Magnetic tape data storage

Magnetic tape data storage is a system for storing digital information on magnetic tape using digital recording. Modern magnetic tape is most commonly packaged in cartridges and cassettes. The device that performs writing or reading of data is a tape drive. Autoloaders and tape libraries automate cartridge handling. For example, a common cassette-based format is Linear Tape-Open, which comes in a variety of densities and is manufactured by several companies.

Contents

Open reels

UNIVAC
- IBM formats
  - Bytes per inch
- DEC format

Cartridges and cassettes

Technical details
- Tape width
- Recording method
  - Linear
  - Scanning
- Block layout and speed matching
- Sequential access to data
- Access time
- Data compression
- Encryption
- Cartridge memory and self-identification

Viability

High density magnetic media

Chronological list of tape formats

See also

References

External links

Open reels

Initially, magnetic tape for data storage was wound on 10.5-inch (27 cm) reels.[1] This de facto standard for large computer systems persisted through the late 1980s, with steadily increasing capacity due to thinner substrates and changes in encoding. Tape cartridges and cassettes were available starting in the mid-1970s and were frequently used with small computer systems. With the introduction of the IBM 3480 cartridge in 1984, large computer systems started to move away from open reel tapes and towards cartridges.

UNIVAC

Magnetic tape was first used to record computer data in 1951 on the Eckert-Mauchly UNIVAC I. The UNISERVO drive recording medium was a thin metal strip of 0.5-inch (12.7 mm) wide nickel-plated phosphor bronze. Recording density was 128 characters per inch (198 micrometre/character) on eight tracks at a linear speed of 100 in/s (2.54 m/s), yielding a data rate of 12,800 characters per
second. Of the eight tracks, six were data, one was a parity track, and one was a clock, or timing track. Making allowances for the empty space between tape blocks, the actual transfer rate was around 7,200 characters per second. A small reel of mylar tape provided separation from the metal tape and the read/write head.

**IBM formats**

IBM computers from the 1950s used ferrous-oxide coated tape similar to that used in audio recording. IBM's technology soon became the de facto industry standard. Magnetic tape dimensions were 0.5-inch (12.7 mm) wide and wound on removable reels up to 10.5 inches (267 mm) in diameter. Different tape lengths were available with 1,200 feet (370 m) and 2,400 feet (730 m) on mil and one half thickness being somewhat standard. During the 1980s, longer tape lengths such as 3,600 feet (1,100 m) became available using a much thinner PET film. Most tape drives could support a maximum reel size of 10.5 inches (267 mm). CDC used IBM compatible 1/2 inch magnetic tapes, but also offered a 1 inch wide variant, with 14 tracks (12 data tracks corresponding to the 12 bit word of CDC 6000 series peripheral processors, plus two parity bits) in the CDC 626 drive.

A so-called mini-reel was common for smaller data sets, such as for software distribution. These were 7-inch (18 cm) reels, often with no fixed length—the tape was sized to fit the amount of data recorded on it as a cost-saving measure.

Early IBM tape drives, such as the IBM 727 and IBM 729, were mechanically sophisticated floor-standing drives that used vacuum columns to buffer long u-shaped loops of tape. Between servo control of powerful reel motors, a low-mass capstan drive, and the low-friction and controlled tension of the vacuum columns, fast start and stop of the tape at the tape-to-head interface could be achieved: 1.5 ms from stopped tape to full speed of 112.5 inches per second (2.86 m/s). The fast acceleration is possible because the tape mass in the vacuum columns is small; the length of tape buffered in the columns provides time to spin the high inertia reels. When active, the two tape reels thus fed tape into or pulled tape out of the vacuum columns, intermittently spinning in rapid, unsynchronized bursts resulting in visually striking action. Stock shots of such vacuum-column tape drives in motion were widely used to represent “the computer” in movies and television.

Early half-inch tape had seven parallel tracks of data along the length of the tape, allowing six-bit characters plus one bit of parity written across the tape. This was known as seven-track tape. With the introduction of the IBM System/360 mainframe, nine-track tapes were introduced to support the new 8-bit characters that it used.

Recording density increased over time. Common seven-track densities started at 200 six-bit characters per inch (CPI), then 556, and finally 800. Nine-track tapes had densities of 800 (using NRZI), then 1600 (using PE), and finally 6250 (using GCR). This translates into about 5 megabytes to 140 megabytes per standard length (2400 ft) reel of tape. The end of a file was designated by a special recorded pattern called a tape mark, and end of the recorded data on a tape by two successive tape marks. The physical beginning and end of usable tape was indicated by reflective adhesive strips of aluminum foil placed on the back side.

At least partly due to the success of the S/360, and the resultant standardization on 8-bit character codes and byte addressing, nine-track tapes were very widely used throughout the computer industry during the 1970s and 1980s.

