

I/O and Storage: Hard Disk Drives

CS 571: Operating Systems (Spring 2020)

Lecture 9b

Yue Cheng

Hard Disk Drives (HDDs)

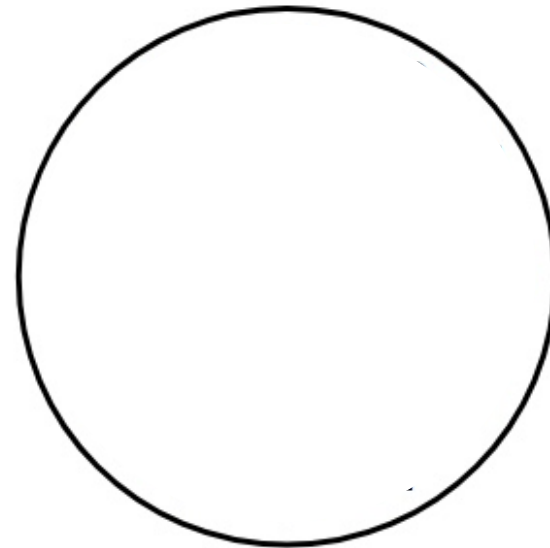
Basic Interface

- A magnetic disk has a **sector-addressable** address space
 - You can think of a disk as an array of sectors
 - Each sector (logical block) is the smallest unit of transfer
- Sectors are typically 512 or 4096 bytes
- Main operations
 - Read from sectors (blocks)
 - Write to sectors (blocks)

Disk Structure

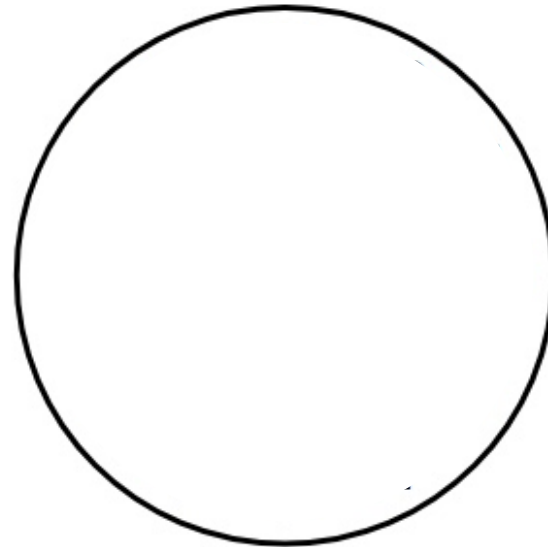
- The 1-dimensional array of logical blocks is mapped into the sectors of the disk sequentially
 - Sector 0 is the first sector of the first track on the outermost cylinder
 - Mapping proceeds in order through that track, then the rest of the tracks in that cylinder, and then through the rest of the cylinders from outermost to innermost
 - Logical to physical address should be easy
 - Except for bad sectors

Internals of Hard Disk Drive (HDD)



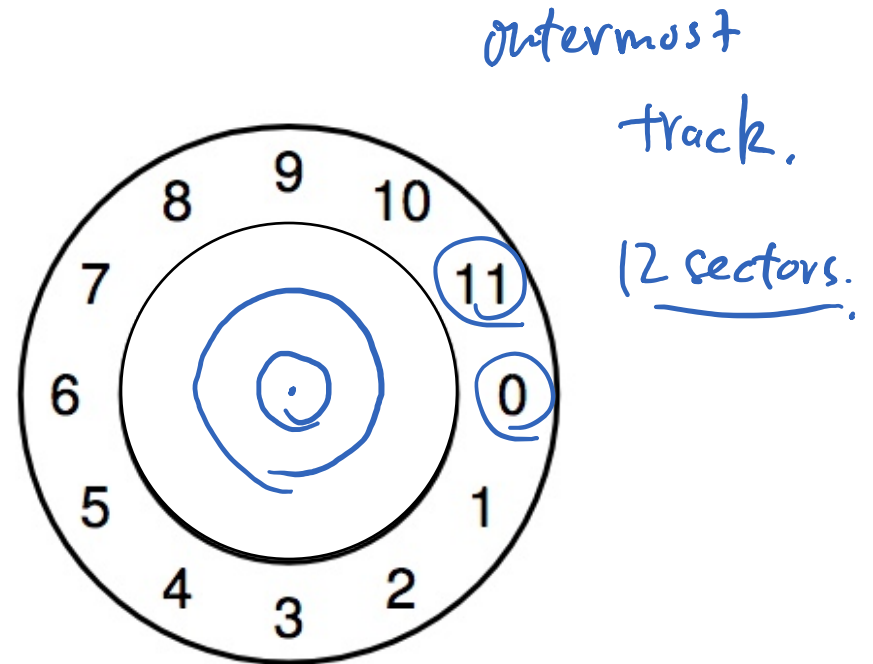
Internals of Hard Disk Drive (HDD)

