CS240 Operating Systems, Communications and Concurrency

Magnetic Hard Disk Operation and Scheduling

Physical Characteristics of Hard Disk Drives

Data Organisation & Performance

Disk Scheduling Algorithms

CS240 Operating Systems, Communications and Concurrency

Disk Operation and Scheduling

Magnetic storage device

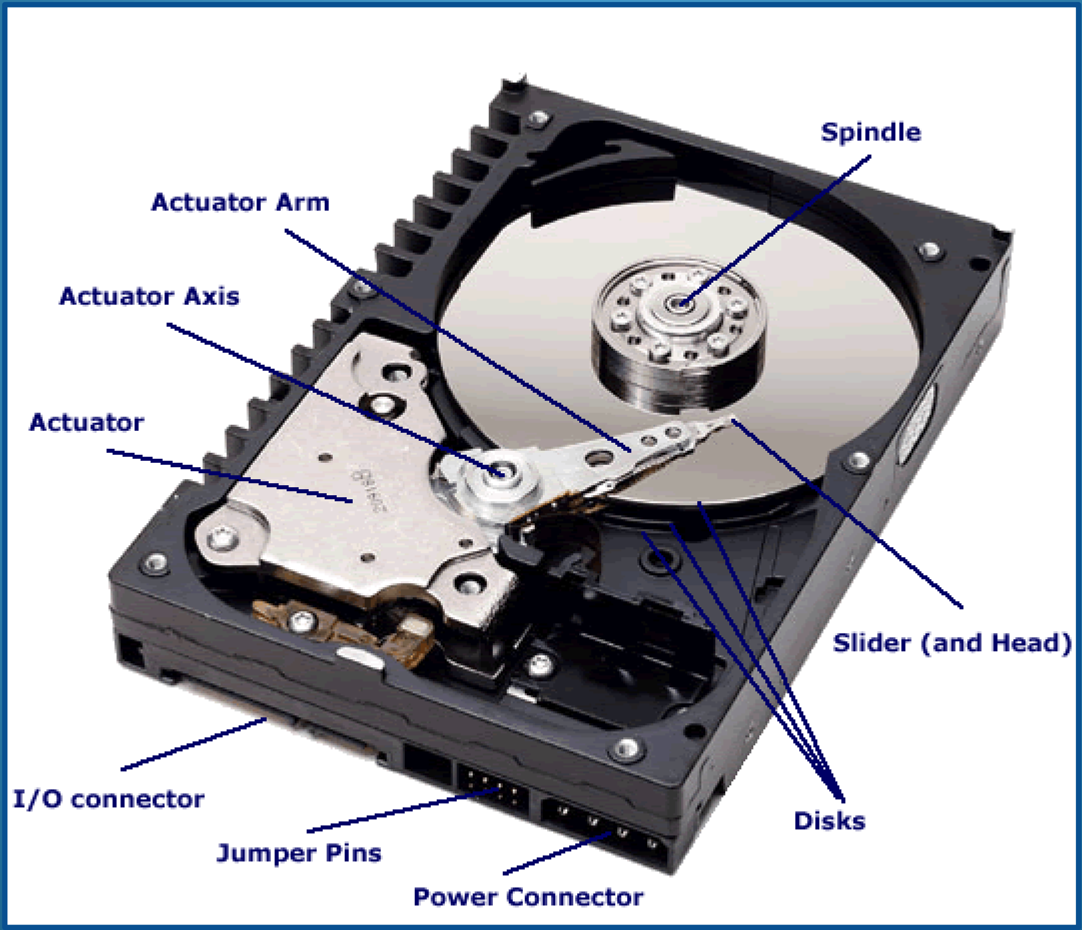
Main form of persistent data storage for decades

Introduced in 1956 by IBM

It is hard to understand the factors that affect performance, reliability and interfacing without knowing a little about how the drive works internally.

File system technology is predicated also on their behaviour.

