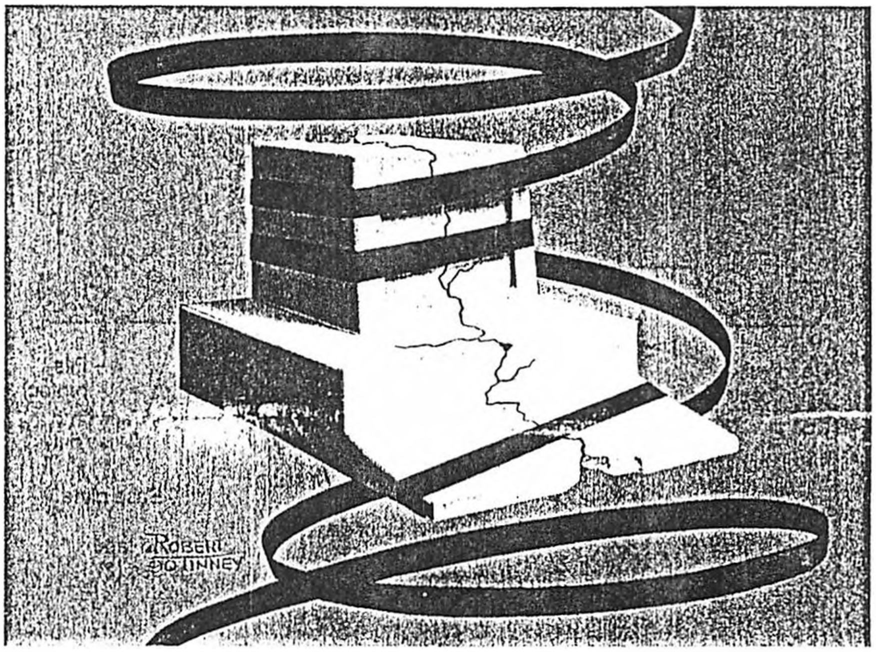
REELING IN THE DATA

Tape backup systems are no longer just for the wealthy or the overanxious



One day it will surely happen: Your system's hard disk drive will die of old age, mysteriously lose data, or succumb to power glitches or (worse yet) a fire. And, if you're like the majority of microcomputer users, you'll lose precious data because you didn't make a recent backup.

This situation is, of course, preventable if you're willing to spend a few minutes each day backing up your data, either to floppy disks (using a high-speed floppy disk backup utility) or to tape. In this installment of Under the Hood, I'll focus on the latter medium and describe how the most popular types of tape backup units work.

Why Tape?

Magnetic tape is ideal for situations in which you need to store large amounts of data without changing the medium—typically, unattended backup or the distribution of large programs or databases. The lowest-capacity tape drive you're likely to find will store 40 megabytes of data on a single cartridge. High-end helical-scan units from Exabyte store 5 gigabytes per cartridge and will probably increase to 8 gigabytes before the end of the year.

By contrast, a standard AT-type 5¼-inch floppy disk can hold only 1.2 MB of data; its 3½-inch counterpart can hold only 1.44 MB. Backing up a large hard disk drive with a utility like Fastback from Fifth Generation Systems can be a true test of one's disk-swapping patience.

Extra-high-density floppy disks—like those from Insite Peripherals and Brier Technology—are available, and you may find them useful as a backup medium. However, these disks will hold at most 40 MB, the same as the least capacious tape, and aren't widely available yet. Tape drives, on the other hand, are available from numerous manufacturers, and media for them are available at virtually all computer stores. And tape is generally less expensive, megabyte for megabyte, than floppy disks.

One area where tape is *not* superior to floppy disks is durability. Floppy disks, which have both a thicker substrate and a heavier magnetic coating, have a better chance of coming through extremes of temperature (e.g., fire) intact.

What about WORM (write once, read many times) drives and read/write optical disk drives? So far, the main disadvantage of these drives is cost: $4000 to $8000 per unit, plus $50 or more per disk. The advantage of such drives is that they're useful for other kinds of storage between backups.

Disks of any kind—magnetic or optical—will always be far superior to tape when random access is required. It's much easier and faster to step from track to track on a disk than to fast-forward and rewind a tape drive. But in backup applications, which primarily involve the transfer of long, continuous streams of data, tape truly comes into its own.

Varieties of Tape Formats

The four tape formats most often used for data storage are half-inch reel-to-reel (most often used on minicomputers and mainframes), quarter-inch cartridge, 8-mm cartridge, and 4-mm DAT (digital audio tape).

