Chapter 1

Tertiary Storage Technologies

The fundamental things apply,
As time goes by.
Herman Hupfeld

1.1 Introduction

In the last chapter, we described a typical computer system storage hierarchy. In this chapter, we describe the workings of tertiary storage, the lowest level of the hierarchy. In Section 1.2, we discuss magnetic tape drive technology, including measurements of tape drive performance. In Section 1.3, we do the same for optical disk technology. We give examples of both magnetic tape and optical disk drive and robot products and describe technology trends. We finish the chapter with a discussion of two more exotic tertiary storage technologies: holographic storage and optical tape.

1.2 Magnetic Tape Technology

1.2.1 The Tape Drive

Magnetic tape drives [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13] store data as small magnetized regions on a tape composed of magnetic material deposited on a thin, flexible substrate such as plastic. Magnetic tapes are written by inducing a current proportional to an input signal in the coil of an inductive write head; this current induces a magnetic field in the tape below the write head that magnetizes a small tape region. After data have been written, a read head can detect magnetic flux on the tape; the rate of change of this flux results in a voltage on the read head that is used to deduce the original input signal [8] [10]. A tape drive mechanism consists of reels, motors, gears, brakes and belts that thread the tape, control tape reeling, guide the tape to maintain its position relative to the heads, control the acceleration and steady-state velocity of the tape, and control tape tension [9].

The substrate for magnetic tapes is a poly-ethylene terepthalate (PET) base film from 12 to 36 microns thick [14]. There is a tradeoff between reliability and capacity in choosing the thickness of the substrate. A thicker substrate is more durable and able to withstand high accelerations and start and stop operations without excessive distortions [9]. A thinner substrate allows a single reel to store more tape, with greater resulting capacity.

On top of the substrate is the magnetic layer that stores data. Ideally, the magnetic layer should be very smooth to allow the minimum spacing, or air bearing, between the read/write head and the tape; the smaller this air bearing, the greater the density of magnetic information that can be stored on the tape. There are two types of magnetic layers: Magnetic Particle (MP) and Metal Evaporated (ME). Historically, most tape systems have used magnetic particle tapes in which the PET substrate is covered with a particulate magnetic coating that is 2 to 5 microns thick. The coating contains magnetic particles

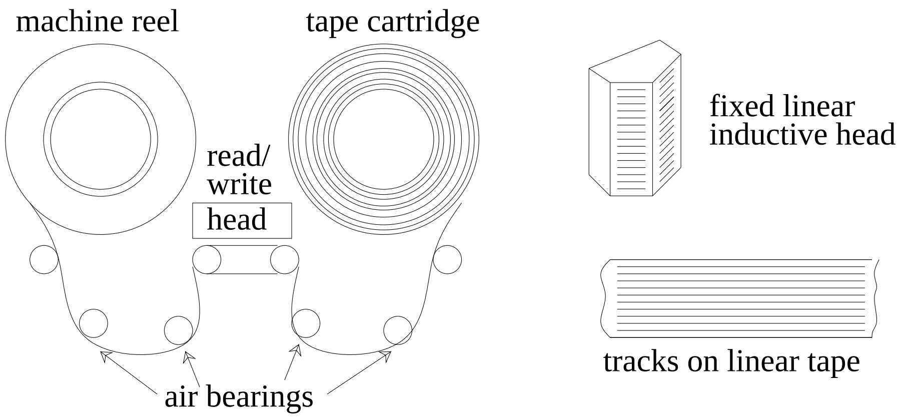


Figure 1.1: *A Linear Recording Magnetic Tape Drive*

such as CrO_2 or Fe_2O_3, polymeric binders and lubricants. In the last five years, metal evaporated tape systems have overcome earlier difficulties with corrosion and tribology (the head/tape interface) and now offer superior recording qualities. Metal evaporated tapes have a thin film of metal (usually Co-Ni) evaporated onto the PET substrate in a layer about 100 nanometers thick. Because the thin film magnetic layer is much thinner than a particulate magnetic coating, thin film tapes hold greater capacity. In addition, thin film media have recording characteristics superior to those of particulate media, including higher effective magnetization, lower error rates, greater smoothness, highly isolated asperities, and better signal-to-noise ratios [8], [15], [14].

Besides the plastic substrate and the magnetic coating, most magnetic tapes have a back coating for protection against static. Electrostatic attraction can lure and embed dust particles from the atmosphere into the tape, causing read and write errors.

Figures 1.1 and 1.2 show the two most common tape formats and drive mechanisms for magnetic recording of computer data: linear and helical-scan.

In linear recording magnetic tape drives, data are written in tracks parallel to the edges of the tape. Many linear recording devices use stationary, multi-track heads to read or write the entire width of the tape simultaneously. Figure 1.1 shows a simplified drawing of a recording head for a linear tape drive. The recording head is made up of a coil and a

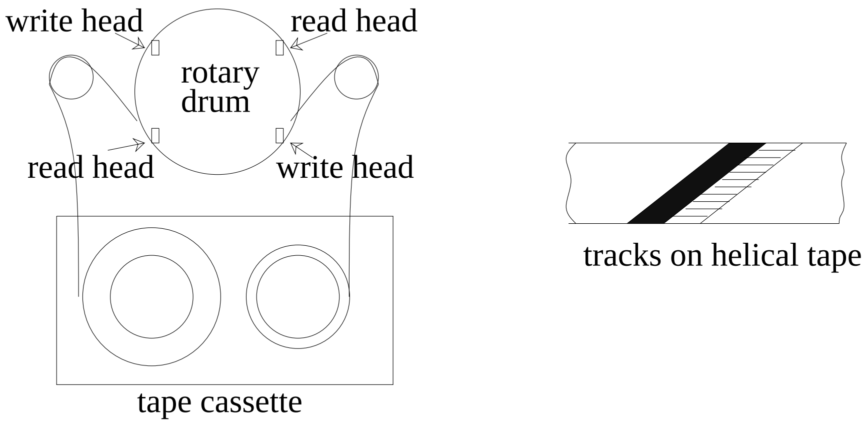


Figure 1.2: *A Helical-Scan Magnetic Tape Drive*

gapped magnetic structure that intensifies and localizes the magnetic field during writing [10]. The read and write elements are of different size, orientation and composition; the read elements are magnetoresistive, while the write elements are inductive. Before writing the tape, a separate erase head erases the tape with a magnetic field. The read/write head is narrower than the erase head, so when new data are written, gaps or "write guard bands" are left between adjacent tracks to minimize cross talk. Further, the read elements on the head are narrower than the write elements, creating a "read guard band" to further reduce the risk of reading erroneous information. Data are read immediately after being written to check for errors and rewritten if errors occur. In older linear tape drives, the write guard bands are separations between adjacent tracks is large; as a result, the areal density of these drives is relatively low compared to helical scan tapes. More recent linear tape drives have reduced or eliminated the separation between tracks and greatly increased the areal density. Another type of linear recording drive is a serpentine drive that reads or writes a few tracks at a time down the entire length of the tape; then tape motion switches directions, and the heads themselves are shifted to read or write a different set of tracks.

