De-Mystifying LTO, or: LTO for the Masses

Reto Kromer • AV Preservation by reto.ch

AMIA Conference

Pittsburgh, PA, United States of America 7–9 December 2022

1

Write and Read

1st part by Linda Tadic

2

Formatting

TΔR

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

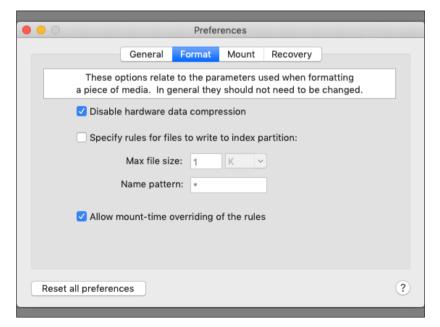
LTFS

• possible (and recommended) since LTO-5

TAR

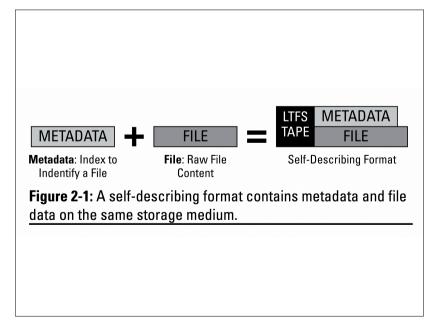
- standard TAR
 - → bloc size
 - > number of archives per cartridge
 - → archives needing more than one cartridge
- TAR with a proprietary data encoding (e.g. BRU, Retrospect)

5



LTFS

- different versions (current 2.5.1)
- almost one implementation per vendor, but ...
- ... "Itfs" and "mkltfs" common commands
- lossless compression (default) or uncompressed data
- unencrypted (default) or encrypted data





Linear Tape File System (LTFS) Format Specification

Version 2.5.1

ABSTRACT: This document defines a Linear Tape File System (LTFS) Format separate from any implementation on data storage media. Using this format, data is stored in LTFS Volumes. An LTFS Volume holds data files and corresponding metadata to completely describe the directory and file structures stored on the volume.

This document has been released and approved by the SNIA. The SNIA believes that the ideas, methodologies and technologies described in this document accurately represent the SNIA goals and are appropriate for widespread distribution. Suggestions for revisions should be directed to https://www.snia.org/deedback/.

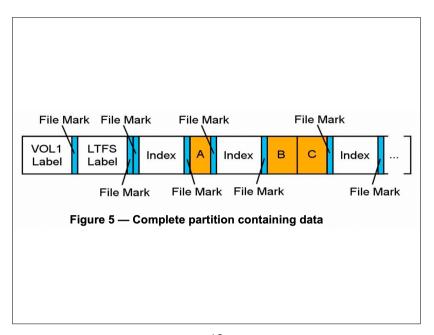
SNIA Standard

August 18, 2020

9

Storage of Tapes

- shelf
- tape library
- fireproof cabinet



10

Operating System

- Unix/Linux
- macOS
- Windows

Software

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

13

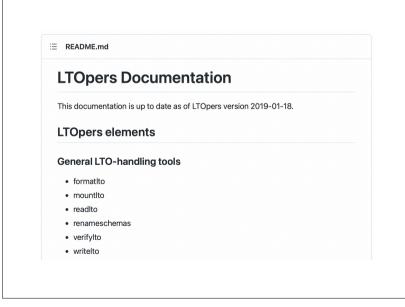
AMIA Open Source

LTOpers (CLI)

→ github.com/amiaopensource/Itopers

Workflow

- prepare the data for archiving
- format an LTO cartridge with LTFS
- mount the formatted LTO cartridge
- write the data onto the LTO cartridge



Formatting LTO

LTO tapes need to be formatted before they can be used. Run the following command:

formatlto

The script will ask you select a deck (if you have more than one attached) and to provide a tape serial your tape. The tape serial format is exactly 6 capital letters and/or numbers.

Example:

Annnnn

where nnnnn is a 5-digit sequential number, e.g. A00001, A00002, etc.

The script will format your tape in LTFS format, with no compression, and sets rules that allow smaller types of files to live in the index partition of the tape, which gives you faster access. If you need to change this, edit the mkltfs line in the script to suit your preferences.

Remember to physically label the tape with the tape serial!

17

Writing to LTO

To write data onto a tape, run the following command:

writelto

When prompted, enter the tape serial and the source directory.

The script will run \mbox{gcp} to transfer the data to the LTO tape (\mbox{rsync} is used as a backup check for \mbox{gcp}).

To read back and create checksums for the contents of a tape, and write checksums to a file named with the tape serial and date, run the following command:

writelto -v

The checksum file will be written to the LTO logs directory: LTO_INDEX_DIR (also set in mmconfig) or \$HOME/Documents/lto_indexes (default if not set in mmconfig).

This command uses the tool md5deep, which has several flags to set options. To customize which flags you want to append to md5deep, set the variable LTO_MD5_FLAGS in mmconfig (type mmconfig to access the configuration GUI).

Mounting LTO

Put the tape you want to mount in the deck (you do not have to push it in all the way).

To mount the tape run the following command:

mountlto

The script will check for attached tape drive, and will prompt for deck name if there are multiple drives.

The tape will load in deck and mount.

In Finder, go to Go > Go to Folder, and enter /Volumes/\$tape_serial to navigate to tape in Finder.



When?

 tape drives for a generation you use are not longer manufactured

21

Barcodes

BARCODE_REGEX="^[A-Z0-9]{6}(L[5-9]|M8)\$"

```
#!/usr/bin/env bash
RED='\033[1;31m'
BUUE='\033[1;34m'
NC='\033[0m'

BARCODE_REGEX="^[A-Z0-9]{6}{L[5-9]|M8}$"

if [[ ! "${1}" ]]; then
    echo -e "Usage:\n $(basename "$0") <LTO_barcode_to_check>"
    exit 0

elif [[ ! $(echo "${1}" | grep -E "${BARCODE_REGEX}") ]]; then
    echo -e "${RED}The barcode '${1}' is not valid.${NC}"

else
    echo -e "${BLUE}The barcode '${1}' is valid.${NC}"
```

Plan the Next Migration

- file naming
- barcodes
- checksums
- write cartridge memory chip data also on the tape
- technical metadata
- code to retrieve the files

22

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 decks

25

Example: Video

from:

- AVI / 8-bit and 10-bit uncompressed
- QuickTime (.mov) / 8-bit and 10-bit uncompressed
- MP4 / 8-bit and 10-bit uncompressed

to:

Matroska (.mkv) / FFV1

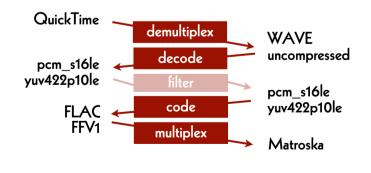
read | script | write

script to modify

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

26

Modify Container and Codec



Container and Codec

- → read file from source LTO
- → demultiplex file
- → decode file
 - Y'CBCR, 4:2:2, 10 bit, "raw" [yuv422p10le]
- → encode file
- → multiplex file

January 2023

→ write file to destination LTO

29

AV Preservation by reto.ch

zone industrielle Le Trési 3 1028 Préverenges Switzerland

> Web: reto.ch Email: info@reto.ch



31

3rd part by Larry Blake

30

AV Preservation by reto.ch

Sandrainstrasse 3 3007 Bern Switzerland

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



from 1st March 2023