IFLA WLIC 2018
KUALA LUMPUR

Submitted on: 24.06.2018

Cost and security improvement of the Long Time Preservation by the use of open source software and new ISO standard in the National Library repository

Dariusz Paradowski
Information Technology Department, National Library of Poland, Warsaw, Poland.
E-mail address: d.paradowski@bn.org.pl

Abstract:
Long Time Preservation (LTP) Policy besides high-level organizational elements should comprise technological foundation based on available technologies. Omitting this element can undermine rationality of the former.

The LTP System should:

- *be independent from a single manufacturer and have known data path,*
- *be permanently ready for total disaster (The Doomsday Clock is now as close to midnight as it was in the fifties),*
- *be scalable in capacity and performance,*
- *keep costs within available budget.*

The carrier and format themselves should be sufficiently durable to limit the frequency of migration.
The analysis of these criteria led to the selection of well-established technologies based on open or well-documented standards, independent from a single manufacturer:

- *Magnetic tape - decades of experience and predictable shelf life,*
- *Linear Tape Open (LTO) - many manufacturers produce tapes compatible with drives from various manufacturers,*
- *Linear Tape File System (LTFS) - allows exchanging data between compatible systems from different producers using open-source software. In the year 2016 LTFS became adopted as ISO/IEC standard,*
- *Metadata Encoding and Transmission Standard (METS) – XML container of digital objects.*

Above technologies and some our inventions were used to develop the software of new LTP System.

- *The System has since worked productively for 10 months and archived almost petabyte of data.*
- *Cost per byte (3 copies) is more than order of magnitude lower than in the former solution.*
- *The System uses standard tape library, is scalable and redundant.*

- *Digital objects saved (each integrally) on every single tape are readable and may be migrated without any knowledge about the LTP System.*
- *The entire data resource is easily transportable*

Keywords: Long Time Preservation, Digitization, Digital libraries, Magnetic Tape.

1. Introduction

Long Time Preservation (LTP) Policy should, besides high-level organizational elements comprise technological foundation based on available equipment. Omitting this element can undermine rationality of the former.

These dependencies may be illustrated by a few facts from history. At the very beginning of digitalization, the resulting files were recorded straight on DVD discs as it was the only directly available carrier of a sufficient capacity. Soon it turned very inefficient and unsafe. Creation of the first Repository system with LTP function in the Library in 2004 required quick overcome of both the financial and organizational barriers. It was achieved by building network server and a software that was converting incoming large source files into small presentation copies (jpg) to be kept on a modest disc array. Original source files were transferred to HDD disks (numbered by the system), which were put in a hot swap bay for the time of access and then stored off-line on a shelf. This solution needed human action whenever source file was needed so access time was very long. It also exposed digital assets to considerable risk due to the undefined shelf life of HDD and hazard of mechanical shock, however, in those times appropriate disc array was too expensive. Analyzed backup solutions also could not serve as archiving tools to keep large files outside expensive media.

In the next generation of the System, the commercial appliance unifying the functions of the current management of resources with their long-term archiving was used. It assigned responsibility to the producer – a stable international corporation. It was supplied with sizeable disc array but, above all, used a very efficient compression and deduplication saving disc space.

Some years later a rapid improvement in the quality and the rate of digitization resulted in the enormous growth of data. It exposed performance issues of the appliance. Analysis of the whole arrangement revealed that the dependence of the LTP system on commercial black-box solution subjected to whims of the producer should be considered as a risk. Moreover, due to necessary additional licenses, a significant increase in the cost of further appliance exploitation was expected. Construction of the new archiving subsystem became inevitable.

When conducting digitization, cultural institutions become similar to large corporations in terms of the size of digital content, but unfortunately not in terms of financial resources. Therefore, it is necessary to develop an own approach to managing the assets.