Figure 1.2 shows a helical scan tape drive. Helical scan tape drives achieve high areal densities and, potentially, high data rates [16] [17] [18] [19]. Read and write heads are situated on a drum that rotates at high speed (around 2000 r.p.m.). Tracks are written at a small angle (10 to 20 degrees) with respect to the direction of tape movement onto a slow-moving (around 1/2-inch per second) magnetic tape. On some drives, a stationary erase head erases the data before writes occur. In addition, helical scan drives usually contain a servo head that can read servo information written at the edges of the tracks. The presence of the servo head makes it possible for the heads to be positioned precisely, so that the gaps between adjacent tracks can be small. If two sets of read/write heads are set at opposite angles, then two tracks can be written simultaneously. These adjacent tracks may overlap slightly, writing a herringbone pattern as shown in Figure 1.2. Using different write angles on adjacent tracks minimizes crosstalk [19]. Because adjacent-track gaps are small or nonexistent in helical scan drives, areal density on these drives is high. As for linear tape drives, in helical-scan drives, data are read immediately after being written; the drive continues to write the data in subsequent locations until the write completes without errors.

There is substantial controversy over the relative merits of linear and helical scan technologies. Stationary head, linear drives are thought to have a less abrasive head-to-tape interface than helical scan systems, resulting in longer-lived heads and tapes that tolerate more recording and play passes. Helical, rotating-head systems boast a relatively slow normal tape speed that allows substantial speedup during fast search and rewind operations; for example, the 4mm DAT drives have a fast positioning operation that is 200 times faster than the normal tape speed. By contrast, the tape typically moves quickly in a linear drive, since the tape speed along with the track density and number of heads determines the data rate of the tape drive. In addition, Bhushan argues that helical systems potentially have higher data rates, since it should be easier to speed up the rotating drum in a helical tape drive than to quickly move heavy rolls of tape in a linear drive [14].

Even among helical scan drives, there is controversy over which design is superior. For example, the “wrap angle,” or amount of tape that is wrapped around the rotating drum contains the recording heads, is 90 degrees for a 4mm DAT drive and 220 degrees for an 8mm drive. Advocates of 4mm technology claim the shorter wrap angle is less stressful, reducing wear and improving access times [17], while 8mm proponents claim their longer wrap angle is actually less stressful, providing better tape guidance [16].

Several features of magnetic tape drives are notable. Most tapes are append-only, although a few update-in-place magnetic tape drives do exist [19]. When data are written on any portion of a magnetic tape, whatever had been written beyond the new write point is irretrievable. There are several reasons for the append-only nature of tapes. First, tape mechanisms are not well-suited for the large numbers of forward and reverse positioning operations that are typical of update-in-place devices like magnetic disk drives. The mechanical stop and start operations are time-consuming and cause wear on the tape, so they are avoided. Instead, applications strive to use tape drives in a *streaming* mode of operation in which data are sent at a constant rate to or from the tape head. Second, update-in-place is difficult because the high rate of errors on the media make it difficult to predict exactly where data will lie on the tape. When a tape drive writes data, it immediately reads back what it has written to check for correctness. When there is an error, the tape drive re-writes the data at a later position on the tape, repeating the read-after-write checking until the data are written without errors. Unlike magnetic disks, where data are mapped into particular sectors at predictable locations on a disk platter, it is impossible to know before a write operation exactly where data will physically reside on a magnetic tape. There have been a few update-in-place magnetic tape drives, notably the recent Data/DAT format for 4mm tape drives. This format has not found favor compared to its more traditional, append-only competing format, DST [19]. Update-in-place tape drives work by zoning the tape into sectors, leaving wide buffers of empty tape between the sectors to allow for problems in writing the data. These buffer zones must also be large to protect good data from the tape drive's large erase head. The result of these large buffer zones is that tapes have much lower total storage capacity than tapes operated in append-only mode.

Many tape drives require that tape cartridges be rewound before they can be ejected from a tape drive. There are several reasons for this requirement. Some tapes are stored on a single tape reel and must be rewound before they can be physically removed. Even for double-reel cassettes, rewind is often required because there may be servo information on the start of the tape that must be read on insertion. Also, because inserting and ejecting a tape can cause considerable wear, some manufacturers prefer that those operations occur on a strip of tape where no data are stored. Some tape drives include load/eject "zones" scattered throughout the tape; these zones contain servo information to allow the tapes to be inserted at those points without rewinding to the start of tape. Finally, keeping tapes in streaming operation is essential to achieving good performance. In order to minimize tape and head wear, tape drives will release tension on the tape if there is a lapse of a certain length of time after an operation; if the delay goes on even longer, the drum in a helical scan tape drive will stop spinning. Subsequent requests will have to wait for the drum to spin up and for tape tension to be reapplied. These delays reduce throughput.

1.2.2 Existing Products

There are a number of magnetic tape products available that provide a variety of tape capacities and speeds for a range of prices. Table 1.1 compares several tape drives based on cost, cartridge capacity and data transfer rate [20], [16], [21], [17], [22], [23], [18], [19], [24], [25], [26]. At the low end, an inexpensive 4mm DAT drive stores several gigabytes of data but has a low transfer rate of 366 kilobytes/second. At the high end, the very expensive Sony D-1 drive stores about 25 times the capacity stored by the DAT drive and transfers data at approximately 100 times the transfer rate of the DAT drive.

In the next chapter, we present detailed measurements of three tape drives and

Tape Drive	Price ($)	Media Cost ($/MB)	Capacity (GBytes)	Data Transfer Rate	Drive Type
Sony SDT-4000	$ 1,695	$0.015	4	183-366 KB/sec	4mm helical
Exabyte EXB8500	$ 2,315	$0.008	5	500 KB/sec	8mm helical
Storage Tek 4220	$ 19,000	$0.013	0.2-0.8	712 KB/sec	1/2-inch linear
Metrum 2150	$ 32,900	$0.0013	18	2 MB/sec	1/2-inch helical
Ampex D-2	$100,000	$0.0018	25	15 MB/sec	19mm helical
Sony D-1	$270,000	$0.0016	90	32 MB/sec	19mm helical

Table 1.1: *Price, cartridge capacity and data transfer rates of a variety of magnetic tape products.*

show that access times for tape requests can take on average several minutes. An access to a tape library or robot may first involve a rewind of the currently-loaded tape. Next, the old tape must be physically ejected from the tape drive. The robot arm moves to unload an old tape and load a new one. Then the tape drive physically loads the new tape, including reading servo information at the start of the tape. The tape drive then performs a forward search operation to position the tape. Finally the tape drive performs the read or write data transfer operation.