Platter
Covered with a magnetic film



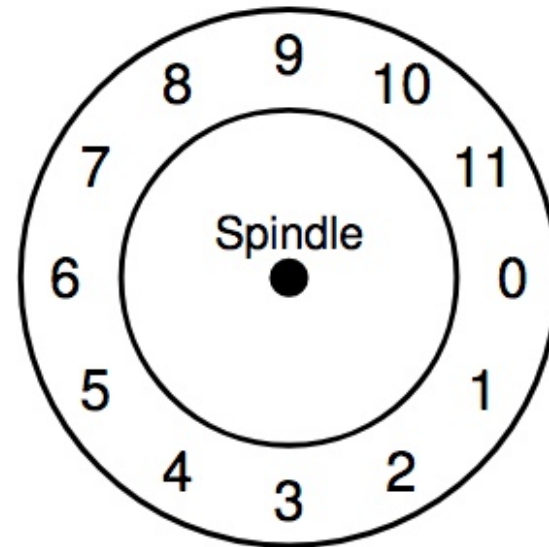
Internals of Hard Disk Drive (HDD)

A single track example



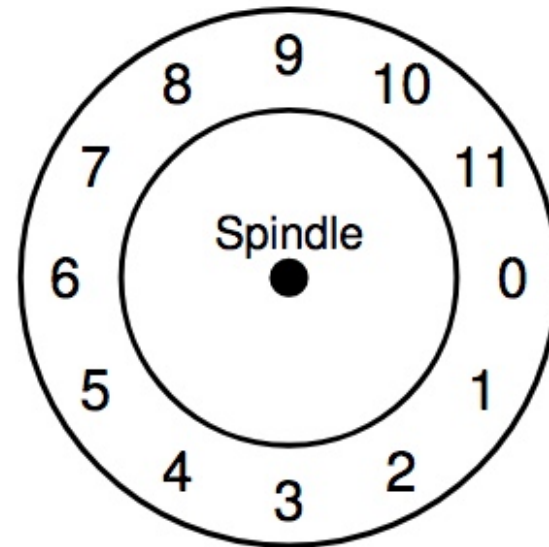
Internals of Hard Disk Drive (HDD)

Spindle in the center of the surface



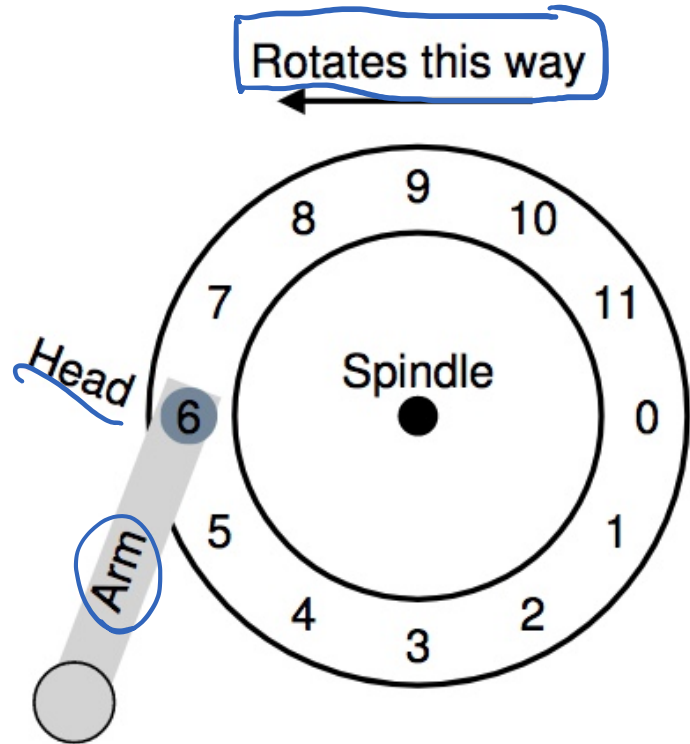
Internals of Hard Disk Drive (HDD)

