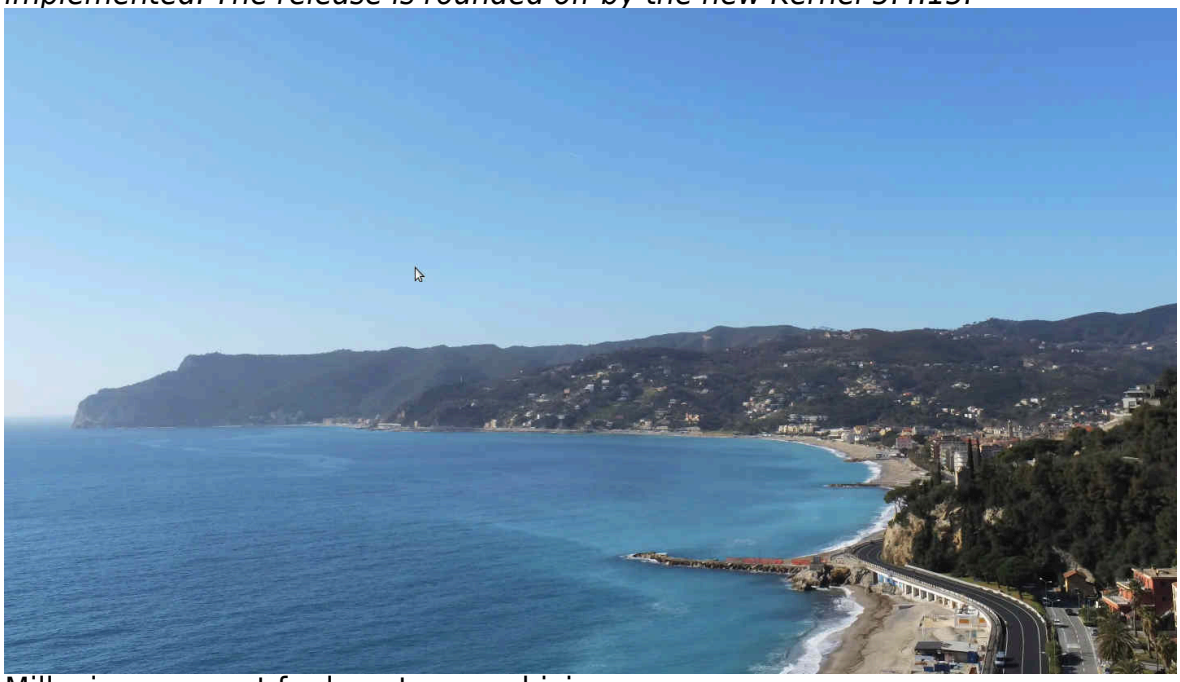
ArchivistaBox & AVMultimedia with long-term archiving
Egg, 31 January 2020: *Both AVMultimedia and the ArchivistaBox 2020/II are now able to create optical data media with up to 100 GByte (BDR-XL). This means (this is particularly important for multimedia content) that even the most extensive projects can now be archived on optical media. In addition, the new release includes support for the latest AMD processors (e.g. Ryzen 3950X) with up to 32 threads and support for data media with up to 16 TByte has also been implemented. The release is rounded off by the new Kernel 5.4.15.*