**Bytes per inch**

The term BPI can mean bits per inch of a single track, or when referring to nine-track tape, to bytes per inch. Since a byte plus its parity bit occupies all nine tracks, the numeric values of the two measurements are the same.

**DEC format**
LINCTape, and its derivative, DECtape, were variations on this "round tape". They were essentially a personal storage medium. The tape was 0.75 inches (19 mm) wide and featured a fixed formatting track which, unlike standard tape, made it feasible to read and rewrite blocks repeatedly in place. LINCTapes and DECTapes had similar capacity and data transfer rate to the diskettes that displaced them, but their "seek times" were on the order of thirty seconds to a minute.

**Cartridges and cassettes**

In the context of magnetic tape, the term cassette usually refers to an enclosure that holds two reels with a single span of magnetic tape. The term cartridge is more generic, but frequently means a single reel of tape in a plastic enclosure.

The type of packaging is a large determinant of the load and unload times as well as the length of tape that can be held. A tape drive that uses a single reel cartridge has a takeup reel in the drive while cassettes have the take up reel in the cassette. A tape drive (or "transport" or "deck") uses precisely controlled motors to wind the tape from one reel to the other passing a read/write head as it does.

A different type of tape cartridge has a continuous loop of tape wound on a special reel that allows tape to be withdrawn from the center of the reel and then wrapped up around the edge. This type is similar to a cassette in that there is no take-up reel inside the tape drive.

The IBM 7340 Hypertape drive, introduced in 1961, used a cassette with a 1 inch (2.5 cm) wide tape capable of holding 2 million six-bit characters per cassette.

In the 1970s and 1980s, audio Compact Cassettes were frequently used as an inexpensive data storage system for home computers, or in some cases for diagnostics or boot code for larger systems such as the Burroughs B1700. Compact cassettes were logically, as well as physically, sequential; they had to be rewound and read from the start to load data. Early cartridges were available before personal computers had affordable disk drives, and could be used as random access devices, automatically winding and positioning the tape, albeit with access times of many seconds.

As of 2014 most magnetic tape systems used reels fixed inside a cartridge to protect the tape and facilitate handling. Cartridge formats include DDS/DAT, DLT and LTO with capacities in the tens to thousands of gigabytes.

**Technical details**

**Tape width**

Medium width is the primary classification criterion for tape technologies. Half inch has historically been the most common width of tape for high capacity data storage.[4] Many other sizes exist and most were developed to either have smaller packaging or higher capacity.

**Recording method**

Recording method is also an important way to classify tape technologies, generally falling into two categories:

Linear
The **linear** method arranges data in long parallel tracks that span the length of the tape. Multiple tape heads simultaneously write parallel tape tracks on a single medium. This method was used in early tape drives. It is the simplest recording method, but also has the lowest data density.

A variation on linear technology is **linear serpentine** recording, which uses more tracks than tape heads. Each head still writes one track at a time. After making a pass over the whole length of the tape, all heads shift slightly and make another pass in the reverse direction, writing another set of tracks. This procedure is repeated until all tracks have been read or written. By using the linear serpentine method, the tape medium can have many more tracks than read/write heads. Compared to simple linear recording, using the same tape length and the same number of heads, data storage capacity is substantially higher.

**Scanning**

Scanning recording methods write short dense tracks across the width of the tape medium, not along the length. Tape heads are placed on a drum or disk which rapidly rotates while the relatively slowly moving tape passes it.

An early method used to get a higher data rate than the prevailing linear method was **transverse scan**. In this method a spinning disk, with the tape heads embedded in the outer edge, is placed perpendicular to the path of the tape. This method is used in Ampex's DCRsi instrumentation data recorders and the old Ampex quadruplex videotape system. Another early method was **arcuate scan**. In this method, the heads are on the face of a spinning disk which is laid flat against the tape. The path of the tape heads forms an arc.

**Helical scan** recording writes short dense tracks in a **diagonal** manner. This method is used by virtually all current videotape systems and several data tape formats.

**Block layout and speed matching**

In a typical format, data is written to tape in blocks with inter-block gaps between them, and each block is written in a single operation with the tape running continuously during the write. However, since the rate at which data is written or read to the tape drive is not deterministic, a tape drive usually has to cope with a difference between the rate at which data goes on and off the tape and the rate at which data is supplied or demanded by its host.

Various methods have been used alone and in combination to cope with this difference. If the host cannot keep up with the tape drive transfer rate, the tape drive can be stopped, backed up, and restarted (known as **shoe-shining**, with the restart optionally occurring at a lower speed). A large memory buffer can be used to queue the data. In the past, the host block size affected the data density on tape, but on modern drives, data is typically organized into fixed sized blocks which may or may not be compressed and/or encrypted, and host block size no longer affects data density on tape. The Linear_Tape-Open article covers this. Modern tape drives offer a speed matching feature, where the drive can dynamically decrease the physical tape speed as needed to avoid shoe-shining.

