Systems and Networking I

Applied Computer Science and Artificial Intelligence

2023-2024

Gabriele Tolomei

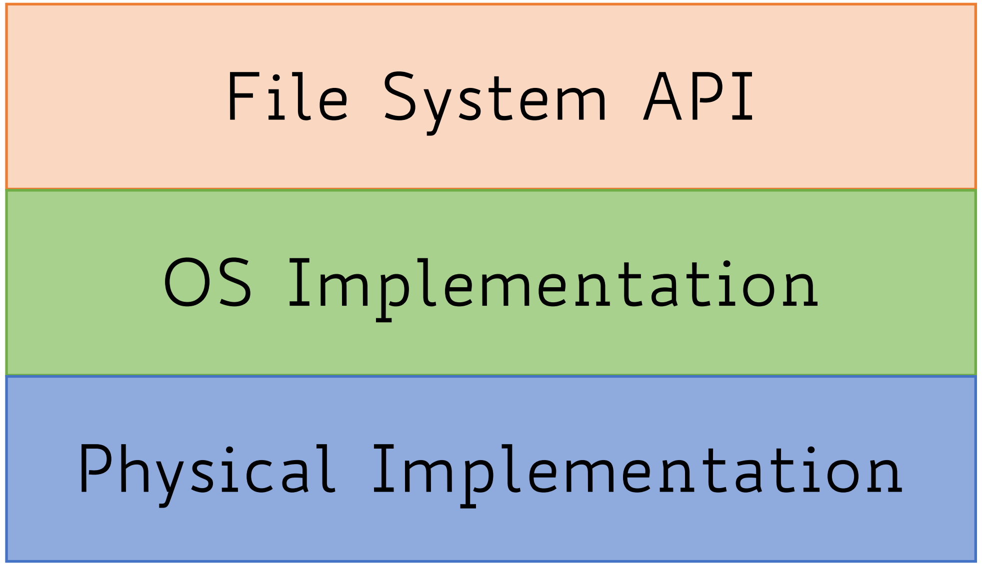
Computer Science Department

Sapienza University of Rome

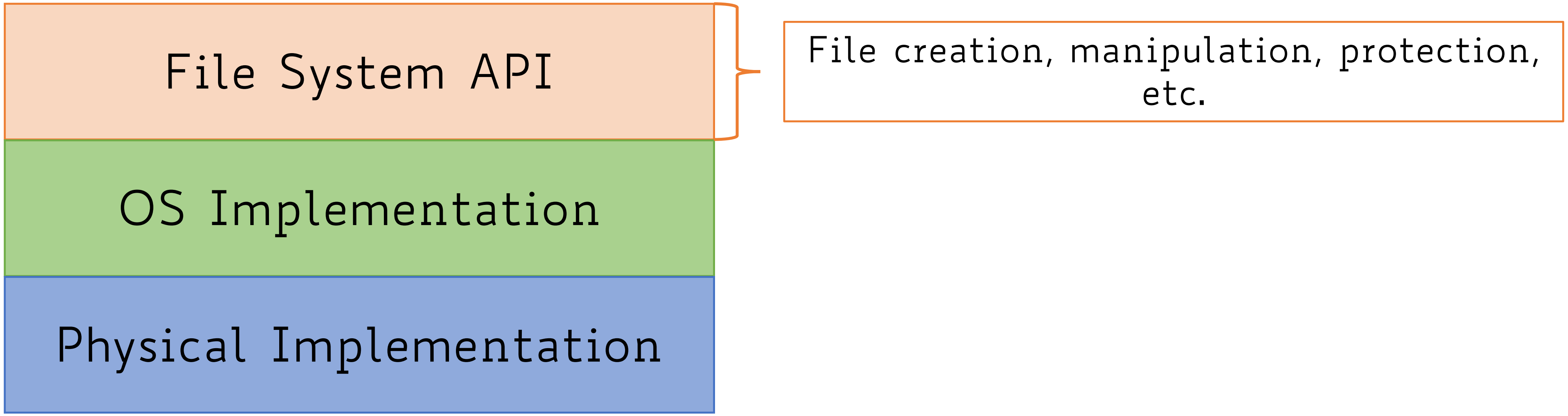
tolomei@di.uniroma1.it

SAPIENZA UNIVERSITÀ DI ROMA

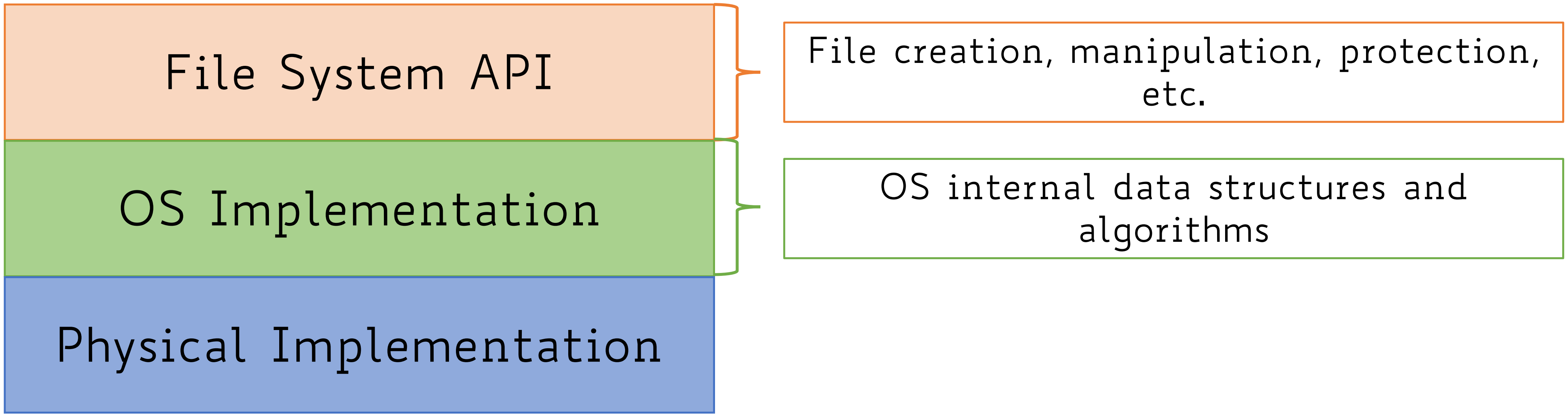
File System's Logical View



File System's Logical View



File System's Logical View



File System's Logical View

File System API	File creation, manipulation, protection, etc.
OS Implementation	OS internal data structures and algorithms
Physical Implementation	Second storage structure, disk scheduling algorithms

File System's Logical View

File System API	File creation, manipulation, protection, etc.
OS Implementation	OS internal data structures and algorithms
Physical Implementation	Second storage structure, disk scheduling algorithms

