



Upon completing this chapter, you will be able to

- Participate in discussions with vendors, work colleagues, and other professionals about the advantages and disadvantages of disk drive implementations being used in or planned for storage networks
- Analyze and compare the specifications and capabilities of different disk drives and their suitability for different types of applications
- Discuss and plan deployments of tape drives for storage networks

## Storage Devices

---

Storage devices perform the fundamental function of storage networks, which is the reading and writing of data stored on nonvolatile media. They operate in the microscopic realm, combining advanced magnetic physics, chemistry, and electronics. Demands to increase capacity and performance continue to force the industry to conduct fundamental scientific research on the microscopic characteristics of materials.

Storage devices are the building blocks of storage in disk subsystems as well as being used as standalone products in server systems. This chapter mostly examines disk drive technology as the device that is used far more than any other. Tape drives are looked at at the end of the chapter.

### Disk Drives

Several types of storage devices are used in storage networking, but by far the most important is the disk drive. If Harry Truman worked in the storage network industry, he would have said the buck stops at the disk drive. By now we have almost come to take disk drives for granted as ubiquitous gadgets that are readily available at volume discounts. The fact is, disk drive technology is as impressive as any other technology in all of IT, with amazing capabilities and physical characteristics.

In the sections that follow, we'll examine the various subassemblies that make up a disk drive, discuss the strengths and limitations of these amazing machines, and point out what they mean for storage network applications.

### Major Parts of a Disk Drive

Disk drives are constructed from several highly specialized parts and subassemblies designed to optimally perform a very narrowly defined function within the disk drive. These components are

- Disk platters
- Read and write heads
- Read/write channel

- Arms and actuators
- Drive spindle motor and servo control electronics
- Buffer memory
- Disk controller

Now we'll discuss each of these subassemblies briefly.

## Disk Platters

The physical media where data is stored in a disk drive is called a *platter*. Disk platters are rigid, thin circles that spin under the power of the drive spindle motor. Platters are built out of three basic layers:

- The substrate, which gives the platter its rigid form
- The magnetic layer, where data is stored
- A protective overcoat layer that helps minimize damage to the disk drive from microscopically sized dust particles

The three different layers within a disk platter are illustrated in Figure 4-1, which shows both the top and bottom sides of a platter.

**Figure 4-1** *Material Layers in a Disk Media Platter*



Substrates are made from a variety of materials, including aluminum/magnesium alloys, glass, and ceramic materials. Considering the microscopic nature of disk recording and how close the heads are to the surface, they must be amazingly flat and relatively inelastic to thermal expansion and contraction. In addition, they have to be almost completely uniform in density and free from material defects that could result in balance imperfections, which cause vibration and friction (heat) problems when spinning at high revolutions per minute (rpm).

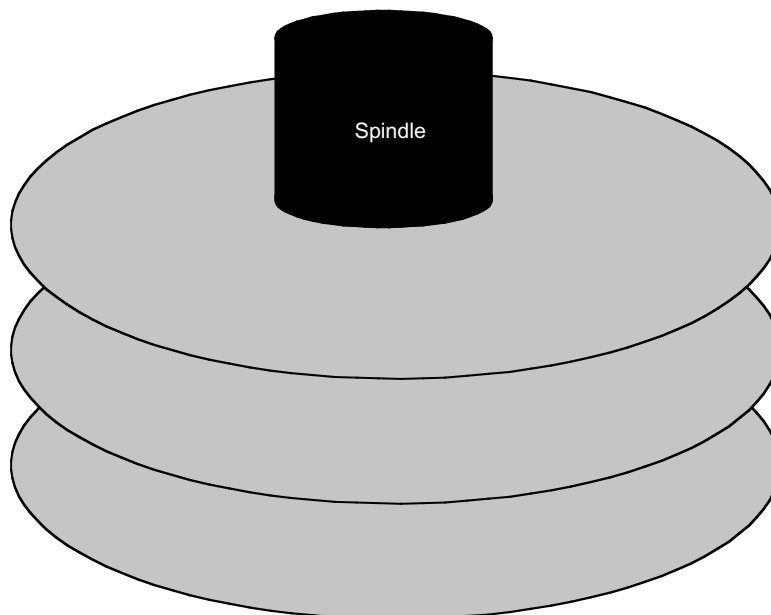
The magnetic layer in most disk drives today uses thin film technology, which is very smooth and only a few millionths of an inch in thickness. The thin film layer is made by spraying vapor molecules of the magnetic materials on the surface of the substrate. The magnetic characteristics of the magnetic materials give the platter its areal density—the measurement of how many bits can be written per square inch.

