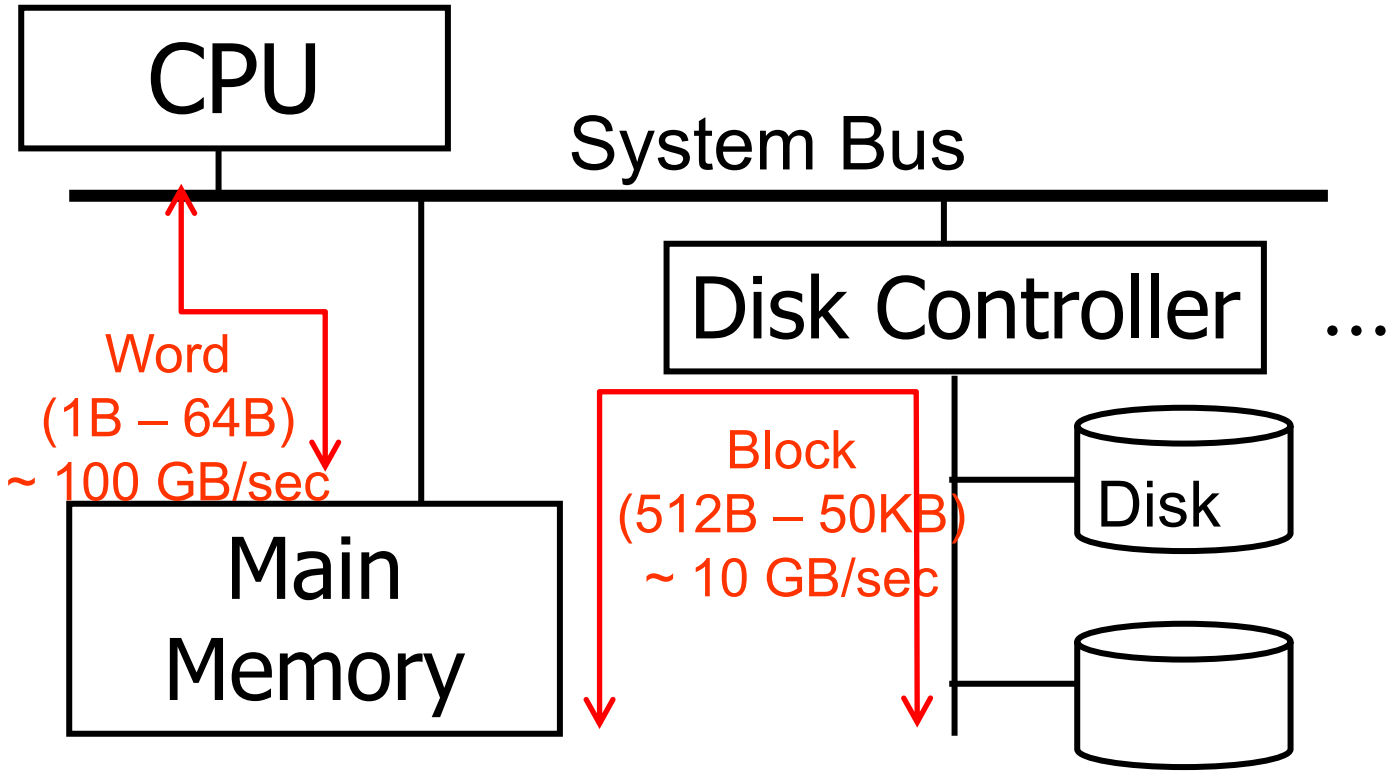
CS143: Disk

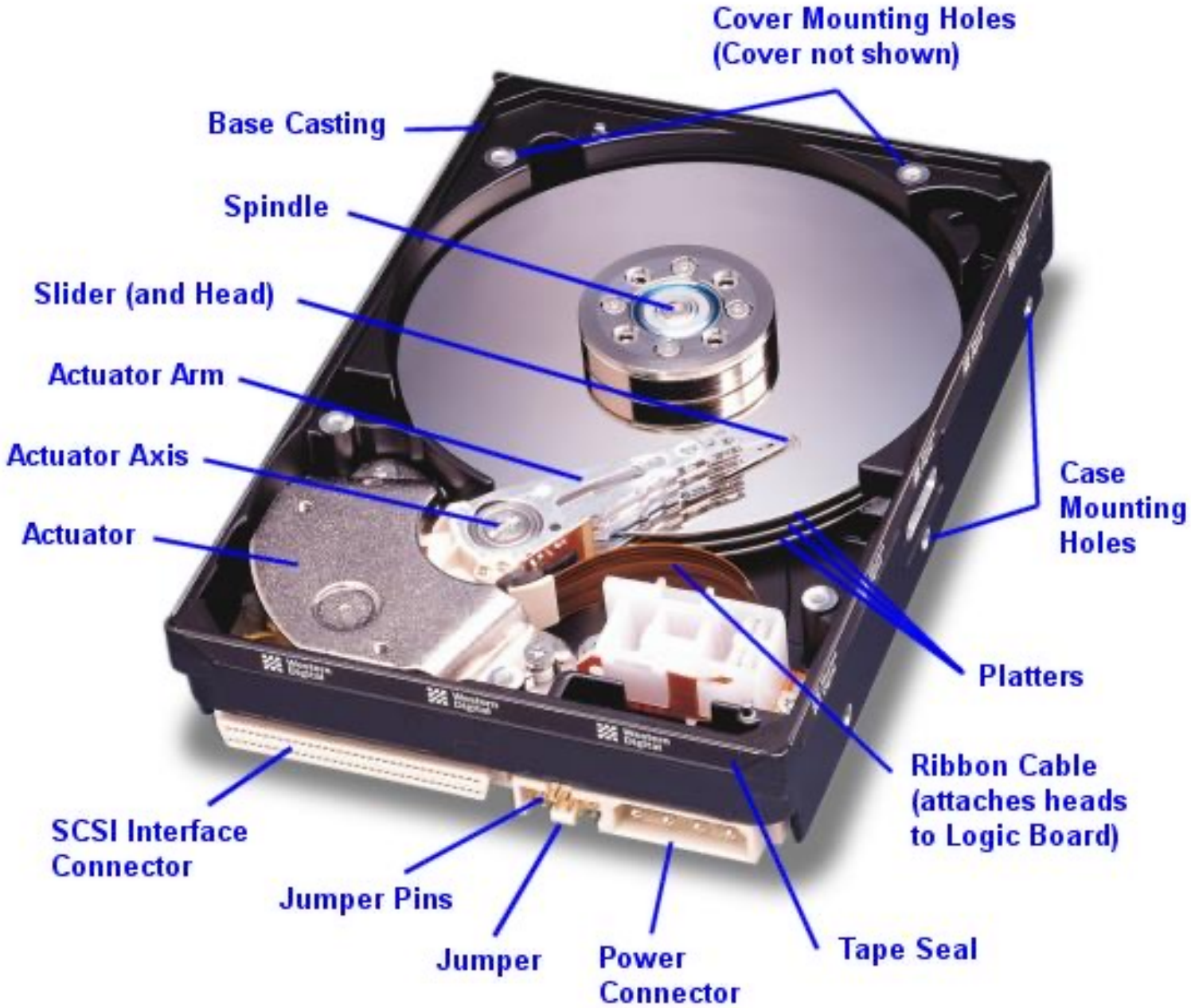
Professor Junghoo “John” Cho

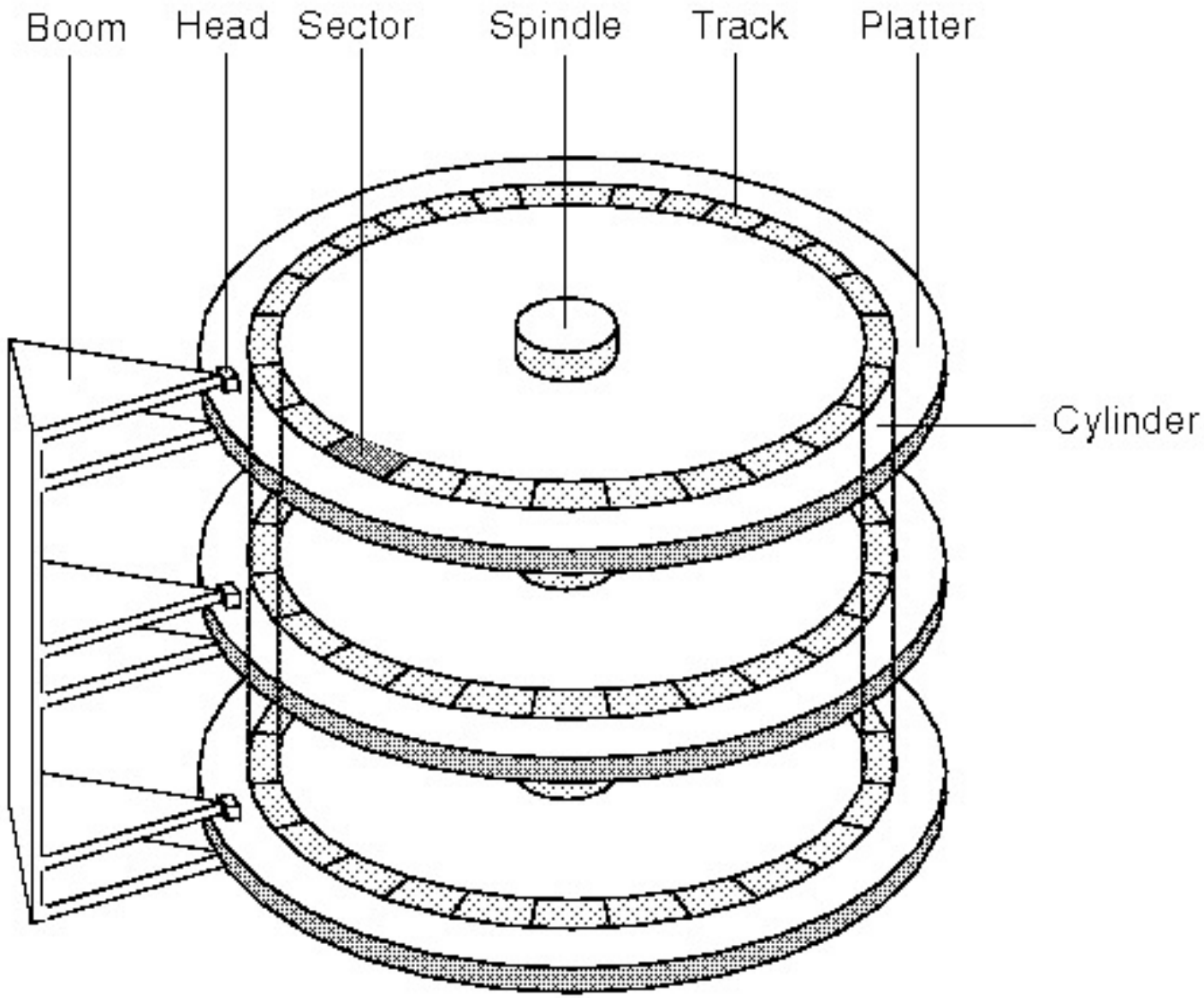
System Architecture



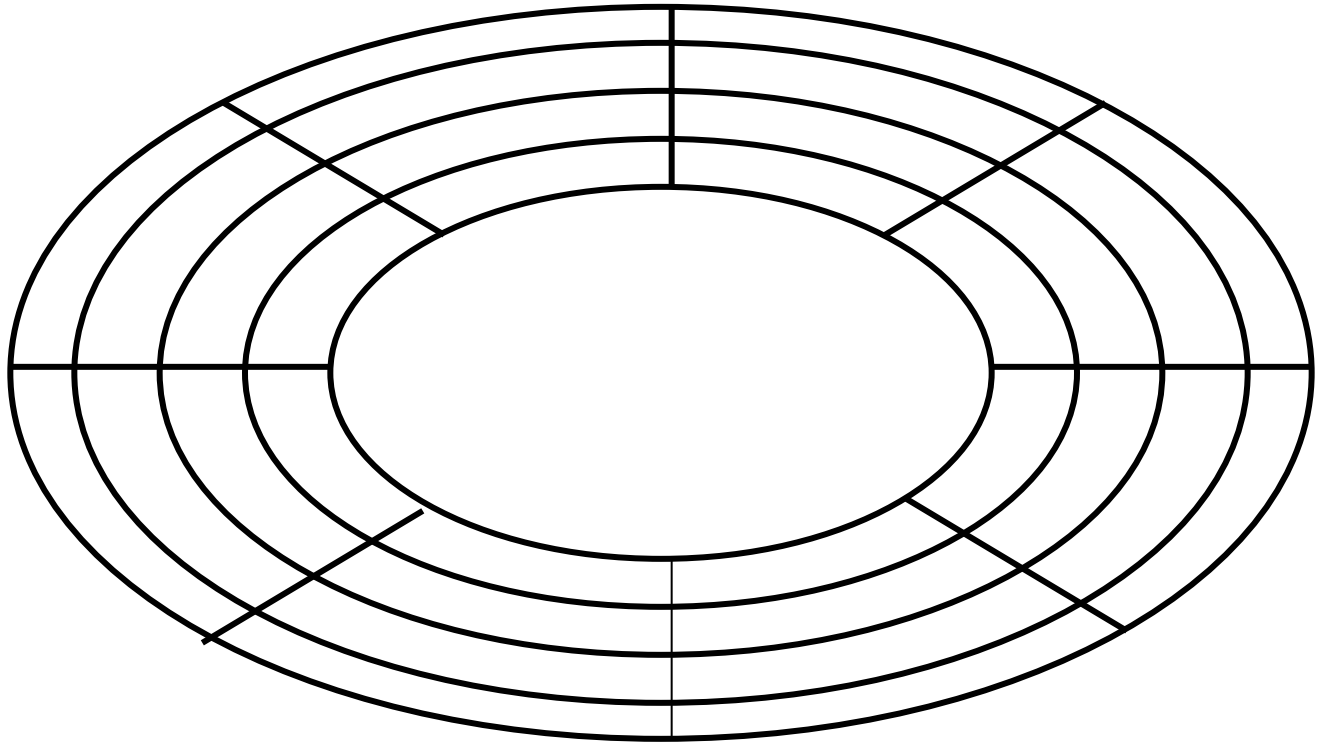
Magnetic disk vs SSD

- Magnetic Disk
  - Stores data on a magnetic disk
  - Typical capacity: 1TB – 20TB
- Solid State Drive (SSD)
  - Stores data in NAND flash memory
  - Typical capacity: 100GB – 10TB
  - Faster than magnetic disk
    - Particularly random disk access
  - But x5 more expensive and limited write cycles (~2000)


Cover Mounting Holes
(Cover not shown)
Base Casting
Spindle
Slider (and Head)
Actuator Arm
Actuator Axis
Actuator
Case
Mounting
Holes
Platters
Ribbon Cable
(attaches heads
to Logic Board)
SCSI Interface
Connector
Jumper Pins
Jumper
Power
Connector
Tape Seal