Of these four, half-inch reel-to-reel (often called nine-track) is the most mature technology. As the name suggests, data is laid down in nine parallel tracks at 800, 1600, 3200, or 6250 bits per inch. The tracks are read in parallel using a head with nine distinct elements, making data transfers extremely fast, even at relatively low tape speeds. A full

continued

ILLUSTRATION: ROBERT TINNEY © 1990

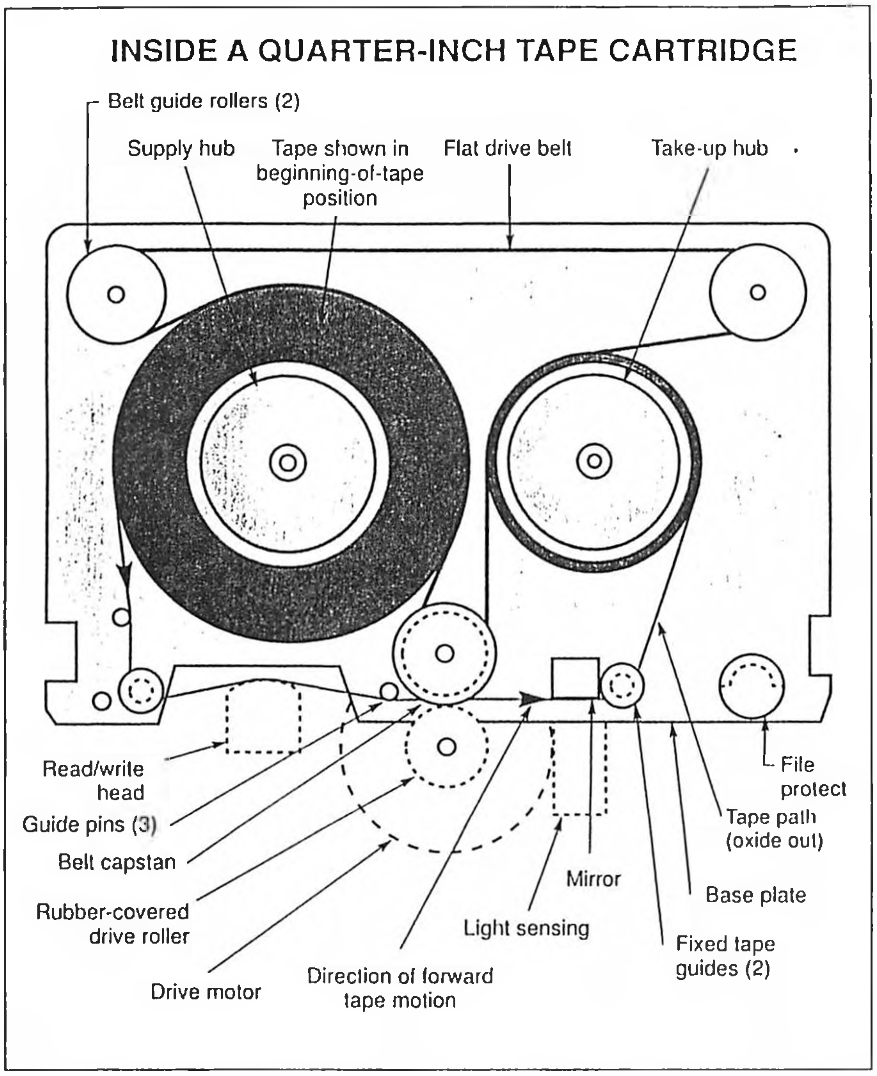


Figure 1: *Quarter-inch tape cartridges use a clever drive mechanism that moves the tape directly, eliminating the need for separate drives and clutches for the hubs.*

reel of tape recorded at the highest density typically holds about 145 MB.

Nine-track tape is time-tested and almost 100 percent interchangeable between virtually all the drives that handle it. It's frequently used to transport data between machines as well as to back up disks. But nine-track drives—especially good ones—aren't cheap, and the big reels of tape are awkward and not likely to be available at your local computer store. Also, while they offered more capacity than cartridge units until a few years ago, they're now falling behind. Perhaps this is the reason you don't see nine-track tape on many personal computers nowadays.

Clever Cartridges

The most common medium used by tape backup units is quarter-inch tape, which is enclosed in one of two sizes of cartridges: a DC2000-class minicartridge, which is about 2½ by 3 inches, or a DC600-class cartridge, which is the same shape but about 30 percent bigger in each dimension. (A new class of cartridge, DC9135, will look like a DC600 but will have 900-oersted tape, which can handle denser recording schemes, inside.) The exact specifications for the cartridges, originally defined by 3M, are now embodied in ANSI standards.

