

Astronomical Data Storage and Distribution in the next five years

B. Pirenne¹

*European Southern Observatory, Data Management Division, Garching,
Email: bpirenne@eso.org*

Abstract. In this review, the current status and expected evolution of data storage technologies in the next few years is considered in the light of the expected needs of constantly growing astronomical data volumes. Questions such as “should we abandon tapes?” and “why don’t we just transfer all data over the net?”, or “Why do we give data out in the first place? The VO will provide results!” will be discussed. The answers to the above questions have led to the definition of a new data distribution policy for the ESO/ST-ECF archive. This policy will affect both ESO program principal investigators as well as general archive users.

1. Astronomical Data Distribution: the accelerating evolution

1.1. Antiquity

Over 2000 years ago, man was already looking at the starry sky and was trying to build a model of the Universe based on the observations he made. Yet data recording and data distribution were major issues in those days as the only recording tools were eyes and hands. For data distribution, other pairs of eyes and hands had to patiently work to copy the works of others. This situation essentially prevailed from Antiquity to the invention of print. We owe copies of say, Ptolemy’s “*Almagest*” to the patient labor of Middle-age monks. It is important to note that there are still relatively many such documents, available today in libraries, musea and private collections. Those copies are mostly readable, proving that good quality paper and ink, together with the human eye allow for a very long life time of that medium. The order of magnitude for the lifetime of paper is therefore about 10^3 years.

1.2. Renaissance

With Gutenberg’s invention of print, the distribution of written material was suddenly a lot more efficient and enabled -for what concerns astronomical data- to easily equip a large amount of open-sea sailing ships with tables containing star positions. This was essential for navigation and has probably been instrumental to the rapid development of high sea navigation at the time. For example the Alfonsine tables created in the 14th century before the invention of print were

¹also at Space Telescope – European Coordinating Facility

to help sailors stay on course. None of the original ones have survived to this day. However, later copies printed in larger quantities are still available and readable. In 1543 and 1566, two editions of Copernicus' book *De revolutionibus orbium cælestium* were published in approximately 500 copies each and over one half survive to this day. We could therefore talk of the "half-life" of printed material to be of the order of 500 years.

The other important consequence of the invention of print around 1450 is that the printed book enabled awareness of one-another's work much faster than before. For astronomy, this meant that Kepler and Tycho Brahe and other scholars of their time could exchange theories and ideas much more rapidly, over large distances. I will dare to claim here that this major technological improvement was instrumental in allowing science to progress much more rapidly than it did before.

1.3. Industrial Times

In the 19th century, another major technology was born: photography. Its importance to astronomy is dramatic as it suddenly allowed the replacement of the subjective human eye by a fast, objective method to record the position and brightness of nighttime objects. The first good quality picture of a night sky object (the Moon) is due to John William Draper in New York, 1840.¹

Taken together with printing techniques, photography also allowed the distribution of material which could now be analyzed simultaneously by several people. The density of information of a photographic plate was also tremendous, multiplying the efficiency of observations by a large factor. Towards the end of the 19th century, as observatories around the world were starting to accumulate photographic plates, we witnessed the birth of astronomical data archives.

1.4. Contemporary Period

The following (and so far last) major technological improvement with tremendous influence on astronomy in recent times is the advent of digital data acquisition. Astronomical space missions trying to detect faint high frequency signals had to have a way to easily downlink the results of measurements. Moreover, the nature of the signals they were trying to measure made it appropriate for devices detecting photon incidence rates (photometers). Photon counters were born and with them the era of digital data acquisition. The other big advantage of digital values coming from a detector is of course that they could immediately be processed by the ever improving computer.

What was initially useful in space at very short wavelengths was also interesting on the ground in the optical. Semi-conductor technologies had in the mean time produced very sensitive photo diodes which could be used with suitable fast low-noise amplifiers as photon counters in the optical and near infrared domain. If the photon counter is a device with 0 dimensions, soon 1D detectors followed (e.g., photo diode bars placed behind a prism) and finally 2D devices

¹I recommend the reading of the very nice book "Star Struck" by Brashear and Lewis for very nice reproductions of some of the ancient material described here.

such as CCDs. Since a few years, 3D detectors try to acquire photon location and energy at once (the STJ detectors), but the technology is still in its infancy. Since the early days of photon counting, the amount of numbers coming out of our detectors per unit of time has increased exponentially and (fortunately) almost always following quite closely the computing power available.