CS240 Operating Systems, Communications and Concurrency



CS240 Operating Systems, Communications and Concurrency

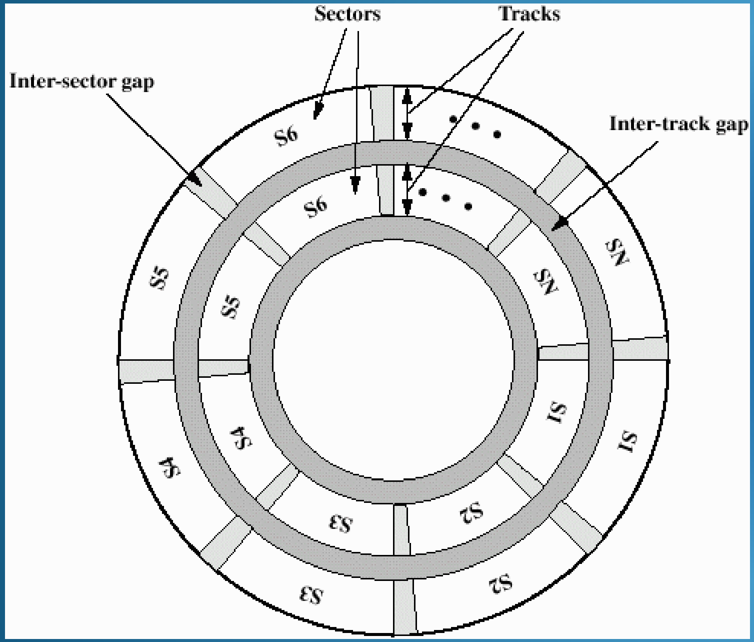
Engineering Developments

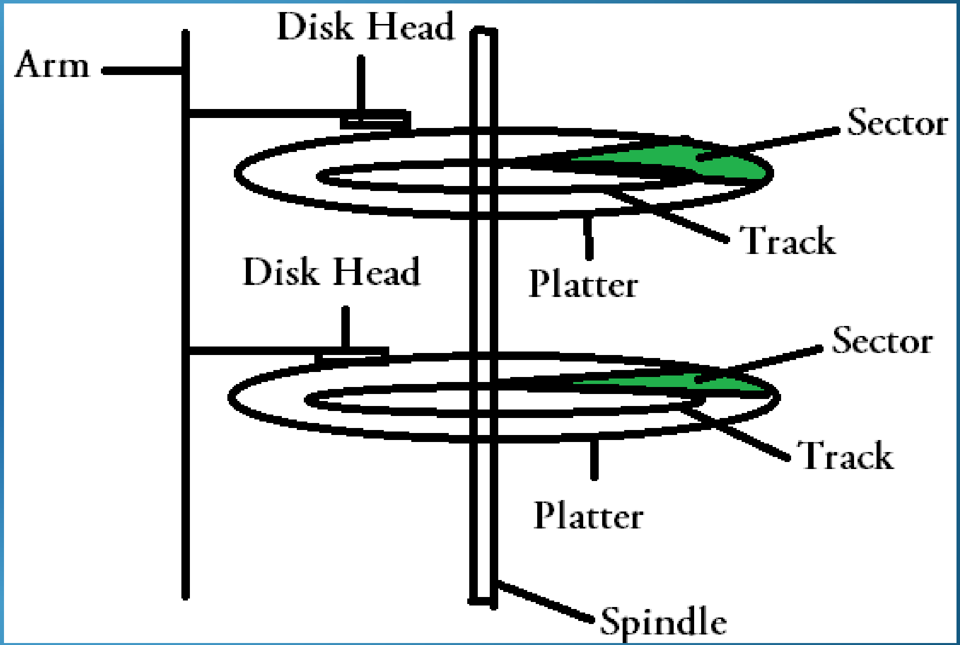
Smaller platter diameters offer better rigidity, weigh less and require less power to spin, meaning more useful with portable devices/usb power, less noise and heat and improved seek performance.

Smaller platters mean a smaller recording surface, but the bit density technology has increased more than sufficiently to compensate.

Faster Signaling Interfaces improving transfer speed.

CS240 Operating Systems, Communications and Concurrency

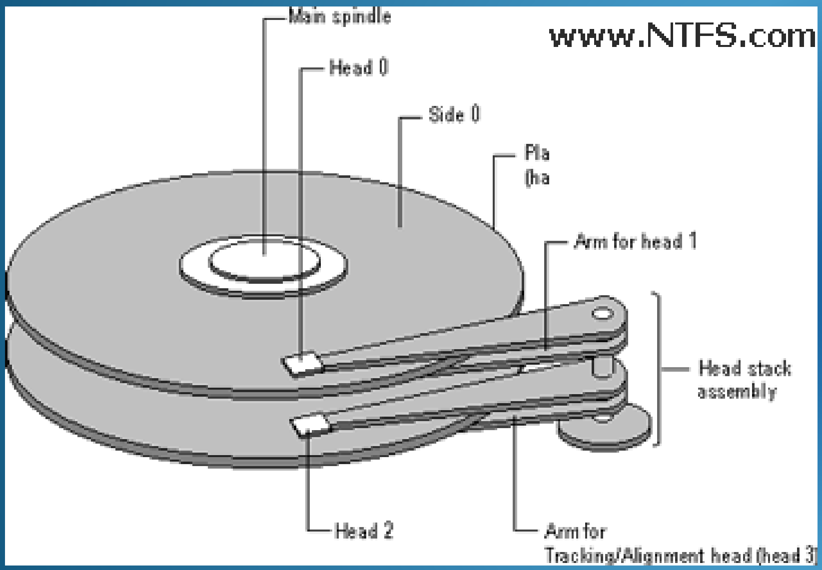




The surface of each platter is organised as a **concentric group of magnetic tracks** , perhaps more than a thousand, on which data can be stored.

