

Hard disk drive performance characteristics

Higher performance in hard disk drives comes from devices which have better performance characteristics.^{[1][2]} These devices include those with rotating media, hereby called *rotating drives*, i.e., hard disk drives (HDD), floppy disk drives (FDD), optical discs (DVD-RW / CD-RW), and it also covers devices without moving parts like solid-state drives (SSD). For SSDs, most of the attributes related to the movement of mechanical components are not applicable, but the device is actually affected by some other electrically based element that still causes a measurable delay when isolating and measuring that attribute.^[3] These performance characteristics can be grouped into two categories: access time and data transfer time (or rate).^[4]

Contents

Access time

- Seek time
 - Short stroking
 - Effect of audible noise and vibration control
- Rotational latency
 - Effect of reduced power consumption
- Other

Data transfer rate

- Effect of file system
- Effect of areal density
- Interleave

Power consumption

Shock resistance

See also

References

Access time

The *access time* or *response time* of a rotating drive is a measure of the time it takes before the drive can actually transfer data. The factors that control this time on a rotating drive are mostly related to the mechanical nature of the rotating disks and moving heads. It is composed of a few independently measurable elements that are added together to get a single value when evaluating the performance of a storage device. The access time can vary significantly, so it is typically provided by manufacturers or measured in benchmarks as an average.^{[4][5]} For SSDs this time is not dependent on moving parts, but rather electrical connections to solid state memory, so the access time is very quick and consistent.^[6] Most testing and benchmark applications do not draw a distinction between rotating drives and SSDs so they both go through the same measurement process.

The key components that are typically added together to obtain the access time are:^{[2][7]}

- Seek time
- Rotational latency
- Command processing time
- Settle time

Seek time

With rotating drives, the *seek time* measures the time it takes the head assembly on the actuator arm to travel to the track of the disk where the data will be read or written.^[7] The data on the media is stored in sectors which are arranged in parallel circular tracks (concentric or spiral depending upon the device type) and there is an actuator with an arm that suspends a head that can transfer data with that media. When the drive needs to read or write a certain sector it determines in which track the sector is located.^[8] It then uses the actuator to move the head to that particular track. If the initial location of the head was the desired track then the seek time would be zero. If the initial track was the outermost edge of the media and the desired track was at the innermost edge then the seek time would be the maximum for that drive.^{[9][10]} Seek times are not linear compared with the seek distance traveled because of factors of acceleration and deceleration of the actuator arm.^[11]

A rotating drive's *average seek time* is the average of all possible seek times which technically is the time to do all possible seeks divided by the number of all possible seeks, but in practice it is determined by statistical methods or simply approximated as the time of a seek over one-third of the number of tracks.^{[7][9][12]} Average seek time ranges from under 4 ms for high-end server drives,^[13] to 15 ms for mobile drives, with the most common mobile drives at about 12 ms^[14] and the most common desktop drives typically being around 9 ms.

The first HDD^[15] had an average seek time of about 600 ms, and by the middle 1970s, HDDs were available with seek times of about 25 ms.^[16] Some early PC drives used a stepper motor to move the heads, and as a result had seek times as short as 80–120 ms, but this was quickly improved by voice coil type actuation in the 1980s, reducing seek times to around 20 ms. Seek time has continued to improve slowly over time.

The other two less commonly referenced seek measurements are *track-to-track* and *full stroke*. The track-to-track measurement is the time required to move from one track to an adjacent track.^[7] This is the shortest (fastest) possible seek time. In HDDs this is typically between 0.2 and 0.8 ms.^[6] The full stroke measurement is the time required to move from the outermost track to the innermost track. This is the longest (slowest) possible seek time.^[9]

With SSDs there are no moving parts, so a measurement of the seek time is only testing electronic circuits preparing a particular location on the memory in the storage device. Typical SSDs will have a seek time between 0.08 and 0.16 ms.^[6]

Short stroking

Short stroking is a term used in enterprise storage environments to describe an HDD that is purposely restricted in total capacity so that the actuator only has to move the heads across a smaller number of total tracks.^[17] This limits the maximum distance the heads can be from any point on the drive thereby reducing its average seek time, but also restricts the total capacity of the drive. This reduced seek time enables the HDD to increase the number of IOPS available from the drive. The cost and power per usable byte of storage rises as the maximum track range is reduced.^{[18][19]}

