The Hipster Backup

BY ZENO KOLLER - BACKS UP TO GLOBI KASSETTLI NOW

My working definition for the modifier 'hipster' for this article: Favoring aesthetics over practicality. Taking pictures with an analogue camera, riding a single-gear bycicle, listening to vinyl records. So - what would be the most hipster way to store data? You guessed it: the cassette tape. Not only does it look pretty awesome - I'd say: Blade Runner-esque - but it is also a pain to handle. As a proof of concept, I wanted to store this very article on a cassette tape. Before getting into the details of that, let me tell you a few things about magnetic tape storage.

Today, magnetic tapes are still deployed on a large scale. This is due to three killer features.

First, they are much cheaper than hard drives. Not only are tape drives cheaper than HDDs for the equivalent amount of storage, the price also grows more linearly as you move towards higher-capacity drives. To make a bigger drive, just add more tape! That's why tapes are ideal for low-cost, warehouse-style archival where industrial robots access the tapes. An example could be Amazon's Glacier cloud storage service - could, because Amazon does not disclose the technology behind the service. Priced at $0.004 per uploaded GB, it's extremely affordable. However, access latency may be up to 5 hours and downloads may incur costs.[1]

The Hipster Backup in action

Second, tapes have a very high linear write speed. This makes them a good candidate for scenarios with HUGE amounts of data and linear writes. Huge as in a petabyte per day, as CERN's LHC accumulates - mind that this is after filtering out most of the data.[2] Not only do single tapes have a huge capacity, there also exist autoloaders, devices that will swap a full tape with the next one. As far as I know, there is no autoloader equivalent for hard disk drives. This means you can just write away a huge dump - you don't need to partition it into disk-sized portions.

Magnetic tape was invented by ze Germans

Third, tapes are much more durable than hard disk drives. They're more shock resistant, and while they tend not to get hot, they can also stand higher temperatures. Compared to a HDD, a tape is a crude device - there's much less stuff that can break. Manufacturers of tape storage media state a lifetime between 15 and 30 years. This makes them ideal for long-term, sporadic-access archival.

However, if we want to talk about real longevity, tapes are probably still not the right option. The tapes will eventually degrade, and at some point, not only the medium, but also the format will become unreadable. Meanwhile, as part of my civil service last summer, I worked in a museum, happily taking inventory of ≥100-year-old postcards that looked brand new.

Now, let's rewind back (pun intended) to the time when some of these postcards have been printed.

A Little Bit of History

Magnetic tapes were not just invented out of thin air. They were an improvement over an existing technology. That year, the dane Valdemar Poulsen invented magnetic wire recording. It works like this: A steel wire is pulled along a recording head which magnetizes it. The magnetization creates a signal which later can be read and interpreted in some way - be it analogue or digital.

The wire moves across the recording head with a speed of 610mm/s, thus, for one hour of recording, 2200m of wire would be needed. Because the wire is only about 0.15mm thin, the spool holding the wire would still be fairly small.

Magnetic tape, which is tape with a magnetizable coating, was invented by ze Germans in the 1920s. It works the same as magnetic wire, but was initially more difficult to produce. Consumer-level recording devices first became available in the 1940s. They were not like the cassettes that we know. Rather, the tape moved between two separate reels. At that time, wire recording devices were still more popular than tape recorders due to their lower price. It was only in the mid-1950s that tape recording took over.

Magnetic wire recording did not completely vanish for quite some time. Because steel wire is more heat-resistant than plastic, it was used in air- and spacecraft up to the 1970s[3].

One more thing - obvious, but still fascinatingly weird: How do you edit a wire recording if you don't have two devices? You just cut the wire and tie the ends. Because the wire speed is so high, the skips due to the knot are barely noticeable.

➔

With the advent of the computer, tape reels similar to the audio ones were also used for storing data.[4] In early IBM supercomputers, these reels were kept in a vacuum and would move in rapid bursts (youtube for "Vacuum Column Tape Drive" to see what I mean).