The protective overcoat layer provides protection from microscopic elements such as dust and water vapor, as well as from disk head crashes. Considering the physics involved in high-speed disk drives, this coating is necessarily thin and provides, at best, light-duty protection. The best way to protect disk drive platters is to operate them in clean, dust-free, temperature-controlled environments.

Storage capacity of a single platter varies from drive to drive, but recent developments in commercially available products have resulted in platter capacities in excess of 100 GB per platter.

Disk drives are usually made by arranging multiple platters on top of each other in a stack where the platters are separated by spacers to allow the disk arms and heads to access both sides of the platters, as shown in Figure 4-2.

**Figure 4-2** *Disk Platters Connected to the Spindle Motor in a Stack*



## Read and Write Heads

The recording heads used for transmitting data to and from the platter are called *read and write heads*. Read/write heads are responsible for recording and playing back data stored on the magnetic layer of disk platters. When writing, they induce magnetic signals to be imprinted on the magnetic molecules in the media, and when reading, they detect the presence of those signals.

The performance and capacity characteristics of disk drives depend heavily on the technology used in the heads. Disk heads in most drives today implement giant magnetoresistive (GMR) technology, which uses the detection of resistance variances within the magnetic layer to read data. GMR recording is based on writing very low-strength signals to accommodate high areal density. This also impacts the height at which the heads “fly” over the platter.

The distance between the platter and the heads is called the *flying height*, or head gap, and is measured at approximately 15 nanometers in most drives today. This is much smaller than the diameter of most microscopic dust particles. Considering that head gap tolerances are so incredibly close, it is obviously a good idea to provide a clean and stable environment for the tens, hundreds, or thousands of disk drives that are running in a server room or data center. Disk drives can run in a wide variety of environments, but the reliability numbers improve with the air quality: in other words, relatively cool and free from humidity and airborne contaminants.

The reference to “flying” with disk heads comes from the aerodynamic physics at work in disk drives: air movement caused by the rapidly spinning platters passes over the heads, providing lift to the heads in much the same way airplane wings are lifted by the difference in air pressure above and below them.

## Read/Write Channel

While we tend to think about data purely in the digital realm, the physical recording is an analog signal. Somehow, the 0s and 1s of digital logic have to be converted to something that makes an impression on magnetic media. In other words, data on disk does not resemble written language at all but is expressed by the pattern of a magnetic signal on moving media. The *read/write channel* is the disk drive subassembly that provides a specialized digital/analog conversion.

The read/write channel is implemented in small high-speed integrated circuits that utilize sophisticated signal processing techniques and signal amplifiers. The magnetoresistive phenomenon that is detected by the read heads is very faint and requires significant amplification. Readers might find it interesting to ponder how data read from disk is not actually based on detecting the magnetic signal that was written to media. Instead, it is done by detecting minute differences in the electrical *resistance* of the media, caused by the presence of different magnetic signals. Amazingly, the resistance is somehow detected by a microscopically thin head that does not make contact with the media but floats over it at very high speeds.

## Arms and Actuators

The read and write heads have to be precisely positioned over specific tracks. As heads are very small, they are connected to disk arms that are thin, rigid, triangular pieces of lightweight alloys. Like everything else inside a disk drive, the disk arms are made with microscopic precision so that the read/write heads can be precisely positioned next to the platters quickly and accurately.

The disk arms are connected at the base to the drive actuator, which is responsible for positioning the arms. The actuator's movements are controlled by voice-coil drivers; the name is derived from voice coil technology used to make audio speakers. Considering that some speakers have to vibrate at very high frequencies to reproduce sounds, it's easy to see how disk actuators can be designed with voice coils to move very quickly. The clicking sounds you sometimes hear in a disk drive are the sounds of the actuator being moved back and forth.

## Drive Spindle Motor and Servo Control Electronics

The drive platters rotate under power of the drive spindle motor, which is designed to maintain constant speeds with minimal vibration over long periods of time, sometimes measured in the tens of thousands of hours.

Most drive failures are related to motor failures. This is not to say the motors are poorly designed or designed to fail, because they clearly are not. However, they are always moving toward higher speeds with less power consumption and less noise, and the tolerances are thin.