Boom
Head
Sector
Spindle
Track
Platter
Cylinder

Structure of a Platter

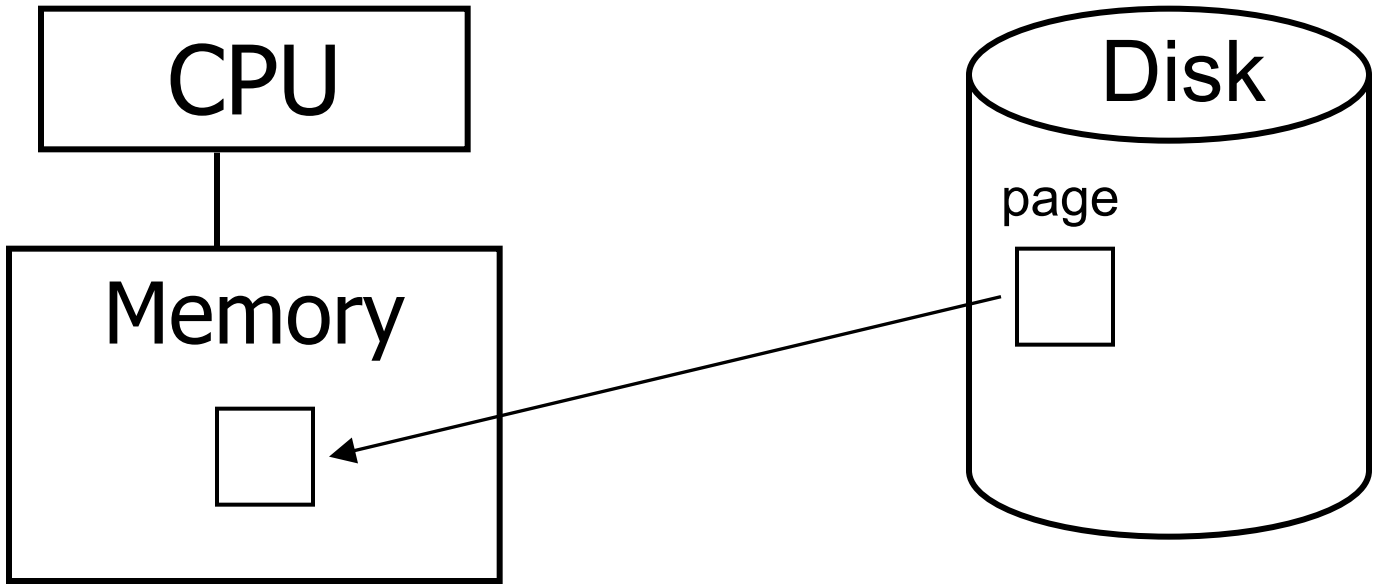


- Track, cylinder, sector (=block, page)
- Data is transferred in the unit of “block” (not bytes) to amortize high access delay

Typical Magnetic Disk

- Platter diameter: 2.5-5.25 in
- Platters: 1 – 20
- Tracks: 1000 – 50,000
- Sectors per track: 1000 – 50,000
- Sector size: 512 – 50K
- Rotation speed: 1000 – 15000 rpm
- Overall capacity: 1TB – 20TB

Access Time

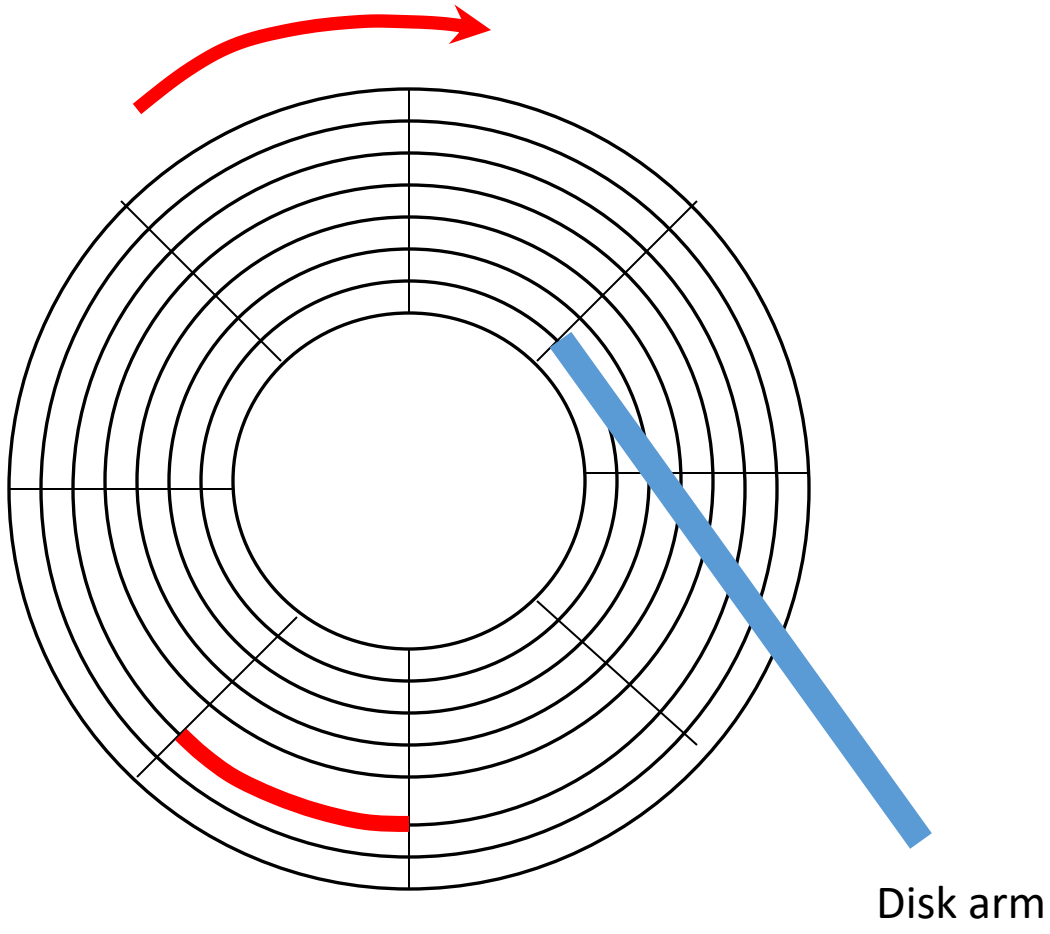


- Q: How long does it take to read a page of a disk to memory?

- Q: What needs to be done to read a page?