1.2.3 Reliability of Tape Drives and Media

Magnetic tape systems face difficult reliability challenges [27]. Tape wear, head wear and rates of errors uncorrectable by error correction codes (ECC) may make errors more frequent in large tape systems than in disk arrays. It is an open research question how much error correction will be required to make tape arrays adequately reliable.

Tape Media Reliability

The rate of raw errors (i.e., errors before any error correction has been performed) is quite high on magnetic tape media: approximately one error in every 10^5 bits read. Most of these errors are caused by "dropouts," in which the signal being sensed by the tape

Drive	Corrected Bit Error Rate
DAT	10^{-15}
Exabyte EXB8500	10^{-13}
1/4"	10^{-14}
Metrum 1/2"	10^{-13}
19mm D-2	10^{-12}

Table 1.2: *Compares the error rates per bits read or written for a variety of tape drives.*

head drops below a readable value. Dropouts are most commonly caused by protrusions on the tape surface that temporarily increase the separation between the head and the tape, causing a drop in signal intensity [9]. There are several ways that debris can become embedded in the tape and cause dropouts. When the tape is originally sliced, loose pieces of tape substrate and coating are left at the edges of the tape; these may become embedded in the tape surface. Second, as the tape passes through the drive mechanism, it can become charged and attract particles in the atmosphere. Another way debris accumulate on the tape is through the wear that occurs when the tape drive starts and stops. At very low velocities, the separation between the head and tape surface normally caused by air flow is not maintained. The head contacts the tape coating, producing a fine, dry powder. Pushed by the head, this powder can accumulate and cause dropouts. Dropouts are also caused by inhomogeneities in the tape's magnetic coating.

Because of the high raw bit error rates on all magnetic tape devices, all drives incorporate large amounts of error-correction code. However, some errors will occur that the ECC cannot correct. The rate of such errors is called the uncorrectable bit error rate. Table 1.2 shows that uncorrectable bit error rates for the different tape technologies vary from one error in every 10^{12} bits read to one in every 10^{15} bits read. Taking the Exabyte 8mm drive as an example, suppose a small tape library contained eight Exabyte drives that each transfer data at 500 KBytes/sec. Assuming that the drives operated on just a 10% duty cycle, an uncorrectable bit error will occur on average every 36.2 days.

It is notable that one of the most expensive tape drives, the Ampex DST800, has one of the highest error rates, while the inexpensive DAT drive has the lowest rate. The DAT drive achieves this low error rate by including 3 layers of Reed-Solomon error correction. The extensive encoding required to perform this error correction consumes a great deal of capacity. For example, the first two ECC layers consume 30% of the bits available for data storage; depending on the amount of encoding performed for the third ECC level, another 4% to 18% of tape capacity may be consumed [19]. Thus, there is a tradeoff between data reliability and tape capacity available for data storage. Incorporating the third ECC level is also expensive, requiring more complex electronics and larger buffers. DAT manufacturers recommend the use of the third level of ECC only for sensitive data.

Magnetic tapes that are frequently read or written eventually wear out. In a traditional archival system, where data are written and probably never read again, this is not a serious concern. However, in applications like digital libraries, there is no limit on the number of times a tape may be read. Tapes last on average several hundred passes [14], [28], [29]. However, they wear out sooner if a particular segment of the tape is accessed repeatedly. A Hitachi study of DAT tape drives showed that the raw error rate (before error correction) after 2500 passes to a single segment of a tape was over one error in ten ECC blocks read [30]. Tapes written by linear recording drives do not suffer so quickly from tape wear-out as tapes written by helical scan drives because the interface with the head is less abrasive, but wear is still a concern. In large tape libraries, controllers will be required to monitor the number of passes to a tape cartridge and replace it before wearout occurs.

In an interactive library application, wear due to stops and starts on the tape is likely to be severe, since accesses to the library will not be sequential. Severe wear is manifested by large portions of the magnetic binding material flaking away from the tape backing. Such problems make large sections of a tape unreadable.

Another set of reliability concerns involves the long-term storage of data on tape. Over time, the metal pigments in tape are subject to corrosion; this problem is eliminated to a large extent by the use of appropriate binders. The tape also undergoes mechanical changes including tape shrinkage, creasing of the edges, peeling of the magnetic layer, and deterioration of surface smoothness [28]. Back coating transfer can also occur, in which the magnetic coating and the back coating from adjacent tape layers are pressed together; when shrinkage occurs during storage, the roughness of the back coating can transfer onto the magnetic layer and cause a deterioration in tape smoothness [28]. Many manufacturers recommend rewinding tapes every 6 months to avoid such problems. Finally, the proliferation of incompatible recording formats threatens the future accessibility of archival data [29].

Tape Head Wear

Tape heads undergo considerable wear in all tape systems. In helical scan systems, they last for a few thousand hours of actual contact between the head and medium. Linear tapes are thought to produce less wear because the interface between the tape and the head is less abrasive. Some tape wear is necessary in order to keep the heads in optimum condition [31]. Tape wear helps remove particles from the head that may have been transferred there from the tape surface or the atmosphere, or that came from the tape coating under conditions of friction or extremely high or low humidity. All tape drive manufacturers recommend periodic use of a cleaning cartridge to remove debris from the tape head. Extensive wear occurs when new tapes are used, since new tapes tend to have a lot of surface debris that is removed by the heads during the first few passes. One way to extend the life of drive heads is to use burnished tapes, from which much of the surface debris has been removed.

Eventually, the head wear becomes extreme. Tape library controllers will need to schedule both cleaning and replacement of the heads to assure adequate reliability. This may require keeping statistics on how many hours particular drives have run.

Repair Type	%
Replace heads	44
Tape mechanism (reel motors, tape tension, etc.)	21
Card failure	17
Other (firmware, power supply, etc.)	14
No defect found	4

Table 1.3: *1991 repair statistics for Exabyte 8mm drives. (Source: Megatape)*

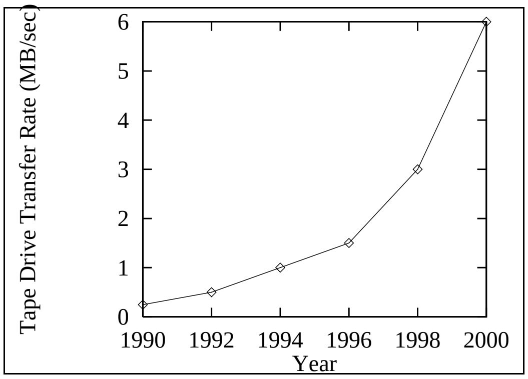
Mechanical Reliability

Head failure is the main cause of tape drive failure; however, the drive may also have other mechanical or electrical failures. Figure 1.3 shows repair statistics for Megatape, an OEM of Exabyte 8mm drives. 44% of the time, tape drive failures were due to failed heads. 21% of the time, some other component of the tape drive mechanism failed. 17% of the time a failure with the electronics was to blame, while 14% of the time other components such as power supplies caused the failures.