The actual spindle that the platters connect to is directly fixed to the motor's drive shaft. The spindle looks a bit like the inner core of some old 45-rpm record players, except the platters do not drop or slide over the core, but are fixed in place. Separator rings are used to space the platters precisely so their surfaces can be traversed by the disk arms and heads accurately.

Among the many parts of a disk drive, the bearings in the motor see constant wear and tear. While many other things can make a disk drive fail, it is inevitable that the bearings will eventually wear out.

The speed of the spindle motor must be constantly monitored to make sure it remains consistent hour after hour, day after day, month after month. The type of technology used to maintain constant speed is called a *servo-controlled closed loop*, and it is used for many different applications to fine-tune automated systems. Disk drives are designed with sophisticated feedback control circuits that detect minute speed variations in the rotating platter by reading tracking and timing data on the disk. If the speed varies too far one way or another, the servo feedback circuit slightly changes the voltage supplied to the spindle motor to counteract the change.

## Buffer Memory

The mechanical nature of reading and writing data on rotating platters limits the performance of disk drives to approximately three orders of magnitude (1000 times) less than the performance of data transfers to memory chips. For that reason, disk drives have internal buffer memory to accelerate data transmissions between the drive and the storage controller using it.

Buffer memory might not have a significant performance impact for a single disk drive system, such as a desktop or laptop system, but buffer memory can make a big difference in storage subsystems that support high-throughput applications. When multiple drives are assembled together in an array, the controller, such as a subsystem controller, can overlap I/Os across multiple drives, using buffer memory transfers whenever possible. The drive can make internal transfers of data between buffer memory and its media platters while the subsystem controller is working with another drive. In general, buffer memory in disk drives can improve I/O performance for applications that read and write small chunks of randomly accessed data. Alternatively, streaming applications with large files stored in contiguous storage locations do not realize many benefits from buffer memory.

Buffer memory sizes have increased over time, although not at the same rate as the areal density of platters. Today, disk drives typically have buffer capacities between 2 MB and 16 MB.

## Disk Controller

Disk drives all have internal target controllers that respond to commands from host or subsystem initiators. In addition to interoperating with the external initiator, the storage controller in a disk drive is responsible for executing the command within the drive. The software component of a disk drive controller is referred to as firmware and is typically stored in e-prom chips in the drive's circuit board.

Processor chips used as disk drive controllers are constantly being improved with faster cores and more memory. Strangely enough, one of the challenges for disk drive manufacturers is how to make the best use of the additional intelligence at their disposal. It's a tougher question than it first appears because disk drives have traditionally been used as relatively stupid slave devices responding to I/O requests from host and subsystem controllers. Storage applications such as NAS can be added to disk drives, but that creates competition between the disk drive manufacturers and their customers—the system and subsystem vendors. So far, the disk drive manufacturers have been at the wrong end of the pecking order and their attempts to integrate higher-level storage functions in the drives have mostly failed.

Instead, what has been successful is the use of processor intelligence to increase reliability and ease of use. If you stop to consider how much easier it is to install disk drives today than it was ten years ago, the improvement has been remarkable. When you think about these improvements in terms of installed capacity, the advancements have been truly incredible. For example, the reliability provided by a handful of today's high-capacity disk drives used together to form a terabyte (TB) of storage is far, far better than the reliability of the approximately 500 disk drives that were needed ten years ago to build the same terabyte of capacity. Obviously, improvements in areal density make this possible, but a significant amount of work has also been done in the drive's internal controllers.

## Data Structures on Disk Drives

The mechanisms of disk drive technology are only half of the story; the other half is the way data is structured on the disk. There is no way to plan for optimal storage configurations without understanding how data is structured on the surface of disk drive platters. This section discusses the following data structures used in disk drives:

- Tracks, sectors, and cylinders
- Disk partitions
- Logical block addressing
- Geometry of disk drives and zoned-bit recording

### Tracks, Sectors, and Cylinders

Disk platters are formatted in a system of concentric circles, or rings, called *tracks*. Within each track are sectors, which subdivide the circle into a system of arcs, each formatted to hold the same amount of data—typically 512 bytes. Once upon a time, the block size of file systems was coupled with the sector size of a disk. Today the block size of a file system can range considerably, but it is usually some multiple of 512 bytes.