Each cartridge (see figure 1) contains two tape hubs, a flexible drive belt that passes partway around each hub, tape guides, and a small mirror assembly (used by the drive to detect holes at the beginning and end of each tape). A small door (not shown) protects the cavity where the read/write head enters the cartridge from dust and debris.

The mechanism inside a quarter-inch cartridge is clever and unique. Even though it's been around for more than a decade, you won't see anything like it in an audiocassette or a videocassette, probably for patent reasons. Note that the drive belt actually contacts the tape and moves with it partway around each hub, ensuring that the linear velocity of the tape is equal to that of the belt no matter how much tape is on each hub. The angular velocities of the hubs, which aren't directly driven, vary as necessary. This makes the drive extremely simple at the expense of a small amount of additional hardware in the cartridge.

All the drive needs to do to position the tape is rotate a pulley called the *belt capstan*; the pulley moves the belt, which in turn moves the tape. There's no need for a tape capstan, a pinch roller, or independent drive mechanisms and clutches for the hubs. (An ordinary Philips-type audiocassette deck, by contrast, must have all these things.)

Cartridges within a class may have different lengths. Many users have been pleasantly surprised to find extra-long tapes appearing on computer store shelves, letting them increase the capacity of an existing drive by up to 50 percent without buying new hardware. A QIC-40 drive (nominal capacity of 40 MB) can often record 60 MB on such tapes, and a QIC-80 drive (80 MB) can get 120 MB on one tape.

QIC and Easy

Standards that describe how to record on quarter-inch cartridges are published by the organization Quarter Inch Cartridge Drive Standards (311 East Cabrillo St., Santa Barbara, CA 93101). So far, the organization (QIC for short) has adopted 36 standards, specifying such things as interfaces between computers and tape drives, tape formats, recording-head properties, error-correction codes, data-compression algorithms (see the text box "Data Compression Doubles Tape Capacity" at right), and SCSI command sets for tape drives (see the table for a partial list; if you'd like a complete list, contact QIC).

Some of these standards have become pervasive enough that makers of drives for other tape formats have adopted them. For instance, most 4-mm DAT drives used for tape backup implement the set of SCSI commands specified in the QIC-104 standard and will probably move to QIC-121 (a SCSI-2 command set) in the future.

As a result of this standardization

continued

Data Compression Doubles Tape Capacity

Data compression is a highly effective way of increasing the capacity of a tape drive. Unfortunately, most software algorithms with good compression ratios can process at most 64K bytes per second—and would therefore constitute a major bottleneck in a tape backup system.

To solve this problem, a company called Stac Electronics (Pasadena, CA) has developed a data-compression/decompression chip with an average throughput of 750K bytes per second and an average compression ratio of 2-to-1. The chip, which uses a modified Ziv-Lempel algorithm, achieves this ratio using only 16K bytes of RAM, only 2K bytes of which is used to hold the string table for the compression algorithm.

Even though the company refuses to divulge exactly how the proprietary algorithm works, the organization Quarter Inch Cartridge Drive Standards (Santa Barbara, CA) has adopted it as the QIC-122 standard. Should vendors wish to implement the algorithm in software only, rather than buying a $50 chip to put in every drive, Stac Electronics will license object code for the cool sum of $25,000.

Because other manufacturers may wish to go their own way rather than embrace one manufacturer's proprietary solution, the QIC committee has established a registry of "data algorithm identifiers" that can be written on a tape to show how the data is encoded. A tape backup system can use this code to apply the correct algorithm—if it knows it—to the data on the tape while reading it.

SOME KEY QIC STANDARDS

A partial list of the 36 standards adopted by Quarter Inch Cartridge Drive Standards.

Interfaces

QIC-02	Quarter-inch cartridge tape drive intelligent interface
QIC-36	Quarter-inch cartridge tape drive basic interface
QIC-104	Implementation of SCSI for QIC-compatible sequential storage devices
QIC-121	Implementation of SCSI-2 for QIC-compatible sequential storage devices

Commonly used tape formats

QIC-24	Serial-recorded magnetic tape cartridge for information interchange (60 MB)
QIC-40	Flexible-disk-drive-controller-compatible recording format for information interchange (40 MB)
QIC-80	Flexible-disk-drive-controller-compatible recording format for information interchange (80 MB)
QIC-120	Serial-recorded magnetic tape cartridge for information interchange (125 MB)
QIC-150	Serial-recorded magnetic tape cartridge for information interchange (150 MB)
QIC-525	Serial-recorded magnetic tape cartridge for information interchange (525 MB)
QIC-1350	Serial-recorded magnetic tape cartridge for information interchange (1.35 GB)