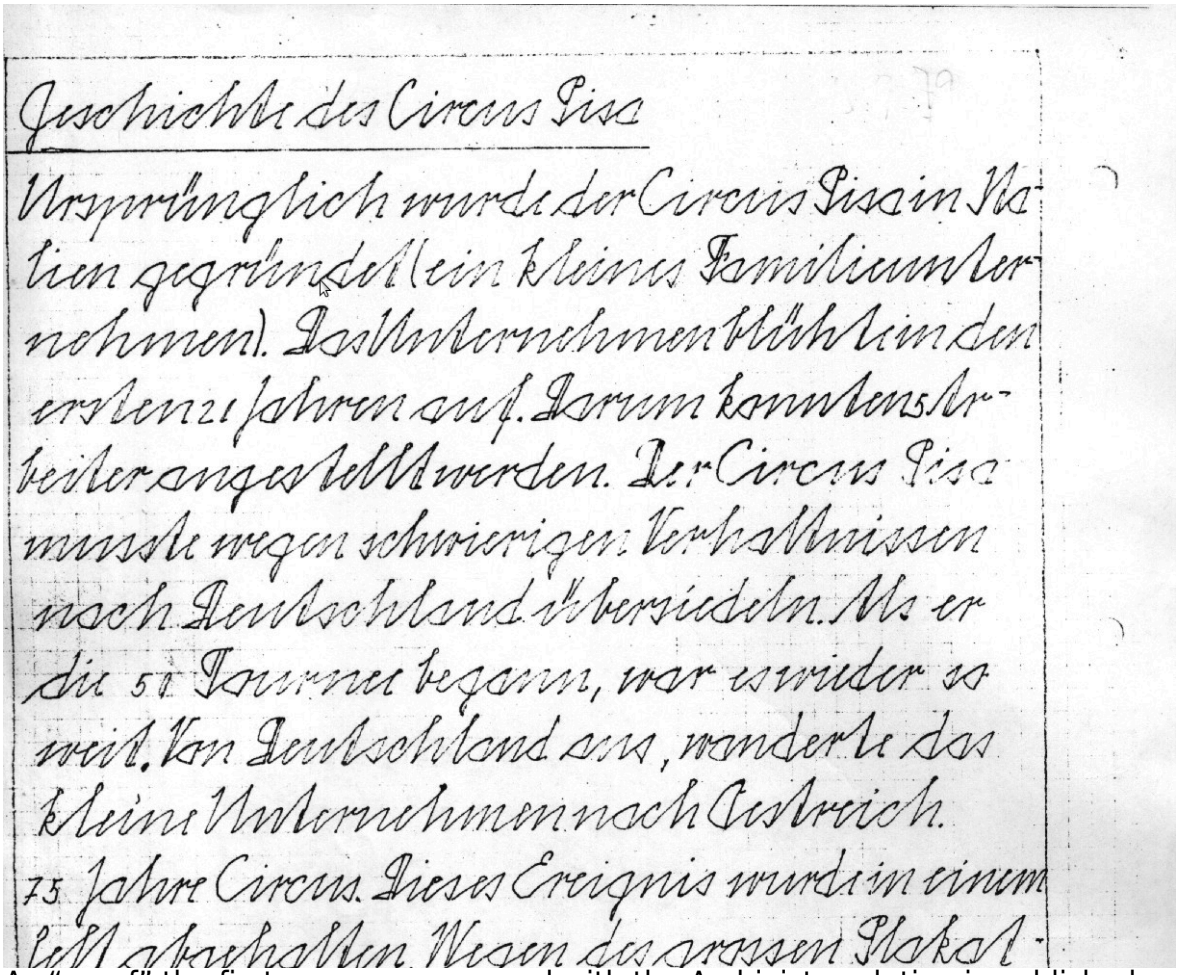
Millenium concept for long-term archiving
Anyone who stores data today uses only hard disks. Modern SSD disks offer extremely fast access times. However, it is often forgotten that data on hard disks is “only volatilely” stored. Just as fast as they can be written to the data media, they are also deleted again, or even worse, even changed unnoticed. Especially for systems in the field of document management (DMS) as well as for archiving, test procedures have been used for some time to ensure that data has not been changed.



Whether it is a question of (not) official certificates or simple cross-checking procedures (keyword MD5), does not play a role in that. The test data can be used to determine whether data has been changed or not, but it cannot prevent the data from being changed unintentionally and/or deliberately.
Many users are not really aware of the consequences of this difference. What good is the best method of checking to find out later on that the data has been changed after all, the digital “originals” are irrevocably lost when they are changed.
Ultimately, hard disk publications are of no help here, because conventional data carriers “empty” themselves over a period of several years. In other words, the contents of hard disks that are saved in the Save for 10 years, for example, are in great danger of no longer being able to read these data.
The only effective protection is offered by optical media where the data is backed up writable only once. These data carriers offer (if stored correctly) the protection that important information deserves. The ArchivistaBox has therefore been offering exactly the concept of backing up data on optical data storage media for more than 20 years.