Cylinders are the system of identical tracks on multiple platters within the drive. The multiple arms of a drive move together in lockstep, positioning the heads in the same relative location on all platters simultaneously.

The complete system of cylinders, tracks, and sectors is shown in Figure 4-3.

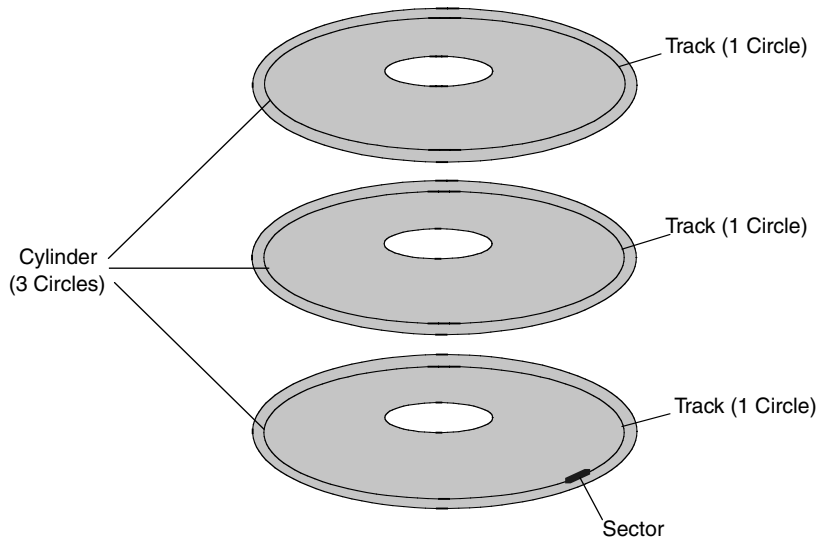
### Disk Partitions

Disk partitions divide the capacity of physical disk drives into logical containers. A disk drive can have one or more partitions, providing a way for users to flexibly create different virtual disks that can be used for different purposes.

For instance, a system could have different partitions to reserve storage capacity for different users of the system or for different applications. A common reason for using multiple partitions is to store data for operating systems or file systems. Machines that are capable of running two different operating systems, such as Linux and Windows, could have their respective data on different disk partitions.

Disk partitions are created as a contiguous collection of tracks and cylinders. Visually, you can imagine partitions looking like the concentric rings of an archery target with the bull's-eye being replaced by the disk motor's spindle. Partitions are established starting at the outer edge of the platters and working toward the center. For instance, if a disk has three partitions, numbered 0, 1, and 2, partition 0 would be on the outside and partition 2 would be closest to the center.



**Figure 4-3** *Cylinders, Tracks, and Sectors in a Disk Drive*

## Logical Block Addressing

While the internal system of cylinders, tracks, and sectors is interesting, it is also not used much anymore by the systems and subsystems that use disk drives. Cylinder, track, and sector addresses have been replaced by a method called logical block addressing (LBA), which makes disks much easier to work with by presenting a single flat address space. To a large degree, logical block addressing facilitates the flexibility of storage networks by allowing many different types of disk drives to be integrated more easily in a large heterogeneous storage environment.

With logical block addressing, the disk drive controller maintains the complete mapping of the location of all tracks, sectors, and blocks in the disk drive. There is no way for an external entity like an operating system or subsystem controller to know which sector its data is being placed in by the disk drive. At first glance this might seem risky—letting a tiny chip in a disk drive be responsible for such an important function. But, in fact, it increases reliability by allowing the disk drive to remap sectors that have failed or might be headed in that direction.

Considering the areal density and the microscopic nature of disk recording, there are always going to be bad sectors on any disk drive manufactured. Disk manufacturers compensate for this by reserving spare sectors for remapping other sectors that go bad. Because manufacturers anticipate the need for spare sectors, the physical capacity of a disk drive always exceeds the logical, usable capacity. Reserving spare sectors for remapping bad sectors is an important, reliability-boosting by-product of LBA technology. Disk drives can be manufactured with spare sectors placed throughout the platter's surface that minimize the performance hit of seeking to remapped sectors.

## Geometry of Disk Drives and Zoned-Bit Recording

There is no way to escape radial geometry when working with disk drives. One of the more interesting aspects of this radial geometry is that the amount of recording material in a track increases as you move away from the center of the disk platter. Disk drive tracks can be thought of as media rings having a circumference that is determined by the mathematical expression  $2\pi r$ , where  $r$  is the radius for the track. The amount of recording material in a track is determined by radial length. This means that the outermost tracks can hold more data than the inside tracks. In fact, they can hold a lot more data than inside tracks.