Part V: Storage Management

Overview of Mass-Storage Structure

3 categories of mass-storage devices

Overview of Mass-Storage Structure

3 categories of mass-storage devices

Magnetic Disks

Overview of Mass-Storage Structure

3 categories of mass-storage devices

Magnetic Disks

Solid-State Disks

Overview of Mass-Storage Structure

3 categories of mass-storage devices

Magnetic Disks

Solid-State Disks

Magnetic Tapes

Overview of Mass-Storage Structure

3 categories of mass-storage devices

Magnetic Disks

Solid-State Disks

Magnetic Tapes

Secondary Storage

Overview of Mass-Storage Structure

3 categories of mass-storage devices

Magnetic Disks

Solid-State Disks

Magnetic Tapes

Secondary Storage

Tertiary Storage

Overview of Mass-Storage Structure

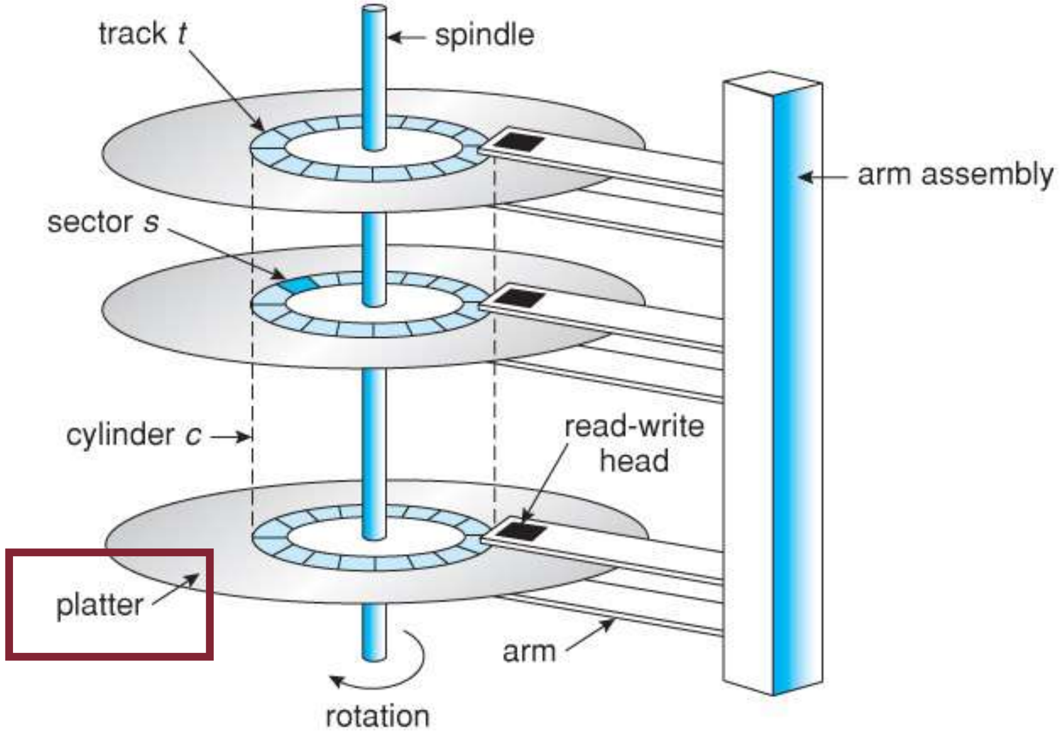
3 categories of mass-storage devices

Magnetic Disks

Solid-State Disks

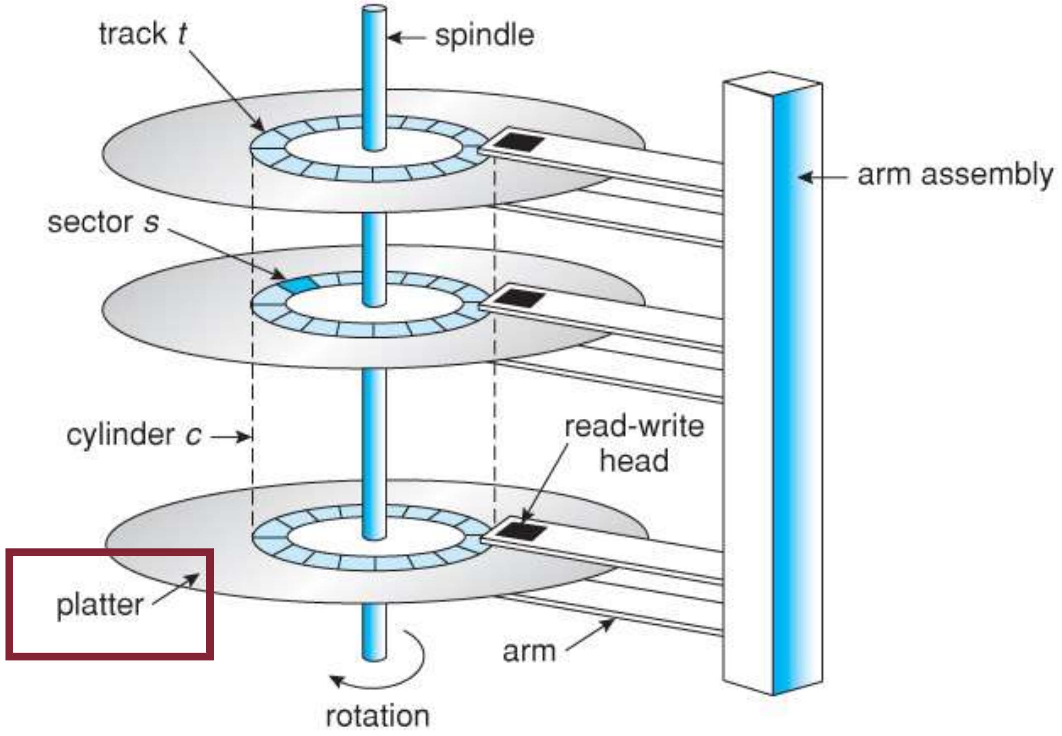
Magnetic Tapes

Magnetic Disks: Structure



One or more platters covered with magnetic media

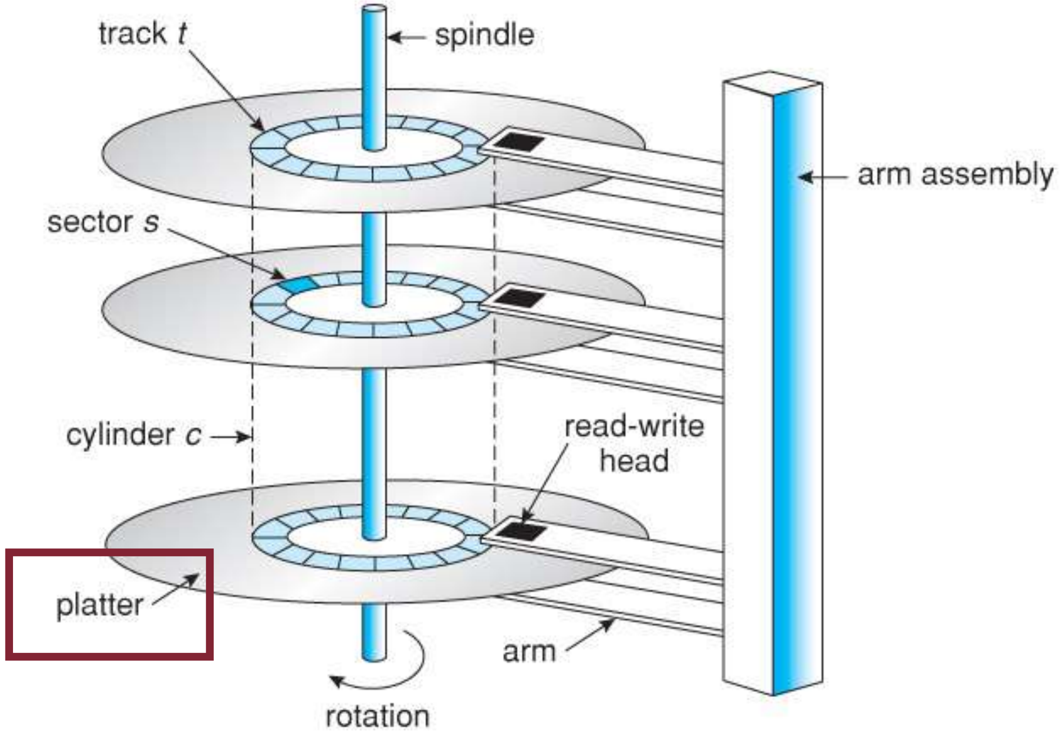
Magnetic Disks: Structure



One or more platters covered with magnetic media

Hard disk
rigid metal

Magnetic Disks: Structure

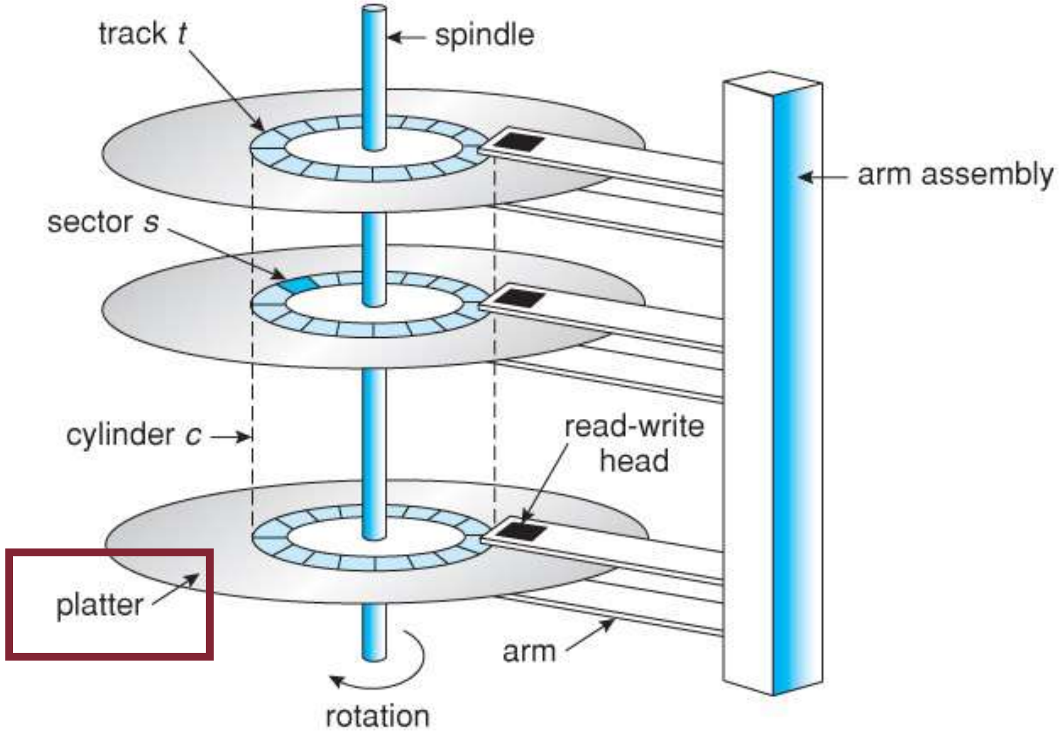


One or more platters covered with magnetic media

Hard disk
rigid metal

Floppy disk
flexible plastic

Magnetic Disks: Structure



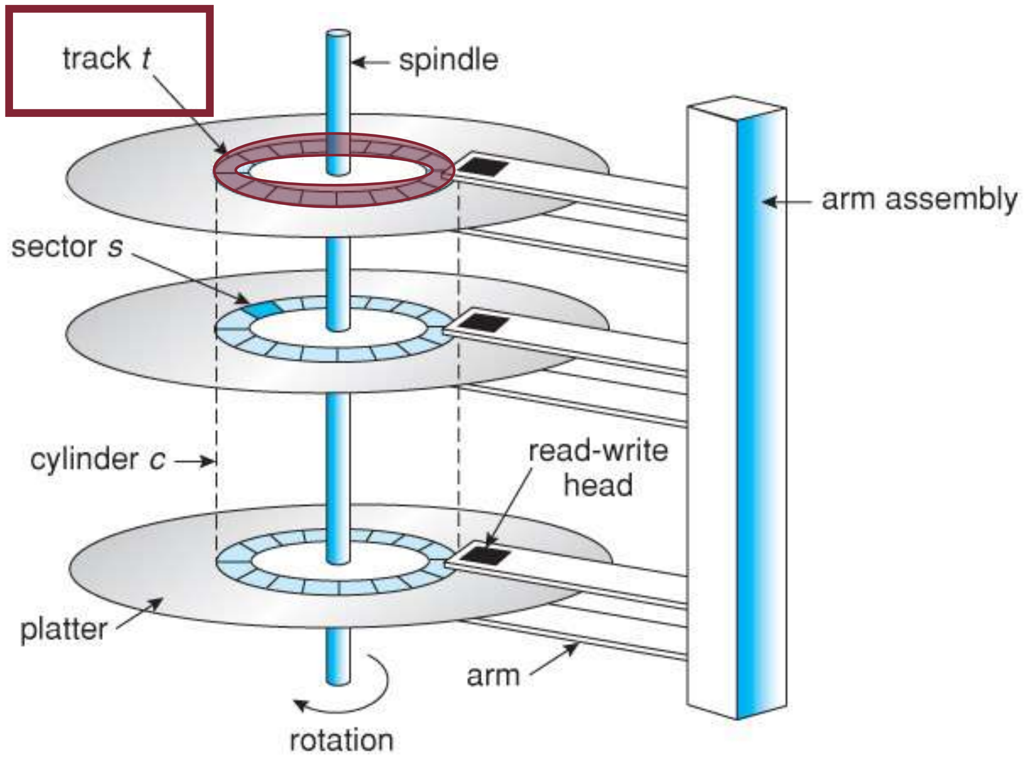
One or more platters covered with magnetic media

