9-Track Tape Data Extraction

Prof. Dr. Bruce Nikkel, Bern University of Applied Sciences

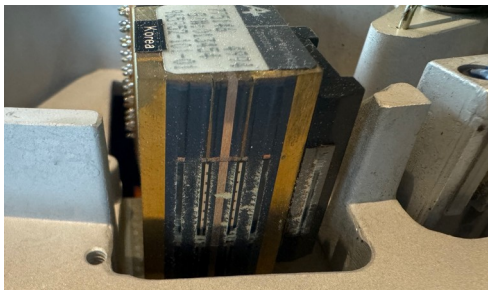
In 1964 the iconic IBM System/360 mainframe computer was announced. This included half-inch (12.7mm) 9-track magnetic reel-to-reel tape storage technology which would eventually become an industry standard. Common tape reel sizes were 7, 8.5, and 10.5 inches in diameter, each with different lengths of tape (the longest being 1.1km). Tapes were physically protected with a plastic container, or an edge cover that needed to be removed (either manually or automatically) before use. On the bottom/back of the tape was a unique identifier and removable write-enable rings (read-only if removed). Tapes also had a round dimple on the front/top for your finger to manually wind the tape mounted in the drive.

Different tape sizes

Protective cover

Write enable rings

The 9-track head could read and write one 8-bit byte and a parity bit. Recording density was measured in characters or bytes per inch (cpi or bpi) and the encoding scheme improved over time: NRZI (Non-Return-to-Zero Inverted) for 800bpi, PE (Phase Encoding) for 1600bpi, and GCR (Group Coded Recording) for 6250bpi. Tapes had reflective markers indicating the BOT (Beginning Of Tape) and EOT (End Of Tape) which were detected with light sensors inside the drive.



9-track tape head

Reflective marker

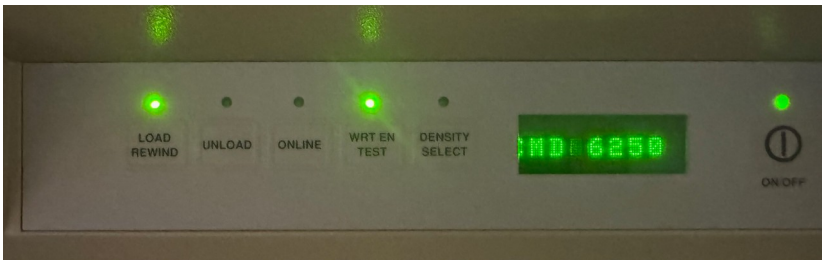
Light sensors

Data on the tape was saved in blocks or records separated by inter-block or inter-record gaps (typically 8-15mm). The encoding, length, and block size determined tape's total storage capacity, commonly ranging from 40MB to 160MB. A set of consecutive tape blocks form a tape file. There can be multiple tape files, separated by file markers, stored in sequential order on a tape. A complete data recovery includes checking for multiple tape files and extracting them all.

Recovering data from a tape that is clean and not damaged, requires a tape drive, host system, and software to restore the data. Using the original large vintage tape drives like the IBM 2400 or 3400 Series is possible, but this involves having a working tape drive and host system, with the original operating system and software, and the knowledge to operate it. Newer SCSI-based tape drives can still read old 9-track tapes, and allow the use of modern Linux systems, open source tools, and basic shell command-line knowledge to perform the data extraction.

For this article I will use a DEC TSZ07 table-top tape drive from the early 1990s to read 9-track tapes. The drive features a front loading door where the tape is placed. To thread the tape, airflow suction is used to guide the fluttering tape-end through the tape heads and sensors to the take-up wheel where it is securely wound. Airflow and pressure chambers were also used to improve tape speed and tension in early IBM drives, but this is not done with the TSZ07. The drive can be configured and operated through the front panel, which also has an LED display and indicators.

Table-top DEC TSZ07



Front control and indicator panel

Adaptec 2944 SCSI card

Rear of TSZ07

Internal tape chamber

The PC and SCSI board does not need to be fast or powerful to read 9-track tapes. I used a recent version of Debian Linux on a PC with an older mainboard, 8GB RAM, and parallel PCI slots. The SCSI board is an Adaptec AHA-2944UW with an adapter to SCSI-2. Choosing the correct SCSI interface is important. This drive uses a differential SCSI interface with higher signal voltages, which is not compatible with common consumer PC SCSI boards.

The software used for any generic SCSI tape drives can be also be used for SCSI-based 9-track tape drives. The Linux kernel uses character devices /dev/st0 and /dev/nst0 to access a single attached tape drive. For data recovery the non-rewinding device /dev/nst0 is more practical.

The necessary tools originate from the UNIX world and typically include dd, mt, and tar. The mt command controls the drive, issuing SCSI commands to rewind, eject, move the tape forward or backward, get status information, and much more. There are two common versions, mt-gnu and mt-st (mt-st has more low level features). The dd command reads sequential blocks of data from the tape and saves them to the local PC filesystem. Conversion from IBM's EBCDIC encoding can be done with dd during the data extraction from tape, or as a second step on already extracted data. If the files stored on tape are known to be tar (Tape ARchive), then the usual tar command can be used. The paramaters and usage of these tools is documented in the manual pages.

Once the tape files have been extracted from tape (and possibly converted from EBCDIC), the next step is to identify the format or structure of the data. The original software used to write the tape files may have this format documented. If unkown, it may take some reverse engineering to reconstruct the file contents. The goal is to convert or transform the tape file data into a modern usable format such as regular text or CSV. Once this is achieved the tape data extraction has been successful. If you need support recovering data from 9-track tapes, contact me at bruce.nikkel@bfh.ch.

References:

The original English version of this article: https://digitalforensics.ch/enter/

TSZ07 Tape Drive Technical Manual: https://bitsavers.org/pdf/dec/magtape/tsz07/

ANSI standard: https://nvlpubs.nist.gov/nistpubs/Legacy/FIPS/fipspub50.pdf

EMCA standard: https://ecma-international.org/publications-and-standards/standards/ecma-62/

mt-st manpage: https://www.unix.com/man_page/debian/1/mt-st

mt-gnu manpage: https://www.unix.com/man_page/linux/1/mt-gnu/

dd manpage: https://www.unix.com/man_page/linux/1/dd/

tar manpage: https://www.unix.com/man_page/linux/1/tar