2. Data Storage Technologies

In this section, the existing digital data storage technologies are reviewed. The description is structured according to the different digital storage method techniques, starting with sequential and finishing with direct access methods.

2.1. Sequential Methods

Several families of sequential devices exist. They all vary in data carrier format, recording technique, physical size etc. The different classes as described below.

Helical Scan Devices: Helical Scan means that the tape and the write head are presenting each other at a certain angle (inclination) such that while the tape is moving, the data is recorded as diagonal stripes on the medium. The advantages include a cheap recorder and good data density. These tapes have appeared many years ago and thanks to their low price and high capacity, have been very popular. The technology came from non-computer fields: analog or digital video or digital audio. Among the various players in the field, we can cite:

- the DAT (or Digital Audio Tape) is presently in its fifth generation already. It offers today a capacity of 36GB per 170-meter tape. It is still a cheap and reliable system, used mostly for backups and data exchange of smaller amounts.
- the Exabyte tape seems to be slowly disappearing from the data backup landscape -at least in Europe. Due to the lack of demand, its use as a data exchange medium was discontinued several years ago already at the ESO/ST-ECF archive.
- the AIT tape is a product of Sony. Sony has been going through 3 generations AIT. The AIT-3 technology now offers capacities of 260GB (uncompressed) on a 55 EUR cassette. The most specific feature of AIT tapes is a special on-cassette chip which records -among other information- file location, allowing for a speedy file access.

Serpentine Track Devices: This technology proposes to record data linearly on parallel tracks. The tape is mounted on a single reel and runs continuously as it does not have a physical end. Once one track is fully written, the device continues with the next parallel track. The advantages are:

- average file access speed is reduced by a factor equal to the number of parallel tracks
- recording speed is a function of recording density and tape speed.
- un-interrupted streaming is easily achieved.

Tape drives using this technology typically require larger tape cartridges and therefore more robust mechanics, which will make the price of the devices significantly higher than those of the helical scan technology. Presently, three main contenders are competing on the market.

- S-DLT (Super Digital Linear Tape) technology can accommodate about 160GB of uncompressed data written at the rate of 16MB/s. The drive costs around 3.7KEUR and a tape is about 65 EUR.
- LTO (Linear Tape Open) is a technology which was introduced fairly recently, in response to the one-vendor “standard” that the DLT represents. The “O” in LTO stands for “Open” and means that the specifications of the technology are public. The initiative to create such a new device stems from 3 major companies active in the field, including HP and IBM. The nominal capacity is 200GB per tape and the drive can be acquired for 4.2KEUR. It is therefore located in the same price range as the Super DLT.
- The new SAIT system from Sony is in fact a mix of helical scan technology and serpentine track systems: thanks to this combination, the highest capacity on a single volume can be offered: 500GB of uncompressed data can be stored on one 600-m cassette. The price of the recording unit is however very high (about 10 KEUR) and will therefore not make this technology well suited for data distribution.

Parallel Track Devices: The best representative of the parallel track tape system is the good old 9-track tape for which data is written linearly, in one pass, on multiple track running alongside the tape direction. The format has been abandoned for a long time already and this system will not be covered in this study. It is just mentioned for completeness.

2.2. Direct Access Methods

Solid-State memory: Solid-state memory comes nowadays in two flavor, either:

- A lot of memory chips on an adapter that can be installed in the computer and which the operating system can be configured to see as a disk drive. This technique allows for a very fast access but is transient (the content of the memory disappears once the computer is turned off).
- A simple USB adapter with memory that can be plugged in the computer and viewed by the OS as a small disk drive. This “memory stick” is slow but permanent.

Besides being both based on silicon and not having any moving parts, both systems share another feature: they are very expensive.