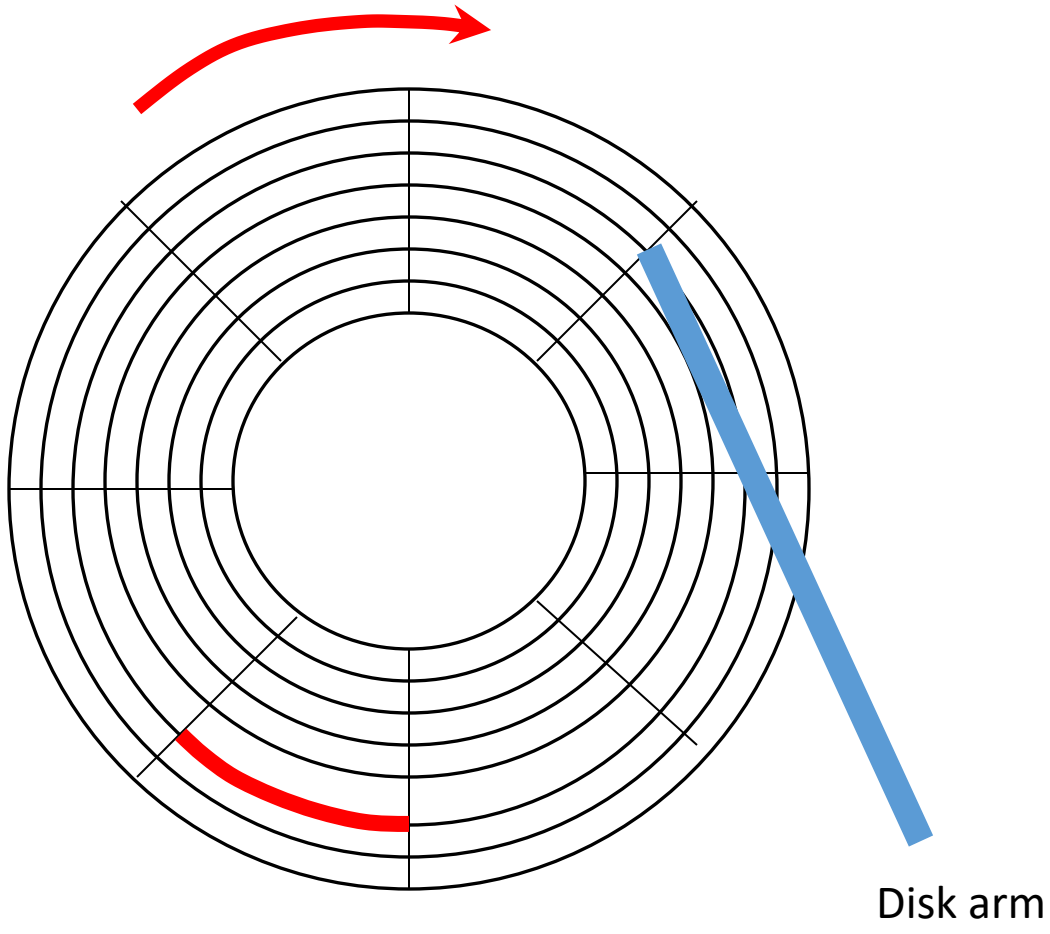
Reading a Page From Disk

- Q: What should happen to read the highlighted sector from disk?



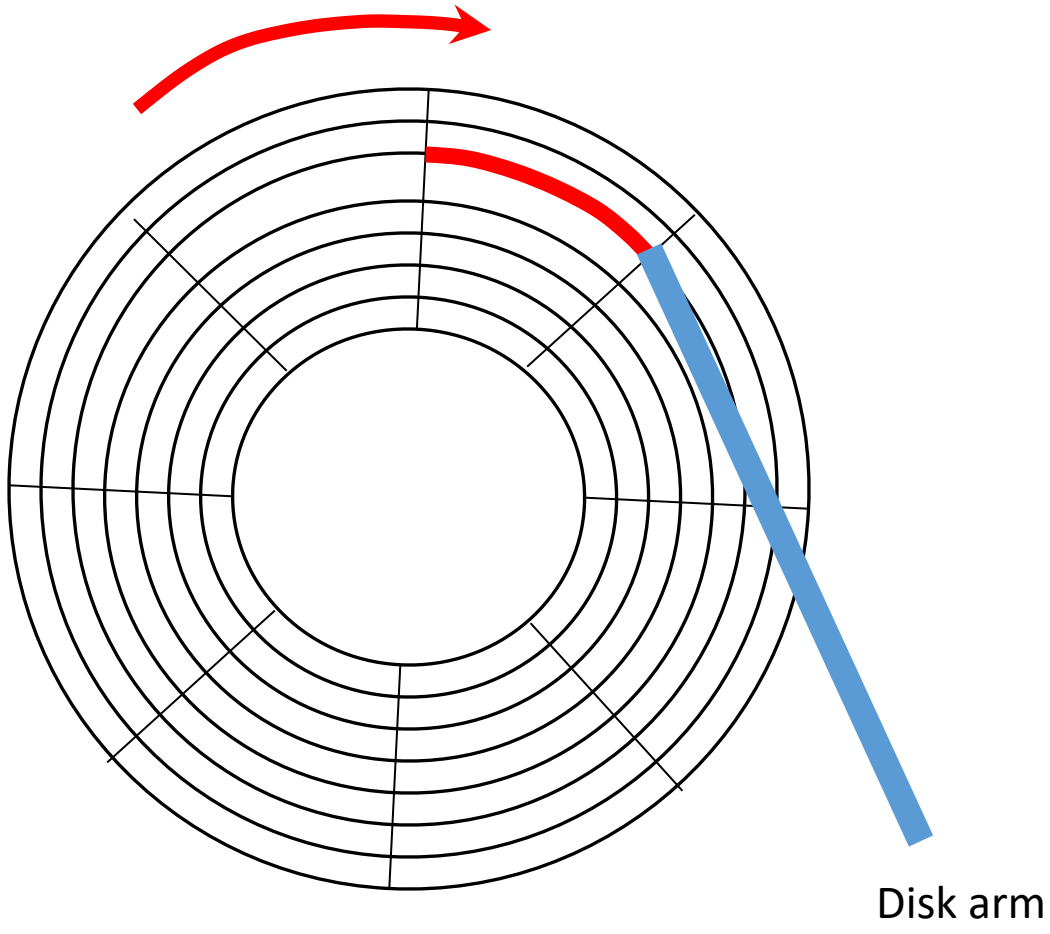
Reading a Page From Disk

- Q: What should happen to read the highlighted sector from disk?