“History” of archiving with optical data storage media

In 1998, CDR data media, which have a capacity of around 700 Mbytes, were well suited for backing up data held at the time. With black/white, at 700 MByte (or 50 KByte per page), 14’000 pages per optical disc can be archived.
Even though DVD drives or the corresponding media were already available at that time, the DVD media of that time did not have the best reputation for longevity, see also the good article at **https://en.wikipedia.org/wiki/DVD.**
Compared to DVD media, CDR media offered a very good prognosis. The manufacturers gave 30 to 50 years of readability to the corresponding data media. The first version of the Archivista software dates from 1998, the first created data media that could be found were from April 1996, which corresponds to around 25 years. In short, the data could be read without any problems (incidentally even silver-coated CDRs).


Geschichte des Circus Pisa

Ursprünglich wurde der Circus Pisa in Italien gegründet (ein kleines Familienunternehmen). Das Unternehmen blühte in den ersten 12 Jahren auf. Darum konnten Mitarbeiter angestellt werden. Der Circus Pisa musste wegen schwierigen Verhältnissen nach Deutschland übersiedeln. Als er die 50 Tournee begann, war es wieder so weit. Von Deutschland aus, wanderte das kleine Unternehmen nach Oestreich.
75 Jahre Circus. Dieses Ereignis wurde in einem [illegible] Wesen das grossen Plakat-

As “proof” the first page ever secured with the Archivista solution is published here. This is the first page of a story that the current managing director submitted in 1978 in a writing competition for a guest performance by the well-known cabaret artist Emil Steinberges at Circus Knie in 1977. The prize was a book on the guest performance of the entertainer and a recording of the performance. The manuscript was digitized in 1996; the record itself could not be digitized at that time. In the meantime even this would be possible, for **about 100 francs or Euro there are players for the old vinyl records.**
As mentioned above, only a few self-burned DVD media were produced internally, but even these approx. 15 year old DVDs could be read without any difficulties. Another decisive factor for readability is that the data was saved in a generally readable format. With the Archivista solution this has always been the case because rasterised image data is always created from all data.
If the data had been created with hard disks, tape drives and/or WORM data carriers (all three formats were considered the measure of all things in the industry at the time), then this data would hardly be so easily available, if at all, because the optical data carriers in 5.25 inch format have an undeniable advantage, the size of the data carriers (5.2 inches) remains the same for decades and the data is only backed up once in a writable form.

ArchivistaBox and AVMultimedia 2020/II with M-Disc

The biggest disadvantage of the previous backup to optical data carriers is the very limited storage capacity of CDR (700 MByte) and DVD media (4.2 GByte). If hard disk sizes of 16 TByte are available, this would result in 3900 discs even with DVD media. At a weighted 40 grams per DVD, this would result in 156 kilograms.
This would allow over 335 billion documents to be archived, which in turn would weigh a whopping 1.3 million tons. But in the year 2020 it is not only about