Effect of audible noise and vibration control



A hard disk head on an access arm resting on a hard disk platter

Measured in dBA, audible noise is significant for certain applications, such as DVRs, digital audio recording and quiet computers. Low noise disks typically use fluid bearings, lower rotational speeds (usually 5,400 rpm) and reduce the seek speed under load (AAM) to reduce audible clicks and crunching sounds. Drives in smaller form factors (e.g. 2.5 inch) are often quieter than larger drives.^[20]

Some desktop- and laptop-class disk drives allow the user to make a trade-off between seek performance and drive noise. For example, Seagate offers a set of features in some drives called Sound Barrier Technology that include some user or system controlled noise and vibration reduction capability. Shorter seek times typically require more energy usage to quickly move the heads across the platter, causing loud noises from the pivot bearing and greater device vibrations as the heads are rapidly accelerated during the start of the seek motion and decelerated at the end of the seek motion. Quiet operation reduces movement speed and acceleration rates, but at a cost of reduced seek performance.^[21]

Rotational latency

Rotational latency (sometimes called *rotational delay* or just *latency*) is the delay waiting for the rotation of the disk to bring the required disk sector under the read-write head.^[22] It depends on the rotational speed of a disk (or spindle motor), measured in revolutions per minute (RPM).^{[7][23]} For most magnetic media-based drives, the *average rotational latency* is typically based on the empirical relation that the average latency in milliseconds for such a drive is one-half the rotational period. *Maximum rotational latency* is the time it takes to do a full rotation excluding any spin-up time (as the relevant part of the disk may have just passed the head when the request arrived).^[24]

Typical HDD figures

HDD spindle speed [rpm]	Average rotational latency [ms]
4,200	7.14
5,400	5.56
7,200	4.17
10,000	3.00
15,000	2.00

- **Maximum latency** = 60/rpm
- **Average latency** = 0.5*Latency Time

Therefore, the rotational latency and resulting access time can be improved (decreased) by increasing the rotational speed of the disks.^[7] This also has the benefit of improving (increasing) the throughput (discussed later in this article).

The spindle motor speed can use one of two types of disk rotation methods: 1) constant linear velocity (CLV), used mainly in optical storage, varies the rotational speed of the optical disc depending upon the position of the head, and 2) constant angular velocity (CAV), used in HDDs, standard FDDs, a few optical disc systems, and vinyl audio records, spins the media at one constant speed regardless of where the head is positioned.

Another wrinkle occurs depending on whether surface bit densities are constant. Usually, with a CAV spin rate, the densities are not constant so that the long outside tracks have the same number of bits as the shorter inside tracks. When the bit density is constant, outside tracks have more bits than inside tracks and is generally combined with a CLV spin rate. In both these schemes contiguous bit transfer rates are constant. This is not the case with other schemes such as using constant bit density with a CAV spin rate.

Effect of reduced power consumption

Power consumption has become increasingly important, not only in mobile devices such as laptops but also in server and desktop markets. Increasing data center machine density has led to problems delivering sufficient power to devices (especially for spin-up), and getting rid of the waste heat subsequently produced, as well as environmental and electrical cost concerns (see green computing). Most hard disk drives today support some form of power management which uses a number of specific power modes that save energy by reducing performance. When implemented, an HDD will change between a full power mode to one or more power saving modes as a function of drive usage. Recovery from the deepest

mode, typically called Sleep where the drive is stopped or spun down, may take as long as several seconds to be fully operational thereby increasing the resulting latency.^[25] The drive manufacturers are also now producing *green drives* that include some additional features that do reduce power, but can adversely affect the latency including lower spindle speeds and parking heads off the media to reduce friction.^[26]

