get confused about it at some point, or you'll start using your backup space for some other purpose, which will result in a very unpleasant situation of having your backups and your originals mixed-up, effectively putting a part of your data outside the backup system.

Automating the routine

The simplest way to make a process more efficient, once you've figured out what it is that you want to do, exactly, is to automate those parts of it that are easy to automate. For example, if you want to make a copy of your documents to a backup disk (internal, external, or networked) daily, there are tools available that will allow you to have the computer do this for you, as well as maintain a backup history. The "reference" Windows and Mac OS X backup automation system is probably Retrospect [<http://www.emcinsignia.com/>]. There is a quite a variety of cheaper and even free solutions available as well; they offer varying levels of security, features, and convenience. I have heard good things about Allwaysync [<http://www.allwaysync.com/>] and 12Ghosts, [<http://www.12ghosts.com/ghosts/backup.htm>] however, I have no personal experience with either of them (nor Retrospect, for that matter), so you're going to have to do your own research about them. Look for the following characteristics in whichever system you pick:

- **It doesn't do completely stupid stuff from a security point of view.** For example, a backup system that nukes your last full backup *before* starting on a new one is fundamentally insecure.
- **It maintains a history.** You will want to roll back to an older version of a file at some point, and a backup history will allow you to do that.
- **It consumes the minimum of system resources.** You don't want it to bog down your computer when doing the backup (or, worse, when *not* doing the backup) -- it'll just make you switch it off.
- **It's easy to administer.** You'll want simple, flexible configuration and simple, fast restore functionality.

The rest is largely a matter of taste. The first two points are particularly important -- if your backup automation solution doesn't fail on them, it's already good enough for most purposes and a major step up from all-manual backups both from a security and from a convenience point of view.

Sample processes

To make things easier, here are three backup regimens based on three different systems, all of which provide very good protection against data loss. The two simplest ones are fully manual and require no special skills or effort to set up, but do require some discipline to maintain. The third automates the routine of daily backups, but is somewhat more expensive and difficult to set up.

Sample process 1: hard disk + optical

Technology: Original on internal hard disk labeled "Documents." On-line backup on second internal hard disk labeled "Backup." Off-line backup on DVD's.

Process, daily:

- Drag-and-drop My Documents to drive "Backup," directory "Petteri," overwrite everything.

Process, on tax day:

1. Burn copy of My Documents onto DVD's, using a low write speed.
2. Format "Backups."
3. Restore "Backups" from DVD's just burned above.
4. Take DVD's off-site and destroy the last but one DVD backup. That is, keep the newest and the one before it. (The older ones will degrade anyway, so might as well destroy it now when you know you have backups.)

Strengths: No special skills or technology needed, inexpensive, good protection against all types of data loss.

Weaknesses: Lots of manual work, which is inherently risky and requires discipline, gets progressively more difficult as amount of data grows.

Upgrade paths: Automate daily backups with backup software. Switch to using hard disk for off-site backups.

Sample process 2: hard disk + hard disk

Technology: Original on internal hard disk labeled "Documents." On-line backup on second internal hard disk labeled "Backup." Off-site backup on external hard-disk labeled "Off-site backup." Daily backups automated with backup software running on workstation.

Process, daily:

- Nothing much.

Process, on tax day:

1. Bring "Off-site backup" on-site, and format it (not quick format).
2. Copy originals onto "Off-site backup."
3. Take "Off-site backup" off-site.

Strengths: No daily manual routine. Excellent protection against mishaps since a files deleted last week can be recovered. Good protection against other types of data loss..

Weaknesses: Automated backup system may cause complacency, which can lead to the Chernobyl effect.

Upgrade paths: Switch to using networked disk or file server for the on-line backup. This will increase security somewhat and use less system resources on the local computer..

Sample process 3: file server + hard disk

Technology: Original on internal hard disk labeled "Documents." On-line backup on networked file-server running automated backup software and maintaining a short backup history. Off-site backup on external hard-disk labeled "Off-site backup."

Process, daily:

- Nothing much.

Process, on tax day:

1. Bring "Off-site backup" on-site, and format it (not quick format).
2. Copy originals onto "Off-site backup."
3. Take "Off-site backup" off-site.

Strengths: No daily manual routine. Excellent protection against mishaps since a files deleted last week can be recovered. Good protection against other types of data loss. Pulling the backups from the server rather than pushing them from the workstation makes for simpler administration and uses up fewer resources on the workstation.

