Secure the Future of your History

Long term data archiving on Disc Archiving Systems Blu-ray storage solutions.

The purpose of this document is to provide answers to questions concerning the use of Blu Ray for long term archival such as:

1. Why archive on Blu-ray?
2. Why not use LTO (magnetic storage) for archiving?
3. If we archive using Blu-ray, how does this look on a systems level?
4. What are the realistic (practical) read and write speeds and what is their impact on other systems parameters.

This document is drafted with a few specific cases in mind. Moreover it contains company confidential information provided by DISC. Hence this document should be treated as company confidential under NDA. The content of the document is DISC's sole responsibility.

Outline of the use-case under consideration.

The end client is looking for a long term 50 year archiving solution for both "near-line and off-line". They need a system that will accommodate 2.6Petabytes every year for 50 years.

The client generates data mostly from (1300) camera's (hence the data is mostly video). Each camera generates about 2 TB of data per year. (Hence 1300 * 2TB = 2.6 PB). Obviously the client captures and stores all video on HDD first. However as a second tier, they want to move the data off HDD to a permanent media that can be exported and stored off-line for long term archival.

This use case resembles very much a current DISC project with a large US Government Agency. This project is in its starting phase; the first system has been delivered and accepted. The completed project is expected to roll out in 2014. The Disc Archiving Blu Ray Systems will be located at ten to twenty remote sites all equipped with DISC's ArXtor 1000-105-02 (small system with two drives; one smartPack holding 15 discs that will store data for two to three days and then removed from the system as off line storage with key location identifiers that make retrieval of the files very easy.

DISC Archiving Systems B.V.
Waarderweg 52G, 2031 BP Haarlem, The Netherlands
Phone: + 31-(0)23-5 172 635, Fax+ 31-(0)23-5 172 639
Chamber of Commerce: 34361892, NL8213.86.773.B.01

Archiving data; a historical perspective

The IT industry has a history of some 30+ years. The seventies and eighties of the 20th century should be regarded as the digital pre-history. Like in real history, we do know facts (i.e we have artifacts) from the pre-history, but we don't have written documentation. Getting data from the era of the digital pre-history is similar, it will be difficult or impossible to read data today of files from the first IBM compatible PC's (floppy disks?). The problem with these pre-historical data formats is not only that the actual media may be corrupted (i.e. are the magnetically recorded bits still there?), there are also two other problems :

1. Are there drives available to read this data?
2. Do we have drivers (i.e. under current operating systems) that support these devices?

It is exactly this perspective that we use, assessing the subject of digital archiving.

Archiving on magnetic tape (LTO)?

Today many people claim to be archiving on magnetic tape (LTO). LTO media vendors claim that their products have a media life of 30 years. These numbers are derived from accelerated media life testing (in climate chambers) and focus on issues like 'media corruption' and stability. They are trying to answer the question whether or not the bits will still be on the tape after 30 years. In our opinion this is not a very relevant matter. LTO tape formats change every year. LTO 3 (released 2005) tapes can be read in LTO 5 (released 2010) drives, but LTO 2 (released 2003) tapes cannot be read by them. This means that data stored on tape must be migrated at least every 4-5 years. Otherwise it will be lost automatically because generation changes imply inherent incompatibility and eventually yield data loss. The rationale for using Blu Ray instead of tape is similar as the case outlined above as we cannot read data from floppy disks any more, not because the media is corrupted (there is a good chance on that), but simply because the equipment to read the data is no longer available.

So in reality people don't archive on tape, they store data on tape. Obviously if one doesn't want to lose the stored data, one can retrieve and re-store data on new media. This is called the process of data migration, a strategy deployed by many tape library companies. However, people should be aware of the costs associated.

1. Migration implies people touching old media, and re-storing to new media, however there is a labor cost factor that has to be considered and can only be derived from customers past experience.
2. The old tapes can be thrown away generating chemical waste after a shelf life of only 4-5 years. This is 6 times earlier than the 30 years media life that the vendor originally advertised and not very environmentally friendly.

DISC Archiving Systems B.V.
Waarderweg 52G, 2031 BP Haarlem, The Netherlands
Phone: + 31-(0)23-5 172 630, Fax+ 31-(0)23-5 172 639
Chamber of Commerce: 34361892, VAT: NL8213.86.773.B.01

3. New tapes must be purchased and corporations have to allocate budgets for the data migration just to preserve old data that might be important and that could add a significant factor to the original cost? .
4. Last but not least, there is a risk element associated with re-storing. Are we absolutely sure that the old files are correctly read and that they are written correctly? If we are working with TB's of data we can't check each and every file.

In summary, tape is an excellent medium to store digital data for a relatively short time (i.e. 4-5 years). However using tape for long term archiving is simply an ostrich-strategy creating a significant problem after 5~ 8 years. The key to most economic decisions is the ROI question. Let us assess this, in the context outlined above.

Currently LTO6 tapes can store 2.4TB and cost 250$. Using the 30 years life expectancy that media vendors claim, this yields:

A cost per TB of 250 $ / 2.4 TB = 104 $/TB
Archiving the data for 30 years implies about $3.50 /TB/year.
Obviously this is a number that tape vendors market and would like people to believe.