Each track is divided into a number of blocks of fixed size called **sectors** in which the data is stored.
(Historically 512 bytes of data per sector but 4k bytes in newer large drives)

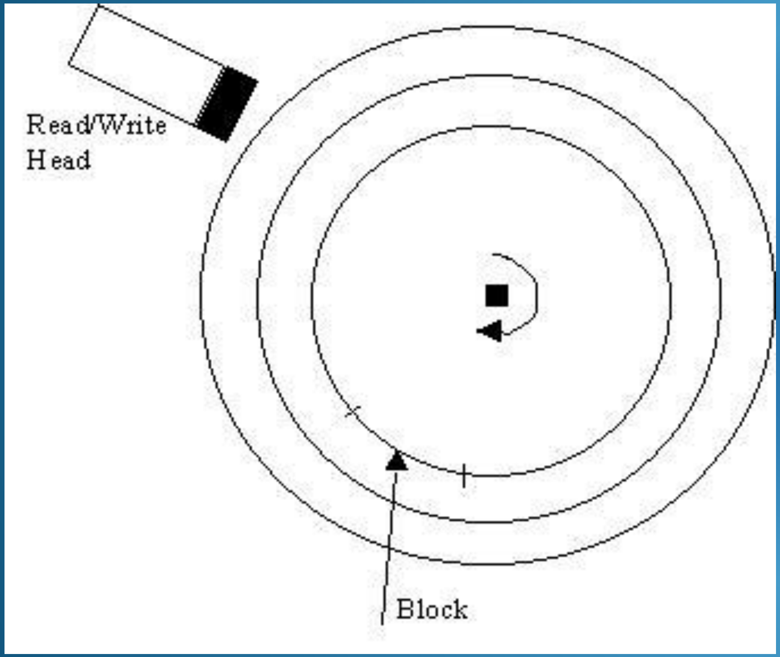
CS240 Operating Systems, Communications and Concurrency



One side of one platter contains space reserved for hardware track-positioning information and is not available to the operating system.

A disk assembly containing two platters has three sides available for data. Track-positioning data is written to the disk during assembly at the factory. The system disk controller reads this data to settle the drive heads in the correct track position.

CS240 Operating Systems, Communications and Concurrency

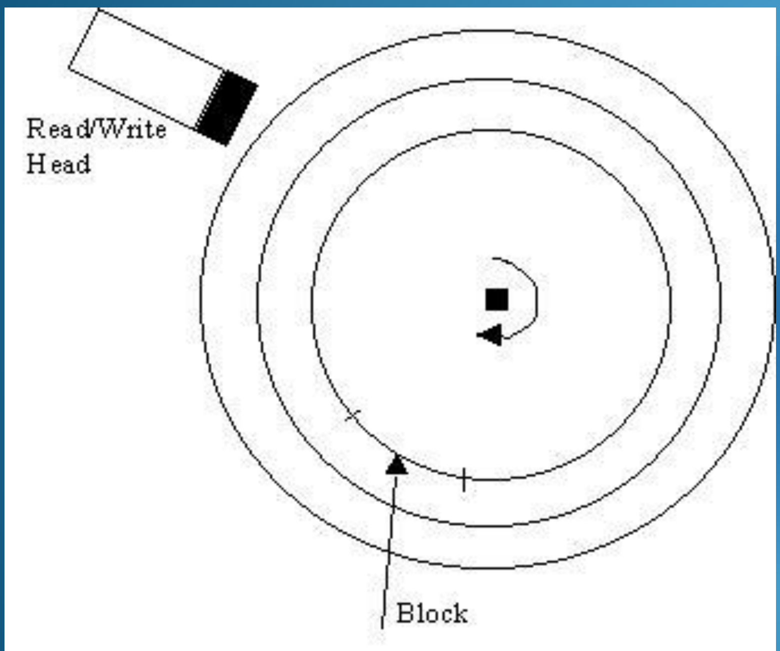


A sector/block is the **smallest storage unit of data** that can be addressed, read from or written to the disk.