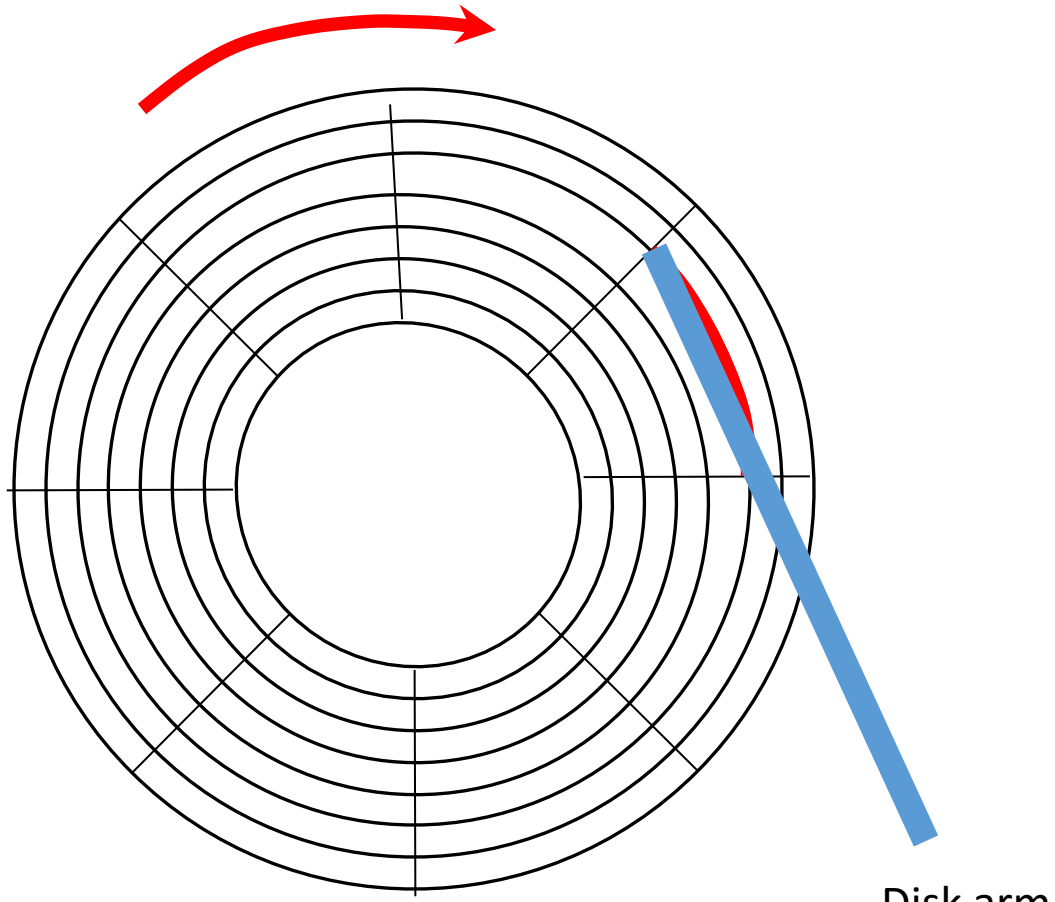
Reading a Page From Disk

- Q: What should happen to read the highlighted sector from disk?



Reading a Page From Disk

- Q: What should happen to read the highlighted sector from disk?