To take advantage of this geometry, disk drive designers developed zoned-bit recording, which places more sectors inside tracks as the radius increases. The general idea is to segment the drive into “sector/track density” zones, where the tracks within that zone all have the same number of sectors. The outermost zone, zone 0, has the most sectors per track, while the innermost zone has the fewest.

Logical block addressing facilitates the use of zoned-bit recording by allowing disk drive manufacturers to establish whatever zones they want to without worrying about the impact on host/subsystem controller logic and operations. As platters are never exchanged between disk drives, there is no need to worry about standardized zone configurations.

Table 4-1 shows the zones for a hypothetical disk drive with 13 zones. The number of tracks in a zone indicates the relative physical area of the zone. Notice how the media transfer rates change as the zones move closer to the spindle. This is why the first partitions created on disk drives tend to have better performance characteristics than partitions that are located closer to the center of the drive.

**Table 4-1** *Disk Drive Zones*

Zone	Number of Tracks	Sectors in Each Track	Media Transfer Rate in Mbps
0	1700	2140	1000
1	3845	2105	990
2	4535	2050	965
3	4365	2000	940
4	7430	1945	915
5	7775	1835	860
6	5140	1780	835
7	6435	1700	800
8	8985	1620	760
9	11,965	1460	685
10	12,225	1295	610
11	5920	1190	560
12	4320	1135	530

## Disk Drive Specifications

Disk drive specifications can be confusing and difficult to interpret. This section highlights some of the most important specs used with disk drives in storage networking applications, including the following:

- Mean time between failures
- Rotational speed and latency
- Average seek time
- Media transfer rate
- Sustained transfer rate

### Mean Time Between Failures

*Mean time between failure (MTBF)* indicates the expected reliability of disk drives. MTBF specifications are derived using well-defined statistical methods and tests run on a large number of disk drives over a relatively short period of time. The results are extrapolated and are expressed as a very large number of hours—usually in the range of 500,000 to 1.25 million hours. These numbers are unthinkably high for individual disk drives—1.25 million hours is approximately 135 years.

MTBF specifications help create expectations for how often disk drive failures will occur when there are many drives in an environment. Using the MTBF specification of 1.25 million hours (135 years), if you have 135 disk drives, you can expect to experience a drive failure once a year. In a storage network environment with a large number of disk drives—for instance, over 1000 drives—it's easy to see that spare drives should be available because there will almost certainly be drive failures that need to be managed. This also underlines the importance of using disk device redundancy techniques, such as mirroring or RAID.

### Rotational Speed and Latency

One of the most common ways to describe the capabilities of any disk drive is to state its rotational speed in rpm. The faster a disk drive spins, the faster data can be written to and read from the disk's media. The performance differences can be enormous. All other things being equal, a 15,000-rpm disk drive can do more than twice the amount of work as a 7200-rpm disk drive. If 50 or more disk drives are being used by a transaction processing system, it's easy to see why somebody would want to use higher-speed drives.

Related to rotation speed is a specification called rotational latency. After the drive's heads are located over the proper track in a disk drive platter, they must wait for the proper sector to pass underneath before the data transfer can be made. The time spent waiting for the right sector is called the rotational latency and is directly linked to the rotational speed of the disk drive.

Essentially, rotational latency is given as the average amount of time to wait for any random I/O operation and is calculated as the time it takes for a platter to complete a half-revolution.

Rotational latencies are on the range of 2 to 6 milliseconds. This might not seem like a very long time. But it is very slow compared to processor and memory device speeds. Applications that tend to suffer from I/O bottlenecks such as transaction processing, data warehousing, and multimedia streaming require disk drives with high rotation speeds and sizable buffers.

Table 4-2 shows the rotational latency for several common rotational speeds.

**Table 4-2** *The Inverse Relationship Between Rotational Speed and Rotational Latency in Disk Drives*

Rotational Speed	Rotational Latency (in ms)
5400	5.6
7200	4.2
10000	3.0
12000	2.5
15000	2.0

### Average Seek Time

Along with rotational speed, seek time is the most important performance specification for a disk drive. *Seek time* measures the time it takes the actuator to reposition the read/write heads from one track to another over a platter. Average seek times represent a performance average over many I/O operations and are relatively similar to rotational latency—in the range of 4 to 8 milliseconds.