Hard disk
rigid metal

Floppy disk
flexible plastic

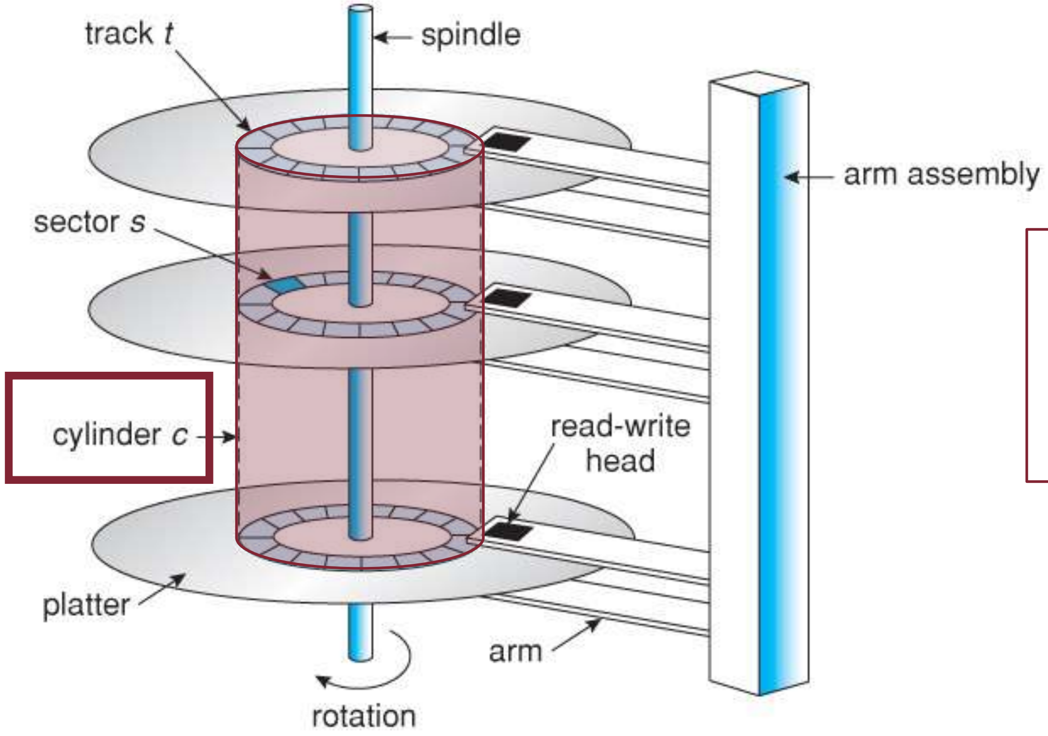
Each platter has 2 working surfaces

Magnetic Disks: Tracks and Cylinders



Each surface is divided into a number of concentric rings, called tracks

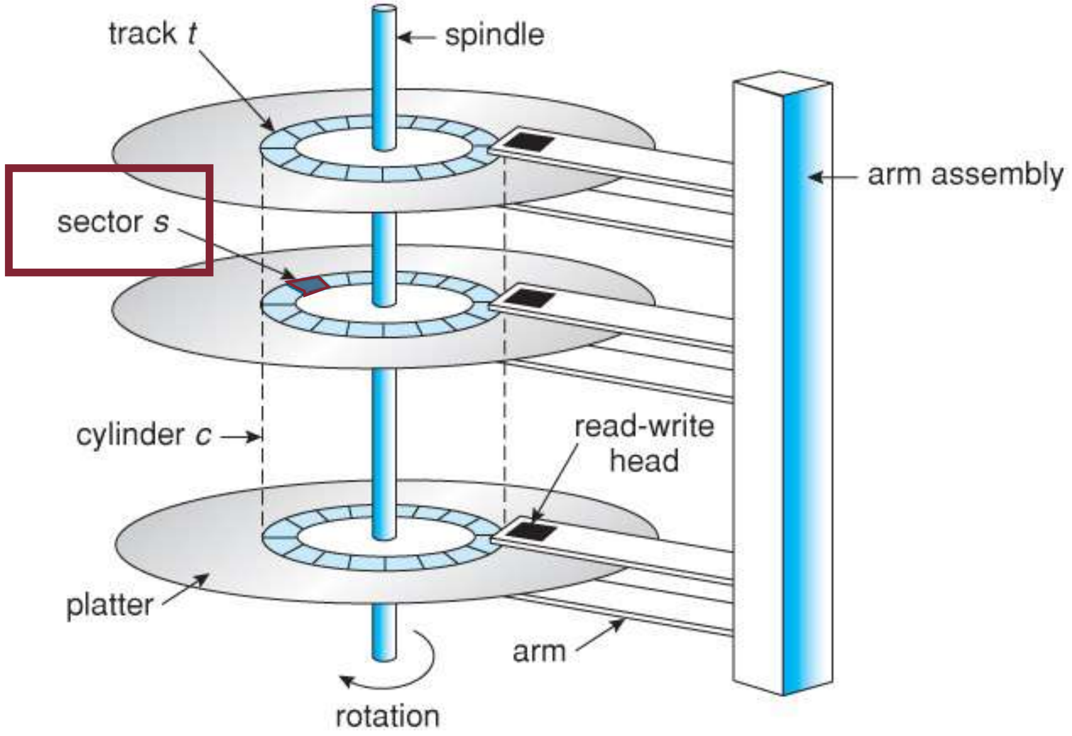
Magnetic Disks: Tracks and Cylinders



Each surface is divided into a number of concentric rings, called tracks

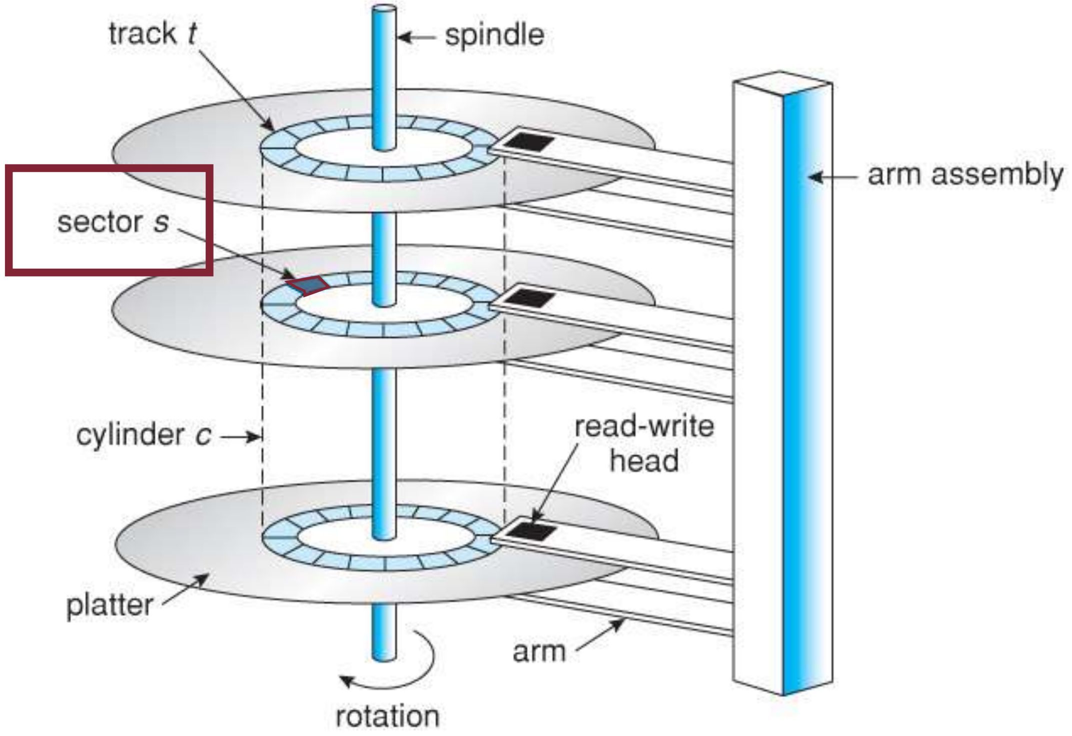
The set of all tracks that are the same distance from the edge of the platter is called a cylinder

Magnetic Disks: Sectors



Each track is further divided into sectors

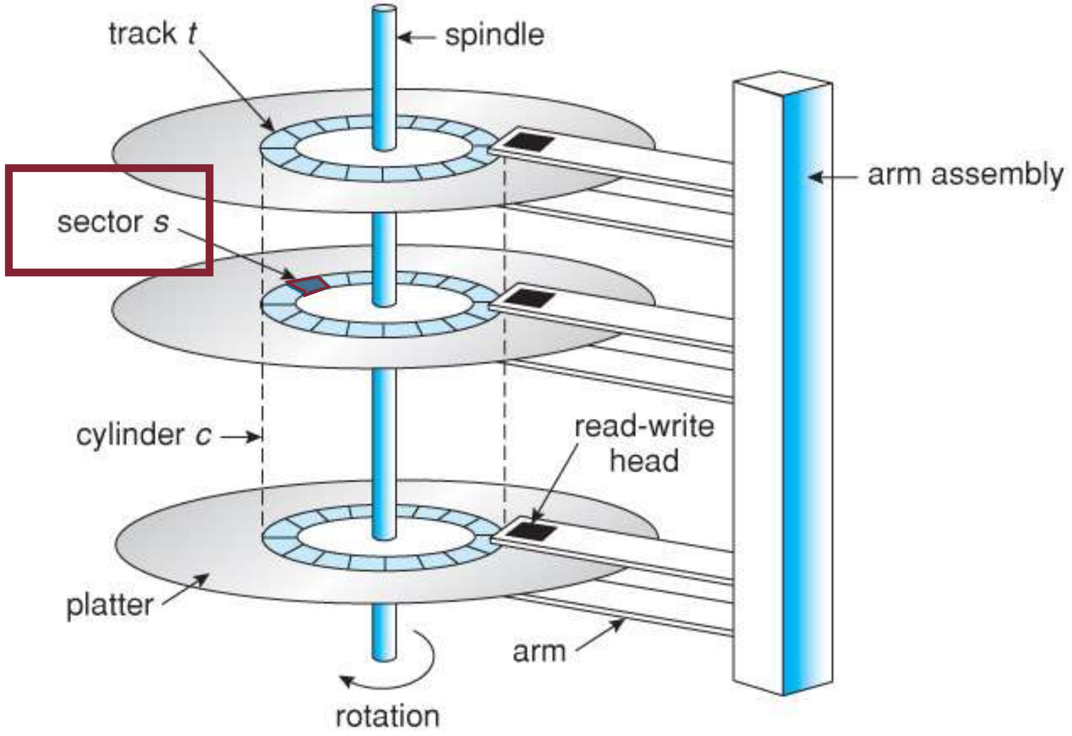
Magnetic Disks: Sectors



Each track is further divided into sectors

Each sector usually contains 512 bytes

Magnetic Disks: Sectors

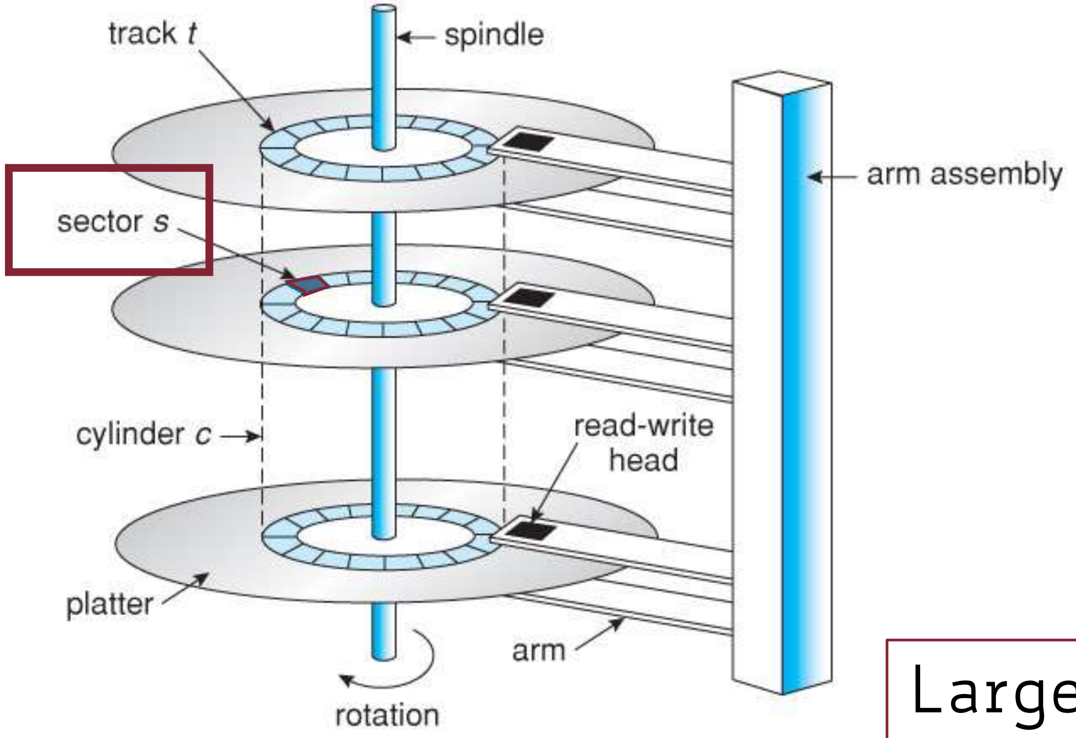


Each track is further divided into sectors

Each sector usually contains 512 bytes

Sectors also include a header and a trailer, and checksum information

Magnetic Disks: Sectors



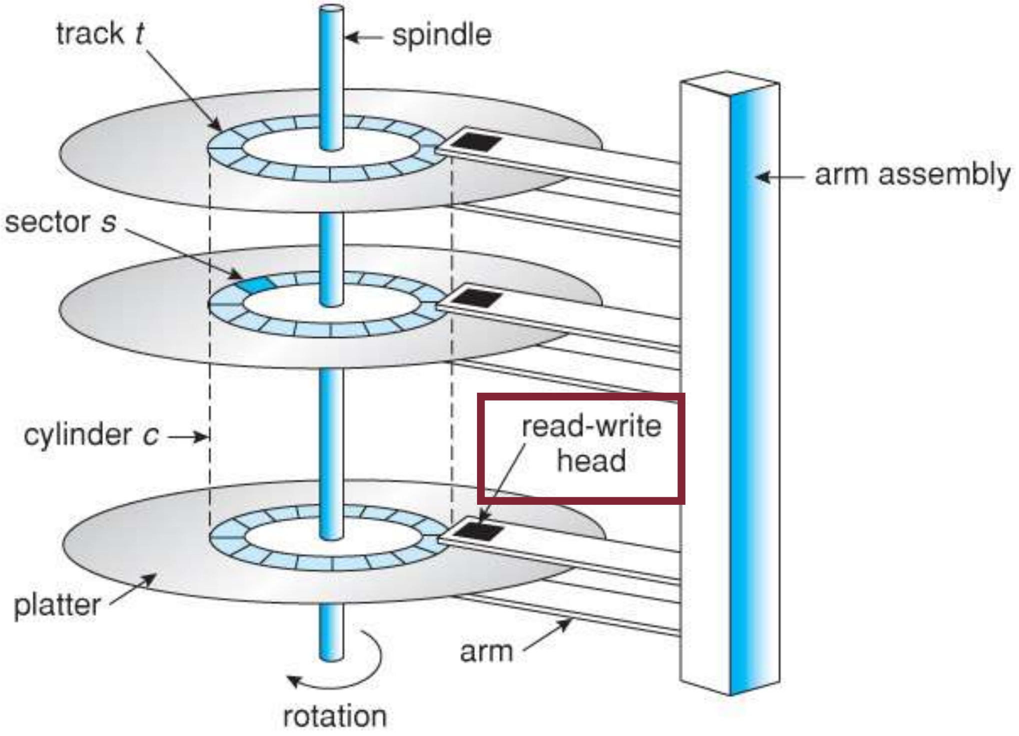
Each track is further divided into sectors