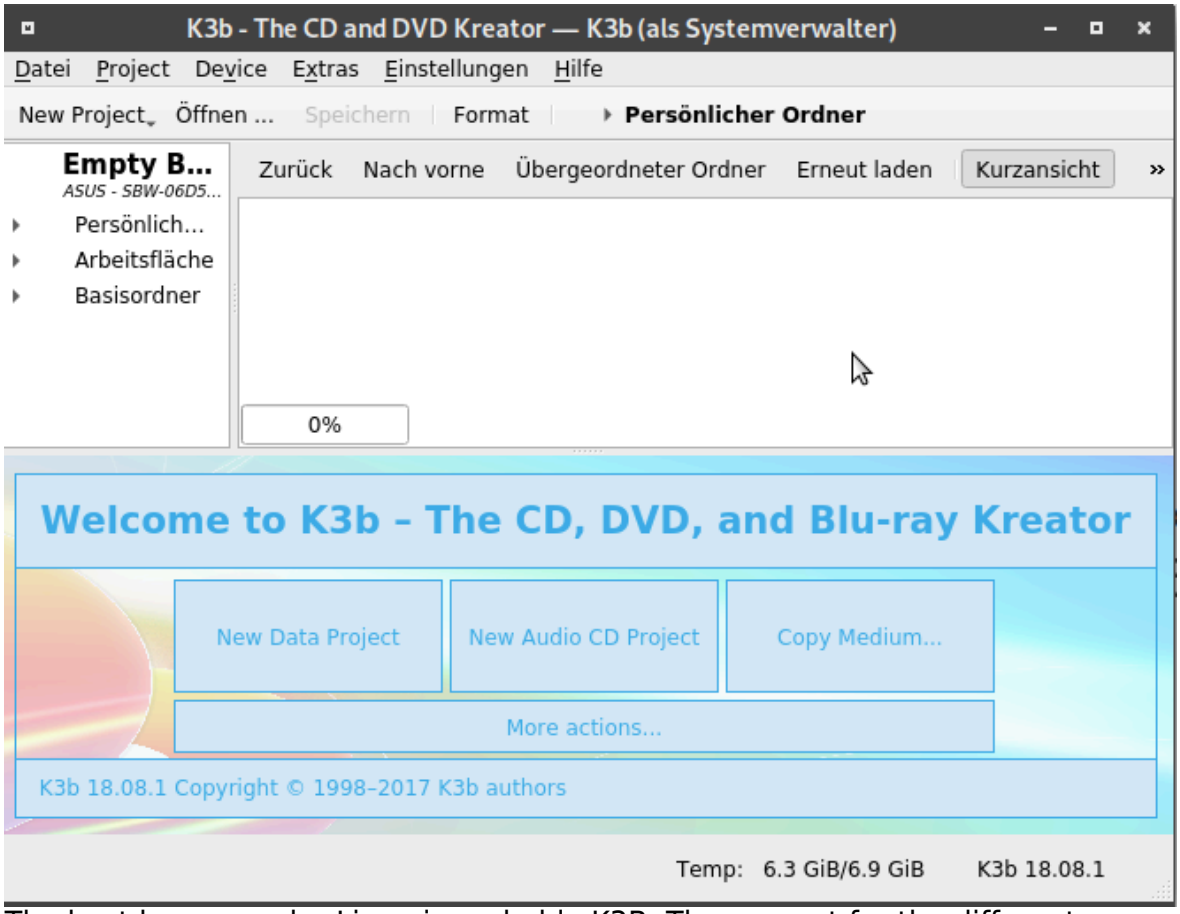
archiving documents. Rather, it is also a question of keeping audio and video material available in the long term. One hour of 4K film material in good quality requires about 10 to 20 GByte of data.
With 16 TByte capacity (at 10 GB 4K material per hour) “only” 1600 hours of film material can be archived. And yes, it is not possible that 1 hour of material in 4K fills more than 2 DVDs. For some years now, BlueRay discs have therefore been available which can store approx. 25 GByte. In the meantime, BlueRay can also be written in several layers (BDR-XL), with which up to 100 GByte of data can be saved per disc.
When using BDR-XL, there are still 160 data carriers for 16 TB, and at 40 grams per disc, this adds up to a weight of 6.4 kilograms. This is still about six to eight times more than a comparable 3.5-inch hard disk, but overall it is still very “bearable”.
It is also important to note that in addition to the BDR-XL discs, there are also those in M-disk format. The former are less expensive, the latter offer a significantly longer lifespan. According to the manufacturer’s specifications the durability is stated to be up to 1000 years (M-Disk=Millenium-Disk), this in contrast to the normal BDR-XL data media, where the lifetime is said to be about 20 to 25 years.
Of course these specifications are only valid if the media are stored correctly, i.e. at low humidity, constant low temperature and without sunlight, but these requirements apply to all media (including hard disks).