We can view the disk as a virtual storage device with an array of sectors in the range 0 to n – 1 , essentially the **address space** of the drive.

Some operating systems may group adjacent sectors into clusters. A cluster is the minimum amount of space on a disk that a file can occupy. Reserving a cluster allows the file some growth while maintaining efficiency of unfragmented access.

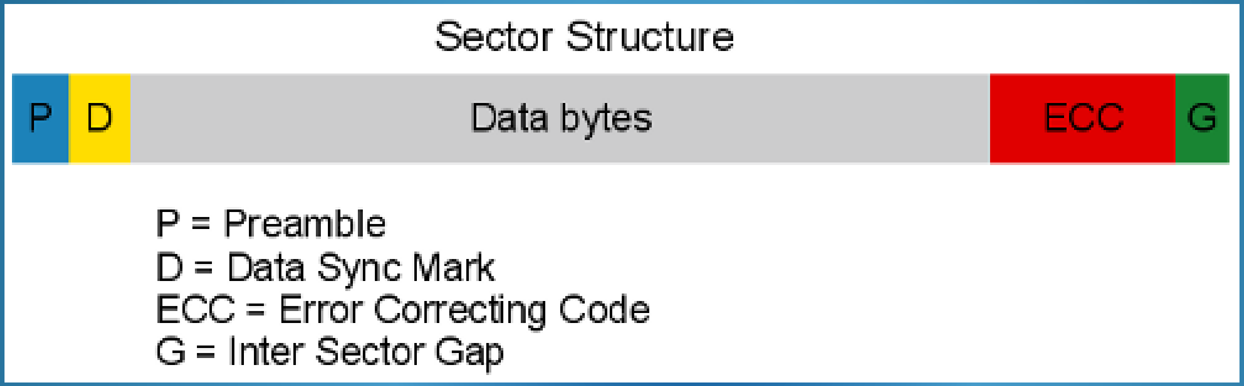
CS240 Operating Systems, Communications and Concurrency



Each platter surface has a **read/write head** which can move linearly across it.

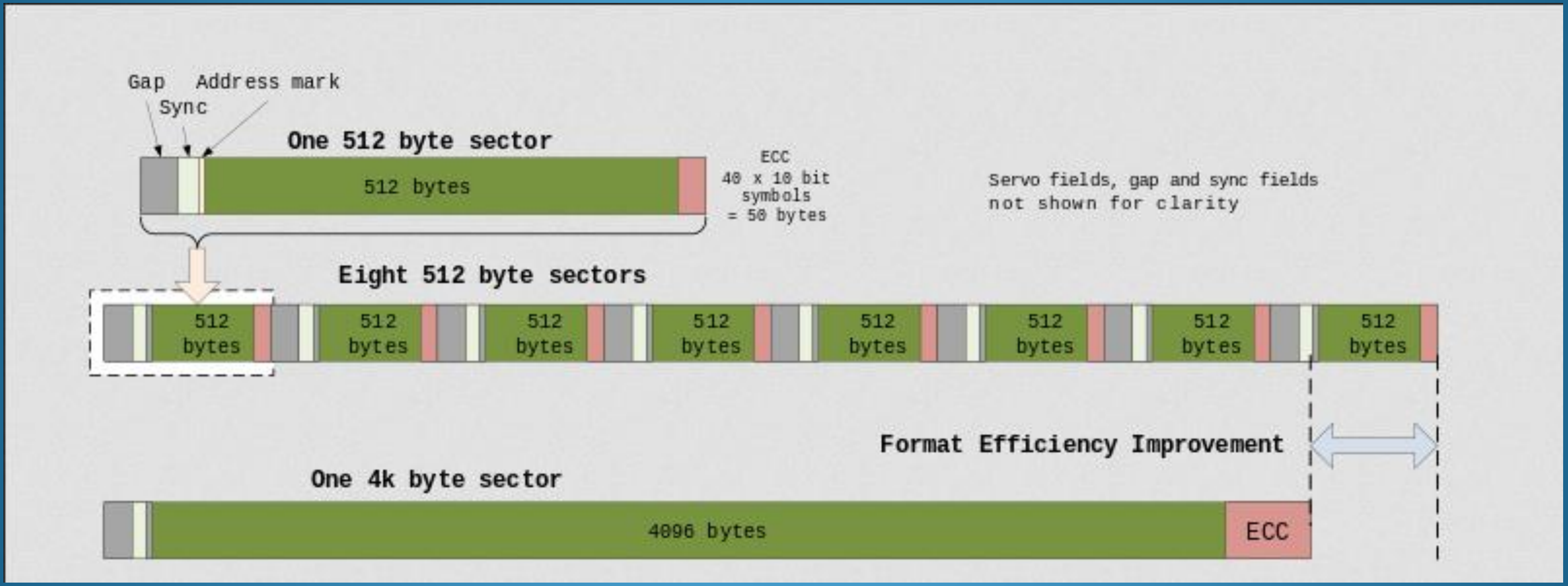
The actuator moves the head to the correct track on a given platter and then waits for the correct block to **pass underneath** it for access.