Each sector usually contains 512 bytes

Sectors also include a header and a trailer, and checksum information

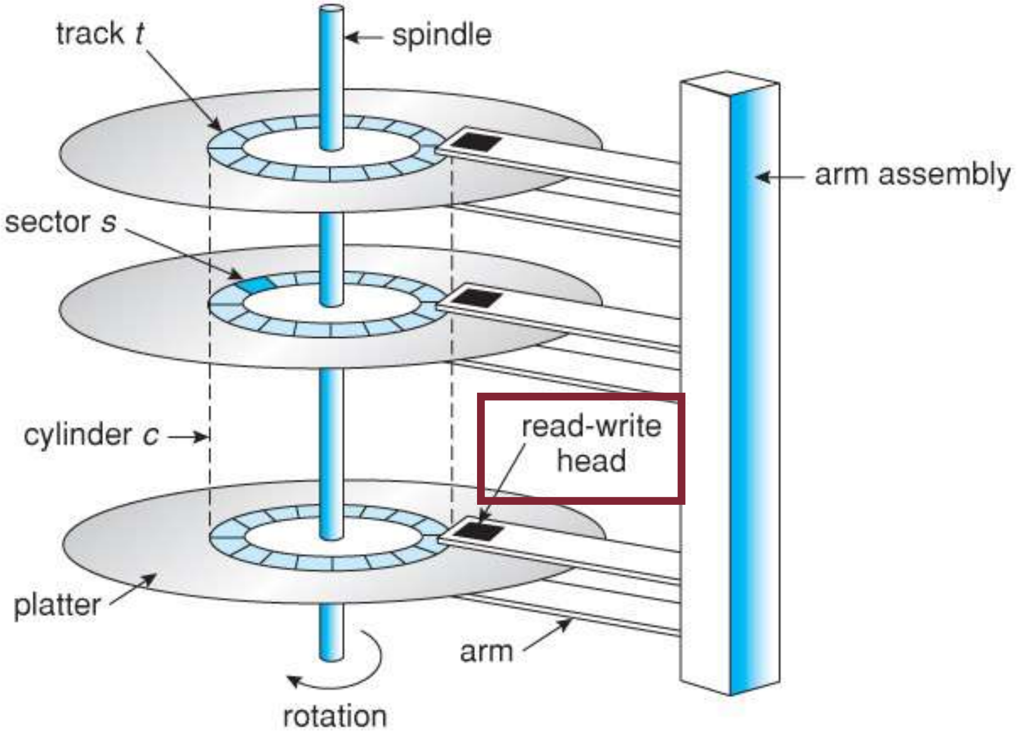
Larger sector sizes reduce the space wasted by headers and trailers, but increase internal fragmentation

Magnetic Disks: Heads



Data on hard drive is read by read-write heads

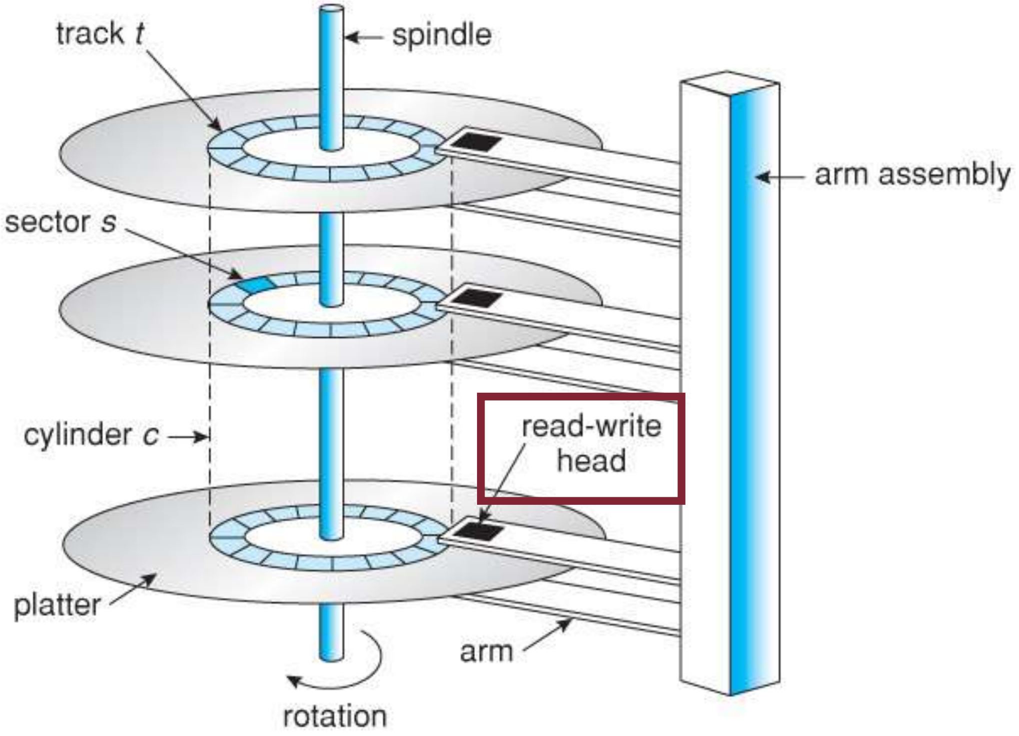
Magnetic Disks: Heads



Data on hard drive is read by read-write heads

Standard configuration uses one head per surface

Magnetic Disks: Heads

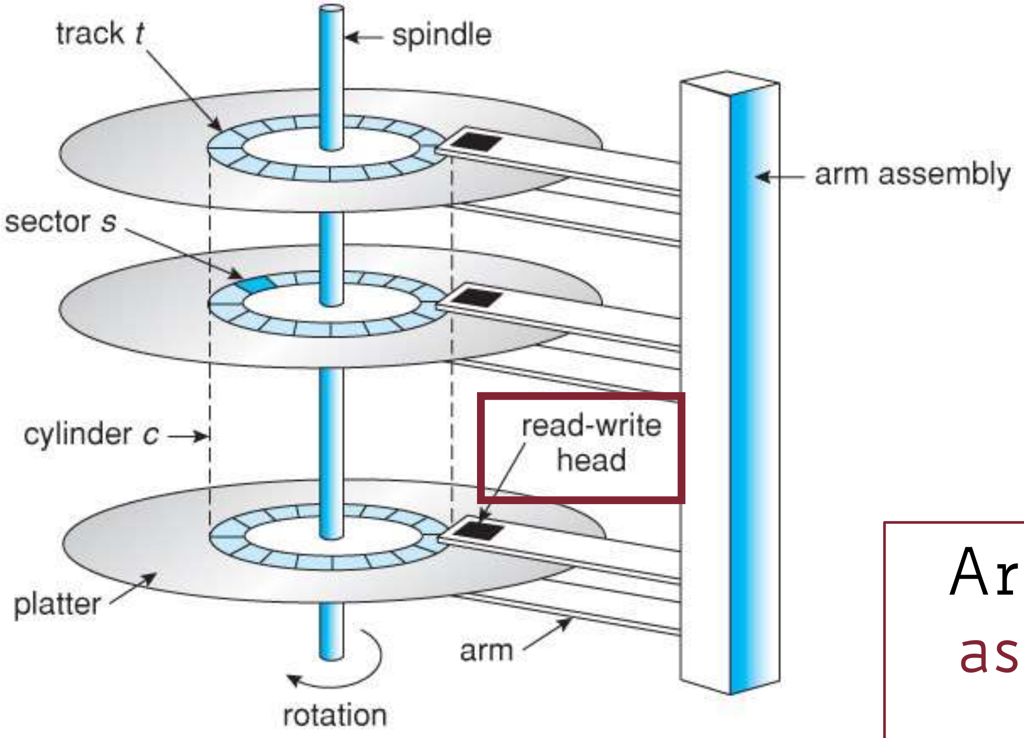


Data on hard drive is read by read-write heads

Standard configuration uses one head per surface

Each head is placed on a separate arm

Magnetic Disks: Heads



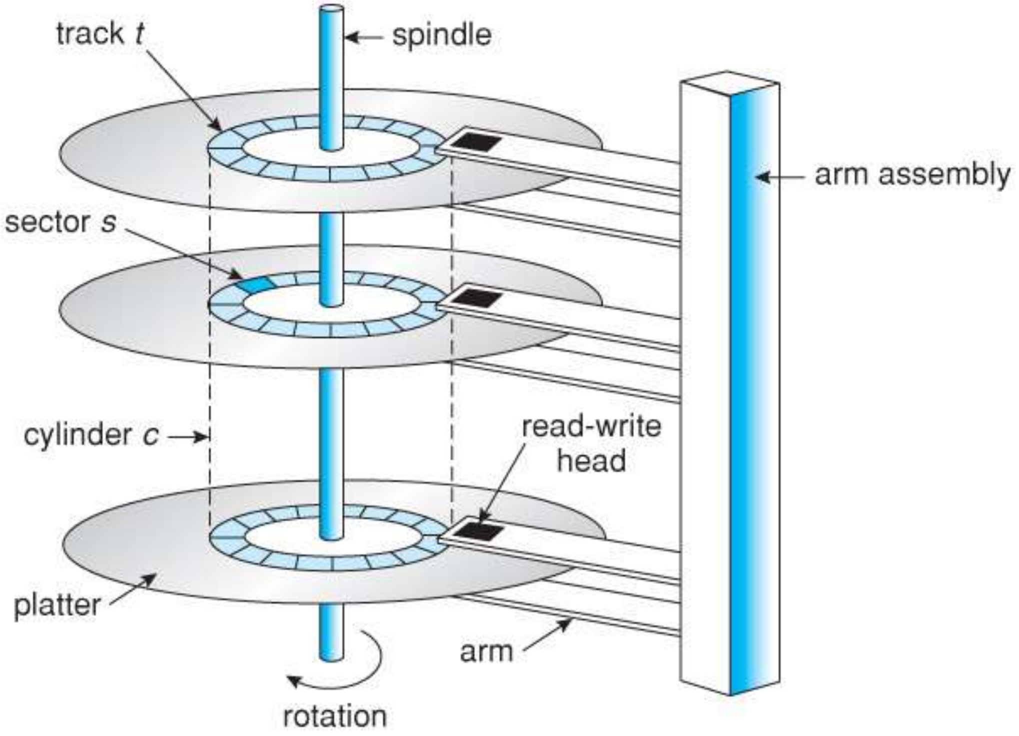
Data on hard drive is read by read-write heads

