

Backups

Last modified: 2011-11-12 (needs work). Epistemic state: [believed](#).

backup ALL THE THINGS!

So, I spent two hours today and further [gvernified](#) my life. Seems like a good time to summarize my backup strategy. Maybe it inspires someone to save more of their own data. As the saying goes, nobody wants backup, but everybody wants restore.

Strategies

1. All data in /home is synced between my laptop and desktop machine. That's the most basic level of redundancy.
2. Everything except TV shows etc. is backed up every 6 hours on a dedicated backup drive. I use a custom [rsync script](#) for incremental backups. That way, each snapshot is self-contained, but snapshots share hardlinks and save on space. A new snapshot of ~200GB of data takes ~150MB and takes ~6 minutes. I do it this way because I often restore stuff and found rdiff-backup horribly slow. I have 2 weeks of snapshots for everything, plus unlimited monthly snapshots for most partitions.
3. Most TV shows, music etc. is also mirrored on the backup drive. (This isn't really crucial because Piratebay is a good backup strategy in itself.)
4. The whole backup drive itself is mirrored on a second backup drive. (To preserve the backup history.)
5. I have a third backup drive that I update every couple of months or so and store at my dad's workplace. (I never throw away old drives. I just put a backup on them and let them rest.)
6. I [take a screenshot](#) every 5 minutes and store it on the backup drive. Useful to reconstruct days or restore content other strategies miss.
7. Everything that *can* be open and online, [is so](#). I'm currently transitioning my notes to my website and some of my daily task tracking to [Beeminder](#). As Linus said, "Only wimps use tape backup: *real* men just upload their important stuff on ftp, and let the rest of the world mirror it."
8. I put everything I edit in a git repo to preserve its history. Especially my private notes, tracking data and so on. I have a cronjob that commits my notes every 20 minutes so I don't have to think about it.
9. I backup my notes, mails and Anki deck on Amazon's S3 every month in case my house

burns down or the police raid me. Takes up about 1.5GB and costs me ~30 cents a month. (I don't sync my Anki deck with ankiweb.net because it's pretty large (>14k cards, 900MB of media) and I don't want to burden Damien. Once he allows me to pay for my account, I'll sync again.)

10. I [mirror all videos](#) I link to on my website because they have a bad habit of getting DMCA'd out of existence. I run linkchecker once a month to fix broken links.
11. I also let Google track me completely. Hey, they aren't more evil than future me and my search history has saved my ass a lot in the past. If they also profit from my data, good for them.
12. I [track](#) all useful daily activities (and [time spent](#)) so I can see how much time I waste. (My task suggestion script also balances activities.)
13. I log all chat communication. This was my very first backup setup and is tremendously important. "Huh, didn't I talk about this with him before? ...*grep*... Yup, 7 weeks ago. *quote*". In fact, if anything has a log option, I use it and never throw away the logs. Text is easy to compress.
14. I use Gentoo, and so save the sources and binaries for all packages I use (and back them up as described above). Every once in a while, a library breaks something and I need a clean package *now*. Or an obscure program disappears and no-one mirrored the sources. Sucks. Don't let it happen to you. Don't clean your cache. (Or at least, have monthly snapshots).
15. I try to put my beliefs and predictions on [PredictionBook](#). Keeps me honest and forces me to turn empty beliefs into ones that actually predict something. And I now have proof whenever I say "Told you so!". Good for my hipster cred.

That's about it. This is all more-or-less automatic, so no effort on my side and it's all cheap. You only really notice how valuable backups are when you have them and can constantly restore stuff. "Oh, that pdf from last week I thought wasn't useful? Need to quote it.", "Crap, deleted the wrong file.", "Nah, that paragraph sucks, let's get the first version.", "I watched this amazing prOn a month ago, but the link is dead. What's the name?", ...

Future

There are a few things I'd really love to store in the future.

1. A webcam in my room. I already have one and modded it to record infrared as well. (Most chips do, but have a filter, typically on the lens. Just scrape it off.) So it also works reasonably well in the dark. I just need to set it up and can get IRL screenshots as well.
2. I really need to record my thoughts more, but I don't know how. I already try to write them down as much as I can, but that's cumbersome. I thought about using an audio recorder, but that isn't as automatic as I want either.
3. Similarly, I'd really love to record IRL conversations, but current tech still sucks too much. Luckily, I'm enough of a loner that I barely have any non-text conversations, but you know. Need to win some more debates with my mother. ("No, I didn't say that at all! Here, listen!")

Rules

So to summarize the summary, I think the most important rules to backup are:

1. Use it. You *will* need your old data at some point and hate yourself if you don't have it. Life is already horrible enough. Don't increase your suffering through laziness.
2. Automate it. The less you have to think about it, the better. When in doubt, just backup it. Space is cheap.
3. Histories matter. Don't just save your files. Save your histories, ideally in something like a git repo. Keep old backups.

DISQUS 

[muflax](#) » [software](#) » [backups](#)