Transaction processing and other database applications that perform large numbers of random I/O operations in quick succession require disk drives with minimal seek times. Although it is possible to spread the workload over many drives, transaction application performance also depends significantly on the ability of an individual disk drive to process an I/O operation quickly. This translates into a combination of low seek times and high rotational speeds.

### Media Transfer Rate

The *media transfer rate* of a disk drive measures the performance of bit read/write operations on drive platters. Unlike most storage specifications, which are listed in terms of bytes, the media transfer rate is given in terms of bits. The media transfer rate measures read/write performance on a single track, which depends on the radial length the track is positioned at. In other words, tracks in zone 0 have the fastest media transfer rates in the disk drive. For that reason, media transfer rate specifications are sometimes given using ranges.

## Sustained Transfer Rate

Most I/O operations on a disk drive work across multiple tracks and cylinders, which involves the ability to change the location of the read/write heads. The sustained transfer rate specification takes into account the physical delays of seek time and rotational latency and is much closer to measuring actual user data performance than the media transfer rate.

That said, sustained transfer rates indicate optimal conditions that are difficult to approach with actual applications. There are other important variables such as the size of the average data object and the level of fragmentation in the file system. Nonetheless, sustained transfer rate is a pretty good indication of a drive's overall performance capabilities.

## Optimizing Disk Drive Performance in Storage Networks

Given that disk drives are relatively slow, people have developed several methods and techniques to increase their performance. The performance enhancements discussed in this section include

- Limiting drive contention
- Short-stroking the drive
- Matching rotation speeds
- Aligning zones

### Limiting Drive Contention

Multiple applications that are actively performing I/O operations on a single disk drive can generate a significant amount of seek time and rotation latency, causing performance to deteriorate. Therefore, it makes sense to understand which applications are accessing each disk drive in order to avoid contention for a drive's slow mechanical resources.

Disk bottlenecks can occur in storage area networking (SAN) environments where many disk drives provide storage to many systems and applications. Without good planning, it's possible for different partitions on a single disk drive to be assigned to different applications requiring higher I/O performance levels. In that case, the two applications would wind up competing for the use of a single actuator and a single rotating spindle.

One of the best ways to limit contention for a disk drive is to limit the number of partitions it has. For example, administrators could configure some percentage of high-speed drives to have two partitions—one for a high-performance application and the other for a lower-performance application. Unfortunately, the enforcement of this objective would have to be a manual task, as the technology for automating it is not yet available. A facility that allowed automated disk partition policies to be applied to drives might be very useful. The intelligence to do it is in the drive, but the requirement to provide it is not yet obvious to disk drive manufacturers.

**NOTE**

Most applications have minimal storage I/O requirements and place a minimal load on disk drives. These “I/O-lite” applications probably don’t need to be monitored for disk drive contention problems. These applications also do not deserve to hog the best parts of high-speed disk drives. Most administrators are not used to creating partitions and letting them sit idle, but that could be done to reserve a percentage of the “choicest cuts” of some drives in case they are needed by a high-performance application in the future. The question is, if an application does not require optimal I/O performance, why allocate a high-performance disk partition to it?

### Short-Stroking the Drive

A technique known as short-stroking a disk limits the drive’s capacity by using a subset of the available tracks, typically the drive’s outer tracks. This accomplishes two things: it reduces seek time by limiting the actuator’s range of motion, and it increases the media transfer rate by using the outer tracks, which have the highest densities of data per track.

Short-stroking is normally done by establishing a single partition on a drive that uses some subset of the total capacity. For instance, you can establish an 80-GB partition on a 200-GB disk drive to increase performance.

### Matching Rotation Speeds

There is a tendency in storage networking to treat all storage address spaces equally, which makes them easier to manage. This approach is likely to work very well for the majority of applications, but it falls short for those applications that need the highest performance.

An obvious difference between disk drives is their rotational speed. It would not be a good idea to match slow-speed drives with high-speed drives in support of a high-throughput transaction processing application.

The same concept applies to drive buffer memory. You would not want to use drives with insufficient buffer capacity, slowing down an otherwise fast configuration.