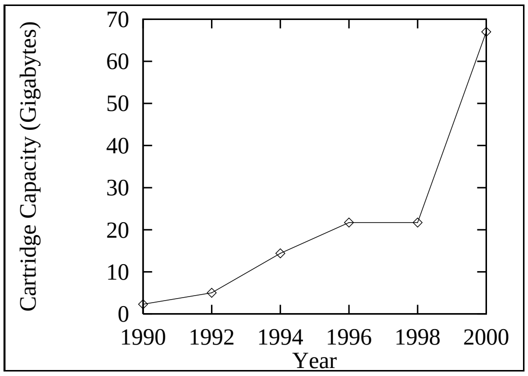
1.2.4 Trends in Magnetic Tape Drives

Magnetic tapes follow the same technology curves [32], [33], [34] as magnetic disks since the magnetic material, whether deposited on a hard disk or a flexible tape, is much the same. Currently, magnetic disk capacity is increasing at a rate of over 50% per year [35], and magnetic tapes should increase in capacity at a similar rate. In 1990, Hinz [36] predicted the growth shown in Figure 1.3 for tape capacity and data transfer rate for 8mm tape drives in this decade. The figure shows both capacity and throughput doubling approximately every two years, reaching 67 GBytes per tape and 6 MBytes/sec by the end of the decade. Actually, the generation of Exabyte tape drive being introduced in 1994 stores greater capacity per tape and transfers at a higher rate than Hinz predicted; each cartridge will hold 20 gigabytes of storage and transfer data at 3 megabytes per second.

Besides data transfer rate, other components of tape drive access time are improv-



1.3a: *Predictions for Bandwidth Improvements.*



1.3b: *Predictions for Capacity Improvements.*

Figure 1.3: *Predictions for bandwidth and capacity improvements for Exabyte 8mm tape drives in this decade. Source: Harry Hinz, Exabyte Corporation. Improvements required to reach these goals include increased track density, decreased track width and pitch, reduced tape thickness and increased rotor speed.*

ing. Several tape drive manufacturers are reducing rewind and search times by implementing periodic zones on the tapes where eject and load operations operations are allowed, rather than requiring that a tape must always be ejected and loaded at the start of the tape [37]. Mechanical operations like load, eject and robot grab and insert will be substantially faster in the next generation of tape drives.

1.2.5 Tape Robots

To provide higher bandwidth and capacity than can be supplied by a single device, several companies have built automated library systems. These libraries hold tens, hundreds or thousands of cartridges that can be loaded by robot arms into a collection of magnetic tape drives.

Table 1.4 shows a classification of some of the robots available for handling magnetic tape cartridges automatically [21], [38]. Table 1.5 describes examples of each type

Type	No. Cartridges	No. Drives	No. Robot Arms	Cost
Large Library	10 to 1000	several	one or two	high
Carousel	around 50	one or two	one (carousel)	moderate
Stacker	around 10	one	one (magazine or arm)	low

Table 1.4: *Classification of tape robots.*

of robot. Large libraries generally contain many cartridges, several drives and one or two robot arms for handling cartridges. The cartridges are often arranged in a rectangular array. Other "large library" configurations include a hexagonal "silo" with cartridges and drives along the walls, and a library consisting of several cylindrical columns holding cartridges that rotate to position them. Usually these large libraries are quite expensive ($500,000 or more), but they often have lower cost per megabyte than less expensive robotic devices. One disadvantage of large tape libraries is the low ratio of drives and robot arms to cartridges. In a heavily-loaded system, there may be contention for both arms and drives.

Carousel devices are moderately priced (around $40,000) and hold around 50 cartridges. The carousel rotates to position the cartridge over a drive, and a robot arm pushes the cartridge into the drive. In most cases, there are one or two drives per carousel.

Finally, the least expensive robotic device ($10,000 or less) is a stacker, which holds approximately 10 cartridges in a magazine and loads a single drive. The magazine may move vertically or horizontally to position a tape in front of the drive slot, or the stacker may have a robot arm that moves across the magazine to pick a cartridge. A storage system composed of stackers would have the highest ratio of robot arms and drives to cartridges.

Robot access times are fairly short compared to tape positioning operations like search and rewind. Removing an old cartridge or inserting a new one may consume between 5 and 20 seconds on typical robots.

Tape Library	Number Drives	Number Tapes	Total Capacity	Price ($)	Tape Drive Technology
ADIC DAT Autochanger 1200C	1	12	24 GB	$ 8,900	4mm helical
Exabyte EXB10i	1	10	50 GB	$ 13,000	8mm helical
Spectralogic STL800 Carousel	2	40	200 GB	$ 31,465	8mm helical
Exabyte EXB120	4	116	580 GB	$ 100,000	8mm helical
Metrum RSS600	5	600	10.8 TB	$ 265,900	1/2-in helical
Ampex DST600	4	256	6.25 TB	$1,000,000	19mm helical

Table 1.5: *Comparison of several tape robots.*

1.3 Optical Disk Technology

1.3.1 Optical Disks

In optical recording, a non-contact optical head uses a laser beam to store information on the disk surface by creating "pits" in the surface material [39], [40], [41], [42]. Optical disks can only be written once and are often called WORM (write once, read many) devices. Figure 1.3.1 shows the typical optical disk structure; a substrate (often plastic) is covered with an aluminum reflective layer, a transparent dielectric, and finally a thin layer of metal. During writing, in response to an electrical input signal, a highly-focused laser beam can melt a small region of the metal layer, opening a hole or "pit." Later, during reading, a lower-intensity, unmodulated laser beam is reflected off the surface of the disk. A photodetector interprets information stored on the disk by detecting differences in reflectivity between pits and the surface of the thin metal layer, called the "land." Data are encoded and stored as alternating regions of pits and land. The encoding may be a simple as a pit representing a zero and the land a one. In CD-ROM disks, a more complicated encoding scheme minimizes the number of consecutive zero or one bits to minimize intersymbol interference on the disk surface [42].