Data compression

QIC-122	Data-compression format for quarter-inch data cartridge tape drives
QIC-123	Registry of data algorithm identifiers for quarter-inch cartridge tape drives

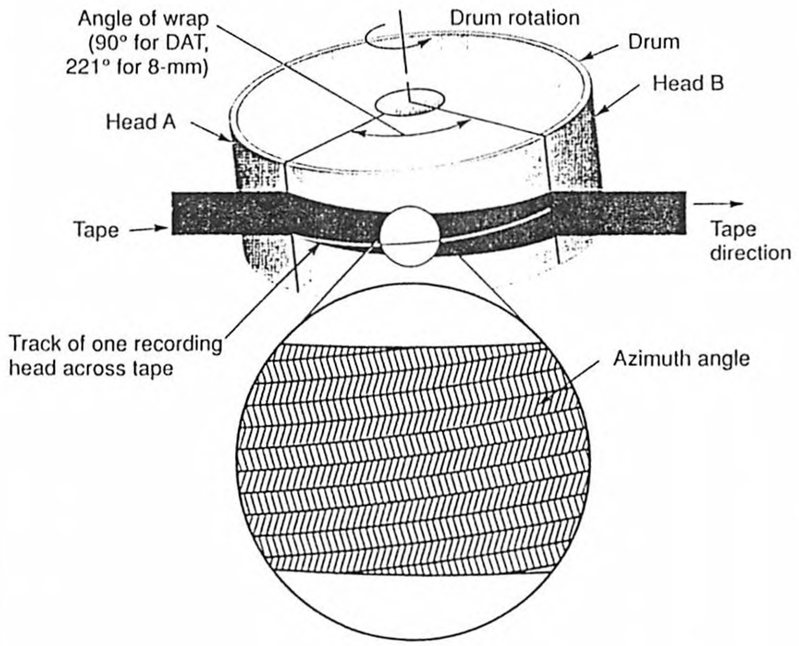


Figure 2: *Slanted tracks maximize use of the tape's surface area; alternate azimuth angles minimize cross talk.*

effort, tapes recorded on one manufacturer's drive can be read in another, provided that the industry-standard format is used. (Almost all manufacturers have their own enhanced, proprietary, and incompatible formats in addition to the standard ones.) In December 1989, an independent laboratory called Pericomp (Natick, MA) certified QIC-40 cartridge drives from three manufacturers—Archive, Mountain Computer, and Colorado Memory Systems—as fully compatible. Other manufacturers (such as Alloy Computer Products and Everex Systems) are expected to follow, and Pericomp has begun to certify QIC-80 drives as well.

"Floppy Tape"

The QIC-40 and QIC-80 standards are sometimes called the "floppy tape" standards, because they're designed to use the computer's existing floppy disk drive controller to read and write tapes. This is less expensive than providing a custom controller and may save backplane slots, which is especially important for users with "baby" motherboards. However, this approach has some drawbacks.

First, since the maximum bit clock rate of many floppy disk drive controllers is either 250,000 bps (as on the IBM XT) or 500,000 bps (as on the IBM AT), there's an upper bound on how fast data can be stored and retrieved without a special controller.

Second, since many computers (e.g., the AT) were designed to have at most two floppy disk drives, they don't provide select lines for any more devices. So, in systems that have two floppy disk drives, clever software and hardware tricks are required to let the tape drive share the interface. One commonly used technique is to tell the floppy disk drive controller to select *no* drive; circuitry on the drive or an interface card detects this condition and activates the tape. Pulses on the floppy disk drive controller's step line are often used to send commands to the tape drive; the drive counts the pulses to determine the command.

If, like me, you have an IBM AT compatible that has high- and low-density floppy disk drives and no more cutouts in the chassis, you may suffer from another problem when buying one of these normally inexpensive tape backup units: sticker shock. An external drive can cost as much as $400 more than an internal drive due to the costs of a case, power supply, and FCC-certified cable.

continued

Still, if you find yourself in this situation, resist the temptation to write (and potentially ruin) 360K-byte floppy disks in the high-density drive. Instead, spend the extra money or consider moving to a high-density 3½-inch drive, which will not have compatibility problems with low-density 3½-inch disks.

In the QIC-40 and QIC-80 standards, the floppy disk drive controller sees blocks of data on the tape as if they were disk sectors 1K byte in length. The standards specify how the tracks of the tape are mapped to the sectors, tracks, and sides of the imaginary floppy disk. The tracks of the tape are read and written sequentially in a serpentine fashion. The cyclic redundancy checks that are generated by the floppy disk drive controller are used for error detection, and a Reed-Solomon code is used to correct errors where possible.