Weaknesses: A file server takes some specialized skills to set up and run, and costs more than a simple hard disk. Automated backup system may cause complacency, which can lead to the Chernobyl effect.

What about hackers, viruses, privacy, all that commotion?

If you're living under a repressive regime where the government has a habit of snooping into people's private lives, or you're up to something that doesn't quite stand the daylight, your privacy requirements will be somewhat higher than average. However, if your main privacy concerns are some photos you'd rather not have your mom see, your situation is simpler. Just keep the damn things in an encrypted ZIP archive for which only you know the password, and don't make any unencrypted copies. Do that and it's extremely unlikely that any will leak out through caches, someone going through your hard disk with a low-level data recovery utility, or similar theoretically conceivable but highly unlikely things. Just don't lose that password, m'kay?

As to viruses, hackers, Trojans, and other external threats, it's also quite simple to protect yourself adequately against them. Unfortunately, most people don't, because it's just that little bit more effort than not doing it. However, if you do the following, you'll be pretty safe:

- **Use a firewall on a separate device.** Unfortunately most home broadband connections are "bridged," which means that your computer is connected directly to the Internet: your door opens directly to the main highway. This means that anyone sauntering by with some lockpicks and an attitude problem can start picking away at it. Securing a computer that's connected that way is a good bit trickier than securing a computer that isn't. It's been claimed that connecting an unpatched, freshly installed Windows XP box to the Internet this way will result in it being infected with viruses, Trojans, and other malware within as little as two minutes -- far too short to download and apply any patches. However, you can buy a firewall/NAT router for a few tens of dollars that's basically plug-and-play: connect your ADSL modem to its WAN port, and your computers to its LAN ports, and you're immediately much safer than before.
- **If you use WiFi, secure it.** It's easy, but unfortunately not securing it is even easier, so many people don't. (My parents' house has, at last count, four unsecured WiFi networks visible from their apartment.) WiFi is nice, but if you've set up a nice home network with your file server, shares, and so on, and you put it on an insecure WiFi network, anyone passing by can just waltz in and do whatever they like. I'm serious -- if you don't feel up to securing your WiFi network, then *don't use one*.
- **Use automatic updates.** Whether it's Windows, Mac OS X, or Linux, do it. Security holes are discovered and patched all the time (yes, even on Mac OS X), and if you don't apply the patches, your system will be open to attack.
- **Don't screw around.** Basically, if you insist on trolling shady porn, warez, or gambling websites with minimal browser security settings, or click "Agree" "OK" and "Install" to every dialogue that pops up while you're on the net, or open every attachment some spammer sends you, you're asking for trouble.

So don't.

- **Install anti-virus software** if you're on Windows. There are thankfully few Mac OS X and Unix viruses and Trojans around, so you can do without it on them (for the time being). I've used AVG Free [<http://free.grisoft.com/doc/1>] on my home computer for a couple of years now, and it gets the job done very well: it's unobtrusive, effective, and free. However, remember that anti-virus software is your last line of defense rather than your first: I haven't had a single genuine virus alert on my computer since 1989, when the Mac IIci I was using caught a bad case of nVIR B [<http://www.ciac.org/ciac/bulletins/ciac-09.shtml>].

Appendix: Setting up an automated backup system

An automated backup system consists of a *network*, a *file server*, *clients*, and the *backup service*. If you're at all technophobic, forget about it -- but if you're cool with tinkering with your home theater system or administering your own computer, you'll manage it. You will need to use Google a lot, though. What follows is not even a primer let alone a manual on how to do it -- merely a pointer about what's involved in case you want to pursue the topic further.

The network

For the network, you'll need a *router*. It's basically a box that you plug into the Internet, and that the devices on your network connect to, either through cables or WiFi. Nowadays these are mostly plug-and-play. However, do be aware that backing up large volumes of data over WiFi is slow and at least somewhat unreliable: therefore, all the computers you want to automatically back up should connect to your router through cables. Suitable routers are pretty cheap nowadays: you can get perfectly good ones for tens of euros/dollars, and "premium" ones for a little more than a hundred. They also pretty much auto-configure themselves.