In the past, the size of the inter-block gap was constant, while the size of the data block was based on host block size, affecting tape capacity. On most modern drives, this is no longer true. Linear_Tape-Open type drives use a fixed size block for tape, independent of the host block size, and the inter-block gap is variable to assist with speed matching during writes. On drives with compression, the compressibility of the data will affect the capacity.

**Sequential access to data**
Tape is characterized by sequential access to data. While tape can provide fast sequential data transfers, it takes tens of seconds to load a cassette and position the tape head to an arbitrary place. By contrast, hard disk technology can perform the equivalent action in tens of milliseconds (3 orders of magnitude faster) and can be thought of as allowing random access to data.

Logical filesystems require data and metadata to be stored on the data storage medium. Storing metadata in one place and data in another requires lots of slow repositioning activity on most tape systems. As a result, most tape systems use a trivial filesystem in which files are addressed by number not by filename. Metadata such as file name or modification time is typically not stored at all. Tape labels store such metadata, and they are used for interchanging data between systems. File archiver and backup tools have been created to pack multiple files along with the related metadata into a single 'tape file'. Serpentine tape drives (e.g., QIC) can improve access time by switching to the appropriate track; tape partitions were used for directory information. The Linear Tape File System is a method of storing file metadata on a separate part of the tape. This makes it possible to copy and paste files or directories to a tape as if it were just like another disk, but does not change the fundamental sequential access nature of tape.

**Access time**

Tape has quite a long latency for random accesses since the deck must wind an average of one-third the tape length to move from one arbitrary data block to another. Most tape systems attempt to alleviate the intrinsic long latency, either using indexing, where a separate lookup table (tape directory) is maintained which gives the physical tape location for a given data block number (a must for serpentine drives), or by marking blocks with tape mark that can be detected while winding the tape at high speed.

**Data compression**

Most tape drives now include some kind of lossless data compression. There are several algorithms which provide similar results: LZ (most), IDRC (Exabyte), ALDC (IBM, QIC) and DLZ1 (DLT). Embedded in tape drive hardware, these compress a relatively small buffer of data at a time, so cannot achieve extremely high compression even of highly redundant data. A ratio of 2:1 is typical, with some vendors claiming 2.6:1 or 3:1. The ratio actually obtained with real data is often less than the stated figure; the compression ratio cannot be relied upon when specifying the capacity of equipment, e.g., a drive claiming a compressed capacity of 500GB may not be adequate to back up 500GB of real data. Data that is already stored efficiently may not allow any significant compression; a sparse database may offer much larger factors. Software compression can achieve much better results with sparse data, but uses the host computer's processor and can slow the backup if it is unable to compress as fast as the data is written.

The compression algorithms used in low-end products are not the most effective known today, and better results can usually be obtained by turning off hardware compression AND using software compression (and encryption if desired) instead.

Plain text, raw images, and database files (TXT, ASCII, BMP, DBF, etc.) typically compress much better than other types of data stored on computer systems. By contrast, encrypted data and pre-compressed data (PGP, ZIP, JPEG, MPEG, MP3, etc.) would normally increase in size, if data compression was applied. In some cases this data expansion could be as much as 15%.

**Encryption**

Standards exist to encrypt tapes. Encryption is used so that even if a tape is stolen, the thieves cannot use the data on the tape. Key management is crucial to maintain security. Encryption is more efficient if done after compression, as encrypted data cannot be compressed effectively. Some enterprise tape drives can quickly encrypt data. Symmetric streaming encryption algorithms can also provide high performance.

**Cartridge memory and self-identification**

Some tape cartridges, notably LTO cartridges, have small, associated data storage chips built into the cartridges that record metadata about the tape, such as the type of encoding, size of the storage, dates and other information. It is also common for tape cartridges to have bar codes on their labels in order to assist an automated tape library.
As of 2009, Medium and large-sized data centers deployed both tape and disk formats.[9]

The costs of disk storage have decreased faster than that of tapes. Until about the end of the twentieth century, prices and capacities allowed backing up a desktop hard drive to tape more cheaply and more compactly than backing up to an additional external or removable drive. Drive prices offer capacities of hundreds to thousands of megabytes on relatively inexpensive machines. Backing up to an external USB drive became cheaper, and the drive more compact, than tape for a non-networked machine used by a business or serious user.[10]

As a basic comparison, mainframe-class tape drives, such as Oracle’s Sun StorageTek T10000B, are priced at approximately US$37,000 each, excluding tape libraries. (IBM’s TS1130 is also representative of this storage class.) At any single moment in time each T10000C tape drive can read or write (or both read and write) to one tape cartridge which can contain up to 5TB of uncompressed data. Real-world sequential data transfer speeds are high (sustained 240MB/second for the T10000C and 160MB/second for the TS1130) compared to disk. However, PC-class hard disks are priced below $200 for 3TB.[11] One mainframe-class hard disk still has a much lower price than one mainframe-class tape drive, so the economics might favor disk.