The track is divided into numbered sectors

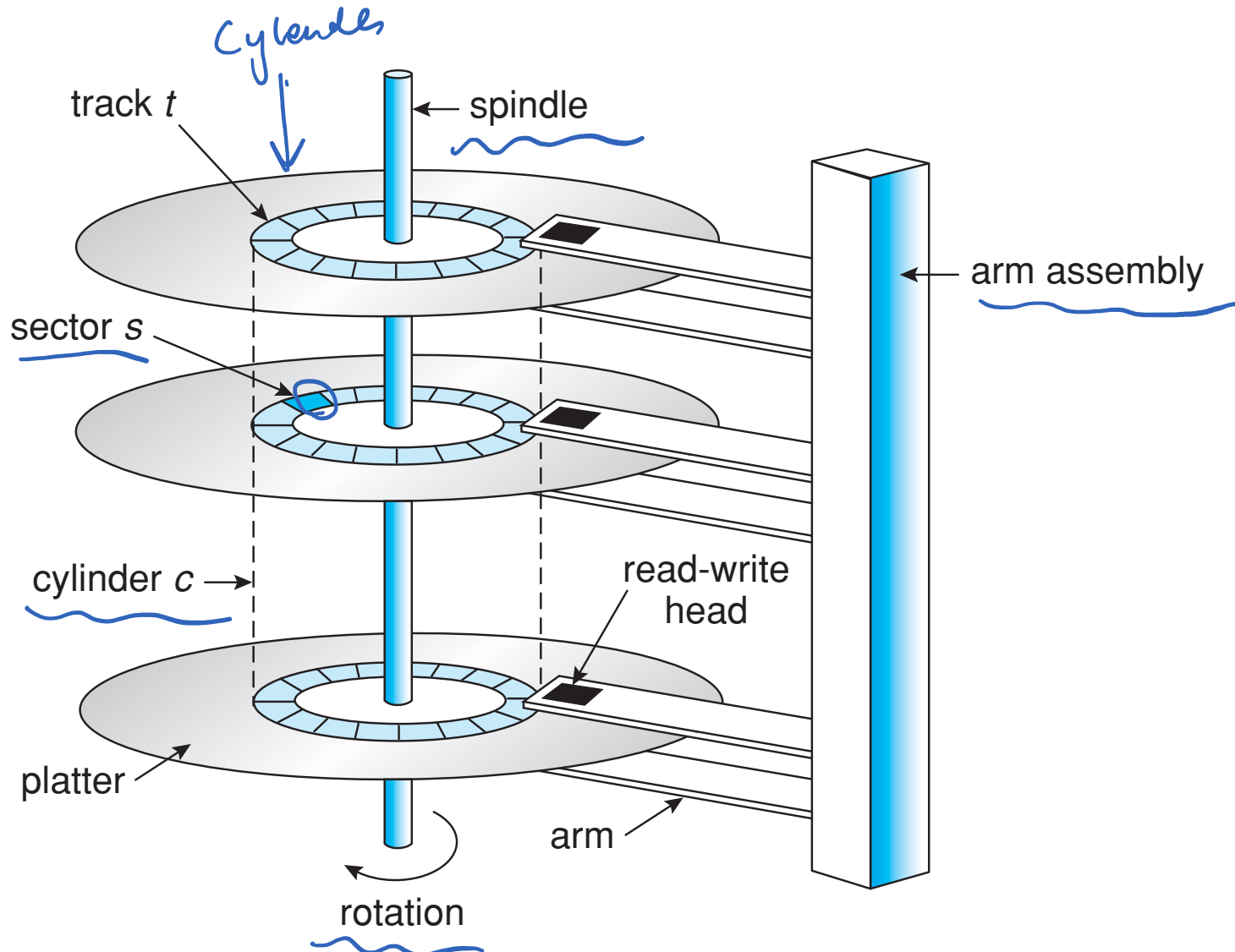


Internals of Hard Disk Drive (HDD)