On optical disks, data may be stored in a single spiral track as in a CD-ROM or as a series of concentric circular tracks. Figure 1.3.1 shows a simplified diagram of an optical

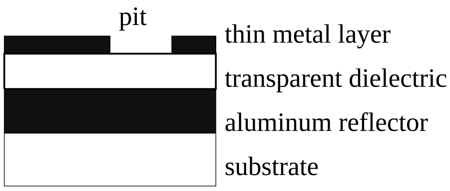


Figure 1.4: *Optical disk structure*

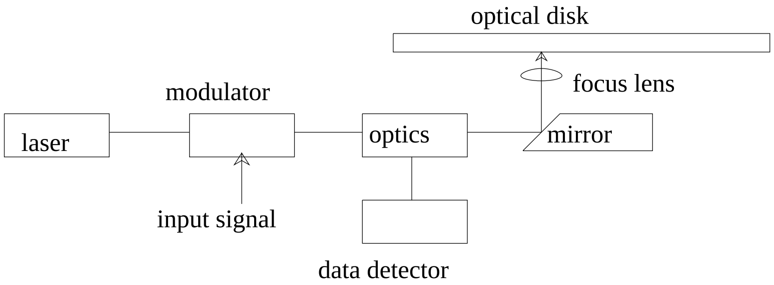


Figure 1.5: *Typical components of an optical disk*

disk.

The optical disk uses several servos, one for tracking, one for focusing and one to control the rotation speed so that the data rate is constant. As the read/write head moves from the center to the outer edge of the disk, the rate of rotation decreases to maintain a constant linear velocity.

Defects in disk material may have several causes. Dust particles may accumulate over time on the disk surface; this effect can be minimized using overcoats to protect the disk surface. Unevenness in evaporated coatings may also cause defects. Finally, intrinsic defects in the substrate disks can be minimized by pre-coating the substrate with a thin layer of plastic.

1.3.2 Magneto-optical Recording

Magneto-optical disks use lasers in a process called thermomagnetic recording to store and read information on magnetic material [43], [44]. An optical recording head that has no direct contact with the recording medium contains a diode laser. During a write operation, the laser heats a spot on the magnetic material above a critical temperature

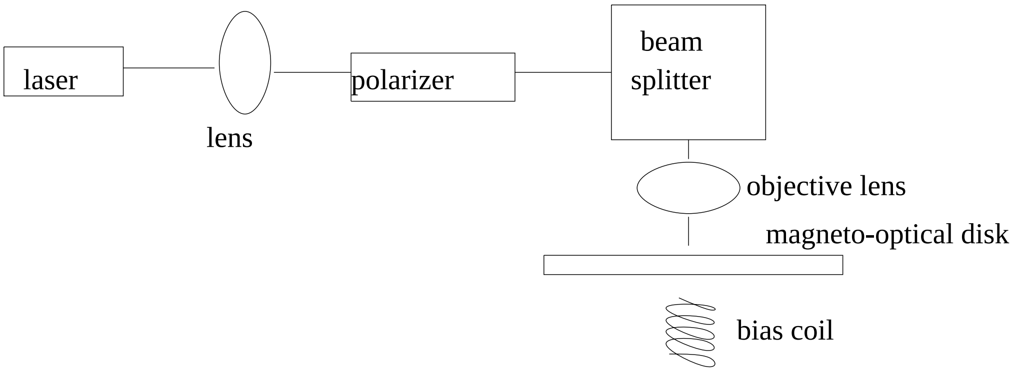


Figure 1.6: *Simplified diagram of magneto-optical disk.*

and then allows the spot to cool while a magnetic field is applied perpendicular to the disk surface. The direction of the magnetic field determines the direction of magnetization in the recording; one direction is used to record a “zero” value, and the opposite, a “one” value. Figure 1.6 shows a simplified diagram of the write mechanism. The magnetic field is generated using a bias coil. The laser beam is focused through a lens and polarizer; a beam splitter separates components of the beam, using the main beam for writing and reading and the other beams for tracking. Because bits are stored on the disk surface in areas of up and down magnetization, the disk can be re-written by repeating the write procedure after a separate erase phase.

Reading the bits off the disk is accomplished by reflecting a linearly-polarized light beam off the magnetic material. An optical principle known as the Kerr effect states that linearly polarized light reflected from a vertically-magnetized material will undergo a rotation in the plane of polarization of the light. The direction of the rotation depends on the direction of magnetization of the magnetic material; by detecting the change in rotation of the light, the magnetization of the material and hence the value of the bits stored is deduced. The read process uses a low-power laser beam that does not heat the magnetic material to its critical temperature, so data stored on the disk are not corrupted during the read process.

Optical Disk Drive	Price ($)	Capacity (megabytes)	Data Transfer Rate (Read)	Drive Type
Sony CDU-541 (CD-ROM)	$ 420	680	150 KB/sec	optical
Sony CDU-561 (CD-ROM)	$ 460	680	300 KB/sec	optical
Hitachi internal CD-ROM	$ 785	682	153.6 KB/sec	optical
Panasonic LF-3600	$1795	128	0.61 MB/sec	magneto-optical
Pioneer DE-UH 7101	$2400	645	635 KB/sec	magneto-optical
Panasonic LF-5014	$2995	940	0.625 MB/sec	optical
Maxoptix 3-1300	$3485	1330	3 MB/sec	magneto-optical
Hewlett-Packard C1617C	$3900	650	1 MB/sec	magneto-optical

Table 1.6: *Price, capacity and data transfer characteristics of a variety of optical and magneto-optical disk devices. The Panasonic drive is 3.5 inches in diameter; all the other drives are 5.25-inch disks.*

The magneto-optic disk is a molded plastic disk onto which a magneto-optic magnetic film is deposited using sputtering or evaporation techniques. The magneto-optic film is approximately 0.1 microns thick and is covered with a protective coating. The film is usually an alloy of rare earth and transition metals, such as gadolinium terbium cobalt (GdTbCo) or gadolinium terbium iron (GdTbFe). The principle method of degradation of the material is oxidation, which can be minimized using over- and undercoats of protective material; these coatings also have the advantage of enhancing the signal-to-noise ratio of the data [43]. Another reliability problem with the medium is the existence and growth of microscopic defects. It is estimated that the raw bit error rates (before error correction) are about one in 10,000 bits. Lifetime of the material is estimated to be at least 8 years.

1.3.3 Existing Optical and Magneto-optical Drives

Table 1.6 compares a number of existing optical and magneto-optical disk drives [45], [21]. The range of capacity and performance is narrower for optical disks than for magnetic tape drives. The low-end CD-ROM drives are priced at about 10% the cost of a high-end drive, have similar capacity per platter and data rates that are lower but on the same order of magnitude as the high-end drives.