Standard configuration uses one head per surface

Each head is placed on a separate arm

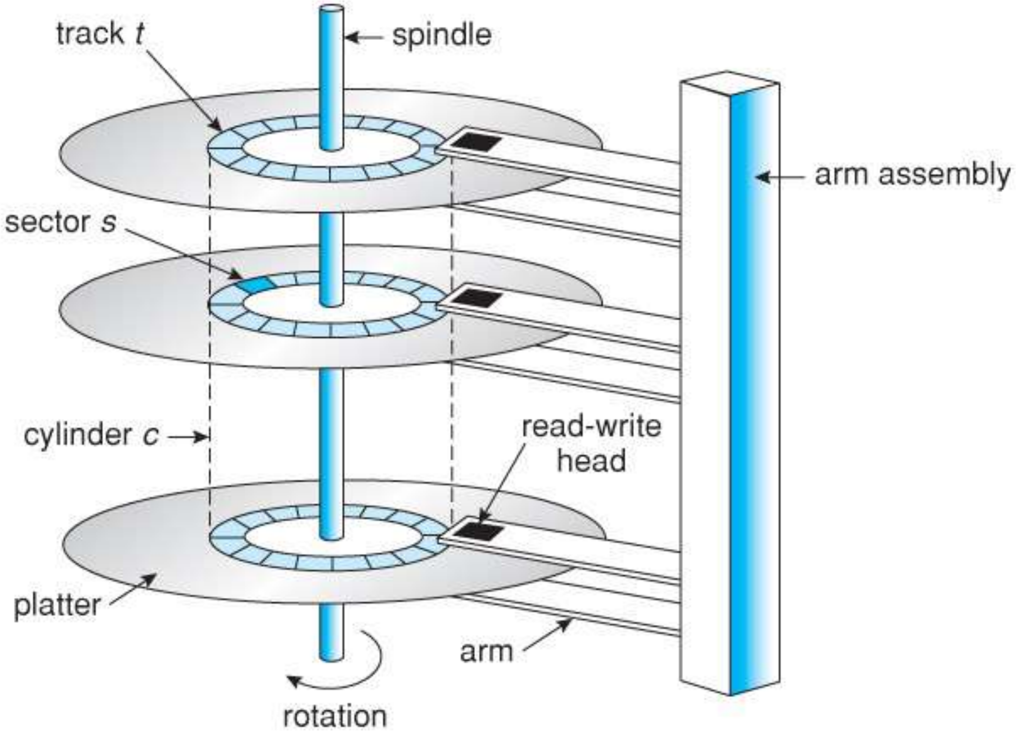
Arms are controlled by a common arm assembly moving simultaneously from one cylinder to another

Magnetic Disks: Storage Capacity



H = number of heads (working surfaces)

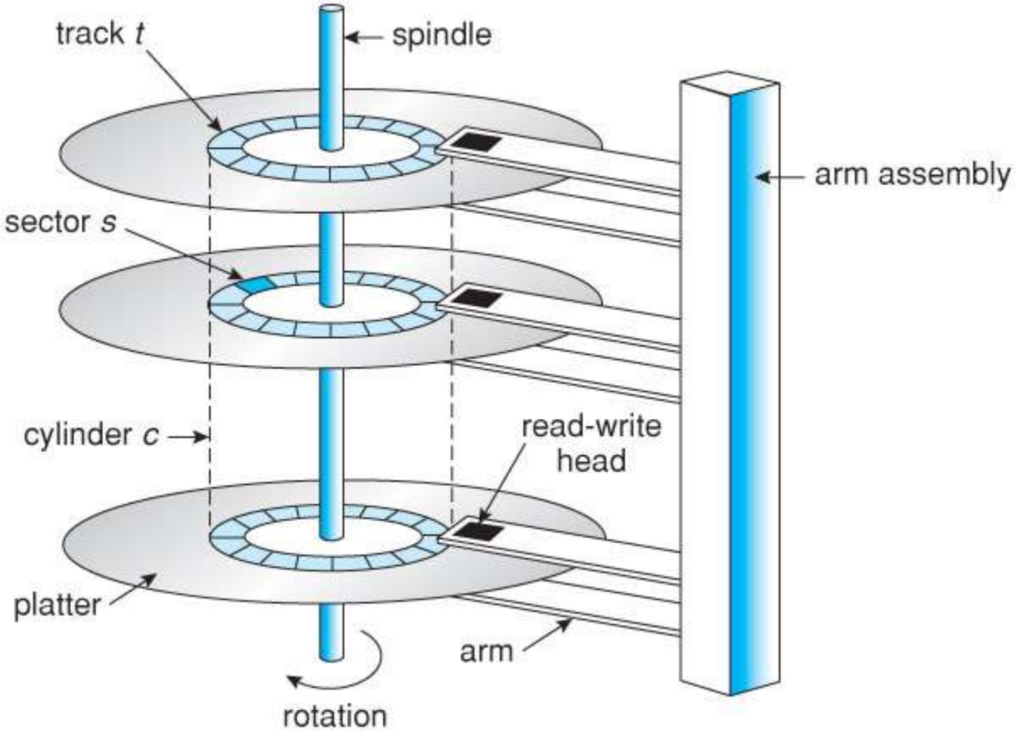
Magnetic Disks: Storage Capacity



H = number of heads (working surfaces)

T = number of tracks per surface

Magnetic Disks: Storage Capacity

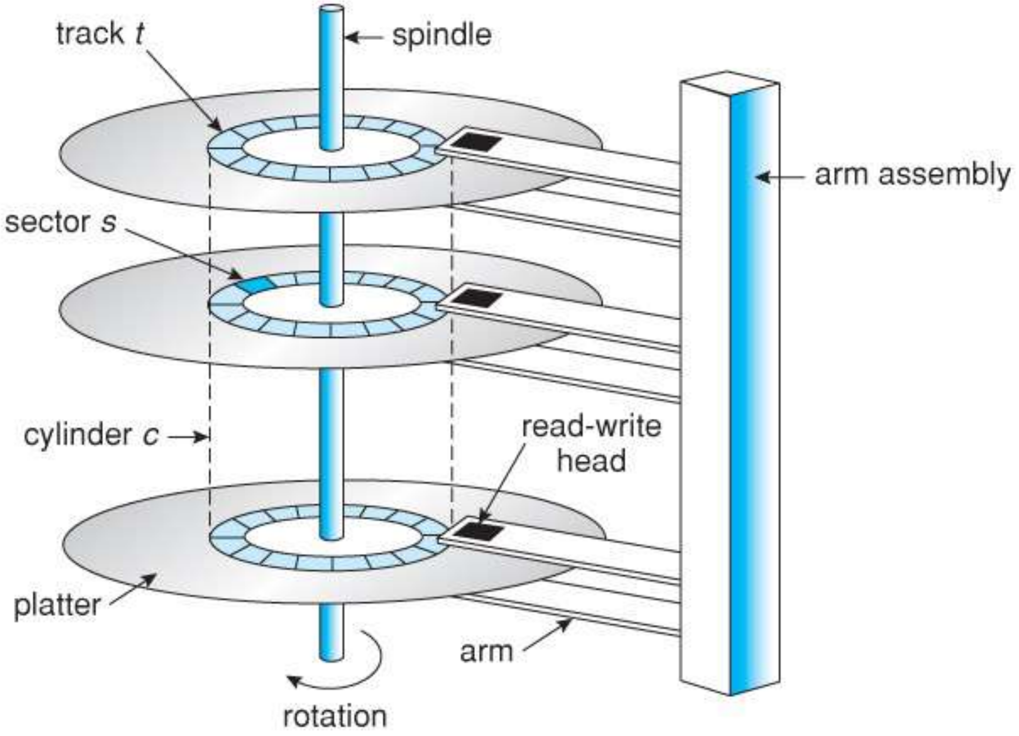


H = number of heads (working surfaces)

T = number of tracks per surface

S = number of sectors per track

Magnetic Disks: Storage Capacity



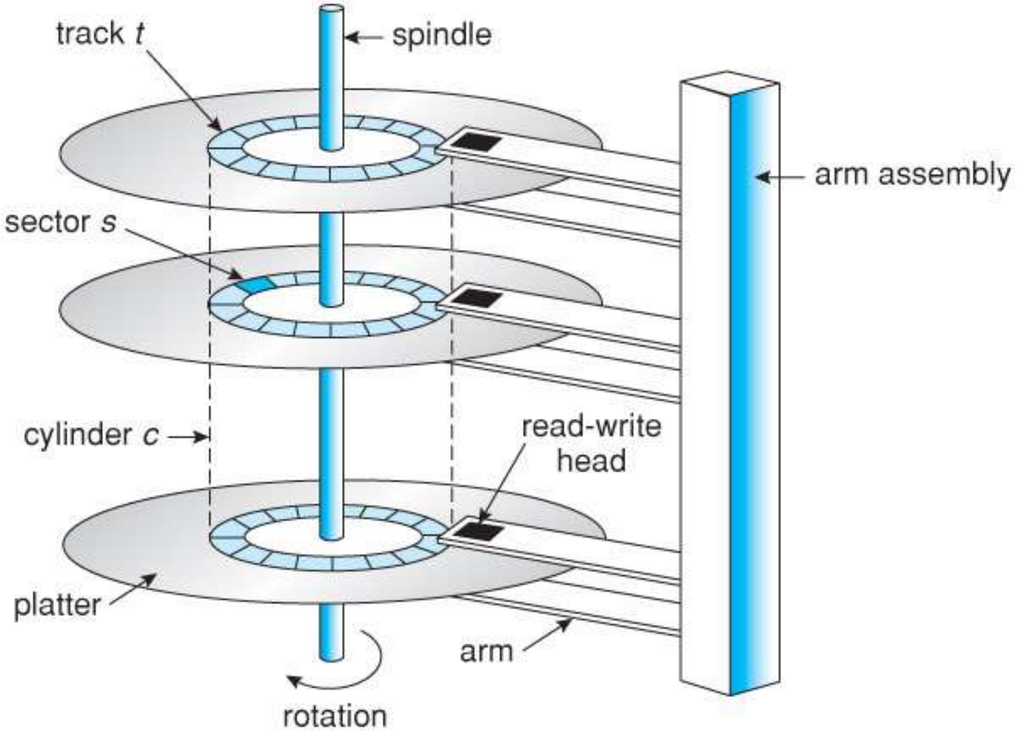
H = number of heads (working surfaces)

T = number of tracks per surface

S = number of sectors per track

B = number of bytes per sector

Magnetic Disks: Storage Capacity



H = number of heads (working surfaces)