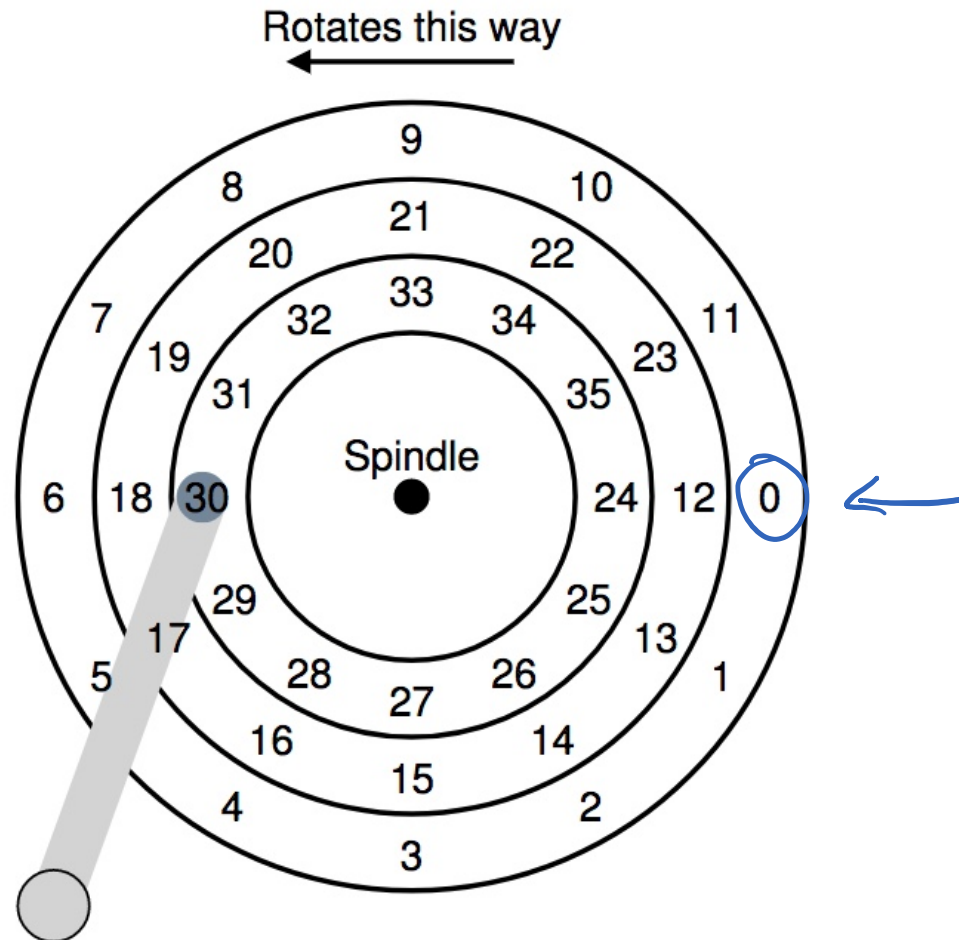
A single track + an arm +
a head



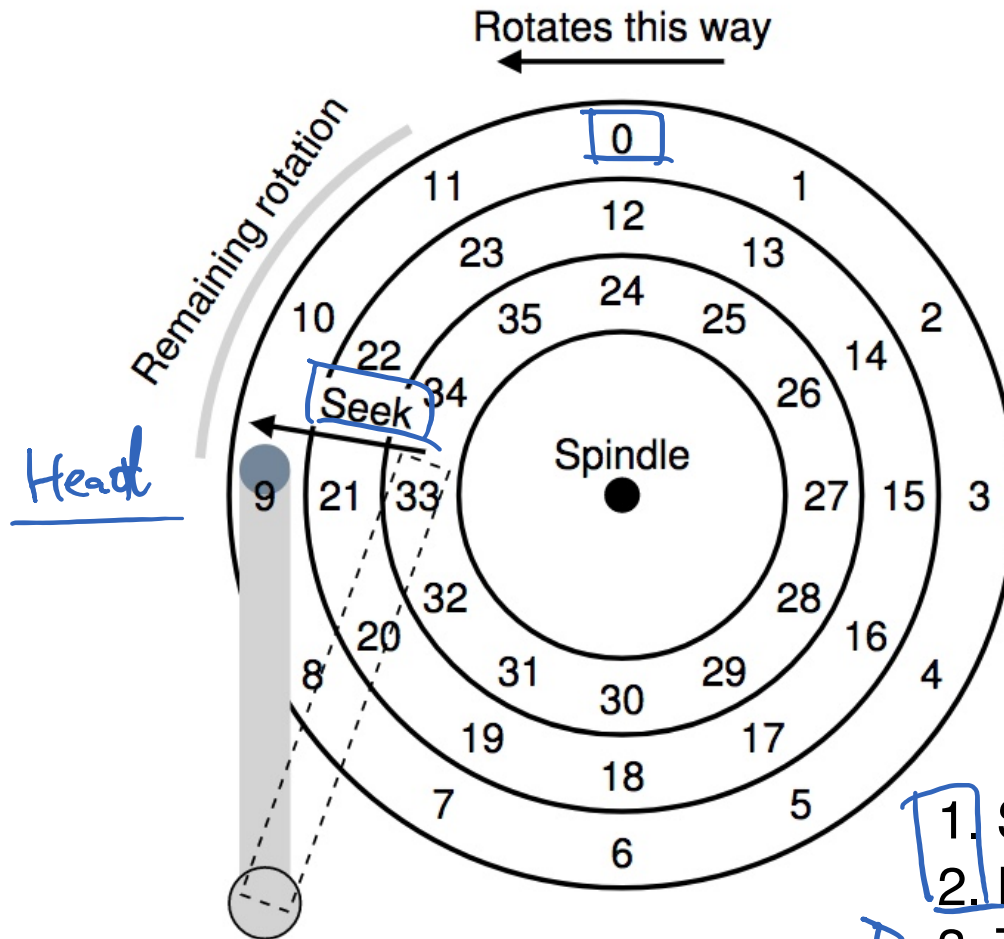
HDD Mechanism (3D view)



Let's Read Sector 0



Let's Read Sector 0



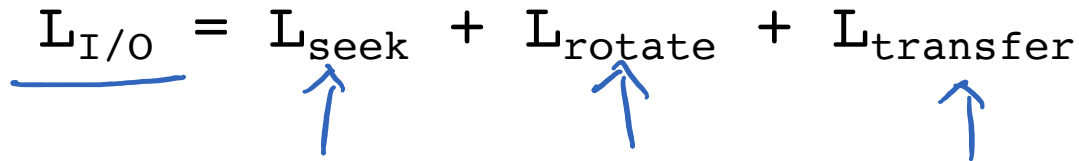
1. Seek for right track
2. Rotate (sector 9 → 0)
3. Transfer data (sector 0)

Don't Try This at Home!

<https://www.youtube.com/watch?v=9eMWG3fwiEU&feature=youtu.be&t=30s>

Disk Performance

- I/O latency of disks

$$\underline{L_{I/O}} = L_{\text{seek}} + L_{\text{rotate}} + L_{\text{transfer}}$$


- Disk access latency at millisecond level

$$1 \text{ sec} = \underline{1000 \text{ ms}}$$

Seek, Rotate, Transfer

- Seek may take several milliseconds (ms)
- Settling along can take 0.5 - 2ms
- Entire seek often takes 4 - 10ms