Magnetic Disk: The old yet better and better magnetic disk has recently made big progress in terms of capacity, data access speed but mostly price per unit volume. So much so that since about 2 years, the cost of a multi terabyte installation based on magnetic disks is smaller than the equivalent capacity provided through say, optical disks and their associated jukeboxes. By the time these proceedings will be made available, the situation will presumably have changed quite a bit still. At the time of this writing and on the average European

market, the price capacity ratio is such that a 300GB ATA internal disk drive has a street price of about 300 EUR. A SCSI disk will have a capacity of 146GB and cost 600 EUR. The equivalent Fiber Channel disk will still cost 730 EUR. It is to be noted that almost any ATA disk can be converted to SCSI or FireWire or USB by simply attaching a 70 EUR adapter to it. So the extra price paid for the genuine SCSI disk is -in principle- for a guarantee of longevity and performance.

Optical Technology: Optical technology, in particular of the “write once – read many” sort was very popular in the 1990’s. This was due to the fact that it was the only contender for direct access, large volume data archives. As mentioned in the section above, optical technology has recently been removed from its throne by magnetic disks, despite the popularity (and hence the low price) of DVDs.

Optical technology is divided into two distinct types:

- *sectorized* media: Magneto-Optical disks as well as DVD-RAM use a medium which is pre-sectorized i.e., markers to help guide the read/write operations are present on the surface. Magnetic disks are also sectorized and this is probably the best way to efficiently address any random portion of the medium to read or write. As far as those technologies in particular are concerned, their aura and popularity has decreased in the face of strong competition from the hard (magnetic) disk drives. Moreover, DVD-RAM never picked up really.
- *non-sectorized* media: the best representatives of this category are the CDs and DVDs. The fact that no sectorization must be physically present on the surface makes the medium very very cheap to produce. However, the medium must typically be written at once, with “sectors” being created on-the-fly, using a specific software.

To be competitive, $5\frac{1}{4}$ inch optical disks should be today in the 50 to 100GB capacity range. And their price should be very competitive so as to convince those in the process of abandoning the technology to stick to it and retain their hardware investment in say, jukeboxes. Since about a year, a company called Plasmon and active in the optical storage field since quite a while is announcing the “UDO” (Ultra Density Optical) device² which should be capable of holding up to 30GB per disk, but at the time of this writing, no product can yet be purchased.

3. Best Data Archive Media Today

3.1. Criteria

In this section, the various criteria to be used for selecting a particular medium type for archival purposes are presented. They consist in a fairly arbitrary list of points, but probably cover most of the requirements one could consider in the selection process. The reader should however keep in mind that such a selection will have to be reviewed every 3 year on average, as available technology and cost evolve very rapidly. As a matter of fact, when facing a continuous increase

²see <http://www.plasmon.com/udo/index.html>

in data archive volume, staying with an old technology is counter-productive and increasingly expensive. The cost of operating an archive is not so much due to the amount of Terabytes to be handled as it is dependent on the number of physical media that have to be kept in proper working order and occasionally migrated. The criteria to consider are based on our experience and include:

- A reasonable cost of the master and backup copy
- At least a 3 year lifetime of the technology
- A high volume efficiency (in GB/m³)
- The total archive volume will fit on the smallest amount of media
- The file access time will be small compared to data processing time

3.2. Comparative Costs

In this part, comparative costs are presented. They will help us select what the best media for a particular archive activity and volume should be. In Table 1 below, only one representative of each major technology has been chosen: for e.g., other types of tapes, the conclusions would have been similar.

Table 1. Cost comparisons between three possible archive storage technologies.

Technology	capacity	Access speed	Volume cost	Manpower	TCO^d
	GB/vol	MB/s	EUR/vol ^a	Hrs/Vol	EUR/GB
LTO-Ultrium2	200	25 ^b	82.2	0.22	0.47
DVD-R	4	3.3	2.8	0.04	1.07
Hard Disk	250	17 ^c	225	0.2	0.51

^aConsiders the price of the medium plus price of the drive divided by 1000. This is obviously only appropriate for tapes and DVDs.