2. Determining principles

The main goals have been set: improvement of security and reduction of costs. The of new LTP system planning took into account experiences and the awareness of particular requirements for the long time of digital preservation. It was specified that the aim will be possible to achieve with the following assumptions:

1) safety
    a. open standards
        i. data must be readable outside environment the archive environment
        ii. the ability to read must not depend on a single manufacturer
    b. recognized standards
        i. metadata must be human-readable
        ii. digital content must be stored in well documented format, readable by many popular readers
    c. storehouse must be easy to transport in case of danger
    d. durable carrier
    e. migration feasible
    f. damage to the part of the data must not prevent access to undamaged data

2) economy
    a. low cost of capacity per byte
    b. no expensive capacity license
    c. low energy consumption
    d. high performance and capacity
    e. scalability of performance and capacity

3. Selection of technologies

Traditional methods of preservation were quickly depreciated in the digital area by short lifespan of carriers, formats and technologies hence leading to the recommendation of frequent migration. The opposite - too frequent migration is cost-generating and increases the risk of data loss. Therefore, the carrier and format themselves should be sufficiently durable [10] to limit the frequency of migration. The analysis of these criteria led to the selection of well-established technologies based on open or well-documented standards, independent from a single manufacturer.

1) Magnetic tape - decades of experience and predictable shelf life, almost immediately ready and easy to transport, easy scalable - capacity by multiplying tapes and speed – by multiplying drives, no electricity cost for storage (except for verification).
2) Linear Tape-Open (LTO) - many manufacturers produce drives and many manufacturers produce tapes - all mutually compatible, high throughput, high capacity. Manufactured from the year 2000, always two generations backward compatible, low cost per byte. [1][2][3]
3) Linear Tape File System (LTFS) - allows exchanging data between compatible systems from different producers. Big hardware manufacturers provide open-source software drivers. In the year 2016 LTFS became adopted as ISO/IEC standard.[4][5]
4) Metadata Encoding and Transmission Standard (METS) – maintained by Library of Congress well-defined XML container for metadata (and digital objects), all documentation is available under the Creative Commons CC0 1.0 Universal Public Domain Dedication.[6]

4. Forming principles

Definition of OAIS Archival Information Package (AIP). [7] [8]

1) Identification of the system (version) that prepared the package.
2) Timestamp.
3) The unique identifier of the digital library object (digital item).
4) All the metadata and structure of the library object in the form of (human readable) XML METS including behavioral metadata, structural metadata, linked collections, relations to other objects and copyright status.
5) Selected digital content files of the object: all source files and those derivatives which are costly to produce.
6) Checksums.

AIP is prepared as TAR file [9] containing a single folder with METS file and subfolders each with different content types.

In order to include physics in policies, the following rules were formulated.

1) Entire AIP package is recorded on a single tape cartridge. This makes each tape independent.
2) API packages fitting single tape are prepared on disc buffer and recorded at one time. The whole tape is filled while a single load to a drive and single pass/rewind. This enhances a lot efficiency and durability of a tape.
3) Once the tape has been written it is never changed. If a change is applied to the object, then the new version of AIP is archived. Only the addition of new content to the object is considered as a change. Any other modification is reflected in the database which is permanently archived separately (or updated in metadata included in AIP when a change occurs). Thanks to this rule if an error appears in the Repository it will not damage archived objects, tape durability is improved and capacity used efficiently.

5 Software

The main idea may be utilized in a small organization at low cost just with a single drive and manual tape operations. However, in the case described here, a tape library was used and a software solution created named Archive subsystem to fully automate operations. Some commercial software systems were prior analyzed but haven't filled above mentioned rules and included costly licenses per volume of data. All the Subystem software was built without any use of commercial software, a lot of open-source and free software was used. All tape library commands are executed using "mtx" - a native Linux media changer tool.

The Archive subsystem consists of the Archiver and a configurable number of Archives. Communication is conducted using REST API and synchronized queues serviced by open-source software Artemis [11] allowing Archives to work asynchronously. Each Archives receives AIPs from the Archiver and may internally replicate received AIP to a configured number by writing on many drives in parallel. All actions on data are transactional, only after all copies have been recorded Archiver sends confirmation of archiving to the Repository.