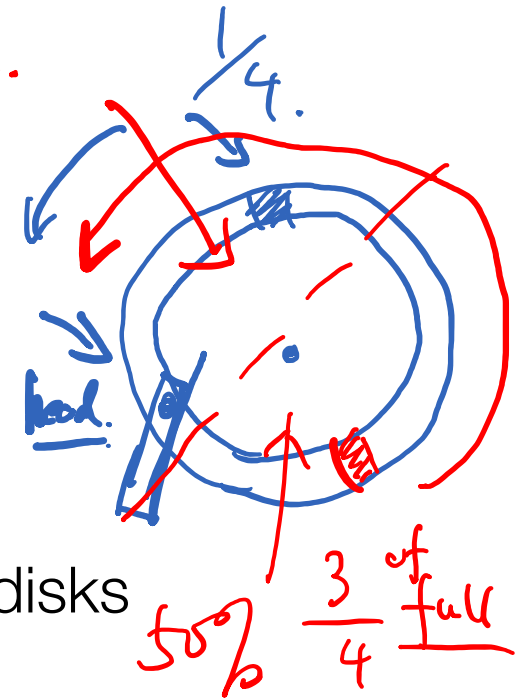
empirical result

Seek, Rotate, Transfer

- Rotation per minute (RPM)
 - 7200 RPM is common nowadays
 - 15000 RPM is high end
 - Old computers may have 5400 RPM disks
- $1 / 7200 \text{ RPM} = \frac{1 \text{ minute}}{7200 \text{ rotations}} = \frac{1 \text{ second}}{120 \text{ rotations}} = \underline{8.3 \text{ ms}} / \text{rotation}$

Seek, Rotate, Transfer

50%



- Rotation per minute (RPM)
 - 7200 RPM is common nowadays
 - 15000 RPM is high end
 - Old computers may have 5400 RPM disks
- $1 / 7200 \text{ RPM} = 1 \text{ minute} / 7200 \text{ rotations} =$
 $1 \text{ second} / 120 \text{ rotations} = 8.3 \text{ ms} / \text{rotation}$
 $8.3 \text{ ms} / 2 = 4.15 \text{ ms}$
- So it may take 4.2 ms on average to rotate to target ($0.5 * 8.3 \text{ ms}$) $\approx 4.2 \text{ ms}$.

Seek, Rotate, Transfer

ms

- Relatively fast
 - Depends on RPM and sector density
- 100+ MB/s is typical for SATA I (1.5Gb/s max)
 - Up to 600MB/s for SATA III (6.0Gb/s)
- 1s / 100MB = 10ms / MB = 4.9us/sector
 - Assuming 512-byte sector




5 us ↑

Workloads

- Seeks and rotations are slow while transfer is relatively fast

- What kind of workload is best suited for disks?
mechanical. magnetic
↓

Workloads

- Seeks and rotations are slow while transfer is relatively fast
- What kind of workload is best suited for disks?
 - **Sequential I/O**: access sectors in order (transfer dominated)  
 - **Random** workloads access sectors in a random order (seek+rotation dominated)
 - Typically slow on disks
 - Never do **random** I/O unless you must! E.g., **Quicksort** is a terrible algorithm for disk! 

Disk Performance Calculation

- Seagate Enterprise SATA III HDD

Metric	Perf
RPM	7200
Avg seek	4.16ms
Max transfer	<u>500MB/s</u>



- How long does an average 4KB read take?

Disk Performance Calculation

- Seagate Enterprise SATA III HDD

Metric	Perf
RPM	7200
Avg seek	4.16ms
Max transfer	500MB/s



- How long does an average 4KB read take?

$$transfer = \frac{1 \text{ sec}}{500 \text{ MB}} \times 4 \text{ KB} \times \frac{1,000,000 \text{ us}}{1 \text{ sec}} = \underline{8 \text{ us}}$$

Disk Performance Calculation

- Seagate Enterprise SATA III HDD

Metric	Perf
RPM	7200
Avg seek	4.16ms
Max transfer	500MB/s



8.3 ms
2

- How long does an average 4KB read take?

$$transfer = \frac{1 \text{ sec}}{500 \text{ MB}} \times 4 \text{ KB} \times \frac{1,000,000 \text{ us}}{1 \text{ sec}} = 8 \text{ us}$$

$$Latency = 4.16 \text{ ms} + 4.2 \text{ ms} + 8 \text{ us} = 8.368 \text{ ms}$$

Avg Seek
Avg Rotate
X