The reality however is different. Best case, every five years the data needs to be migrated. This yields a cost of $21.00 /TB/year which is 6 times higher and you have to add in the labor and new media cost and the organizational burden of migration every 5 years!

About Data Archiving on Blu-ray.

Let us look at the alternative of Data Archiving on optical media.

The first and easy observation that we can make is the following. If we would have started some 20 years ago with CD, the predecessor of Blu-ray.CD has been succeeded by DVD and recently by Blu-ray. CD disks can be read in DVD drives as well as in Blu-ray drives. DVD disks can be read in Blu-ray drives. This media format doesn't have the problems of magnetic tape of periodic format disruption and forced migration enriching the pocket books of the media vendors. Data archived 20 years ago to CD can still reside on that same CD, as we can still read this data today in any current Blu-ray DVD Blu-raydrive. In addition, operating systems provide a native support for these drives making these media readable in any PC or laptop.

In the literature there have been several attempts to discredit CD/DVD/Blu-ray media implying short term media corruption. However Media vendors like Panasonic and Mitsubishi have done numerous life cycle tests proving the opposite. Clearly these accelerated life cycle tests reveal significant differences in quality and data stability directly

related to the media manufacturing process. This has yielded different consumer and archival grade media Blu-ray-disk products. Hence in our opinion, Blu-ray is an excellent choice for data archiving, with the premise that high quality media is being used i.e., typically brands like Pansonic or Mitsubishi. We do not recommend the use of commercial brands that source media from different factories.
Cost of Blu-ay disks depend heavily on volume purchases from the manufacturers. The largest volume manufactured today of Blu Ray disks currently is single layer (yielding 25GB per disk). In the archiving business today we estimate that over 80% of the market uses 50 GB (dual layer) media. Volume (OEM) pricing is around 10$/disk, yielding a cost per TB of 200$/TB. The most recent development of Blu-ray is BD-XL, with 100GB disks (triple layer) volume (OEM) pricing of BD-XL is around 20$/disk yielding again a cost of 200$/TB. It is important to note to these prices that BD-XL manufacturing volumes are still quite low, whereas volume has a major impact on cost prices. We expect the volume to increase over the next couple of years. Hence prices are quite likely to go down.

Comparing these costs with LTO6, one observes

Real cost per TB/year for LTO 6 (5 year) = 21 $/TB/Year
Real cost per TB/year for Blu-ray (50 year) = 200/50= 4 $/TB/Year

This comparison reveals a remarkable cost difference between cost per TB between tape and Blu-ray. This difference is even more favorable for Blu-ray, because the cost of data migrations have been left out of the comparison.

System solution for the use-case assessed.

Now going back to the use-case presented. The DISC Archiving Systems latest and greatest product is the ArXtor series (see enclosed flyers). A key element of the ArXtor series is the smartPack. This is a pack filled with 15 Blu-ray Disks. Hence with BD-XL media, each pack holds 1.5 TB of data. (flyer).

The advantages of using smartPacks for the use case outlined can be summarized as follows:

1. Smartpacks can be burned with data by DISC's ArXtor systems. These systems are easy to connect from any server environment, commission and maintain. DISC provides a complete solution hardware, software and support from a single point of supply.
2. Once archived, ful, smartpacks can be exported from the library system and stored off-line. One smartPack holds 1.5TB of data. Typical write speeds today are up to 6 X BD-XL disks per day per writer taking on average 4 hours to burn one disc. (including a full read cycle for bit-to-bit verification)

3. SmartPacks can be purchased from DISC Archiving Systems filled with media. This implies a complete product delivery including manufacturer's warranty and volume pricing.
4. No need for end-users to handle or touch individual media disks. Smartpacks can be inserted and exported from libraries.
5. A systems solution would keep recent data in a cache on HDD and for long term/deep archiving near-line in a library and off-line in a stored smartpacks.
6. Blu-ray disks can be exported from the SmartPack and be read in any normal Blu-ray drive. Hence in the future, reading of data is assured by the mere existence of Blu-ray drives and their connectivity to PC systems. Obviously retrieval in a library systems is straightforward, but relies on the availability of a functioning library.

Based on the overall statistics of the use-case, this would imply:

- 2.6 PB -> 1730 SmartPacks per year, consuming 97 cubic feet of physical space.
- Approx. 5 SmartPack to be burned per day. This will require several ArXtor systems. DISC would like to optimize this element of the solution. Current requirement would be around 4 ArXtor 4000-145-6 systems (i.e. DISCs mid-range system with 6 drives).
- For reasons of redundancy we would recommend the implementation of 3 ArXtor systems in two different data centers.
- An alternative would be to use 6 - 8 ArXtor 1000-45-02 systems

DISC Archiving Systems B.V.
Waarderweg 52G, 2031 BP Haarlem, The Netherlands
Phone: + 31-(0)23-5 172 630, Fax+ 31-(0)23-5 172 639
Chamber of Commerce: 34361892, VAT: NL8213.86.773.B.01