CS240 Operating Systems, Communications and Concurrency



The contents of each sector area include sector id information, sector status codes, synchronization bit patterns to guide the read head, 512 or 4k bytes of data, error correction codes and sector gaps.

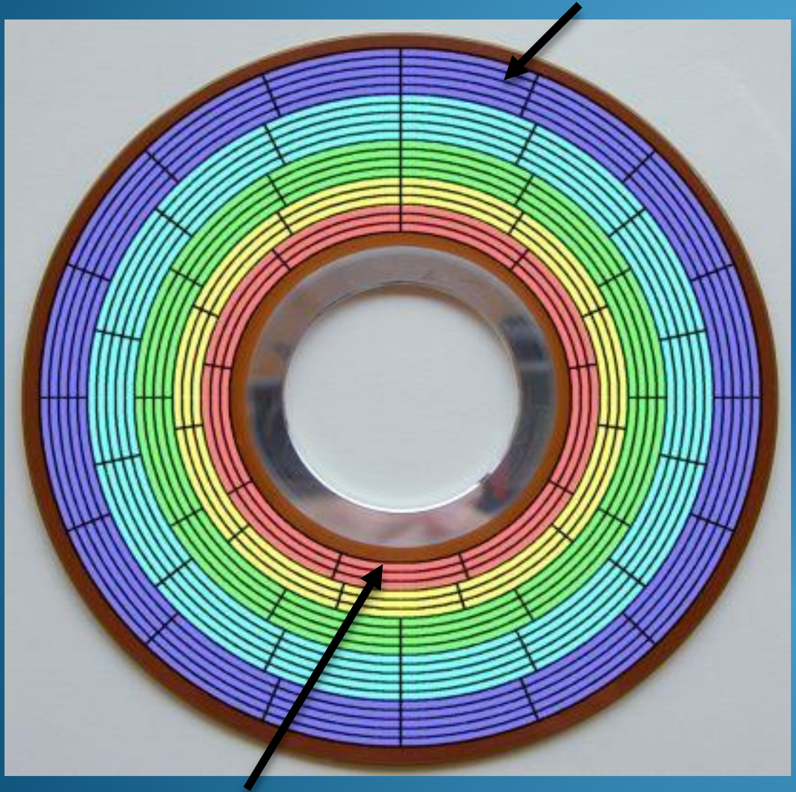
CS240 Operating Systems, Communications and Concurrency



Larger sector size improves overheads efficiency but at cost of increasing minimum unit of space allocation.