In the 1970s, cartridges became more commonplace and the cassette tape (introduced by Phillips in 1962) soon became one of the main formats in music distribution[5]. Cassettes were also used in home computing as a cheap alternative to floppy disks, although they had the disadvantage that the user often needed to manually start and stop the tape. For the Commodore and many other brands, the tape drive that connects to the computer was called Datasette.[7]

In the 1980s, better materials became available, which made thinner and more sensitive tapes possible. The limits of tape are still actively being pushed. This year, IBM and Sony achieved a new record in magnetic storage density by cramming 201 Gigabits onto a square inch of tape[6]. Theoretically, about a kilometer of tape would fit into a palm-sized drive with a capacity of 330 terabytes

Cassettes where a cheap alternative to floppy disks, although they needed to be started and stopped manually

Storing Data on an Audio Cassette

Back to the hipster backup. I wanted to store a file on an audio cassette, so at my parents' place, I dug up some 20-year-old cassette tapes which seemed to be in perfect condition. The setup to write a file to a cassette is pretty simple. First, encode the files into waveform audio in some format. Second, record that audio onto the tape. To read the file from the tape, it needs to be recorded to waveform, from which the file is recreated.

After about 5 minutes of googling, I found a suitable software for encoding and decoding: Minimodem[8], which you can install with the package manager of your choice. Minimodem enables any form of data transfer via sound and supports a wide range of protocols.

Obtaining a cassette recorder was a bit more difficult. After some asking around, I was able to use a Panasonic Ghettoblaster at a friend's home. To transfer the sound waves from my laptop to the plastic monstrosity, I used a 3.5mm cable. As I soon found out, the 3.5mm port of a current MacBook Pro does not double as line-in port, which I wrongly assumed. Thus, I had to record the file back via the laptop's microphone[9].

The protocols for converting data to waveform have originally been developed for data transfer over radio, phone lines, etc. and have different trade-offs between bitrate and robustness (how likely it is that the data can be recovered from a lossy signal). They all use frequency shift keying, which basically means representing different bit patterns with different frequencies. I considered three protocols for my cassette trial: RTTY, Bell103 and Bell202.

Radioteletype (RTTY) is a protocol that was invented by the US military in the 1920s and has been extensively used during World War II[10]. Its purpose was to send text between two stations via a terminal, for instance from an air base to an airplane - sort of like a telegraph, but with keyboard input. RTTY has a bitrate of 45 bits/s - so a short paragraph text that takes 30 seconds to read out loud takes about 75 seconds to transmit via RTTY. While this is quite disappointing, reading the text back in, even via the microphone, worked flawlessly. The characters will be converted to ALL CAPS because RTTY uses Baudot code, a character set with five bits per character that does not support capitalization. So this whole Ghettoblaster setup is essentially a complicated way of implementing `` `string.toUpperCase()` ``.

So this whole Ghettoblaster setup is essentially a complicated way of implementing `string.toUpperCase()`

Next, I tried modem protocols with higher bitrates: Bell103 (300 bits/s) and Bell202 (1200 bit/s). These yield much more practical file lengths (15 respectively 4 seconds for the above mentioned paragraph of text). Sadly, minimodem was not able to recover the data from the waveform recorded via the microphone. Over the wire it should have been possible because the cassettes used in home computing used similar bitrates.

If you want to try for yourself or just know how data converted to different protocols sounds, here's how you can reproduce it yourself (spoiler alert: It sounds like a dial-up modem). To encode a file of your choice to RTTY (The file needs to be encoded in ASCII, you can base64 encode non-text files first):

```
$ cat file.txt | minimodem --write
rtty -f audio.wav
```

To recreate the file from waveform:

```
$ minimodem --read rtty -f audio.wav
> decoded_file.txt
```


Sources

[1] Amazon Glacier - https://en.wikipedia.org/wiki/Amazon_Glacier

[2] Tape storage at CERN - https://home.cern/about/updates/2017/07/cern-data-centre-passes-200-petabyte-milestone

[3] Wire recording (Wikipedia): https://en.wikipedia.org/wiki/Wire_recording

[4] Magnetic tape storage https://en.wikipedia.org/wiki/Magnetic_tape_data_storage

[5] Compact Cassette https://en.wikipedia.org/wiki/Compact_Cassette

[6] Sony and IBM's storage density world record https://arstechnica.com/information-technology/2017/08/ibm-and-sony-cram-up-to-330tb-into-tiny-tape-cartridge/

[7] Datasette: https://de.wikipedia.org/wiki/Datasette

[8] MiniModem: http://www.whence.com/minimodem/

[9] There are actually adapters from 3.5mm headphone jack to headset-style microphone audio in, which would have worked, but I could not obtain one at the time of the experiments)

[10] Radioteletype (RTTY): https://en.wikipedia.org/wiki/Radioteletype

For general information about the Magnetic tape, the English and German Wikipedia entries: https://de.wikipedia.org/wiki/Magnetband respectively https://en.wikipedia.org/wiki/Magnetic_tape