K3B, Brasero and Xfburn in test

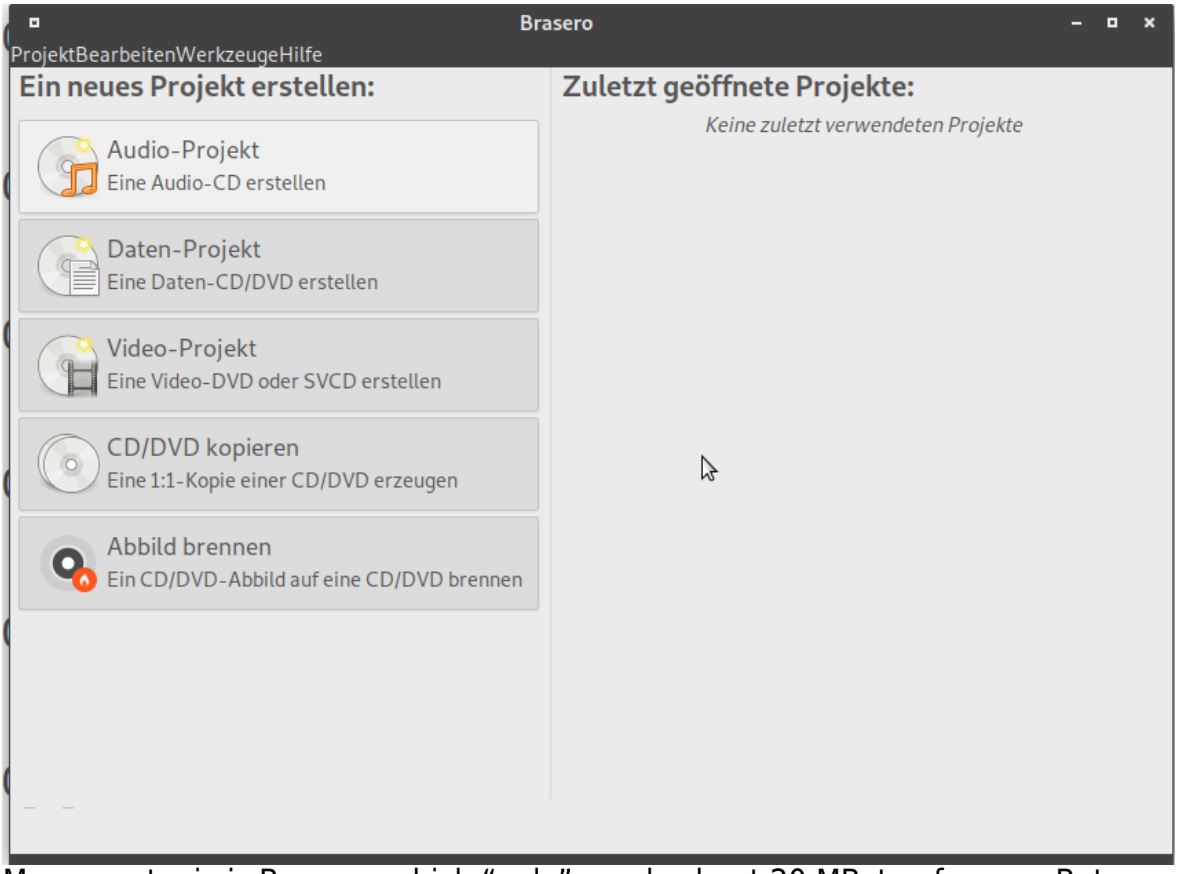
A good support of BlueRay (or BDR-XL) was not or only rudimentarily available under Linux for a long time. Meanwhile (or with current Linux kernels) the corresponding drivers are available. The writing of such media with console programs (e.g. xorriso) succeeds without difficulties. Here is an example for the console:

```
xorriso -as mkisofs -iso-level 3 -o /home/data/data.iso
/home/data/archivista/*
growisofs -Z /dev/sr0=/home/data/data.iso
```

