

Building Archival Storage Component Using New Open Standard Infrastructure

Dariusz Paradowski
National Library of Poland
al. Niepodleglosci 213
02-086 Warsaw Poland
+48 22 608 26 17
d.paradowski@bn.org.pl

ABSTRACT

The National Library of Poland (NLP) makes use of emerging open standard of digital magnetic tape structure - LTFS to build simple, efficient, economic, scalable and safe archival storage component for institutional repository.

Keywords

Digital archival storage; open standard; long time digital preservation; digital library

1. INTRODUCTION

For over ten years the National Library operates comprehensive system for supporting digitization of library materials, processing created digital files, presentation and storage. The system called Repository has a number of functional modules responsible for multitrack workflow, import from various sources (scanning, digital legal deposit), conversion of data formats and metadata, generating derivatives (jpg, OCR, etc.) presentation and access management, storage and finally archiving. The current generation of the repository immediately after the introduction of the source files to the system places them in an object archive. It is a commercial appliance in the form of licensed software running on a multi-node cluster of dedicated servers connected to the disc array.

In the beginning it was a very useful solution:

- it provided a guarantee that extremely important archiving module will be reliable - it was a commercial solution verified in considerable applications,
- cluster equipped with load balancing efficiently acquired and provided files,
- it supplied manageable WORM functions,
- by compression it optimally used disc array capacity, ensuring that the current amount satisfied the needs for a longer period of time.

2. ISSUE

After a few years the rate of digitization in BN, was greatly raised which, accompanied by quality improvement, resulted in a rapid increase of the bytestream of files entering the repository.

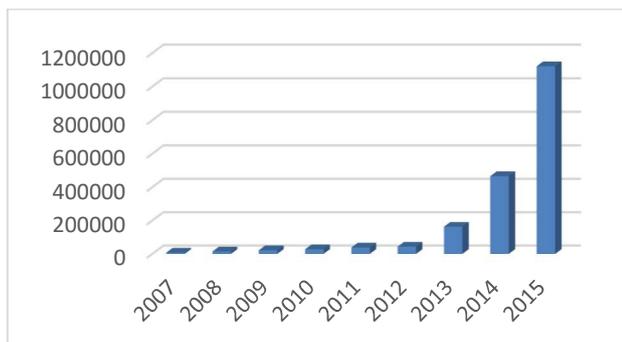


Figure 1. Items available in polona.pl

As a result, there appeared drawbacks of the used archive:

- insufficient performance - expected increase would require changes to the architecture of application transmitting the data to the archive and also hardware upgrade of the appliance (dedicated servers and arrays),
- rapid depletion of capacity - the further operation would require costly expansion of the storage array and also expensive purchase of licenses for the appliance archive
- increasing energy consumption of the solution.

3. CHALLENGE

To solve these problems reasonably, different archive module was needed. Essential requirements of equal importance were defined:

- low cost storage
- high, easily scalable performance
- high, easily scalable capacity
- safety

These were then turned into more practical ones:

- low cost storage
 - low cost of capacity per byte
 - no expensive license capacity
 - low energy consumption
- high, easily scalable performance
 - horizontal scaling possible without rebuilding the infrastructure (just extension) and software
 - vertical scaling possible without software change
- high, easily scalable capacity
 - horizontal scaling possible without rebuilding the infrastructure (just extension) and software
 - vertical scaling possible without software change
- safety
 - open standards of storage writing
 - the system can not depend on single manufacturer
 - data must be readable outside environment of the archive
 - metadata must be human-readable
 - recognized standards
 - damage to any part of the data can not prevent the reading of data undamaged

4. FACTORS CONSIDERED

4.1 Choice of Carrier

Aside from stone and paper the magnetic tape is best recognized carrier that has very long shelf life. Tape has exceptionally low cost per unit of capacity. It is essential to use an open and yet recognized standard. Linear Tape Open (LTO) is an open standard supported by many major manufacturers, it also has defined roadmap for development.

The new LTO7 standard appeared on the market at the end of 2015 and has a capacity of 6TB. This is enough to avoid necessity to purchase another expansion frame for the automatic tape library currently used in NLP in the foreseeable future. It has long time to the end of support and moreover will be readable by next two generations of drives [1].