CS240 Operating Systems, Communications and Concurrency

E.g. 16 Sectors per Track



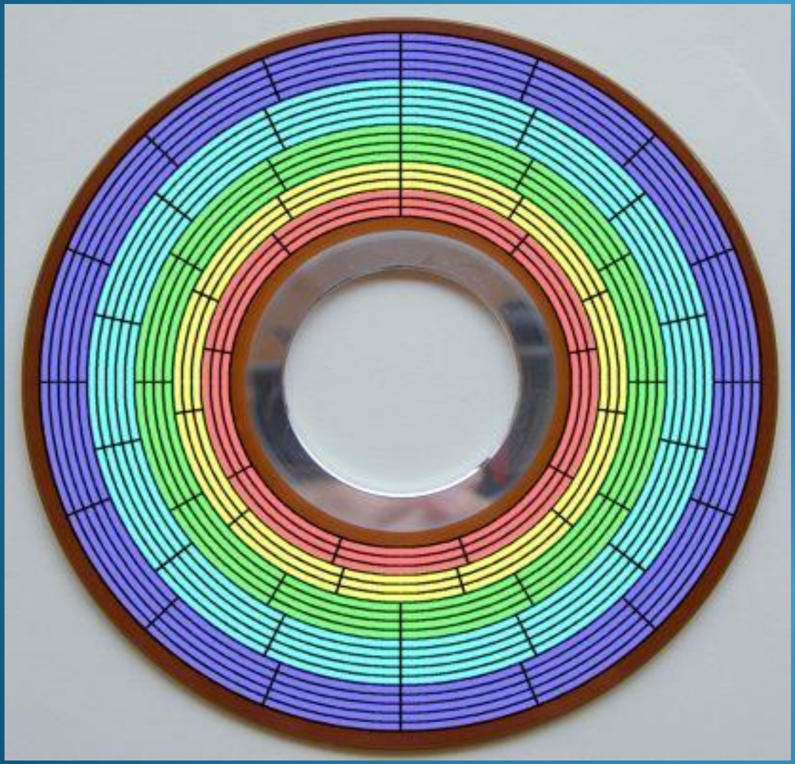
E.g. 9 Sectors per Track

Zoned Bit Recording

The larger outer tracks contain a much greater area of recording material than the inner ones. The number of sectors stored on inner tracks is constrained by the bit density of the magnetic surface.

Tracks are grouped based on their distance from the centre with the outer zones being used to store more sectors than the inner ones.

CS240 Operating Systems, Communications and Concurrency

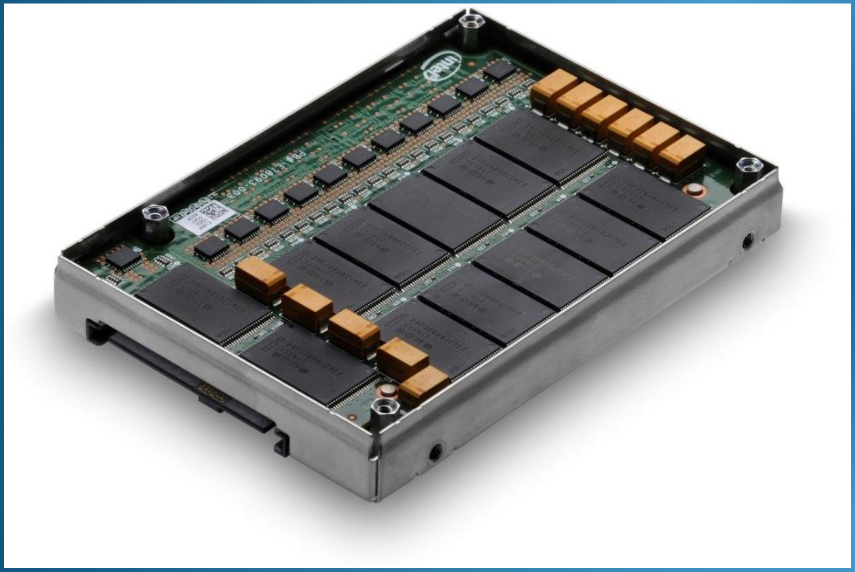


Zoned Bit Recording

This also means that if the rotational velocity is constant, that data can be transferred faster from the outer tracks than the inner tracks.

As the drive is filled from the outside in, it gives its best performance when new and empty.

CS240 Operating Systems, Communications and Concurrency



Similar packaging, designed to be physical replacements for hard drives but access is electrical not mechanical, uses floating gate transistor technology for non-volatile storage.

Solid State Drives

These storage drives have energy and performance benefits over hard drives but are more expensive and generally have smaller capacity.

Operating system may boot faster from SSD & file opening speeds maybe 20-40x faster than a standard HDD.

CS240 Operating Systems, Communications and Concurrency

Hard Disk Drive Performance Metrics

Capacity – Highly reliable high capacity cheap solution

Data Rate – The bandwidth between a drive and the system, connected by a particular interface e.g. SATA 1.5Gb/s, SATA-2 3Gb/s, SATA-3 6Gb/s, SATA 3.2 16Gb/s

Average Latency – Time to access a sector e.g. 4ms-9ms

Reliability – Mean Time to Failure

CS240 Operating Systems, Communications and Concurrency

Scheduling Algorithms - Seek Time

The slowest part of accessing a disk block is physically moving the head to the correct track.

This is called the ***seek time***.

In a multitasking system, requests to access different parts of the disk arrive from different processes or a file's blocks may be scattered across a platter.

Correct scheduling of these requests can reduce the average seek time by reducing the average distance the head has to travel.