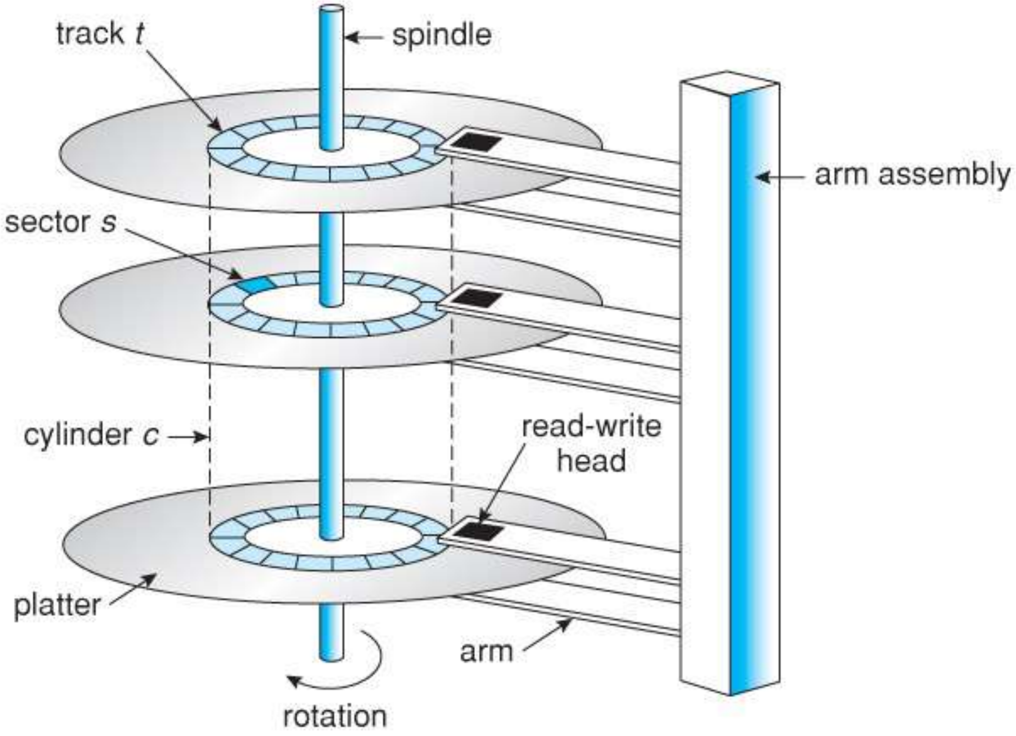
T = number of tracks per surface

S = number of sectors per track

B = number of bytes per sector

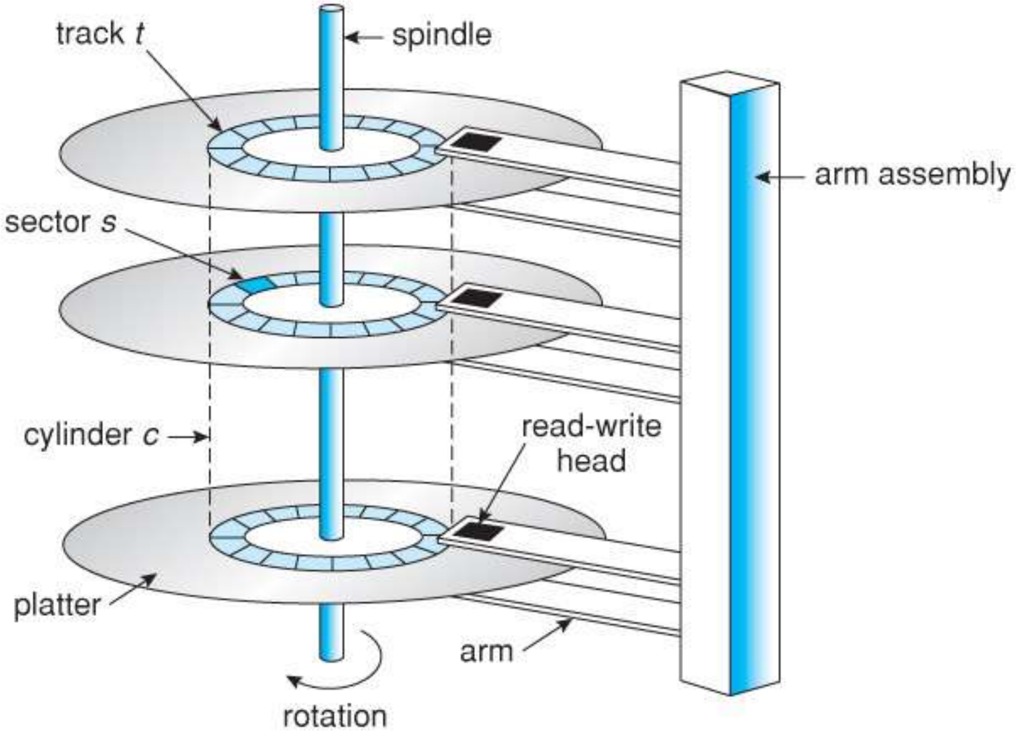
$$\underbrace{C = H * T * S * B}_{\text{OVERALL CAPACITY}}$$

Magnetic Disks: Storage Capacity



Until the end of 1980s, every track had the same number of sectors with the same number of bits

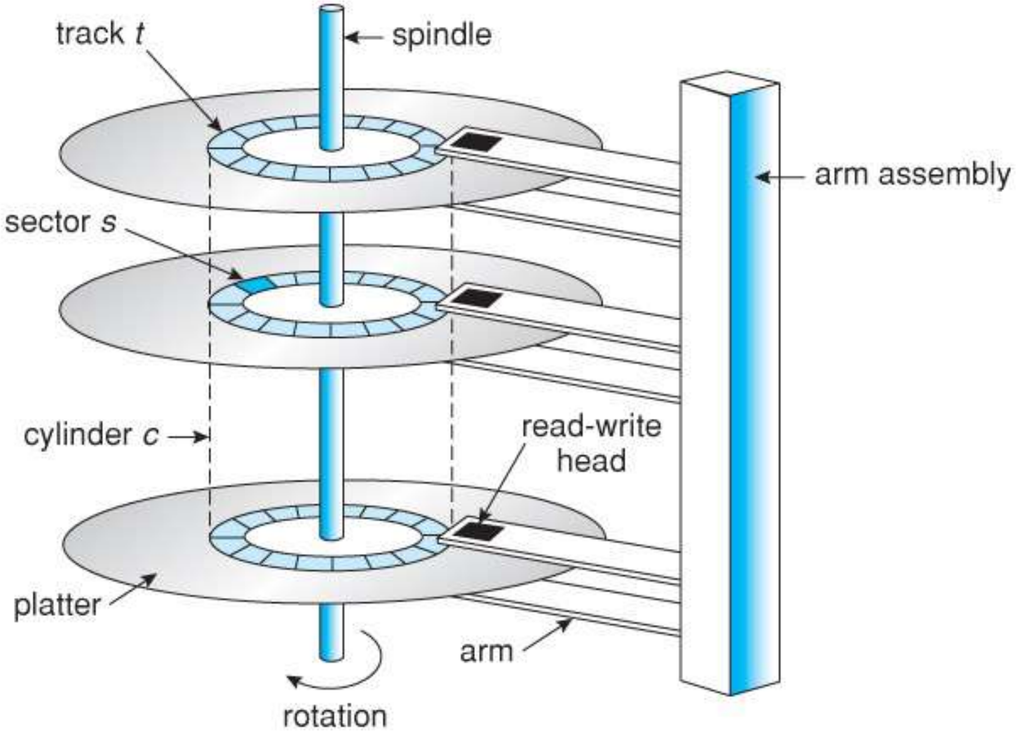
Magnetic Disks: Storage Capacity



Until the end of 1980s, every track had the same number of sectors with the same number of bits

Therefore, the bit density in the inner sectors was much higher than in the outer sectors

Magnetic Disks: Storage Capacity

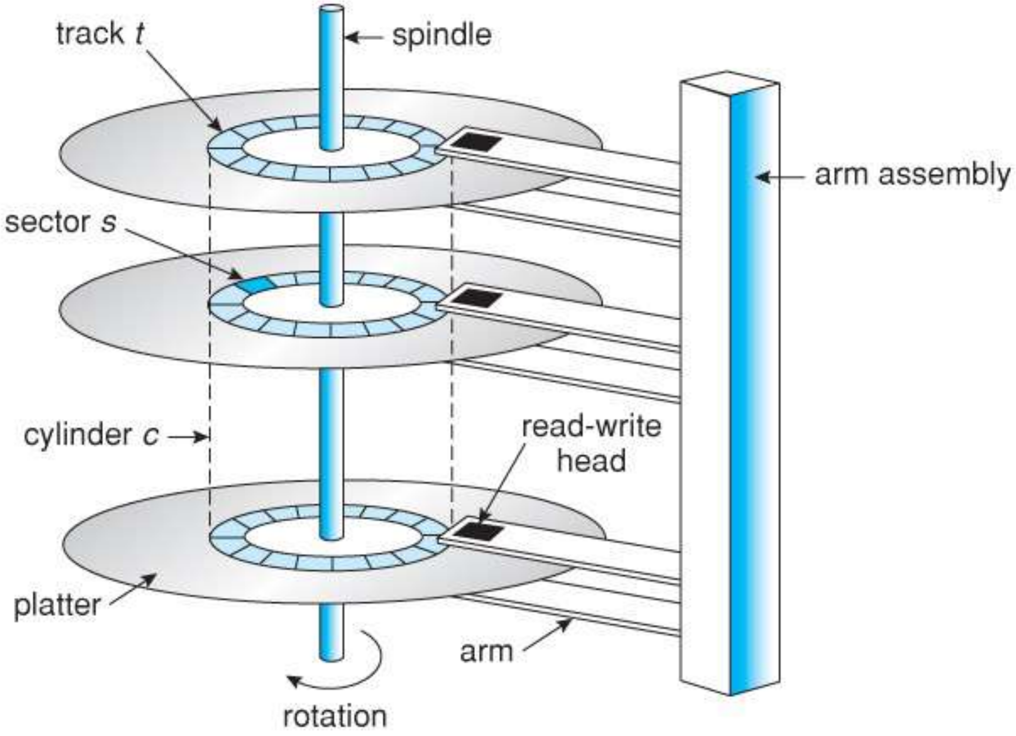


Until the end of 1980s, every track had the same number of sectors with the same number of bits

Therefore, the bit density in the inner sectors was much higher than in the outer sectors

Disk controllers have no "intelligence"

Magnetic Disks: Storage Capacity



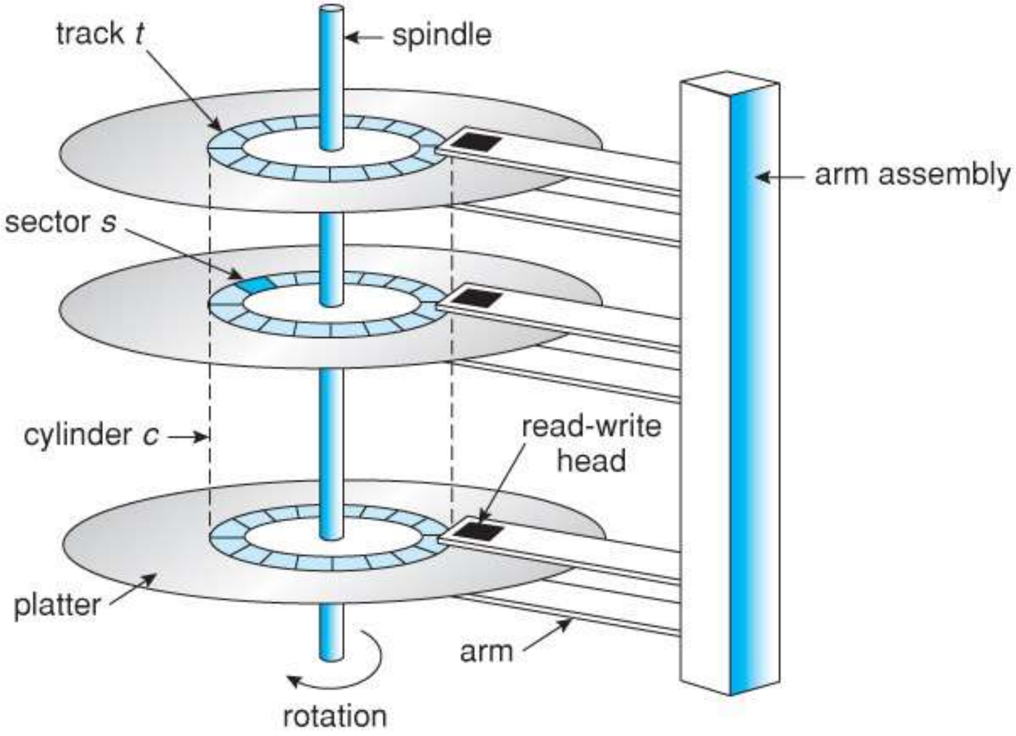
Until the end of 1980s, every track had the same number of sectors with the same number of bits

Therefore, the bit density in the inner sectors was much higher than in the outer sectors

Disk controllers have no "intelligence"

Drawbacks:
- The capacity of the disk was determined by the maximum bit density a controller could handle

Magnetic Disks: Storage Capacity



Until the end of 1980s, every track had the same number of sectors with the same number of bits