To minimize costs source files which consume a lot of space and are hardly ever used they are deleted from the Repository precious fast storage after being archived. In case the Repository requires archival data it sends a request to Archiver and the AIP is received from least loaded Archive.

Archiving subsystem run verification of the stored data in two opportunities:

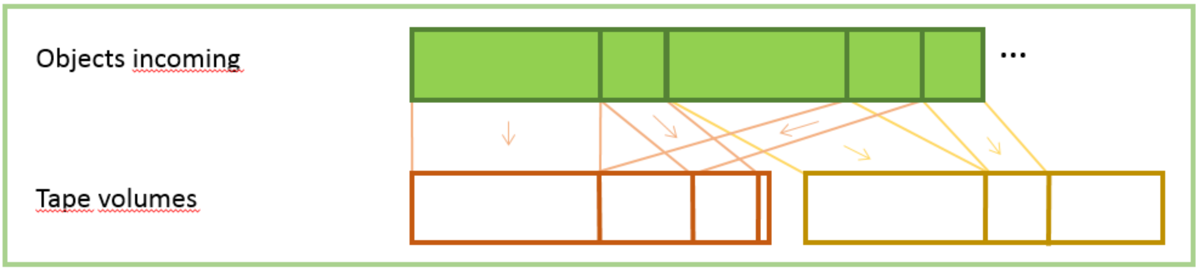
1) permanent sequential check of all the tapes
2) during a reading of any package

and in two ways: by interpretation of data provided by the drive and by verification of AIP checksum. In case of a checksum error, Archives sends a request to Archiver to get another replica of AIPs and automatically creates new tape.

Each part of the Archive subsystem utilizes buffer disc space. This includes dynamic shifts of giant amounts of data by concurrent services. To avoid deadlock, each microservice reserves space in the small fast register before write operation and releases it after the file is removed.

High tape space utilization is achieved by implementation of matching algorithm: AIPs to be recorded are arranged in descending order and then selected to best fit tape, see Figure 1.

Figure 1



However, this algorithm may cause some objects to wait in the buffer for a very long time. To avoid this selection algorithm in the first pass selects form objects that stay too long in the buffer and in the second pass selects from the others.

Taking into account experience of backup software that destroyed annual stock of cleaning tapes during one night trying to clean damaged drive, blocking procedure is implemented that stops drive operation in case of repeating errors.

The result is versatile solution currently utilized in National Library for over a year that already has stored near petabyte of unique data. Current work aims at separation of the Archive Subsystem from the Repository so that it may be used independently. Each unique object is archived in two geographically distant locations. Overall cost of data storage is more than order of magnitude smaller than in the former commercial solution. Every tape may be read in any LTO7 drive with the use of open-source driver and retrieved objects are readable without any information of an Archiving Subsystem.

References

[1] http://www.fujifilmusa.com/shared/bin/LTO_Data_Tape_Seminar_2012.pdf

[2] 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

[3] https://www.lto.org/resources/tcotool/

[4] Pease D. 2016, *Linear Tape File System now an International Standard,* https://www.ibm.com/blogs/research/2016/04/ltfs-now-iso/

[5] *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

[6] http://www.loc.gov/standards/mets/mets-schemadocs.html

[7] http://www.oclc.org/research/publications/library/2000/lavoie-oais.html

[8] https://public.ccsds.org/pubs/650x0m2.pdf

[9] https://www.gnu.org/software/tar/manual/html_node/Standard.html

[10] Lambert J.M. Perdereau J. 2008, *QUALITE DES DVD+R ET DES DVD-R DISPONIBLES SUR LE MARCHE POUR L'ARCHIVAGE DES DONNEES NUMERIQUES NUMERIQUES*, https://francearchives.fr/file/de7f8ea96ceb4ce38eb6d9278b3df49260752449/static_2140.pdf

[11] https://activemq.apache.org/artemis/