Table 1. RAID10 vs independent carriers vulnerability

No of carriers destroyed	hard disc drives, RAID 10, (4 mirrored = 8 drives) best case	Data loss %	
		2 sets of 4 independent tapes worst case	2 sets of 4 independent tapes best case
1	0	0	0
2	0	25	0
3	0	25	0
4	0	50	0
5	100	50	25
6	100	75	50
7	100	75	75
8	100	100	100

Comparison of tape to disc storage. Table 1 shows comparison of the sensitivity of the system disk in a very robust and expensive version of RAID 10 compared with a set of independent carriers comprising two copies of the data. With minor injuries RAID gives greater protection than the worst case for independent media and the same as the best case. It is worth noting that the worst case is relatively unlikely (damage to same data on different tapes) and the best case gives better results than RAID. In particular, the destruction of more than 50% of the media RAID causes a loss of 100% of the data, while the media independent only 25-50%. It is also worth noting that the price of a unit capacity of good quality media is much lower for tape cartridges. It is also important from the point of long time preservation that in case of danger cartridges may be easily removed from tape library and transported which is much more complicated for hard drives.

Sequential recording on tape cartridges, which corresponds to scenarios of archival usage is performed with a very high speed. The performance of the system can be easily multiplied by increasing the number of drives and the capacity by increasing number of tapes.

4.2 Choice of Filesystem

Linear Tape File System (LTFS) meets the requirement of the system to be open, it is supported by several leading manufacturers, developed in the mature form and present on the market for several years. This year – 2016 LTFS became adopted as standard ISO / IEC [2]. Record in LTFS can be read on another device from another manufacturer, without the need to reconstruct an environment where it was saved. Moreover, basic software solutions - allowing the use LTFS on a single drive are available as open source by many hardware manufactures.

4.3 Developed Methods

The system is to serve as a disaster recovery solution, so the assumption that there will survive a random subset of cartridges implicates the requirement that each object will be stored on one medium. A complete object is understood as: a unique object identifier, all the metadata and structure of the object in the form of (human readable) XML METS and all source content files of the object.

A carrier that most of the time is kept offline allows to postulate an idea of avoiding backwards error propagation. Once saved, the object in the archive is never to be changed. Any change will be treated as formation of a new version of the whole object. Information about the location of the next version (barcode of the cartridge) will be placed in the database system. It freed us from designing a complex and unreliable predictions of reserved free space on the tape needed to create new versions of files (that way would be also very inefficient considering linear nature of the tape recording).

The challenge was the metadata, which, in the national library reality are subject to frequent revisions. This problem was solved by independent archiving the updated metadata of all objects through saving the entire database of the system in an XML file (with checksums). Likewise objects each database copy is to be kept without adjustments and with versioning instead (version and time markers apply).

Thus, if after the disaster, a random set of tape cassettes has been discovered, it is enough to find the latest version of the database which allows to quickly find the latest versions of objects.

To economically save objects of random, often large size on the discrete space on the tapes, it is necessary to optimize it. The archive will temporarily gather on it's own disc buffer, objects supplied by the repository that are ready for archiving. Choosing from this pool archive will construct packages of these objects of a size as close as possible to the capacity of a single tape cartridge to write it at once.

Contradictory parameters such as the maximum allowable time of the object in the buffer and the minimal tape capacity loss will be fine-tuned on the basis of statistics. After achieving at optimal thresholds, the tape will be written only once and never changed. This approach ensures maximization of write speed and durability of the tape. This also allows to overcome the incompatibility between WORM and LTFS by switching write protection tab on the tape cartridge further improving safety of the archive.

5. REFERENCES

- [1] Buffington, J. 2016. *Analyzing the Economic Value of LTO Tape for Long-term Data Retention*. http://www.lto.org/wp-content/uploads/2014/06/ESG-WP-LTO-EVV-Feb_2016.pdf
- [2] *ISO/IEC 20919:2016 Information technology — Linear Tape File System (LTFS) Format Specification* (First edition) <https://www.iso.org/obp/ui/#iso:std:iso-iec:20919:ed-1:v1:en>