^bA ceiling of 25MB/s is introduced as practical average transfer rate between machines running Gb Ethernet

^cCurrent maximum practical speed of the single disk drives

^dTotal Cost of Ownership

3.3. Conclusions for archiving

To conclude our search for a suitable archive medium today and based on the cost table 2 below, we could summarize the situation with the following weight table:

To summarize our findings, we could conclude that:

- Optical technology is more expensive than magnetic disks
- Optical technology involves too many volumes and therefore requires more operations manpower
- Magnetic disks involves more fragile hardware and therefore requires more system administration manpower
- Tapes are not always reliable, have poor file access time (in particular for small files)

Table 2. Multi-criteria archive media selection

Medium Criteria	Tape	Optical	Mag disk
Low cost/GB	2	-2	0
Random file access time	-2	-1	2
Technology lifetime > 3 years	-1	2	2
Large density GB/m^3	2	-2	0
Small amount of media	2	-2	1
Small amount of media	2	-2	1
Durability of media	-2	2	0
Sum	1	-3	5

So what is an archive manager to do? The all optical solution today implies an investment in hardware and operations manpower; an all magnetic disks solution implies an investment in lots of computers and disks with obviously added benefits of retrieval and processing speed, but system administration can become a significant burden with so many active computer elements. Maybe the best solution is a mixed one: The main copy of the archive can reside on spinning disks and its backup on tapes. The benefits involve fast access to data, the possibility to process data on-line and the use of a somewhat cheaper tape system as a backup copy.

Again, what is to be kept in mind here is that the technical solution must be designed and built to sustain the archive load for the following 3 years, after which point it will remain necessary to review the decisions taken earlier.

4. Best Data Distribution Media Today

4.1. Criteria

For data distribution media, the selection criteria will be quite different. They can be viewed at several different levels: user level and data provider level. The challenge here is to find the medium which will satisfy as many as possible of the sometimes contradicting requirements.

- The medium has a wide format acceptance
- The medium supports physical transport and manipulation
- The medium currently enjoys a low cost per GB
- The medium can support various capacities
- The medium is disposable

As a matter of fact, data provided to users or collaborators in a physical form will have to be easily readable, with no need for expensive or otherwise inconvenient reading equipment (e.g., high-density tape drives). Moreover, to facilitate data access, a popular format for which software drivers exist on most platforms will be preferred (e.g., the ISO9660 CD format). For obvious reasons of costs, a medium with low production cost will also be preferred, but obviously this will not necessarily be sufficient for very large data transfer. The best system in terms of file reception time and cost for the data provider remains electronic

transfer (e.g., FTP), but remains not interesting for data volumes going beyond a few GB.

4.2. Comparative Costs

In this part, comparative costs are presented. In this respect, Table 3 will help us select what the best media for a data distribution and exchange is.

Table 3. Cost comparisons between some possible data distribution technologies.

Technology	Capacity	Access speed	Volume cost	Manpower	Re-use ^d	TCO ^e
	GB/vol	MB/s	EUR/vol ^a		Hrs/Vol	EUR/GB
DAT-DDS5	36	3.0	24	0.23	1	0.99
LTO-Ultrium2	200	25 ^b	82.2	0.22	1	0.47
S-DLT	160	16	77	0.22	1	0.55
SAIT	500	25	205	0.31	1	0.44
DVD-R	4	2.7	2	0.04	1	1.07
USB/FW Disk	250	17 ^c	265	0.2	10	0.51
FTP	11	1	1.4	0	1	1.5

^aConsiders the price of the medium plus price of the drive divided by 1000. This is obviously only appropriate for tapes and DVDs.

^bA ceiling of 25MB/s is introduced as practical average transfer rate between machines running Gb Ethernet

^cCurrent maximum practical speed of the single disk drives

^dNumber of times we plan to use the medium: 1 means no re-use

^eTotal Cost of Ownership

4.3. Conclusions for data distribution

To conclude our search for a suitable data distribution medium today, we could summarize the situation using a weight table (see Table 4).

Table 4. Multi-criteria archive media selection

Medium Criteria	Tape	Optical	Mag disk	Electronic
Low cost/GB	0	-2	0	2
Technology lifetime > 3 years	1	2	1	2
Large density GB/m ³	2	-2	2	2
Small amount of media	2	-2	2	2
Flexible capacity	2	2	2	2
Disposable	1	2	-2	2
Wide format acceptance	-2	2	2	2
Convenient file structure	-2	2	2	2
Supports transport	1	1	0	2
Sum	5	5	9	18

To summarize, we could conclude that electronic data distribution is a clear winner in all category, in particular for what concerns the cost for the distributing site and delivery speed. However:

- Electronic transfer cannot keep the data (need to store it somewhere local)
- The receiving end has to pay for transfer (Internet provider deal)
- And obviously, the network bandwidth limits maximum data package size

Examining the pros and cons of other media, we come to the following conclusions: tapes still provide a reasonable format convenience and reasonable price provided one uses them a lot. They also have fairly high capacities. Their disadvantages include sometimes very expensive drives, which have to be used a lot to compensate for the purchase price and the fact that the medium is fairly sensitive to environment. The compatibility between different drives do not always guarantee readability from one drive to the next. The biggest disadvantage for those drives is the very inconvenient sequential access and long file access time. So tapes can only be used to copy data to another disk. They are therefore good for backup but not so good for data transport.