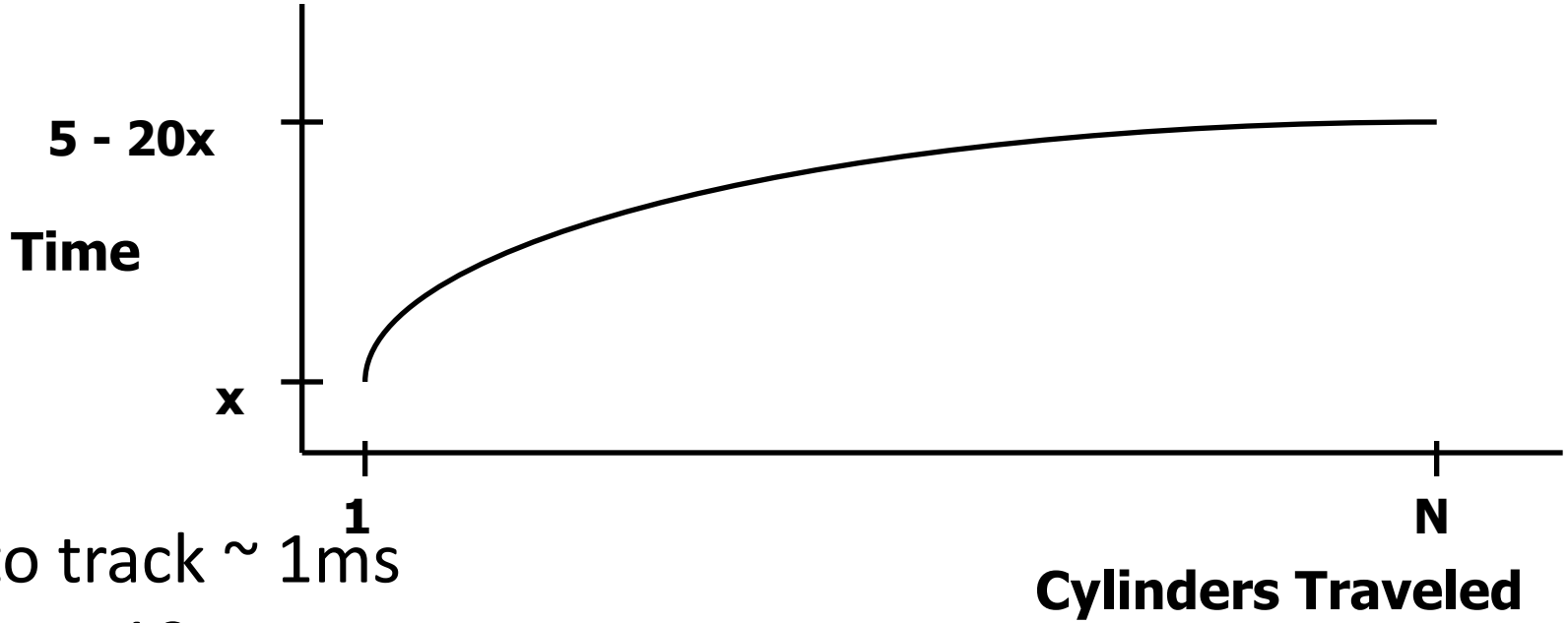


Access Time

- Access time =
  (seek time) + (rotational delay) +
  (transfer time)

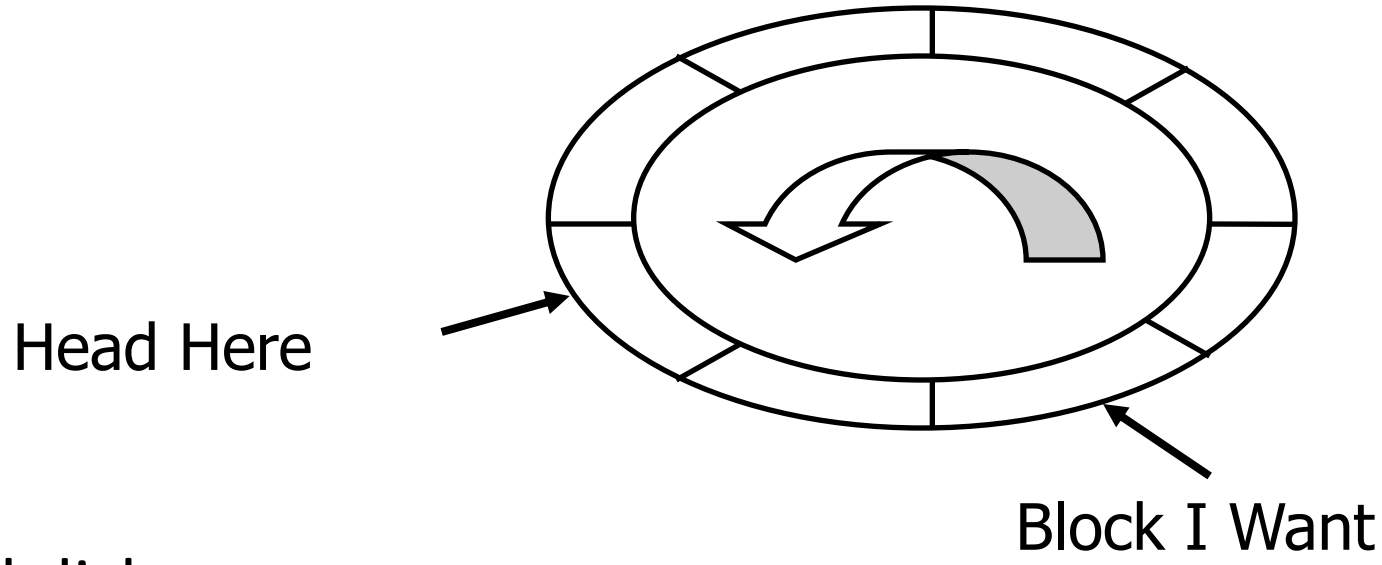
Seek Time

- Time to move a disk head between tracks



- Track to track ~ 1ms
- Average ~ 10 ms
- Full stroke ~ 20 ms

Rotational Delay

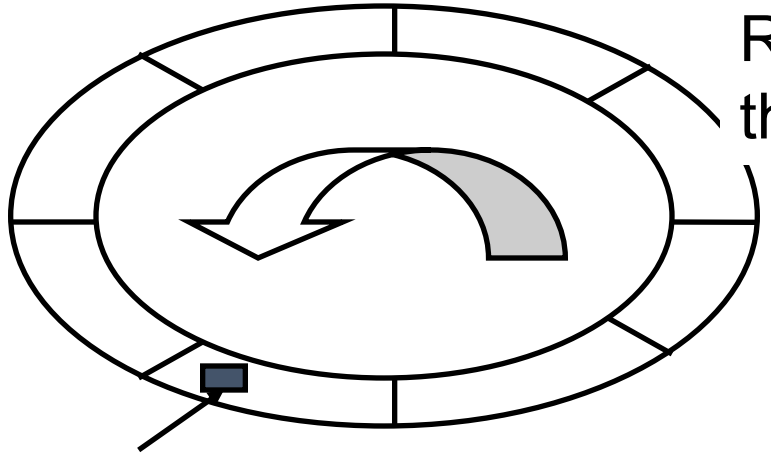


- Typical disk:
  - 1000 rpm – 15000 rpm
- Q: For 6000 RPM, average rotational delay? 5ms