However, the key difference is that tape drives can exchange their magnetic media (the cartridges) frequently, while the magnetic media installed inside each hard disk is fixed and cannot be swapped. (The drives themselves could be moved if installed in swappable caddies at extra cost, with extra cost hot-swappable infrastructure.) Mainframe-class tape drives are almost always installed in robotic tape libraries which are often quite large and can hold thousands of cartridges. The StorageTek SL8500 library is one representative example. The smallest SL8500 library holds up to 1,448 tape cartridges, for 1.4 Petabytes of online uncompressed storage. An equivalent amount of PC-class hard disk storage would be priced at $100,000 or less for the drives. The tape library would likely deliver a higher sustained sequential write speed, the media would be more rugged (for off-site storage), the media would meet or exceed long-term archival storage requirements (for reliable retrieval decades into the future), and the data center power and cooling requirements would be considerably lower. The economics of this comparison are more complicated than a single-spindle versus tape drive comparison.

Whether tape’s characteristics versus disk are useful or not will depend on the particular data center and its data storage requirements. What has tended to happen in recent years is that the amount of data has grown exponentially, with both disk (especially) and tape participating in the growth. In the early twenty-first century solid state storage encroached on disk’s previous near-monopoly in random access non-volatile data storage, while disk pushed into tape’s territory to some extent, particularly in situations where sequential data access is only a relatively small part of a particular data center’s storage requirements.

**High density magnetic media**

Sony announced in 2014, that they had developed a tape storage technology with the highest reported magnetic tape data density, 148 Gbit/in² (23 Gbit/cm²), potentially allowing tape capacity of 185TB.[12] It was further developed by Sony, with announcement in 2017, about reported data density of 201 Gbit/in², giving standard tape capacity of 330TB.[13]

In May 2014, Fujifilm followed Sony and made an announcement that it will develop a 154 TB tape cartridge in conjunction with IBM, which will have an areal data storage density of 85.9 billion bits per in² (13.3 billion bits per cm²) on linear magnetic particulate tape.[14] The technology developed by Fujifilm, called NANOCUBIC, reduces the particulate volume of BaFe magnetic tape, simultaneously increasing the smoothness of the tape, increasing the signal to noise ratio during read and write while enabling high frequency response.

**Chronological list of tape formats**
1951 - UNISERVO
1952 - IBM 7 track
1958 - TX-2 Tape System
1961 - IBM 7340 Hypertape
1962 - LINCTape
1963 - DECTape
1964 - 9 Track
1964 - Magnetic tape selectric typewriter
1966 - 8-track tape[^15]
1972 - Quarter-inch cartridge (QIC)
1975 - KC standard, Compact Cassette
1976 - DC100
1977 - Tarbell Cassette Interface
1977 - Commodore Datasette
1979 - DECTape II cartridge
1979 - Exatron Stringy Floppy
1981 - IBM PC Cassette Interface
1983 - Sinclair ZX Microdrive
1984 - Sinclair QL Microdrive
1984 - Rotronics Wafadrive
1984 - IBM 3480 cartridge
1984 - Digital Linear Tape (DLT)
1986 - SLR
1987 - Data8
1989 - Digital Data Storage (DDS) on Digital Audio Tape (DAT)
1992 - Ampex DST
2000 - LTO-1
2003 - SAIT
2003 - LTO-2
2003 - 3592
2005 - LTO-3
2005 - TS1120
2006 - T10000
2007 - LTO-4
2008 - TS1130
2008 - T10000B
2010 - LTO-5
2011 - TS1140
2011 - T10000C
2012 - LTO-6
2013 - T10000D
2014 - TS1150
2015 - LTO-7
2017 - TS1155
2017 - LTO-8
1992 - Ampex DST
1994 - Mammoth
1995 - IBM 3590
1995 - StorageTek Redwood SD-3
1995 - Travan
1996 - AIT
1997 - IBM 3570 MP
1998 - StorageTek T9840
1999 - VXA
2000 - StorageTek T9940

See also

- Computer data storage
- Magnetic storage
- Tape drive
- Information repository
- Data proliferation
- Tape mark
- Linear Tape-Open

References


7. As illustrated by the pigeonhole principle, every lossless data compression algorithm will end up increasing the size of some inputs.


11. Mainframe-class tape vendors specify terabytes of capacity while hard disk vendors typically specify trillions of bytes of capacity. Two trillion bytes is approximately 1.8 terabytes.


External links

- ISC 35.220.22 Magnetic Tapes
- ISC 35.220.23 Cassettes and cartridges for magnetic tapes


This page was last edited on 6 April 2018, at 05:19.

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.