Data Storage and Data Migration

Reto Kromer • AV Preservation by reto.ch

Open-Source Tools and Resources for Audio-Visual Archives
Elías Querejeta Zine Eskola
Donostia (San Sebastián), Spain
1–4 and 9 June 2021

Summary

- data storage
- data migration

Data Migrations

2014

- our internal archive from LTO-4 to LTO-6 (5.7 PB)

2014–2021

- many migrations for clients

2021

- our internal archive from LTO-6 to LTO-8

Magnetic Tape

- in use since the 1950s by IT
- cartridges are always on polyester base (old open reels can be on triacetate base)

Packaging

- open reel
- cassette
- cartridge

Recording

- linear or diagonally
- analogue or digital

LTO

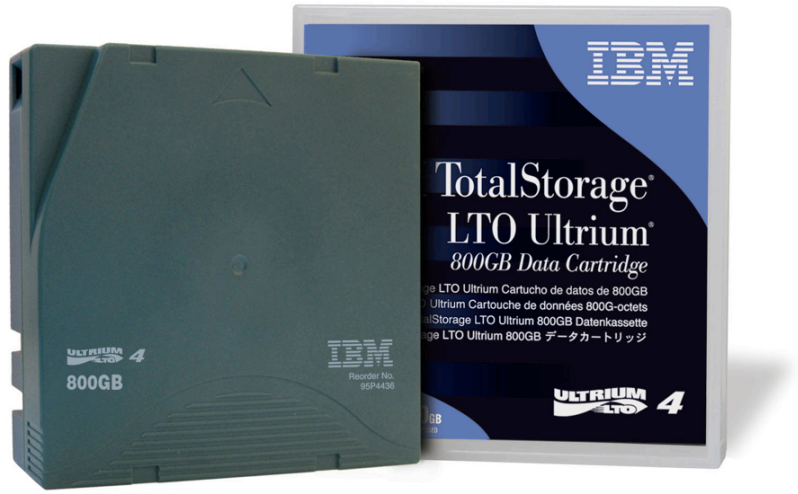
- Linear Tape-Open
- answer from the IT industry to the bank and insurance sector
- in 2000 LTO-1
- currently LTO-8
- currently Hewlett Packard Enterprise, IBM, and Quantum form the LTO Consortium

LTO-8

- only one-generation backward reading capabilities
- format M8 = LTO-7 cartridges formatted as LTO-8
- M8 can be used on LTO-8 drives only

LTO-9

- LTO-9 roll-out pushed to September 2021
- LTO-9 drives manufactured by IBM only
- LTO-9 cartridges manufactured by Fujifilm and Sony only
- backward reading capabilities for regular LTO-8 (L8), but not M8



Formatting

TAR

- from LTO-1 to LTO-4 only possibility
- still possible possible today

LTFS

- possible (and recommended) since LTO-5

TAR

- standard TAR
  - ➔ bloc size
  - ➔ number of archives per cartridge
  - ➔ archives needing more than one cartridge
- TAR with a proprietary data encoding (e.g. BRU, Retrospect)

LTFS

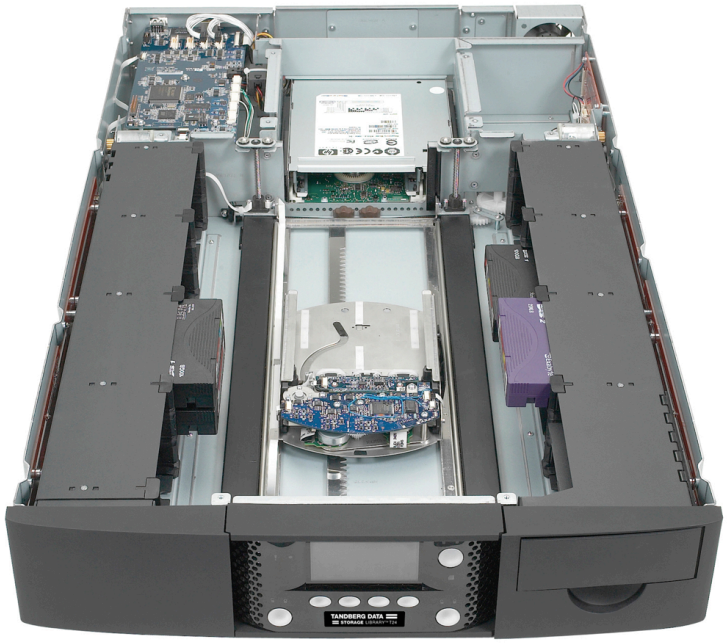
- different versions
- almost one implementation per vendor, but…