### Aligning Zones

The performance differences across different zones on disk drives could be significant for certain applications that expect consistently high I/O rates, such as data warehousing processing. For example, a disk partition on the outside of a disk might have almost twice the performance of a partition on the inside of a disk. In effect, this is the same situation as wanting to match the rotational speeds of disk drives used for high-performance applications, except this variable adds the element of matching the zones or the relative position of the partitions on a drive.

## Tape Drives

Tape drives are used universally for backup and restore operations. While there are far fewer tape drives than disk drives, the function they provide is very important. The primary difference between tape and disk storage is that tape media is removable, which means it can be transported for safekeeping from disasters as well as being a mechanism for sharing data over the “sneaker net.”

**NOTE**

Sneaker net is certainly not the only way tapes have been used to share data. Companies have used courier services, air cargo, and even refrigerated trucks to move tapes from one site to another to share data.

One of the less-obvious differences between disk drives and tape drives is their read/write ratios. While disk drives are most often used to read data, tape drives are usually used to write data.

This section discusses topics concerning tapes and tape drives:

- Tape media
- Caring for tape
- Caring for tape heads
- Tape drive performance
- The tale of two technologies

## A Look at Tape Media

The media where data is stored on a tape drive is—surprise!—magnetic tape. The result of a great deal of chemicals, materials, and manufacturing technology, magnetic tape is pretty amazing and very durable, as long as it is not abused. Magnetic tape is constructed in four basic layers:

- Backing
- Binder
- Magnetic material
- Coating

The backing of a tape is the foundation material that gives the tape its inherent flexibility and strength. In addition, backing provides a magnetic barrier so that signals from one section of tape do not “print through” onto adjacent sections of the tape when tapes are rolled up tightly and stored for long periods of time.

Tape binder is the flexible glue-like material that adheres to the backing as well as the magnetic material.

The magnetic materials in tape are where the action is, of course, and where data is written and read. The magnetic properties are provided by fine metal oxides, which are smooth to the human eye but somewhat jagged and rough at a microscopic level.

The coating layer levels the surface of the tape and provides a smoother surface for running over the tape heads. Without the coating layer, wear and tear on tape heads would be excessive.

## Caring for Tape

In general, tapes deteriorate slowly over time. They develop cracks in the surface, they tear along the edges, and the metal oxides corrode. It is important to use and store data tapes under conditions of moderate temperature and low humidity. This includes tapes that have not been used yet but are being stored for future use. Tape deterioration that starts before data is written to it can lead to data loss later, even if the tapes are well taken care of after they have been used.

## Caring for Tape Heads

Unlike disk drive heads that float at microscopic levels above the platter, tape heads are designed to be in contact with the tape when reading or writing data. As a result, tape heads eventually wear out over time due to the constant friction of regular operations.

Tape heads should be cleaned after every 30 hours of use. Unlike disk drives, which have limited exposure to airborne particles, tape drives are exposed when tapes are removed and inserted. This makes it practically impossible to keep particulate matter away from the tape heads, and that's why it's important to operate tape drives in a clean environment. Also, tapes shed fine pieces inside the tape drive as they are run through the tape transport, especially when they are new. This material sticks to the tape heads and reduces their effectiveness. Dirty tape heads are often reported by backup software as tape failures, even though there might be nothing wrong with the tapes you are using.

## Tape Drive Performance

Tape drives have wide performance ranges based on two variables:

- A sufficient amount of data being transferred
- The compressibility of data



## Streaming and Start-Stop Operations

Unlike disk drives, tape drives run at different speeds. A drive's streaming transfer rate represents the speed at which data is written from buffers onto tape. Obviously, this implies that a sufficient amount of data is being written into the drive's buffers by host or subsystem controllers.

Start-stop operations occur when there is insufficient data to maintain streaming mode operations. Start-stop speeds are typically far less than streaming speeds because the tape has to be stopped, rewound slightly, and started up again. Unlike analog recording, where it is no problem to have "dead air," digital recording cannot have undefined gaps. Therefore, if there is no data to record, the tape drive must stop and reposition the tape appropriately for when there is more data to write. This starting and stopping is obviously detrimental to performance. Larger buffers can help, but they do not solve the problem of having too little data to write.

## Compression

Tape drives typically incorporate compression technology as a feature to boost data transfer rates. Compression can increase performance several times beyond native (uncompressed) data transfer rates, but that depends on how much the data can be compressed. Different types of data vary a great deal in this respect; multimedia data might not compress at all, and database data might compress many times. Tape drives in storage networks should be able to support native streaming transfer rates of at least 10 Mbps.