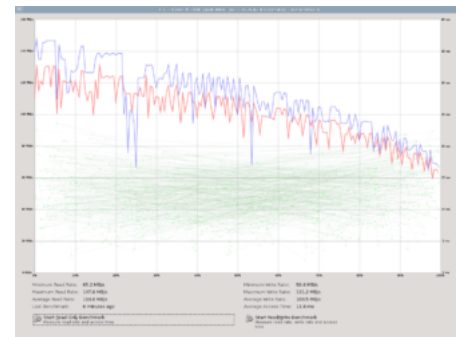
Other

The *command processing time* or *command overhead* is the time it takes for the drive electronics to set up the necessary communication between the various components in the device so it can read or write the data. This is of the order of 3 μ s, very much less than other overhead times, so it is usually ignored when benchmarking hardware.^{[2][27]}

The *settle time* is the time it takes the heads to *settle* on the target track and stop vibrating so they do not read or write *off track*. This time is usually very small, typically less than 100 μ s, and modern HDD manufacturers account for it in their seek time specifications.^[28]

Data transfer rate

The *data transfer rate* of a drive (also called *throughput*) covers both the internal rate (moving data between the disk surface and the controller on the drive) and the external rate (moving data between the controller on the drive and the host system). The measurable data transfer rate will be the lower (slower) of the two rates. The *sustained data transfer rate* or *sustained throughput* of a drive will be the lower of the sustained internal and sustained external rates. The sustained rate is less than or equal to the maximum or burst rate because it does not have the benefit of any cache or buffer memory in the drive. The internal rate is further determined by the media rate, sector overhead time, head switch time, and cylinder switch time. These are not applicable to SSDs.^{[7][29]}



A plot showing dependency of transfer rate on cylinder

Media rate

Rate at which the drive can read bits from the surface of the media.

Sector overhead time

Additional time (bytes between sectors) needed for control structures and other information necessary to manage the drive, locate and validate data and perform other support functions.^[30]

Head switch time

Additional time required to electrically switch from one head to another, re-align the head with the track and begin reading; only applies to multi-head drive and is about 1 to 2 ms.^[30]

Cylinder switch time

Additional time required to move to the first track of the next cylinder and begin reading; the name cylinder is used because typically all the tracks of a drive with more than one head or data surface are read before moving the actuator. This time is typically about twice the track-to-track seek time. As of 2001, it was about 2 to 3 ms.^[31]

Data transfer rate (read/write) can be measured by writing a large file to disk using special file generator tools, then reading back the file.

- As of 2010, a typical 7200 RPM desktop HDD has a "disk-to-buffer" data transfer rate up to 1030 Mbit/s.^[32] This rate depends on the track location, so it will be higher on the outer zones (where there are more data sectors per track) and lower on the inner zones (where there are fewer data sectors per track); and is generally somewhat higher for

10,000 RPM drives.

- Floppy disk drives have sustained "disk-to-buffer" data transfer rates that are one or two orders of magnitude lower than that of HDDs.
- The sustained "disk-to-buffer" data transfer rates varies amongst families of Optical disk drives with the slowest 1x CDs at 1.23 Mbit/s floppy-like while a high performance 12x Blu-ray disc drive at 432 Mbit/s approaches the performance of HDDs.

A current widely used standard for the "buffer-to-computer" interface is 3.0 Gbit/s SATA, which can send about 300 megabyte/s (10-bit encoding) from the buffer to the computer, and thus is still comfortably ahead of today's disk-to-buffer transfer rates.

SSDs do not have the same internal limits of HDDs, so their internal and external transfer rates are often maximizing the capabilities of the drive-to-host interface.

Effect of file system

Transfer rate can be influenced by file system fragmentation and the layout of the files. Defragmentation is a procedure used to minimize delay in retrieving data by moving related items to physically proximate areas on the disk.^[33] Some computer operating systems perform defragmentation automatically. Although automatic defragmentation is intended to reduce access delays, the procedure can slow response when performed while the computer is in use.^[34]

Flash memory-based SSDs do not need defragmentation; however, because SSDs write pages of data that are much larger than the blocks of data managed by the file system, over time, an SSD's write performance can degrade as the drive becomes full of pages which are partial or no longer needed by the file system. This can be ameliorated by a TRIM command from the system or internal garbage collection. Flash memory wears out over time as it is repeatedly written to; the writes required by defragmentation wear the drive for no speed advantage.^[35]