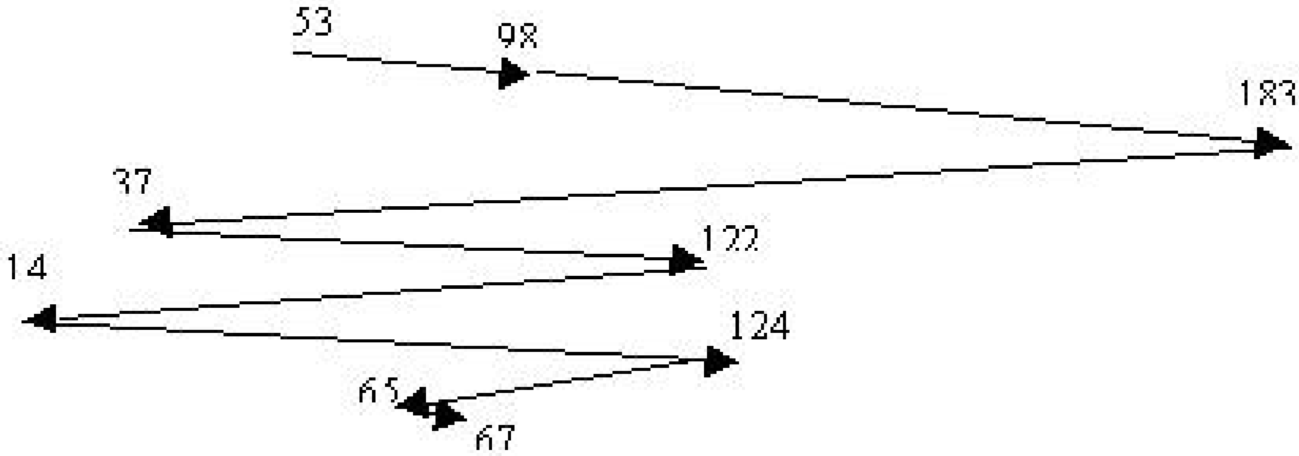
Native Command Queuing (NCQ) is an extension of the SATA protocol allowing hard disk drives to internally optimize the order in which received read and write commands are executed.

CS240 Operating Systems, Communications and Concurrency

Say tracks are numbered outer to inner 0-200.
Head is currently sitting on track 53 with following track requests in queue.
How far does it have to move to service each of the tracks?

98, 183, 37, 122, 14, 124, 65, 67

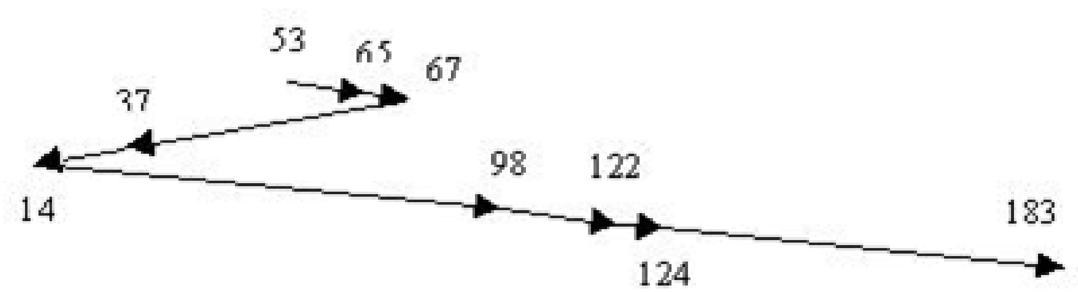
First Come First Served (FCFS)
With this algorithm we would get the following head movements:-



CS240 Operating Systems, Communications and Concurrency

98, 183, 37, 122, 14, 124, 65, 67

Shortest Seek Time First (SSTF)
Choose the request which is closest to the current head position. With this algorithm we would get the following head movements:-



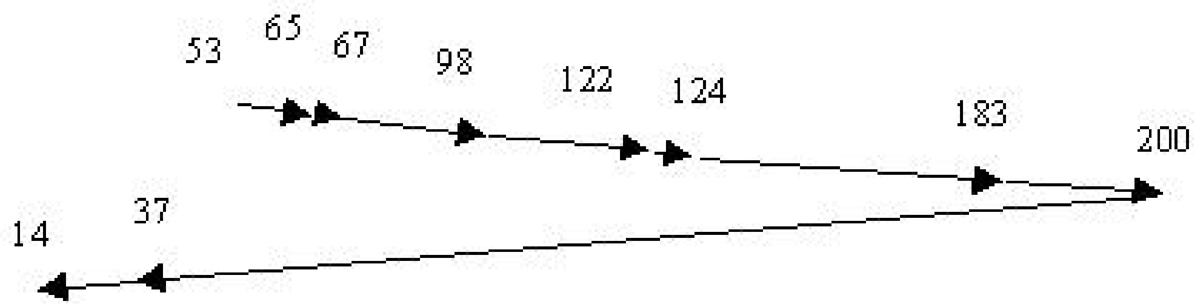
Note that on average, the head will be somewhere around the middle of the disk and that this algorithm therefore favours requests in this area. The outer and inner tracks of the disk may therefore be starved.