60 sec 6000
1 sec 100

$\frac{1}{100}$ sec = 10 ms

Transfer Time



- 6000RPM 10,000 sectors/track
- Q: How long to read one sector?

$$\frac{10}{\cancel{100}\,ns \;/\; 10{,}000} = \cancel{0.01}\,0.001\,ms$$

Access Time

- 6000RPM, 10,000 sectors/track
- Average access time to read one sector =
  (seek time) + (rotational delay) + (transfer time)

10 ms 5 ms ~~0.01~~ ms 0.001

= ~~15.01~~ ms

15.001

Transfer Rate

- The rate at which we can transfer data from disk
  - Measured in bytes/sec
- Q: 6,000 RPM, 10000 sectors/track, 1KB/sector what is the transfer rate?
  - Burst transfer rate vs Sustained transfer rate

1KB × 10,000 = 10MB 10ms 10MB/10ms = 1000MB/sec

- (Burst) Transfer rate =
  (RPM / 60) * (sectors/track) * (bytes/sector)

Random I/O

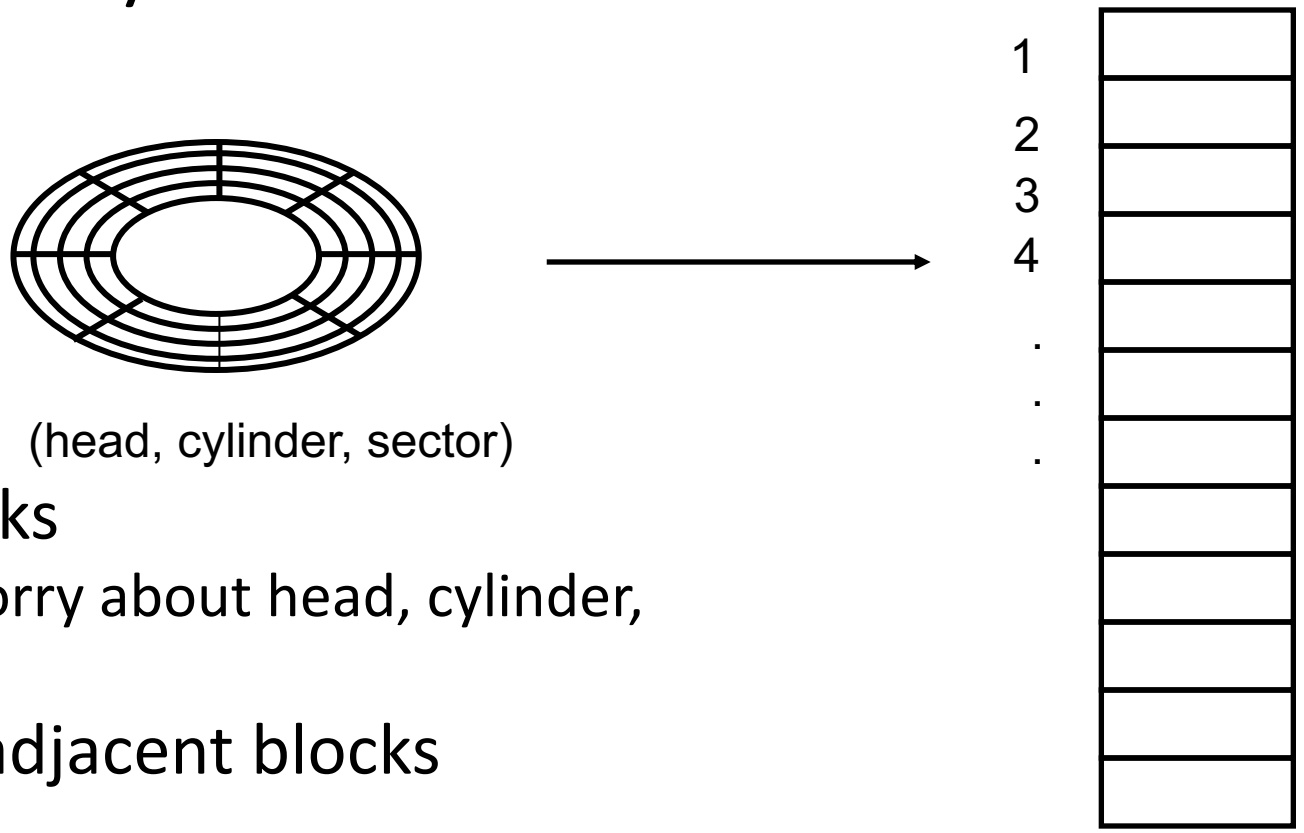
- For magnetic disks:
  - Random I/O is VERY expensive compared to sequential I/O
- For SSD disks:
  - Random I/O is still expensive but not as much as for magnetic disks
- Avoid random I/O to minimize delay

	Magnetic	SSD
Random IO	~100 IOs/sec	~100K IOs/sec
Transfer rate	~ 100MB/sec	~ 10GB/sec

Buffers, Buffer pool

- Temporary main-memory “cache” for disk blocks
  - Avoid future read
  - Hide disk latency
  - Most DBMS let users change buffer pool size

Abstraction by OS



- Sequential blocks
  - No need to worry about head, cylinder, sector
- Access to non-adjacent blocks
  - Random I/O
- Access to adjacent blocks
  - Sequential I/O

Things to Remember

- Platter, track, cylinder, block (sector)
- Access time = seek time + rotational delay + transfer time
- Random I/O vs Sequential I/O