Optical Jukebox	Drives	Platters	Total Capacity	Cost ($)	Drive Type
Pioneer	1	6	4 GB	$ 1,845	optical
Ricoh	1	5	3.25 GB	$ 8,990	optical
Hewlett-Packard 20LT	1	16	20 GB	$ 9,495	magneto-optical
Hewlett-Packard 40LT	1	32	41.6 GB	$ 22,500	magneto-optical
Plasmon FR50JM	2	50	50 GB	$ 36,500	magneto-optical
Hewlett-Packard 120LT	4	88	114 GB	$ 52,500	magneto-optical
Hewlett-Packard 200LT	4	144	187 GB	$ 67,700	magneto-optical
Plasmon FR495JM	4	495	495 GB	$129,500	magneto-optical

Table 1.7: *Comparison of several optical and magneto-optical robot devices. All the optical disk drives used in these robots are 5.25-inches in diameter.*

Media costs for optical disks vary widely. For example, a 650 megabyte magneto-optical platter for the Hewlett-Packard C1617C drive costs approximately $185.00 [46], or $0.28 per megabyte. Since CDROMs are read-only media, their cost may be determined largely by the cost of their contents; for the purposes of comparison, we use the cost of manufacture, about $2000 for 1000 platters containing 680 megabytes each [47], or $0.003 per megabyte.

1.3.4 Optical Disk Jukeboxes

Table 1.7 compares a number of robotic devices or jukeboxes that automatically load optical or magneto-optical platters into disk drives [45], [21], [48], [49]. The jukeboxes range from inexpensive autochangers for CD-ROMs to large jukeboxes with hundreds of platters. Like tape robots, large optical disk jukeboxes have a large number of platters and a small number of disk drives. Usually a single robot arm loads all the drives and robot operations take several seconds.

1.3.5 Trends in Optical Disk Technology

In predicting the future of optical disk drives, it is most important to compare their progress with that of magnetic disks. Magnetic disks are increasing in capacity and dropping in price per megabyte of storage at dramatic rates, currently over 50% per year [35]. Optical disks have not kept up with these dramatic improvements in magnetic disk technology. As a result, the cost per megabyte of magnetic disk drives is now comparable to the price per megabyte of many optical disk jukeboxes, approximately 50 cents per megabyte. Given a choice between an optical disk jukebox involving a robot arm and average response time of several seconds and a collection of magnetic disks with millisecond response times for the same price, the choice is highly likely to be in favor of the magnetic disk system. Thus, today optical disks stand in a precarious position compared to magnetic disks.

CD-ROM devices appear to be a possible exception because of their popularity and convenience as an inexpensive means of distributing software and data. In addition, CD-ROM technology will be driven by the high quantities associated with consumer devices. Therefore, they should drop in price faster than other optical disk technologies and may successfully compete with magnetic disks in the future.

1.4 Other Tertiary Storage Technologies

There are a number of tertiary storage technologies still in the laboratory or in a small number of products that may someday compete with magnetic tape and optical disk. In this section we discuss two: holographic storage and optical tape.

1.4.1 Holographic Storage

Perhaps the highest-density storage medium will be the holographic store, which stores data in a pattern of electronic charges in a crystallite material [50]. Parish [51] describes one such product. The storage material itself is a photorefractive crystalline structure of strontium barium niobate doped with cerium. Each crystal is divided into a series of two-dimensional "pages" or patterns of light. A collection of pages in one crystallite is called a "stack." Figure 1.7 shows a simplified diagram of a holographic store.

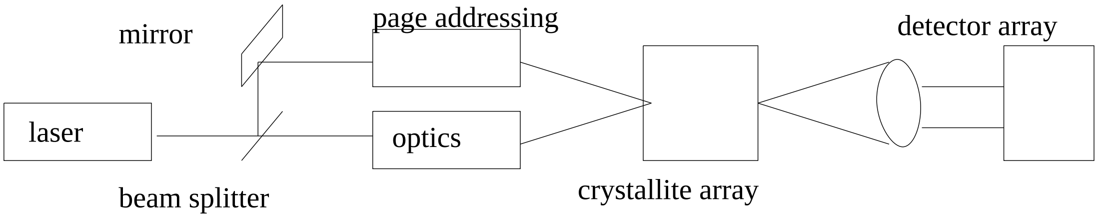


Figure 1.7: *Simplified diagram of holographic store.*

Data are written in the holostore by first creating a pattern of dark and light spots corresponding to the digital data. This pattern is superimposed onto the laser beam. As this data pattern is sent through the crystallite array, an interference grating is written in the photorefractive material. That interference grating is converted to a stored electronic charge pattern that modifies the crystallite, storing an image of the bit pattern from the data beam.

Data are read from the holostore by shining a reference beam through the crystallite at a particular angle corresponding to the address of a page of data. The image of the bit pattern written to the page is reconstructed by focusing the bit pattern onto a detector array that reproduces the original digital signals.

Parish projects the following achievable targets: one megabit per page, 100 pages per stack or crystallite and 1000 stacks per module of size 10x10x0.5 centimeters. A module of that size would have capacity over 100 gigabytes. Page read time is predicted as 100 nanoseconds, page write time 10 microseconds, and sustained transfer rate would be 1 terabyte per second. The projected cost of the storage is less than twice the cost per bit of magnetic disk.

1.4.2 Optical Tape

Optical tape technology [52] is similar to that used in optical disks except that the optical coatings are deposited onto flexible material. On top of a polyester base film like that used in magnetic tapes, ICI Imagedata, one manufacturer of optical tape, deposits a proprietary smoothing layer, a metallic reflective layer, an infrared-absorbing dye-polymer layer that is tuned to the wavelength of the read/write laser, and finally a protective overcoat [53], [54]. A backing coat is also applied to the reverse side of the tape.

During the write operation, an 830 nanometer laser with a power of approximately 10 milliwatts is used to melt instantaneously small sections of the dye-polymer layer. Surface tension forms a physical pit with well-defined edges and a depth of about half the layer thickness; the overcoat remains above the pit as a "protective dome." While the unwritten tape has a reflectivity of about 50%, the pit has lower reflectivity. A low-power laser beam is used during reading to detect changes in reflectivity and determine the value of the original data. Because data are stored as physical pits in the tape, optical tape can only be written once.

Reliability studies on the media suggest lifetimes of over 15 years. The medium does not degrade with exposure to UV radiation and is resistant to hydrolysis. Because the optical head has a non-contact interface to the medium, there is no head wear. Tapes can undergo tens of thousands of passes before unacceptable media wear occurs.

Two companies, Creo [55] [56] and LaserTape [57], have announced products using this optical tape medium. Creo ships a tape drive that uses an optical tape reel that is 12 inches in diameter, 35 millimeters wide and holds 1 terabyte of data.