CS240 Operating Systems, Communications and Concurrency

98, 183, 37, 122, 14, 124, 65, 67

SCAN Algorithm

The head moves in one direction, either to the center or to the outermost track, servicing requests as it passes the corresponding tracks. It then turns around and repeats this in the other direction. With this algorithm, assuming the head is moving in the direction of increasing track numbers and track numbers are from 0..200, we would get the following head movements:-



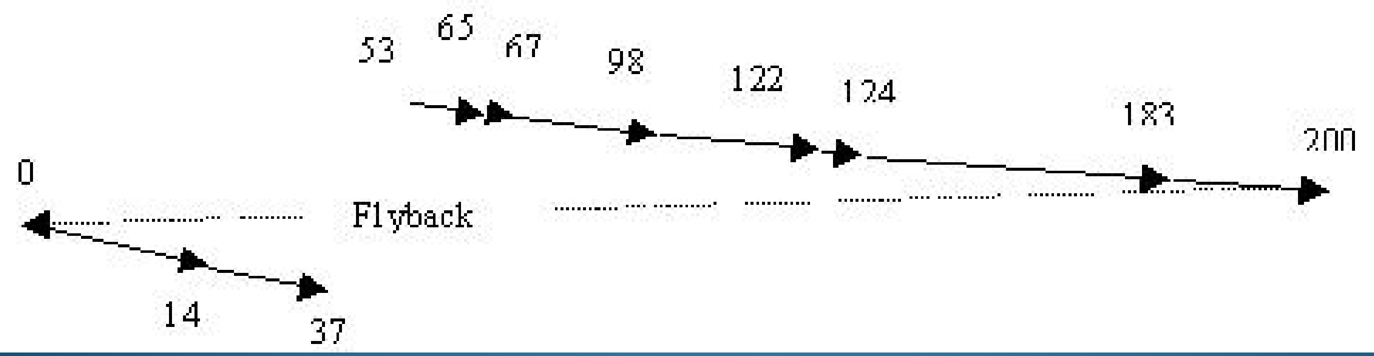
CS240 Operating Systems, Communications and Concurrency

98, 183, 37, 122, 14, 124, 65, 67

C-SCAN Algorithm

This algorithm attempts to regularise the distribution of service offered by SCAN. When the head has completed movement in one direction, it flys back quickly to the start, without scaning the surface, and begins the scan again in the same direction. Each track then has the same time between consecutive visits by the head.

With this algorithm, we get the following head movements:-



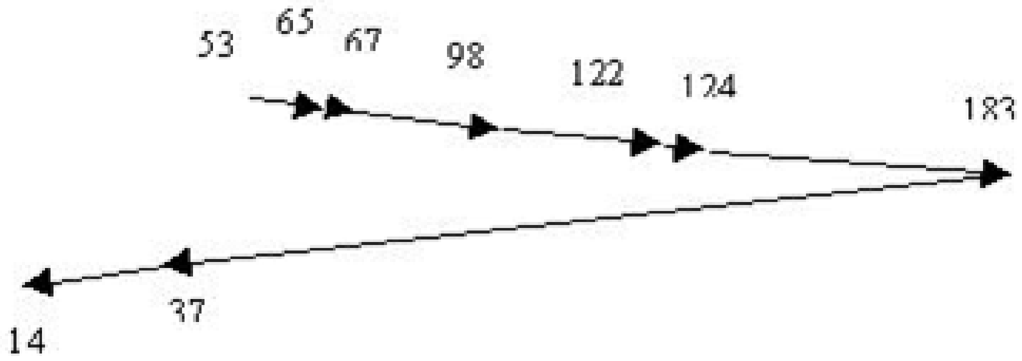
CS240 Operating Systems, Communications and Concurrency

98, 183, 37, 122, 14, 124, 65, 67

LOOK Algorithm (Elevator Algorithm)

The LOOK algorithm is the same as SCAN except that where SCAN always continues to the last track in a particular direction before turning around, the LOOK algorithm simply checks to see if there are anymore requests in that direction before turning around.

With this algorithm we get the following head movements:-



CS240 Operating Systems, Communications and Concurrency

98, 183, 37, 122, 14, 124, 65, 67

C-LOOK Algorithm

The C-LOOK algorithm is the same as C-SCAN except that where C-SCAN always continues to the last track in a particular direction before flying back, the C-LOOK algorithm simply checks to see if there are anymore requests in that direction before flying back.

With this algorithm we get the following head movements:-