  … "ltfs" and "mkltfs" common commands
- lossless compression (default) or uncompressed data
- unencrypted (default) or encrypted data

	advantages	disavantages
TAR		
LTFS		

Drive

- internal or external unit
- library



17

	advantages	disavantages
unit		
library		

18

Storage of the Tapes

- in a tape library
- on a shelf

19

	advantages	disavantages
library		
shelf		

20

Software

- proprietary or open source
- graphical user interface (GUI) and/or command-line interface (CLI)

Plan the Next Migration

- file naming
- barcodes
- checksums
- write the full index to the cartridge
- technical metadata
- code to retrieve the files

#1: Film

FILM

- FILM_DPX/Film_nnnnnn.dpx
- Film_PCM.wav
- Film_ProRes.mov
- Film_H264.mp4

#2: Video

VIDEO

- Video_YUV422.mkv
- Video_ProRes.mov
- Video_H264.mp4

Reading

Reto Kromer: **On the Bright Side of Data Migrations**, in «IASA Journal», n. 49 (December 2018), IASA, p. 18–22

➔ retokromer.ch/publications/IASA_49.html

read | script | write

script to modify

- container
- codec
- both container and codec
- metadata
- filename

#1: ProRes-born Content

from:

- ProRes stored in a QuickTime (.mov) container

to:

- ProRes stored in a Matroska (.mkv) container

Update the Container

➔ read file from source LTO

➔ demultiplex file

- ProRes 422, 10 bit [yuv422p10le]
- ProRes 4444, 10 bit [yuv444p10le or yuva444p10le] or 12 bit [yuv444p12le]

➔ multiplex file

➔ write file to destination LTO

SMPTE RDD 36:2015

SMPTE REGISTERED DISCLOSURE DOCUMENT

Apple ProRes Bitstream Syntax and Decoding Process

Page 1 of 39 pages

The attached document is a Registered Disclosure Document prepared by the sponsor identified below. It has been examined by the appropriate SMPTE Technology Committee and is believed to contain adequate information to satisfy the objectives defined in the Scope, and to be technically consistent.

This document is NOT a Standard, Recommended Practice or Engineering Guideline, and does NOT imply a finding or representation of the Society.

Every attempt has been made to ensure that the information contained in this document is accurate. Errors in this document should be reported to the proponent identified below, with a copy to eng@smpte.org.

#2: Video

from:

- AVI / 8-bit and 10-bit uncompressed
- MOV / 8-bit and 10-bit uncompressed
- MP4 / 8-bit and 10-bit uncompressed

to:

- Matroska / FFV1

Container and Codec

→ read file from source LTO
→ demultiplex file
→ decode file
  - $Y'C_BC_R$, 4:2:2, 8 bit, «raw» [uyvy422]
→ encode file
→ multiplex file
→ write file to destination LTO

Container and Codec

→ read file from source LTO
→ demultiplex file
→ decode file
  - $Y'C_BC_R$, 4:2:2, 10 bit, «raw» [yuv422p10le]
→ encode file
→ multiplex file
→ write file to destination LTO

#3: Filename

from:

- Title_YUV422.mkv

to:

- Title_YCbCr422_9d5084b5b0a08d5022b3 9e0e75241d12.mkv

#3: Filename

from:

- Title_YUV422.mkv

to:

- Title_YCbCr422_9d5084b5b0a08d5022b3 9e0e75241d12.mkv

Plan the Next Migration

- file naming
- code bars
- checksums
- write the full index to the cartridge
- technical metadata
- code to retrieve the files

AV Preservation by reto.ch

chemin du Suchet 5
1024 Écublens
Switzerland

Web: reto.ch
Twitter: @retoch
Email: info@reto.ch