Automated post-processing of videos

Egg, July 3, 2021: *The private film archive of the managing director has meanwhile grown to a considerable size. With 6 TBytes, the question arose: expansion, stop or standstill, or optimization. This blog shows why the optimization option was chosen. At the end of the article there are instructions for putting any ArchivistaBox or AVMultimedia into deep sleep or waking it up again via the network.*



Simply more hard disk space does not lead to the solution

The ArchivistaBox is currently designed for up to 200 TBytes as standard. It would therefore have been simple in itself to expand the archive with additional hard disks. However, it must be noted that this only creates the primary space for storage on hard disks. Only when the materials are stored on write-once media do long-term storage media exist.

The M-Disk and proprietary Sony-Pro formats are available for this purpose. M-Disk drives are inexpensive (approx. 120 francs/euro), while the media currently costs approx. 25 francs/euro and holds 100 GBytes. With 6 TByte, 120 media would already be required for two archive copies. This would result in costs of over 3120 francs/euro (120*25+120). With the proprietary Sony format, the cost of the media is lower (approx. 250 francs/euro for 5.5 TByte), but the drive costs approx. 4000 to 5000 francs/euro.

Doubling the capacity to 12 TBytes would therefore result in costs of 6120 francs (240*25+120) for clean archiving of the data on non-rewritable media. Frankly, this seems a bit much for a private media archive. Especially since the cost of creating the media is not even included. An expansion of the capacity was therefore not an option.



Standstill would mean, you snooze, you rust

The private media archive project came into being precisely because the offerings of the streaming providers would not only have resulted in higher monthly costs in the medium to long term, but also because streaming ultimately corresponds to nothing more than a very one-sided rental relationship.

The private archive was designed to hold about 8 TBytes (equivalent to plus or minus the ArchivistaDom). In the last eight months it became apparent that so many offers are available via **MediathekView** as well as via **wilmaa.ch** that the archive grew much faster than calculated. By the way, one reason for this was that many CDs and DVDs were found while rummaging in our own cellar, which were also recorded.

At the same time, it could be seen that currently many DVD collections can be purchased for a “ridiculous price”, since obviously no one wants to buy and insert DVDs anymore. According to Swiss law, these DVDs may be copied to the hard drive as a private copy even if they are delivered with copy protection. To ‘rip’ DVDs, there are appropriate programs for that.



But: Even if DVDs can be legally “ripped” in Switzerland, it is not free. According to **Suisseimage.ch,** copies are financed via the **blank media levy (in German).** According to **Suisa.ch (in German),** the fee for M-discs is **just over 1 franc.** The **FAQ from Suissimage.ch (in German)** seems paradoxical here, since on the one hand fees are charged on blank media, while on the other hand offering programs for DVD copying is prohibited under **Art. 39a Para. 3 URG (in German).** Nevertheless, according to Art. 39a para. 4, circumventing copy protection for private copying is legal. The length of this paragraph shows, the topic is not simple, a discussion on the topic can be found in the **MakeMKV forum.**

With newly acquired DVDs in the double-digit franc range, about 150 hours of film material were added again. Less money was spent than a corresponding subscription would have cost over eight months with the market-leading streaming provider. But this also meant that the size of the archive grew much faster than planned.

Since expanding the hard drives was out of the question (see above), the second option would have been to “mothball” the media archive with its existing volume. Ultimately, it should not come to rest, because this would have started the archive symbolically speaking to rust digitally.



Solution: Optimize the video files
Over the last few months, it has been noticed that the video files are available in very different sizes depending on the source. For example, when downloading via **MediathekView,** it had to be determined that the content is increasingly no longer provided in HD quality with 25 images, but with FullHD and 50 images. As a result, a one-hour documentary no longer requires about 1 GByte, but somewhere between 4 and 5 GBytes.
Ultimately, the file size of MP4 files is primarily related to the bandwidth used for streaming. Here, too, it had to be noted that the content is increasingly delivered in “higher quality”. This may be technically feasible with the higher bandwidths of Internet providers. For a private media archive, on the other hand, this is not very helpful for cost reasons.
Now the ArchivistaBox offers the entire multimedia stack. Consequently, the video files could simply be re-rendered with 25 frames and HD quality using **Shotcut,** for example. But ultimately this would mean that a lot of manual work would be necessary. The corresponding processes can of course be automated with the **console program ‘ffmpeg’.** This is where the ArchivistaBox 2021/VII comes in, in that these processes can be carried out automatically.
The **‘Manage jobs’** form now exists in **WebConfig** for this purpose. The first two options ‘Check external content’ and ‘Clean up internal check values’ are used to check audio and video files that have already been captured for correctness. The item **‘Optimize video files’,** on the other hand, brings the possibility to optimize already existing MP4 files.



Findings from the optimization

In the last month, all external material (the own created videos were not optimized) was rasterized to HD quality with 25 frames per second. This reduced the size of the archive from about 6 to about 2.4 TBytes, which is plus/minus a factor of 2.5 less data.