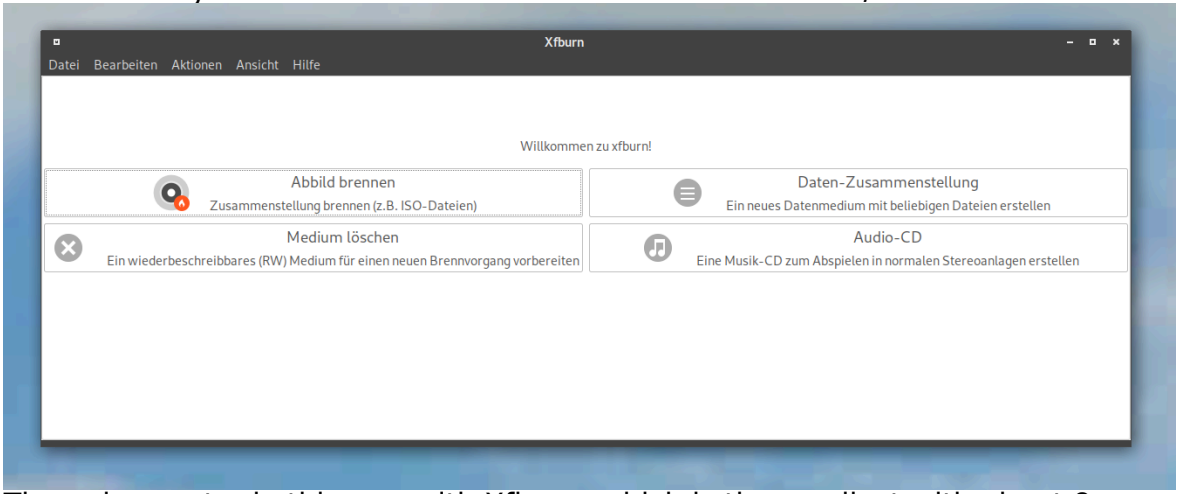
Now working with the console is not what “normal mortals” usually want. For this reason it was necessary to evaluate a suitable graphic tool for the ArchivistaBox 2020/II and AVMultimedia and ultimately to integrate it into the distribution.



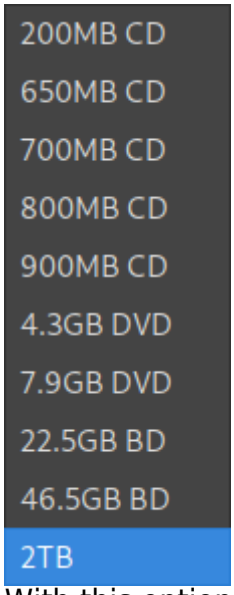
The best known under Linux is probably K3B. The support for the different formats is very good, but K3B needs about 150 MByte of software. Apart from the size, K3B offers an unbelievable amount of parameters, but these can be confusing when getting started.



More spartanic is Brasero, which “only” needs about 20 MByte of space. But Brasero (at least not with the version 3.12.2 tested with AVMultimedia) was not able to create BDR-XL media with 100 GByte. For this reason Brasero also failed to find its way into AVMultimedia or the ArchivistaBox 2020/II.



The only way to do this was with Xfburn, which is the smallest with about 8 MByte. Xfburn offers few but effective options and the tested version 0.5.5 is able to create BDR-XL with 100 GByte. The only thing you have to get used to is that you have to select the option 2 TByte when creating the image file.



With this option BDR-XL media could be created without any problems. It is highly recommended to always create an ISO file (image) first, and only then burn this file onto the optical disc.