1.5 Summary

In this chapter, we have given a brief overview of the workings of magnetic tape and optical disk drives and robots. We surveyed existing products and discussed technology trends. Magnetic tape offers the highest capacity for the lowest cost but suffers from the longest access times of any technology. Optical disk has shorter access times than magnetic tape but is more expensive. Trends show that magnetic tape systems should continue to have dramatic increases in both capacity and transfer rate. Optical disk systems face

Tape or Platter	Capacity per Unit	Media Cost	Units for 1 Terabyte	Cost for 1 Terabyte
8mm helical tape	5 GBytes	$0.008	200	$8389
D-2 helical tape	25 GBytes	$0.0018	41	$1887
3480 linear tape	800 MBytes	$0.013	1310	$13,631
HP magneto-optical disk	650 MBytes	$0.28	1613	$293,601
CDROM optical disk	680 MBytes	$0.003	1542	$3146

Table 1.8: Comparison of media costs for storing a terabyte in different tertiary storage technologies.

increasing competition from magnetic disks, which now match optical disk libraries in cost per megabyte. Finally, we described two emerging tertiary storage systems, holographic storage and optical tape, that may someday compete with magnetic tape and optical disk.

We end this chapter by summarizing the cost and performance characteristics of various technologies. Table 1.8 compares the media costs for a tertiary storage system that stores 1 terabyte of data. The high capacity D-2 helical tape has by far the lowest cost per megabyte. The low manufacturing costs of CD-ROM make it a close second. The highest by far is the magneto-optical disk, which is an order of magnitude more expensive than any other technology.

Table 1.9 shows the total system cost for purchasing various robots to store 1 terabyte of data, and shows the maximum bandwidth of the resulting systems. Note that these bandwidth numbers are normalized for comparison purposes, so that in the case of the Ampex robot, since we use only a small fraction of the total capacity of the robot, we likewise attribute only a portion of the total bandwidth of one robot to the system. Using this normalization, we see that among the magnetic tape systems , the cost of storing a terabyte is approximately the same in each case; the largest library is slightly less expensive than the others. The normalized bandwidth of the systems is highest for the stacker tape system, reflecting its high ratio of tapes to tape drives. The next best bandwidth belongs to the Ampex library, reflecting its high per-drive bandwidth. For the optical disk systems,

Robot	Robot Capacity	Robots to Store 1 Terabyte	Cost to Store 1 Terabyte	Max. Bandwidth of Collection of Robots
EXB10i Tape Stacker	50 GBytes	20.5	$180,359	10.25 MB/sec
EXB120 Tape Library	580 GBytes	1.77	$177,000	3.54 MB/sec
Ampex Tape Library	6400 GBytes	0.16	$160,000	9.6 MB/sec
Pioneer CD Changer	4 GBytes	256	$472,320	76.8 MB/sec
HP120LT M.O. Jukebox	120 GBytes	8.5	$446,250	54.4 MB/sec

Table 1.9: Compares system cost to store 1 terabyte of data and the resulting maximum bandwidth for a collection of robots.

the system cost for purchasing the robots is substantially higher than for the magnetic tape systems; however, the aggregate bandwidth of the systems is substantially higher as well.

Bibliography

[1] Tasuya Adachi, Dazuo Arai, Kinji Kawamoto, Hideshi Taki, and Kazuhiro Murase. A Fast Random Accessing Scheme for R-DAT. *IEEE Transactions on Consumer Electronics*, CE-33(3), August 1987.

[2] J. P. Harris, W. B. Phillips, J. F. Wells, and W. D. Winger. Innovations in the Design of Magnetic Tape Subsystems. *IBM Journal of Research and Development*, 25(5), September 1981.

[3] Akihiko Hitomi and Tetsuya Taki. Servo Technology of R-DAT. *IEEE Transactions on Consumer Electronics*, CE-32(3), August 1986.

[4] Fukuzo Itoh, Haruo Shiba, Masashi Hayama, and Takateru Satoh. Magnetic Tape and Cartridge of R-DAT. *IEEE Transactions on Consumer Electronics*, CE-32(3), August 1986.

[5] Masao Kawagishi and Takashi Niwa. Characterstics and Performance of Metal Particle Tape in a Broadcast Video Recorder. Video Recorder Division, Matsushita Electric Industrial Company Ltd., Kadoma City, Osaka, Japan.

[6] T. Kitamoto, M. Nakamura, and S. Takayama. Magnetic Recording Tapes for Video Recording. Fuji Photo Film Co., Ltd.; SMPTE Private Committee Document.

[7] John C. Mallinson. Tutorial Review of Magnetic Recording. *Proceedings of the IEEE*, 64(2), February 1976.

[8] John C. Mallinson. Achievements in Rotary Head Magnetic Recording. *Proceedings of the IEEE*, 78(6):1004–1016, June 1990.

[9] C. Denis Mee and Eric D. Daniel, editors. *Magnetic Recording, Volume II: Computer Data Storage*. McGraw-Hill, New York, 1988.

[10] Michael P. Sharrock. Particulate Magnetic Recording Media: A Review. *IEEE Transactions on Magnetics*, 25(6), November 1989.

[11] Paul H. Siegel. Recording Codes for Digital Magnetic Storage. *IEEE Transactions on Magnetics*, 21(5), September 1985.

[12] Hiroshi Sugaya. Recent Advances in Video Tape Recording. *IEEE Transactions on Magnetics*, 14(5), September 1978.

[13] Roger W. Wood. Magnetic Recording Systems. *Proceedings of the IEEE*, 74(11), November 1986.

[14] Bharat Bhushan. *Tribology and Mechanics of Magnetic Storage Devices*. Springer-Verlag, New York, 1990.

[15] Klaus J. Peter and Dennis E. Speliotis. Recording and Wear Characteristics of 4 and 8 MM Helical Scan Tapes. Media Logic Inc., 310 South St., Plainville, MA 02762, Sept. 28 1992.

[16] Kelly J. Beavers. Helical Scan Recording Technology: 8mm Evolves. *SunTech Journal*, September/October 1990.

[17] Eugene C. Nagel. Digital Storage Technology: DAT Evolves. *SunTech Journal*, September/October 1990.

[18] Steven F. Small. 8mm and DAT Formats Compete for Tape Technology of Choice. *Computer Technology Review*, Spring 1990.

[19] Eng Tan and Bert Vermeulen. Digital Audio Tape for Data Storage. *IEEE Spectrum*, October 1989.

[20] The Future Today: DAT Technology Computer Peripherals. Background paper, Archive Corporation, 1650 Sunflower Avenue, Costa Mesa, CA 92626, 1989.

[21] Michael McCarthy and Barry Green. Buyers Guide Part 2: Mass Storage OEMs. *SunWorld*, pages 69–90, January 1994.

[22] Tracy Wood. D-1 Through DAT. In *Digest of Papers*. Ninth IEEE Symposium on Mass Storage Systems, October 1988.

[23] Staff of Digital Review Labs. DAT Takes Backup to the Future. *Digital Review*, June 18, 1990.

[24] Bert Vermeulen. Helical Scan and DAT–a Revolution in Computer Tape Technology. In *Systems Design and Networks Conference (SDNC)*, pages 79–86, May 1989.

[25] Gerry Schadegg. Optimizing Digital 8mm Drive Performance. In *Goddard Conference on Mass Storage Systems and Technologies*. NASA/Goddard Space Flight Center, September 1992.