This is not nothing. Because instead of 120 M-disk data carriers, not even 50 M-disks (for 2 copies each) are now needed, which corresponds to a “saving” of approx. 2000 francs for the data carriers. Of course, costs of about 1000 francs for long-term archiving are still not nothing.

However, the M-Disk format has a high physical life expectancy over many decades (even 1000 years are promised). If 10 to 20 years would also be enough, normal Blue-Ray data carriers could be used. The costs would be reduced to about 60 percent.

Whoever now wants to object that any streaming offer can be financed for this amount, should be told that the approx. 600 francs with normal Blue-Ray data carriers are wasted in less than three years with the family subscription. And that doesn’t even include the price increases that are currently under consideration.



Blind test and patience

By the way, a blind test in the family showed that no one could notice a visible loss of quality. The deciding factor in playback is not primarily whether 50 or 25 frames with HD per second are available in the files, but how well the video player can play back the available data or how good the monitor itself is. Since these components were not changed, there was no visible difference in quality visually.

Optimizing video files requires considerable resources. On the ArchivistaDom, it would have taken about 60 days to re-render about 3000 hours. On ArchivistaK2, it would probably have taken about 12 to 15 days, and on ArchivistaEverest, it was ultimately less than 3 days. Or calculated the other way round. With ArchivistaDom, approx. 80 to 100 hours can be optimized per day, with ArchivistaK2 it is approx. 250 hours and with ArchivistaEverest it is a good 1000 hours per day.

Conclusion: Optimization is always worthwhile

Optimizing video files makes a lot of sense, especially for private archives. Not only can considerable monetary resources be saved, but a smaller data stock also leads to a correspondingly better eco-balance. According to a **WDR article from 2019, streaming is even extremely harmful to the climate (in German).** As the article correctly states, corresponding calculations are indeed realtive. However, it is undisputed that the larger the data stocks are, the greater the environmental costs will be, see also the article at **Welt.de (in German).** In this sense, the optimizations of version 2021/VII make a lot of sense not only economically but also ecologically.

Such optimizations naturally also result in a high economic benefit in the professional environment. As an example, we will only mention a school media server, where the content for teaching can be reduced by a factor of three before working with it. This is not just about the three times smaller storage requirements, but also the three times smaller bandwidth required to play the content.



Tip of the month: Deep sleep and waking up

In line with the topic of power consumption, we would like to describe here how an ArchivistaBox or AVMultimedia can be put into deep sleep or can be restarted via the network. To do this, two lines must first be entered in the **/etc/network/interfaces** file. The following is an example:

```
auto lo
iface lo inet loopback

auto vmbr0
iface vmbr0 inet static
  address 192.168.2.222
  netmask 255.255.255.0
  gateway 192.168.2.1
    post-up /sbin/ethtool -s eth0 wol g
    post-down /sbin/ethtool -s eth0 wol g
  bridge_ports eth0
  bridge_stp off
  bridge_fd 0
```

The two colored lines must be added. After that, the network card must be reinitialized. The following command can be used for this:

/etc/init.d/networking restart

Before the ArchivistaBox or AVMultimedia can be put into deep sleep, the Mac address of the computer must be known, as only with this address is it possible to wake it up from deep sleep later. For this purpose the command **'ifconfig eth0'** can be used. The following output is displayed:

```
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
  ether 70:85:c2:db:70:b5 txqueuelen 1000 (Ethernet)
  .... (weitere Zeilen)
```

The colored area behind 'ether' is needed to wake the ArchivistaBox or AVMultimedia from deep sleep.

First, however, the computer must be put into deep sleep. The following command is used for this:

echo mem >/sys/power/state

If the screen is not deactivated, the following command can be used to check whether the sleep mode is available at all:

cat /sys/power/state

Note: *If 'mem' is missing from the output, the function is not available from the hardware used.*
Now the ArchivistaBox or AVMultimedia can be woken up from deep sleep by another computer. To do this, enter the following command:
`wakeonlan 70:85:c2:db:70:b5`
Note: *The Mac address is different on each computer. It goes without saying that the Mac address used on the ArchivistaBox or AVMultimedia must be used. Incidentally, the wakeonlan program has only been included on the ArchivistaBox or AVMultimedia since version 2021/VII.*



Of course, the deep sleep or wake-up only works if at least one other computer is running in the same network. Either the firewall or a small single-board computer (Odroid, Raspberry etc.) can be used here. The result of these efforts is a power consumption (standby) of less than 1 watt. Currently there is no GUI for deep sleep, neither with AVMultimedia nor with the ArchivistaBox. However, we will gladly accept any corresponding requests.
P.S: The pictures in this blog are from the Laugavegur hiking trail in Iceland. If you want to know more about it, you can find a good 70 minutes long film at ***https://azurgo.ch/english.*** *The raw material for this work comprises about 800 GByte of data. This alone shows how valuable media archives are.*