Optical disk (CDs and DVDs) advantages include the very convenient direct, random access to files and a very cheap medium for both the producer and the the receiver. The current capacity makes it appropriate for medium-size data packages. The Universal data format (ISO9660+extensions) is a guarantee of readability on any computer platform. Detrimental to the optical technology acceptance as a distribution medium is their low capacity (around 4 GB) which means that their lifetime is limited by the progress of the internet bandwidth. Another disadvantage of DVDs and CDs with respect to tapes or magnetic disks is the relatively low read rate (3-6MB/s).

Magnetic disks also have pros and cons: a very convenient direct, random access to files, very fast file download and fairly high capacity. A data format such as ISO9660 can be written on them and be readable by many computers transparently. Finally, the new USB and FireWire interfaces make external magnetic disks almost universally connectable on most modern computers. Among the disadvantages, one will recall the expensive units that only makes sense if returned after use - which implies more handling/shipping costs. A magnetic disk remains a fragile device: it needs to be wrapped carefully for shipment. All this means that it is only meaningful at the largest capacities (250 GB and more).

5. Conclusions

5.1. Lifetime

This review has tried to give an idea of where data storage and distribution is coming from, what it can do today and what are its limits. One of the aspects that was not yet mentioned is where it is going. In this respect, the recommendation of not making plans and decision as far as archive storage for more than 3 years is an strong indication as to the practical lifetime of a given modern technology. This does not mean that media written 3 years ago will suddenly become unusable, but that it becomes increasingly expensive to maintain and operate an archive with aging technology.

As far as data distribution is concerned, digital media can practically be read up until about 15 years, but the old claims of having optical disks readable for 100 years is not to be taken seriously as, after a tenth of that period, no reading equipment will be able to decipher their content. One could point to the CD and DVD as a better future investment in this respect: true enough, the devices built to read CD since 15 years can read the old and new media. Conversely a DVD reader purchased today will still read your old CD from 20 years ago –a lot faster even, but DVD-Rs for instance are very sensitive to dust and fingerprints and can be scratched very easily. Recovering lost data on the surfaces of those media can be very difficult and multiple copies of a particular set of data is the only guarantee of data survival. Much more difficult still is the recovery of old magnetic disks...

So digital media will probably have a lifetime of 10 years, given the proper reading equipment remains available. This is a far cry from older technologies which could boast a 1000 year survival and rely on human eyes to decipher them.

5.2. The future of data distribution

Given all those considerations, is it still worth distributing data on tangible, physical media? Shouldn't archives and data centers take care of delivering content in reduced, visual form, providing users with only a *view* on their data?

The upcoming large data production instruments such as the ALMA sub-millimeter observatory and the many large visible and infrared mosaic cameras on survey telescopes should all be good reasons for astronomers not to want raw data delivery to their home bases: they should rather rely on reduced, calibrated results as provided by the GRID and the VO tools. The data volume will be such that no single observatory will probably afford to support the data reduction infrastructure requirements of large survey programs.

Shouldn't we also remember that most of the old books that came to survive until our age were preserved because they represented mostly final, important results? Publications are concerned with finished papers and conclusions rather than early drafts and raw data. Of course to get there, we still need the raw data but we also need the engines to process it and the infrastructure to deliver the results. There is maybe no need to replicate everything everywhere.

References

- Brashear, R., Lewis, D., 2001, Star Struck
Pirrenne, B., Albrecht, M., Schilling, J., 1999, in ASP Conf. Ser., Vol. 172, ADASS VIII, ed. D. M. Mehringer, R. L. Plante, & D. A. Roberts (San Francisco: ASP), "The Prospects of DVD-R for Storing Astronomical Archive Data"
Transtec AG, Herbst/Winter 2003 Produktkatalog