[26] John Berbert, Ben Kobler, P. C. Hariharan, Jen-Jacques Bedet, and Alan M. Dwyer. Magnetic media. Review of magnetic media and recording systems from NASA Earth Observing System Data and Information System (EOSDIS) group at the NASA Goddard Space Flight Center, Greenbelt, MD, and Hughes STX Corporation, Greenbelt, MD., August 1993.

[27] Richard A. Baugh, David J. Bromley, and Bruce F. Spenner. Extremely Low Error Rate Digital Recording with a Helical Scan Recorder. *IEEE Transactions on Magnetics*, 22(5), September 1986.

[28] Toshiyuki Kitahara. On the Long Term Storage of Metal Tapes. Fuji Photo Film Co. Magnetic Recording Lab; FIAT/IFTA Technical Commission in Hilversum (The Netherlands), June 14th 1988.

[29] John C. Mallinson. Magnetic Tape Recording: Archival Considerations. In *Digest of Papers*. Tenth IEEE Symposium on Mass Storage Systems, May 1990.

[30] H. Goto, A. Asada, H. Chiba, T. Sampei, T. Noguchi, and M. Arakawa. A New Concept of Data/DAT System. *IEEE Transactions on Consumer Electronics*, 35(3), August 1989.

[31] C. Denis Mee and Eric D. Daniel, editors. *Magnetic Recording, Volume III: Video, Audio, and Instrumentation Recording*. McGraw-Hill, New York, 1988.

[32] S. B. Luitjens. Magnetic Recording Trends: Media Developments and Future (Video) Recording Systems. *IEEE Transactions on Magnetics*, 26(1), January 1990.

[33] Mark H. Kryder. Data Storage in 2000–Trends in Data Storage Technologies. *IEEE Transactions on Magnetics*, 25(6), November 1989.

[34] Fred Moore. Long-range Storage Perspectives. Technical report, Storage Technology Corporation, 2270 South 88th Street, Louisville, Colorado, 80028-4315, 1991.

[35] David A. Patterson. Trends in magnetic disk technologies., August 1994. Personal communication.

[36] Harry C. Hinz. Magnetic Tape Technology in the 1990s. In *Digest of Papers*. Tenth IEEE Symposium on Mass Storage Systems, May 1990.

[37] Metrum Information Storage, Denver, Colorado. RSS-600 Rotary Storage System product literature.

[38] David D. Larson, James R. Young, Thomas J. Studebaker, and Cynthia L. Kraybill. StorageTek 4400 Automated Cartridge System. In *Digest of Papers.* Eighth IEEE Symposium on Mass Storage Systems, May 1987.

[39] Robert A. Bartolini. Optical Recording: High-Density Information Storage and Retrieval. In *Proceedings of the IEEE*, pages 516–524, June 1982.

[40] Ichiro Yamada, Minoru Saito, Akinori Watanabe, and Kiyoshi Itao. Automated Optical Mass Storage Systems with 3-Beam Magneto-Optical Disk Drives. In *Digest of Papers.* Eleventh IEEE Symposium on Mass Storage Systems, October 1991.

[41] Glenn T. Sincerbox and James M. Zavislan, editors. *Selected Papers on Optical Storage.* SPIE–The International Society for Optical Engineering, Bellingham, Washington, 1992.

[42] Andy Poggio. CD-ROM Technical Summary: From Plastic Pits to 'Fantasia'. Available by ftp from ftp.apple.com, March 1988.

[43] W. H. Meiklejohn. Magnetooptics: A Thermomagnetic Recording Technology. In *Proceedings of the IEEE*, pages 581–592, November 1986.

[44] Peter Hansen and Heinrich Heitmann. Media for Erasable Magnetooptic Recording. *IEEE Transactions on Magnetics*, 25(6), November 1989.

[45] Maureen McKeon. A Sampling of Optical Disk Drives. *SunExpert*, pages 63–71, June 1992.

[46] Michael Jay Tucker. Optical Disks: The Fabulous Closet of Fibber McGee. *SunExpert*, pages 52–62, June 1992.

[47] alt.cd-rom newsgroup. Frequently asked questions about CD-ROMs. Available by ftp from ftp.apple.com, March 1994.

[48] Daniel R. Dauner, Raymond C. Sherman, Michael L. Chreistensen, Jennifer L. Methlie, and Jr. Leslie G. Christie. Mechanical Design of an Optical Disk Autochanger. *Hewlett-Packard Journal*, December 1990.

[49] Donald J. Stavely, Mark E. Wanger, and Kraig A. Proehl. A Rewritable Optical Disk Library System for Direct Access Secondary Storage. *Hewlett-Packard Journal*, December 1990.

[50] Steve Redfield and Jerry Willenbring. Holostore Technology for Higher Levels of Memory Hierarchy. In *Digest of Papers.* Eleventh IEEE Symposium on Mass Storage Systems, October 1991.

[51] Tom Parish. Crystal Clear Storage. *Byte*, pages 283–288, November 1990.

[52] Andrew J. G. Strandjord, Steven P. Webb, Donald J. Perettie, and Robert A. Cipriano. Flexible Storage Medium for Write-Once Optical Tape. In *Goddard Conference on Mass Storage Systems and Technologies.* NASA/Goddard Space Flight Center, September 1992.

[53] Andrew J. Ruddick. ICI Optical Data Stroage Tape – An Archival Mass Storage Media. In *Goddard Conference on Mass Storage Systems and Technologies.* NASA/Goddard Space Flight Center, September 1992.

[54] Andrew Ruddick and Joe Duffy. ICI's Optical Tape Offers Flexible Alternative to Rigid Optical Media. *Optical Memory News*, February 1991.

[55] Ken Spencer. Terabyte Optical Tape Recorder. In *Digest of Papers.* Ninth IEEE Symposium on Mass Storage Systems, October 1988.

[56] Ken Spencer. The 60-second Terabyte. *Canadian Research*, June 1992.

[57] Peter C. Boulay. The LaserTape 'DOTS': Next Digital Paper Product. Technical Report IDC Washington Publication No. W1220, IDC Washington, Inc., July 1990.

[58] Sam Coleman and Steve Miller, editors. *Mass Storage System Reference Model: Version 4*. IEEE Technical Committee on Mass Storage Systems and Technology, May 1990.

[59] Penelope Constanta-Fanourakis, Ken Kaczar, Gene Oleynik, Don Petravick, Margaret Votava, Vicky White, George Hockney, Steve Bracker, and Jussara M. De Miranda. Exabyte Helical Scan Devices at Fermilab. *IEEE Transactions on Nuclear Science*, 36(5), October 1989.

[60] E. I. Cohen, G. M. King, and J. T. Brady. Storage Hierarchies. *IBM Systems Journal*, 28(1), 1989.