## The Tale of Two Technologies

The history of tape technology is littered with many obsolete technologies that faded fast. People have protested about the interoperability and compatibility problems of tape technologies for many years, and the situation is no different today. Tape drives used in storage networks can be divided into two broad technology areas, with two contestants in each area—all of them being incompatible with the others.

In the sections that follow, we will briefly look at both technologies.

## Linear Tape Technology

Linear tape reads and writes data just as it sounds—by placing "lines" of data that run lengthwise on the tape media. Linear tape drives use multiple heads operating in parallel, reading and writing data simultaneously. They tend to have very high transfer rates and capacities.

Tape used in linear tape drives is .5 inches wide, leaving a great deal of room for capacity improvements over time. Linear tape cartridges have only one spool to hold tape. When the tape

is loaded, the outside end of the tape is fed into the drive and wrapped around a take-up spool inside the drive.

There are two primary, competing linear tape technologies:

- Super digital linear tape (SDLT)
- Linear tape open (LTO)

SDLT has its technology roots in the digital linear tape technology developed by Digital Equipment Corporation many years ago for its VAX line of computers. LTO is a relatively new type of tape technology that was jointly developed by IBM, HP, and Seagate.

## Helical Scan Tape Technology

Helical scan tape technology was originally developed for video recording applications. Most helical scan tape drives used for data storage applications use 8 mm tape, which is approximately .25 inches wide. Helical scan drives write data in diagonal strips along the tape.

In general, helical scan tape cartridges are much smaller than linear tape cartridges, even though they have two reels, unlike linear tape. Data density with helical scanning technology is typically better, although helical scan tape cartridges usually hold less data than linear tape cartridges.

Two primary helical scan technologies are used in storage networking environments: Mammoth-2 and AIT-3. Mammoth tape technology was developed by Exabyte Corporation in the mid-1990s in response to DLT's growing popularity and success. At this point it seems to have been eclipsed in functionality by AIT, LTO, and DLT as the preferred technologies for storage network environments.

Sony's Advanced Intelligent Tape (AIT) was also developed as a new data recording technology in the mid 1990s in response to the need for higher reliability, performance, and capacity. AIT included many innovations, including a memory chip embedded in the tape cartridge that could be used by storage applications.

## Comparing Tape Technologies

Table 4-3 compares the leading tape technologies used in storage networks.

**Table 4-3** *Comparison of Tape Technologies Used in Storage Networks*

Technology	Linear or Helical	Capacity (Native/Compressed)	Maximum Transfer Rate (Native/Compressed)
Mammoth-2	Helical	60 GB/150 GB	12 Mbps/30 Mbps
AIT-3	Helical	100 GB/250 GB	12 Mbps/30 Mbps
SuperDLT	Linear	110 GB/220 GB	10 Mbps/20 Mbps
LTO Ultrium	Linear	340 GB/680 GB	20 Mbps/40 Mbps

## Summary

Device technology continues to evolve and is being heavily pushed by the requirements of storage networks. In general, bare devices are not addressed individually in a storage network but are part of larger subsystems that provide centralized management for a larger number of resources.

That said, understanding the capabilities and shortcomings of these devices is important to maintaining a well-run SAN environment—especially where high-performance I/O operations are required. The rotational speed of a drive, its average seek time performance, the relative location of the tracks being used, the number of applications sharing space on the drive, and the amount of buffer memory are all contributing factors that need to be considered for I/O performance tuning in a SAN.

Tape drives have made tremendous strides in the last five years. Capacities and transfer rates have increased by a wide margin, pushed by the enormous amount of data that is being stored. While there are some who believe tape is a dead and dying technology, it is not likely to go away anytime soon because its portability works so well for disaster recovery situations and the ability to restore historical versions of data.

## Q & A

1. Which of the following is not part of a disk drive: platters, arms, controller, or fan?
2. What is the flying height of modern disk heads?
3. True or false: The clicking sound coming from disk drives is caused by the head touching the media.
4. What does logical block addressing provide?
5. Where is disk throughput greater: inside tracks or outside tracks?
6. What is MTBF used for?
7. True or false: Multiple applications can access different partitions on a disk drive without regard for performance.
8. How often should tape heads be cleaned?
9. True or false: Data compression always cuts the time it takes to run backup in half.
10. What are the two major types of tape technologies used in UNIX and Windows environments?