The server

The server can be any ol' computer that's quiet enough that you can bear to listen to it. In fact, old boxes are very nice, since they have cooler processors and smaller heatsinks... although you don't want to get one that's so old it doesn't have connectors for the disks you want to use. If you don't have an extra computer lying around, I'd suggest something based on a late version AMD Socket-A (462) motherboard: being obsoleted for a while now, they're cheap on the second-hand market, but the newer ones come with SATA as well as IDE disk connectors. The main thing is that it has lots and lots of disk space. You can get three hard disks and a CD-ROM on any ol' computer, and if you remove the CD-ROM after you've installed your system, you can even shoehorn in a fourth. If your motherboard has a couple of SATA connectors on it, that makes for six to eight disks. Four 250GB disks is a terabyte of storage. Four 500GB disks is two terabytes. Eight 500 GB disks is four terabytes. Ought to be OK for the duration.

The system you'll want on your server is Linux (or some other flavor of Unix). They run great on slow, old boxes, are very stable, and you can set up any services you like on them quite easily. Also, they're free. I use Debian, because it's conservative, standards-compliant, well documented, has a very wide user and developer community, and has an excellent package manager for installing and updating stuff. It's by no means the only such distribution, of course, and I won't even try to list all the others worth trying -- there are so many. I hear, though, that Ubuntu [www.ubuntu.org] is pretty good for beginners -- as long as you make sure to install the server version if you're making a server.

Learning Unix is a bit like learning a new (although very simple) language. Once you do know the basic grammar and vocabulary, it becomes very fast and easy to live with... and Windows starts to seem really clunky and annoying. For example, I set up a "limited" account on my computer (running Windows XP Home) for my wife, but when I copied her files over, I discovered she doesn't have "ownership" of them. On Linux, this is very simple to change *if* you know the lingo: as superuser, you just type `chown -R joanna:joanna /home/joanna` which means "change owner -Recursively to user **joanna** and group **joanna** on the directory **/home/joanna**". To do the same on Windows, I had to (1) give joanna "administrator" rights, (2) boot up into "safe mode," (3) log in as joanna, (4)

find the directory I wanted to change, right-click Properties, click on the Security tab, surf through a few more tabs and buttons until I finally got to a screen where I could "take ownership" of the files, then reboot the computer, log in as myself, and reset the account "joanna" to "limited." Ouch. However, to be able to work with Linux, you do need to learn the lingo and understand the basics of what goes where -- and that's where the learning curve of the system lies.

The procedure for setting up your network backup server is pretty simple, really.

1. Read up on the basics of Linux/Unix. Buy a book, or read *The Linux Newbie Administrator Guide*, [<http://linux-newbie.dotsrc.org/>] at least sections 1-3 and 5. For a chattier guide that covers similar topics, see *Linux Newbie Guide*. [<http://www.linuxnewbieguide.org/>] Finally, leaf through the installation guide of your chosen distro, such as this one, for Debian. [<http://www.debian.org/releases/stable/i386/>] Yup, you do need to do some homework here.
2. Install a basic system, set up so that you have a directory where you've mounted your backup volume(s) -- configured as a "spanned" JBOD array, for example. I used the default partitions for a multiuser system, more or less, and mounted my backup disk as one big partition in `/var/local/backups`.
3. Install backuppc. [<http://backuppc.sourceforge.net/>] In Debian, it's as simple as typing `apt-get install backuppc` (as root).
4. Set up a "share" on your Windows box of whatever directory you want to back up. For example, "My Documents." Right-click, Properties, Sharing, Share this folder. Give the share a name with no spaces or special characters; it'll save trouble later. For example, "backupme."
5. Configure backuppc to back up your PC's. This involves editing the configuration file. I won't go into that here since it's very well documented with the application itself. Note, though, that backuppc assumes that its backup directory is in `/var/lib/backuppc`. Therefore, you have to go to `/var/lib/backuppc`, move all the files in it to wherever you want your backups to really be (in my case `/var/local/backups`), then go back to `/var/lib` and create a symbolic link to your real backup directory:
`ln -s /var/local/backups backuppc`.

And that's it. Don't forget you can add all kinds of fun stuff onto the box. For example, Samba [<http://www.samba.org/>] will let you set up network disks and share printers. Have fun!

Unless otherwise indicated, all materials on this site are by Petteri Sulonen. They are licensed under the Creative Commons Attribution License [<http://creativecommons.org/licenses/by/1.0/fi/>]. I would appreciate it if you dropped me a line if you want to reproduce them. Any trademarks are property of their respective owners; their use is purely editorial and does not constitute an infringement.