Effect of areal density

HDD data transfer rate depends upon the rotational speed of the disks and the data recording density. Because heat and vibration limit rotational speed, increasing density has become the main method to improve sequential transfer rates.^[36] *Areal density* (the number of bits that can be stored in a certain area of the disk) has been increased over time by increasing both the number of tracks across the disk, and the number of sectors per track. The latter will increase the data transfer rate for a given RPM speed. Improvement of data transfer rate performance is correlated to the areal density only by increasing a track's linear surface bit density (sectors per track). Simply increasing the number of tracks on a disk can affect seek times but not gross transfer rates. According to industry observers and analysts for 2011 to 2016,^{[37][38]} "The current roadmap predicts no more than a 20%/yr improvement in bit density".^[39] Seek times have not kept up with throughput increases, which themselves have not kept up with growth in bit density and storage capacity.

Interleave

Sector interleave is a mostly obsolete device characteristic related to data rate, dating back to when computers were too slow to be able to read large continuous streams of data. Interleaving introduced gaps between data sectors to allow time for slow equipment to get ready to read the next block of data. Without interleaving, the next logical sector would arrive at the read/write head before the equipment was ready, requiring the system to wait for another complete disk revolution before reading could be performed.

However, because interleaving introduces intentional physical delays between blocks of data thereby lowering the data rate, setting the interleave to a ratio higher than required causes unnecessary delays for equipment that has the performance needed to read sectors more quickly. The interleaving ratio was therefore usually chosen by the end-user to

suit their particular computer system's performance capabilities when the drive was first installed in their system.

Modern technology is capable of reading data as fast as it can be obtained from the spinning platters, so hard drives usually have a fixed sector interleave ratio of 1:1, which is effectively no interleaving being used.

Power consumption

Power consumption has become increasingly important, not only in mobile devices such as laptops but also in server and desktop markets. Increasing data center machine density has led to problems delivering sufficient power to devices (especially for spin up), and getting rid of the waste heat subsequently produced, as well as environmental and electrical cost concerns (see green computing). Heat dissipation is tied directly to power consumption, and as drives age, disk failure rates increase at higher drive temperatures.^[40] Similar issues exist for large companies with thousands of desktop PCs. Smaller form factor drives often use less power than larger drives. One interesting development in this area is actively controlling the seek speed so that the head arrives at its destination only just in time to read the sector, rather than arriving as quickly as possible and then having to wait for the sector to come around (i.e. the rotational latency).^[41] Many of the hard drive companies are now producing Green Drives that require much less power and cooling. Many of these Green Drives spin slower (<5,400 rpm compared to 7,200, 10,000 or 15,000 rpm) thereby generating less heat. Power consumption can also be reduced by parking the drive heads when the disk is not in use reducing friction, adjusting spin speeds,^[42] and disabling internal components when not in use.^[43]

Drives use more power, briefly, when starting up (spin-up). Although this has little direct effect on total energy consumption, the maximum power demanded from the power supply, and hence its required rating, can be reduced in systems with several drives by controlling when they spin up.

- On SCSI hard disk drives, the SCSI controller can directly control spin up and spin down of the drives.
- Some Parallel ATA (PATA) and Serial ATA (SATA) hard disk drives support power-up in standby (PUIS): each drive does not spin up until the controller or system BIOS issues a specific command to do so. This allows the system to be set up to stagger disk start-up and limit maximum power demand at switch-on.
- Some SATA II and later hard disk drives support staggered spin-up, allowing the computer to spin up the drives in sequence to reduce load on the power supply when booting.^[44]

Most hard disk drives today support some form of power management which uses a number of specific power modes that save energy by reducing performance. When implemented an HDD will change between a full power mode to one or more power saving modes as a function of drive usage. Recovery from the deepest mode, typically called Sleep, may take as long as several seconds.^[45]

Shock resistance

Shock resistance is especially important for mobile devices. Some laptops now include active hard drive protection that parks the disk heads if the machine is dropped, hopefully before impact, to offer the greatest possible chance of survival in such an event. Maximum shock tolerance to date is 350 g for operating and 1,000 g for non-operating.^[46]

See also

- vRPM



Low-level formatting software from 1987 to find highest performance interleave choice for 10 MB IBM PC XT hard disk drive