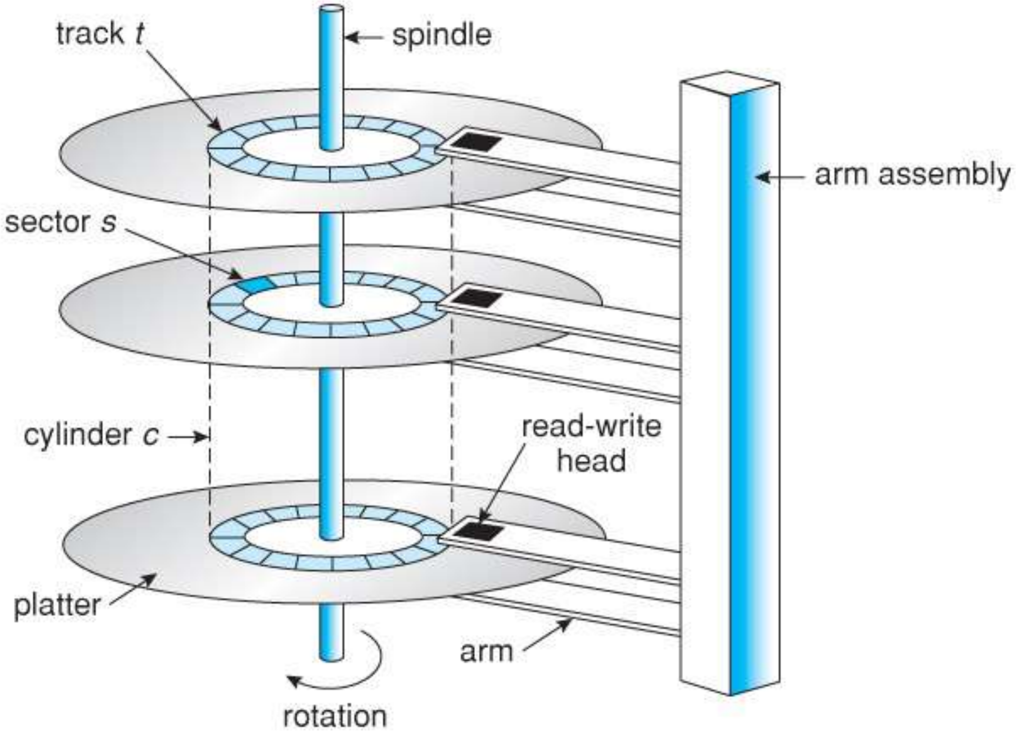
Therefore, the bit density in the inner sectors was much higher than in the outer sectors

Disk controllers have no "intelligence"

Drawbacks:

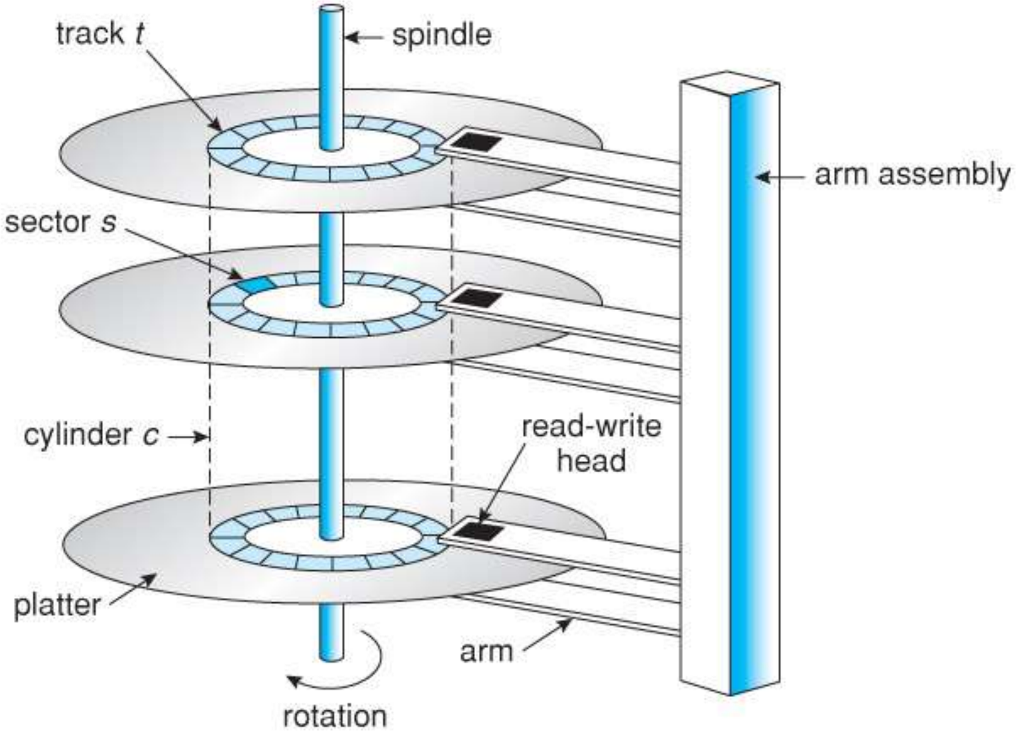
- The capacity of the disk was determined by the maximum bit density a controller could handle
- Different frequencies and timing from innermost to outermost tracks

Magnetic Disks: Storage Capacity



In practice, the number of sectors per track (S) varies with the radius of the track on the platter

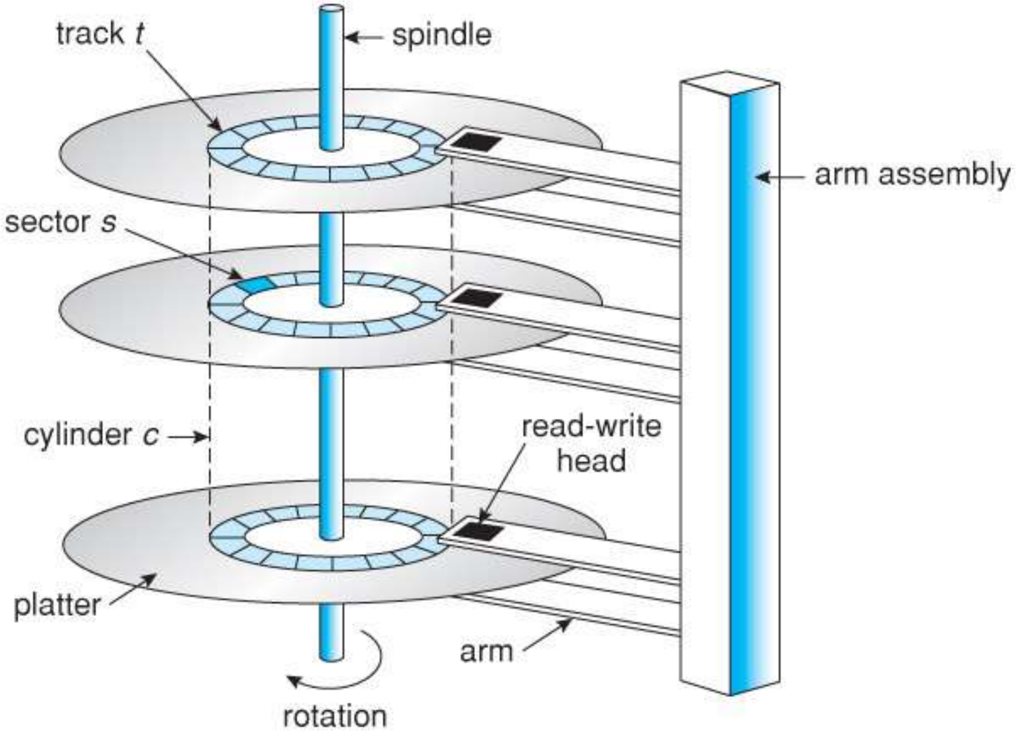
Magnetic Disks: Storage Capacity



In practice, the number of sectors per track (S) varies with the radius of the track on the platter

The outermost track is larger and can hold more sectors than the inner ones

Magnetic Disks: Storage Capacity

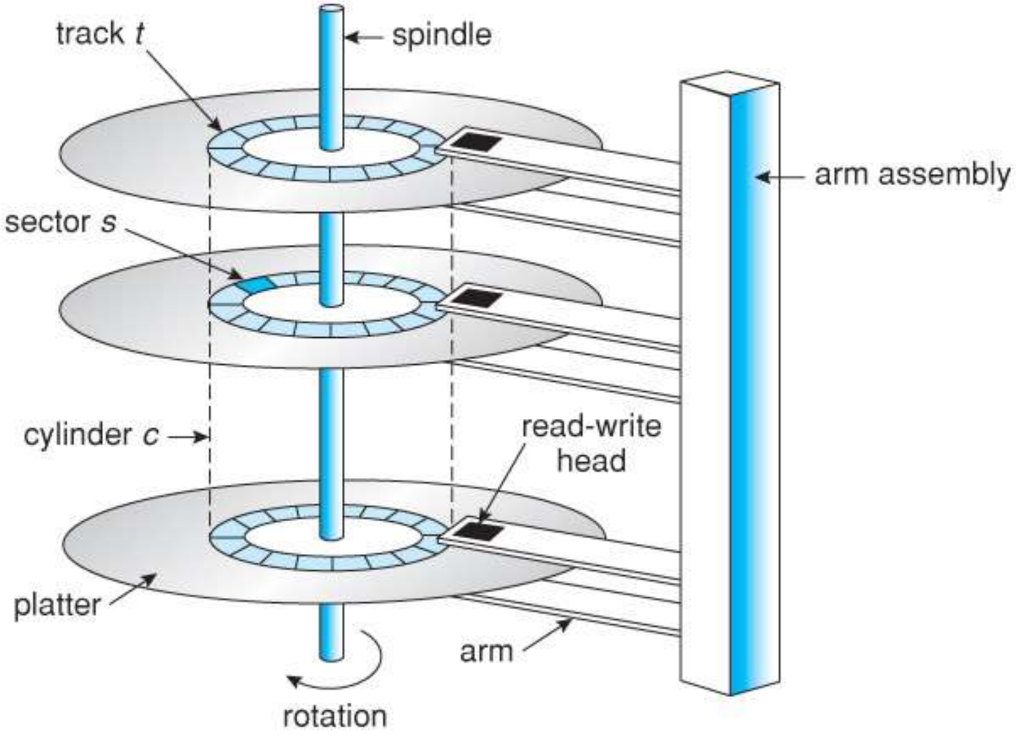


In practice, the number of sectors per track (S) varies with the radius of the track on the platter

The outermost track is larger and can hold more sectors than the inner ones

The bit density is kept almost constant

Magnetic Disks: Storage Capacity



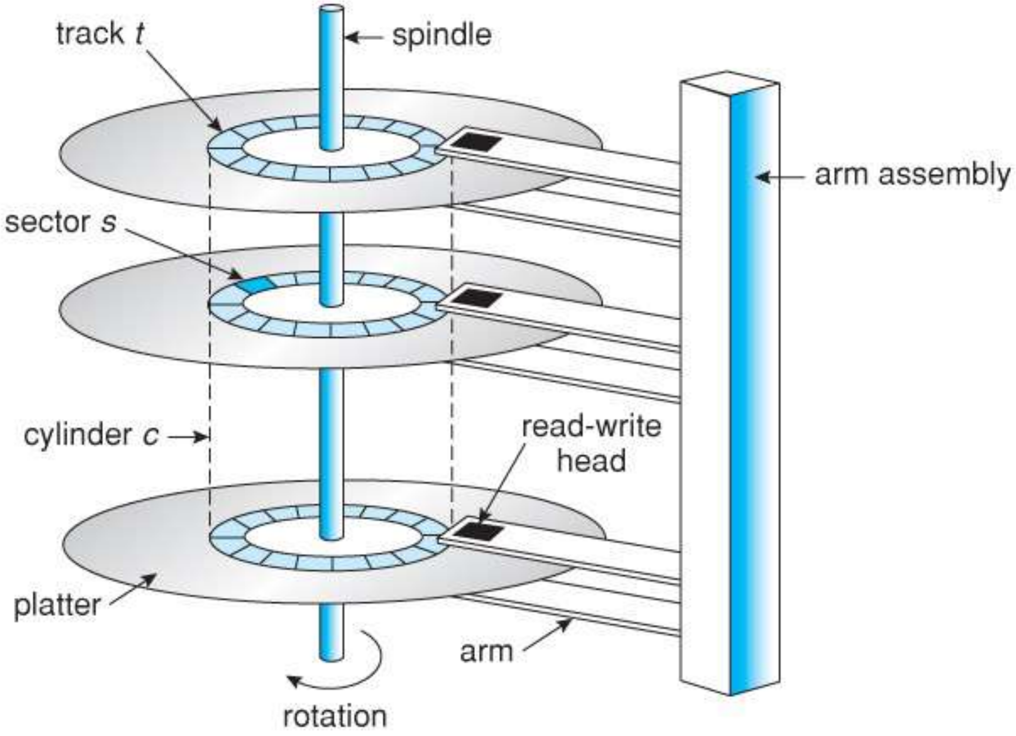
In practice, the number of sectors per track (S) varies with the radius of the track on the platter