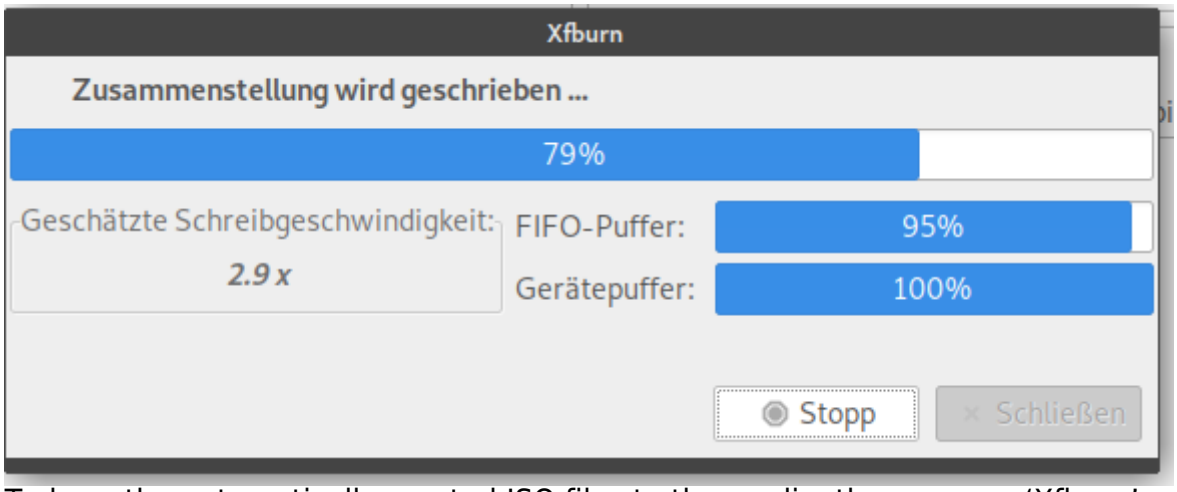
Options for ArchivistaBox 2020/II in WebAdmin

In order to now archive data from the ArchivistaBox with up to 100 GByte, three parameters must be adjusted in WebAdmin:. For ‘swap folder size’ 10000 MByte should make sense, we also recommend limiting the maximum number of files to 10000 and for ‘CD/DVD size’ 10000 MByte. (MByte)’ the value of ‘92000’ must be entered for BDR-XL.



In the end ‘92000’ is not 100 GByte, but first of all the last blocks of an optical medium should never be written completely (statistically seen the error rate is highest there) and secondly the 100 GByte on the invoice are 100 * 1000 MByte * 1000 KByte * 1000 Bytes. If you divide by 1024 (mathematically correct calculation method), the 100 GByte BDR-XL “only” holds a good 93 GByte. However, this should not fundamentally reduce the new possibilities. Thanks to the BDR-XL format it is possible to write far more than 150 CDR disks to a data carrier.

If the above options are entered correctly in WebAdmin, the corresponding ISO files are automatically created during archiving. You can find them in the folder ‘/var/lib/vz/template/iso’, whereby the name of the database and the name of the first folder is always used (example ‘archivista_ARCH0002.iso’).



To burn the automatically created ISO files to the media, the program ‘Xfburn’ can be started under ‘Multimedia’. To do this, a medium should be inserted beforehand, as this is the only way the burner can be correctly recognised. For the first screen, select ‘Burn image’ and the desired file from the folder ‘/var/lib/vz/tempalte/iso’. Afterwards the process can be started with ‘Burn image’. The burning process takes a little more than an hour with 90 GByte. Basically you should always create two disks at a time. Whether two different makes and/or drives are used is debatable. The market leader in terms of data media is Verbatim, but corresponding media is also available from Sony, and there are currently a considerable number of manufacturers for drives (Asus’ drives were tested with USB3.1. However, a USB2 drive is also sufficient for the amount of data to be written.



Other new features in the current release

With release 2020/II the Linux kernel 5.4 is delivered. This kernel branch will be available for the next years. The latest AMD processors are supported. The new version was tested with the Ryzen 3950X with 16 cores or 32 threads. The performance of this processor is remarkable. 4K films can be created almost in real time (approx. 70 minutes for 60 minutes of footage).

Better support for UEFI and legacy bios derivatives was also implemented. The current version tests whether either UEFI or legacy is booted. Accordingly, the software is set up on the hard disk. Those who have worked with legacy can

continue to do so, those who boot with UEFI will always have UEFI set up for installation. New, however, is the complete support for hard disks over 2 TB, both for UEFI and UEFI.
If you want to get detailed information, you can find the corresponding hints in **the Devuan forum.**

Long-term archiving over decades

At the end of this blog we would like to **refer to a contribution on the same topic from 2008 (which is 12 years, sorry, only in German).** In this article, everything is explained that already existed at that time for the Archivista solution and is still valid today. For all those who prefer to look to the future, it is enough to know that with version 2020/II not only can multimedia content be archived, but also that it can be neatly stored on optical data media for decades.