Stockage et migration des données

Reto Kromer • AV Preservation by reto.ch

Atelier Memoriav
Logiciels ouverts dans l'archive
Berne, 11 janvier 2024

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 (25.2 PB)

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-9
- currently the LTO Consortium consists in: Hewlett Packard Enterprise, IBM and Quantum

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 drives manufactured by IBM only
- LTO-9 cartridges manufactured by Fujifilm and Sony Group only
- only one-generation backward reading capabilities
- only 50% capacity increase
- backward reading capabilities for regular LTO-8 (L8), but not M8

LTO-10

- Will there be two-generation backward reading capabilities?
- Is LTFS strong enough?
- Release possibly end of 2024, probably beginning of 2025.

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)

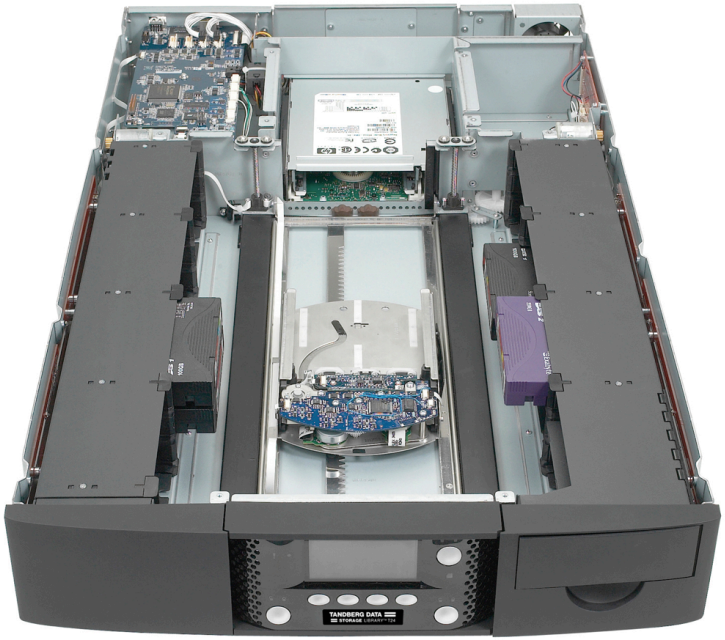
LTFS

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

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

Drive

- internal or external unit
- library



Storage of the Tapes

- in a tape library
- on a shelf
- in a fire-proved cabinet

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

Checksums

cryptographic

- MD5
- SHA-1
- SHA-256
- SHA-512

non-cryptographic

- CRC-32
- xxHash 32
- xxHash 64
- xxHash 128

Longterm

- storage of the cartridges
- three copies…

  … in geographically distant locations
- data integrity check
- data migration
- availability of LTO desks

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

#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_9d5084b5b0a08d5022b39e0e75241d12.mkv

Equipment

- servers
- network
- storage
- ventilation (cooling)

Working Place

- software
- computer with at least one good monitor
- storage
- illumination
- chair

Common Equipment

- LTO desk or library
- software
- network

Consumables

- LTO cartridges

AV Preservation by
reto.ch

Sandrainstrasse 3
3007 Berne
Suisse

reto.ch
info@reto.ch