The outermost track is larger and can hold more sectors than the inner ones

The bit density is kept almost constant

Smarter disk controllers allow for logical addressing of sectors rather than physical

Magnetic Disks: Storage Capacity



In practice, the number of sectors per track (S) varies with the radius of the track on the platter

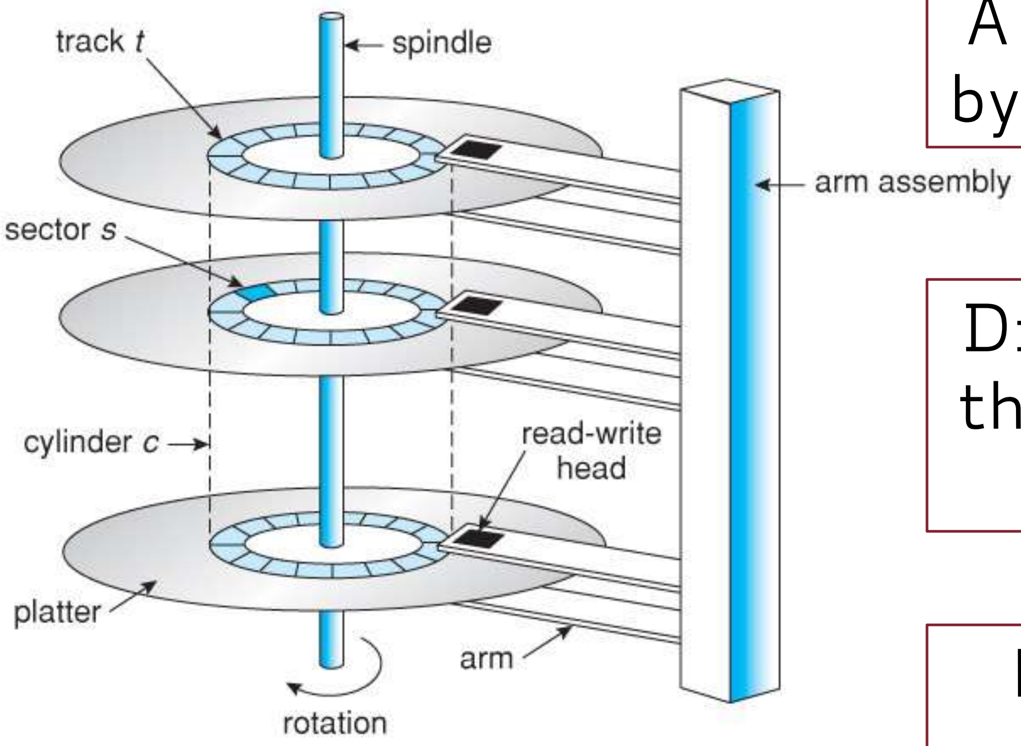
The outermost track is larger and can hold more sectors than the inner ones

The bit density is kept almost constant

Smarter disk controllers allow for logical addressing of sectors rather than physical

Zone Bit Recording (ZBR)

Magnetic Disks: (Logical) Referencing



A physical block of data is specified by the (head, cylinder, sector) number

Disk blocks are numbered starting at the outermost cylinder, identified by 0

Note that cylinder coincides with track

Magnetic Disks: Data Transfer

- The disk rotates at constant angular speed (e.g., 7200 rpm = 120 rps)

Magnetic Disks: Data Transfer

- The disk rotates at **constant angular speed** (e.g., **7200 rpm** = **120 rps**)
- Outer tracks spin **faster** than inner tracks (more sectors traversed in the same amount of time due to larger radius → more sectors per zone in ZBR)

Magnetic Disks: Data Transfer

- Data transfer from the disk to memory is made of 3 **steps**:
  - positioning time (seek time or random access time)
  - rotational delay
  - transfer time

Magnetic Disks: Data Transfer

- Data transfer from the disk to memory is made of 3 **steps**:
  - positioning time (seek time or random access time)
  - rotational delay
  - transfer time

mechanical

Magnetic Disks: Data Transfer

- Data transfer from the disk to memory is made of 3 **steps**:
  - positioning time (seek time or random access time)
  - rotational delay
  - transfer time

mechanical

electronic

Magnetic Disks: Positioning (Seek) Time

- The time required to move the heads to a specific track/cylinder

Magnetic Disks: Positioning (Seek) Time

- The time required to move the heads to a specific track/cylinder
- Includes the time needed for the heads to settle

Magnetic Disks: Positioning (Seek) Time

- The time required to move the heads to a specific track/cylinder
- Includes the time needed for the heads to settle
- Depends on how fast the hardware moves the arm

Magnetic Disks: Positioning (Seek) Time

- The time required to move the heads to a specific track/cylinder
- Includes the time needed for the heads to settle
- Depends on how fast the hardware moves the arm
- Typically, the slowest step in the entire process

Bottleneck of overall disk data transfer

Magnetic Disks: Rotational Delay

- The time required for the desired sector to rotate and come under the read-write head

Magnetic Disks: Rotational Delay

- The time required for the desired sector to rotate and come under the read-write head
- Can range from 0 up to one full revolution
  - 0 → the sector is already underneath the head
  - full revolution → the sector is the one before but in the opposite direction

Magnetic Disks: Rotational Delay

- The time required for the desired sector to rotate and come under the read-write head
- Can range from 0 up to one full revolution
  - 0 → the sector is already underneath the head
  - full revolution → the sector is the one before but in the opposite direction
- On average, **0.5 revolutions** (r)
  - E.g., for a 7200 rpm (120 rps) disk this equals to 0.5 r/120 rps **~4 msec**

The second highest bottleneck

Magnetic Disks: Transfer Time

- The time required to move data (i.e., bytes) electronically from disk to memory

Magnetic Disks: Transfer Time

- The time required to move data (i.e., bytes) electronically from disk to memory
- This is sometimes expressed as **transfer rate** (bandwidth) in bytes per second

Magnetic Disks: Transfer Time

- The time required to move data (i.e., bytes) electronically from disk to memory
- This is sometimes expressed as **transfer rate** (bandwidth) in bytes per second

Data Transfer Time = Seek Time + Rotational Delay + Transfer Time

Sometimes the term transfer rate is used to refer to the overall data transfer time

Magnetic Disks: Structure

- Addressed as large one-dimensional arrays of **logical blocks**

Magnetic Disks: Structure

- Addressed as large one-dimensional arrays of **logical blocks**
- Each logical block is the smallest unit of transfer (e.g., **512 bytes**)

Magnetic Disks: Structure

- Addressed as large one-dimensional arrays of **logical blocks**
- Each logical block is the smallest unit of transfer (e.g., **512 bytes**)
- The array of blocks is mapped onto disk sectors sequentially

Magnetic Disks: Structure

- Sector 0 is the first sector of the first track of the outermost cylinder

Magnetic Disks: Structure

- Sector 0 is the first sector of the first track of the outermost cylinder
- The mapping proceeds in order through that track

Magnetic Disks: Structure

- Sector 0 is the first sector of the first track of the outermost cylinder
- The mapping proceeds in order through that track
- Then through the rest of tracks in the same cylinder

Magnetic Disks: Structure

- Sector 0 is the first sector of the first track of the outermost cylinder
- The mapping proceeds in order through that track
- Then through the rest of tracks in the same cylinder
- Then through other cylinders (from the outermost to innermost)

Magnetic Disks: Risks

- Disk heads "fly" over the surface on a very thin cushion of air

Magnetic Disks: Risks

- Disk heads "fly" over the surface on a very thin cushion of air
- If they accidentally contact the disk then a **head crash** occurs

Magnetic Disks: Risks

- Disk heads "fly" over the surface on a very thin cushion of air
- If they accidentally contact the disk then a **head crash** occurs
- Head crash may permanently damage the disk or even destroy it

Magnetic Disks: Risks

- Disk heads "fly" over the surface on a very thin cushion of air
- If they accidentally contact the disk then a **head crash** occurs
- Head crash may permanently damage the disk or even destroy it
- To avoid such a risk, disk heads are "parked" when the computer is turned off

Magnetic Disks: Interfaces

- Hard drives may be removable as floppy disks, and some are even hot-swappable
  - they can be removed while the computer is running

Magnetic Disks: Interfaces

- Hard drives may be removable as floppy disks, and some are even hot-swappable
  - they can be removed while the computer is running
- Disk drives are connected to the computer via the I/O bus

Magnetic Disks: Interfaces

- Hard drives may be removable as floppy disks, and some are even hot-swappable
  - they can be removed while the computer is running
- Disk drives are connected to the computer via the I/O bus
- Some of the common interface formats include:
  - Enhanced Integrated Drive Electronics (EIDE);
  - Advanced Technology Attachment (ATA) and Serial ATA (SATA);
  - Universal Serial Bus (USB);
  - Fiber Channel (FC);
  - Small Computer Systems Interface (SCSI)

Magnetic Disks: Controllers

- The **host controller** is at the computer's end of the I/O bus

Magnetic Disks: Controllers

- The **host controller** is at the computer's end of the I/O bus
- The **disk controller** is built into the disk itself

Magnetic Disks: Controllers

- The **host controller** is at the computer's end of the I/O bus
- The **disk controller** is built into the disk itself
- The CPU issues commands to the host controller (typically via memory-mapped I/O ports)

Magnetic Disks: Controllers

- The **host controller** is at the computer's end of the I/O bus
- The **disk controller** is built into the disk itself
- The CPU issues commands to the host controller (typically via memory-mapped I/O ports)
- Data is transferred between the magnetic surface and onboard **cache** by the disk controller

Magnetic Disks: Controllers

- The **host controller** is at the computer's end of the I/O bus
- The **disk controller** is built into the disk itself
- The CPU issues commands to the host controller (typically via memory-mapped I/O ports)
- Data is transferred between the magnetic surface and onboard **cache** by the disk controller
- Finally, data is transferred from that cache to the host controller and the motherboard memory at electronic speeds

Minimize Data Transfer Time

- Mechanical components of magnetic disks cause bottleneck
  - Seek Time
  - Rotational Delay

Minimize Data Transfer Time

- Mechanical components of magnetic disks cause bottleneck
  - Seek Time
  - Rotational Delay
- To minimize data transfer time from disk we need to minimize those

Minimize Data Transfer Time

- Smaller disks → lower seek time, since arms have to travel smaller distance

Minimize Data Transfer Time

- Smaller disks → lower seek time, since arms have to travel smaller distance
- Fast-spinning disks → lower rotational delay

Hardware Optimization

Minimize Data Transfer Time

- How can the OS help minimize data transfer time?
- Schedule disk operations so as to minimize head movement
- Lay out data on disk so that related data are located on close tracks
- Place commonly-used data on a specific portion of the disk
- Pick carefully the block size contained on each sector:
  - Too small → more seeks are needed to transfer the same amount of data
  - Too large → more internal fragmentation and space wasted

Summary

- Disks are slow devices compared to CPUs (and main memory)
- Manage those device efficiently is crucial
- Minimize seek and rotational delay on magnetic disks
- HW optimizations are limited → OS needs to take the lead here!