Higher-numbered QIC formats move away from using the floppy disk drive controller to read and write and are usually implemented within SCSI-based tape drives. This results in full SCSI speed and greater interchangeability of drives—but not necessarily tapes—among vendors.

Triple Helix

At this writing, the only available backup systems that can store more than a gigabyte per cartridge use helical-scan technology (i.e., the same kind that is used in VCRs and DAT recorders) to achieve this and higher capacities.

Figure 2 shows how a helical-scan drive reads and writes data. The tape wraps around a spinning drum whose axis is tilted at a small angle (typically 5 to 6 degrees) to the tape. The heads, which are mounted on the drum, traverse the tape in long, diagonal stripes (actually, sections of a helix). Each section is about nine times the width of the tape in length.

This is the point where the 8-mm and 4-mm designs diverge, however. Exabyte's 8-mm helical-scan technology, which is based on an 8-mm analog VCR mechanism supplied by Sony, uses three heads (a servo head, a read head, and a read-after-write head) on the drum, plus a separate erase head that erases the full width of the tape at once.

Tracks are written individually, and they contain 8K bytes of data each. The tape wraps more than halfway around the head—211 degrees, to be exact. The current storage capacity of the Exabyte drive is 2.3 gigabytes per cartridge, with an increase to 5 gigabytes expected sometime this year.

The DDS (Digital Data Storage) format, developed and licensed by Hewlett-Packard and Sony, currently appears to be the dominant data format for 4-mm DAT. DDS uses the innards of a DAT deck, with four heads on the drum: two write and two read-after-write. Tracks are written in pairs (called *frames*) that actually overlap slightly.

The heads can distinguish the tracks, however, because they are recorded at different *azimuth angles* (i.e, with heads that are tilted at different angles relative to the direction of the track). Each frame contains 8K bytes, and the tape wraps only 90 degrees around the drum, which the designers claim helps to reduce tape wear. DDS tapes currently hold 1.3 gigabytes each.

Another format for DAT, called Data/DAT, is now under development by another consortium. While it is similar to DDS in many ways, it adds one important

continued

feature: the ability to update data in place. Since this allows partial updates of a backup tape, it could potentially save a lot of time. It also lets the system use a tape drive as a block-oriented, random-access medium, which is an especially useful feature in systems where inexpensive but not superfast mass storage is required.

However, proponents of DDS argue that using a tape as a random-access device—starting and stopping it rather than streaming data smoothly—may promote tape wear. They also claim that writing a driver for such a device is likely to be more difficult than for a streaming device. Is it worth it? The answer is unclear today, but it may make a difference to some applications in the long run.

All the helical-scan drives offer read-after-write data verification and error correction, and they can correct an error detected during a write without breaking stride.

Can QIC Catch Up?

Makers of helical-scan drives claim that they can pack data more densely than on any equivalent "linear" drive because a helical-scan drive doesn't need a multitrack head, or a single one that must shift its position from track to track, to cover almost the entire surface area of the tape. (Doing either, they claim, would introduce mechanical tolerances that limit data density.)

They also claim that the combined motion of the tape and head, which generates an effective head-to-tape velocity of roughly 150 inches per second, provides higher data transfer rates than can be achieved with a linear drive. A linear drive couldn't move the tape at that speed without generating enough friction to melt it!

Meanwhile, QIC adherents are claiming that QIC-1350, a new development standard based on special heads, 900-oersted tape, and data compression, will soon let quarter-inch cartridges store 1.35 gigabytes at data transfer rates of 600K bytes per second. However, the point may soon be moot. Nearly all the major manufacturers of QIC drives have licensed the DDS technology, and many are expected to release DDS drives in the near future.

If they do so, the cost advantages of a mass-market mechanism and the expense of supporting yet another standard may cause them to abandon quarter-inch cartridges for their high-end products. But even if this occurs, quarter-inch cartridges are likely to remain dominant in lower-capacity applications for a long time to come.

The good news for all of us, however, is that mass production and standardization are forcing down prices—at least on the low end of the tape backup spectrum. It's now possible to purchase an internal QIC-40 tape backup unit for around $300 from any of several mail-order dealers. Hopefully, this will be enough to convince most users that there's no sense leaving one's data in danger. ■

ACKNOWLEDGMENT

Special thanks to Dan Stromska of Mountain Computer and Mike Barton of HP Ltd. for help in preparing this article.

L. Brett Glass is a freelance programmer, author, and hardware designer residing in Palo Alto, California. He can be reached on BIX as "glass."

Your questions and comments are welcome. Write to: Editor, BYTE, One Phoenix Mill Lane, Peterborough, NH 03458.