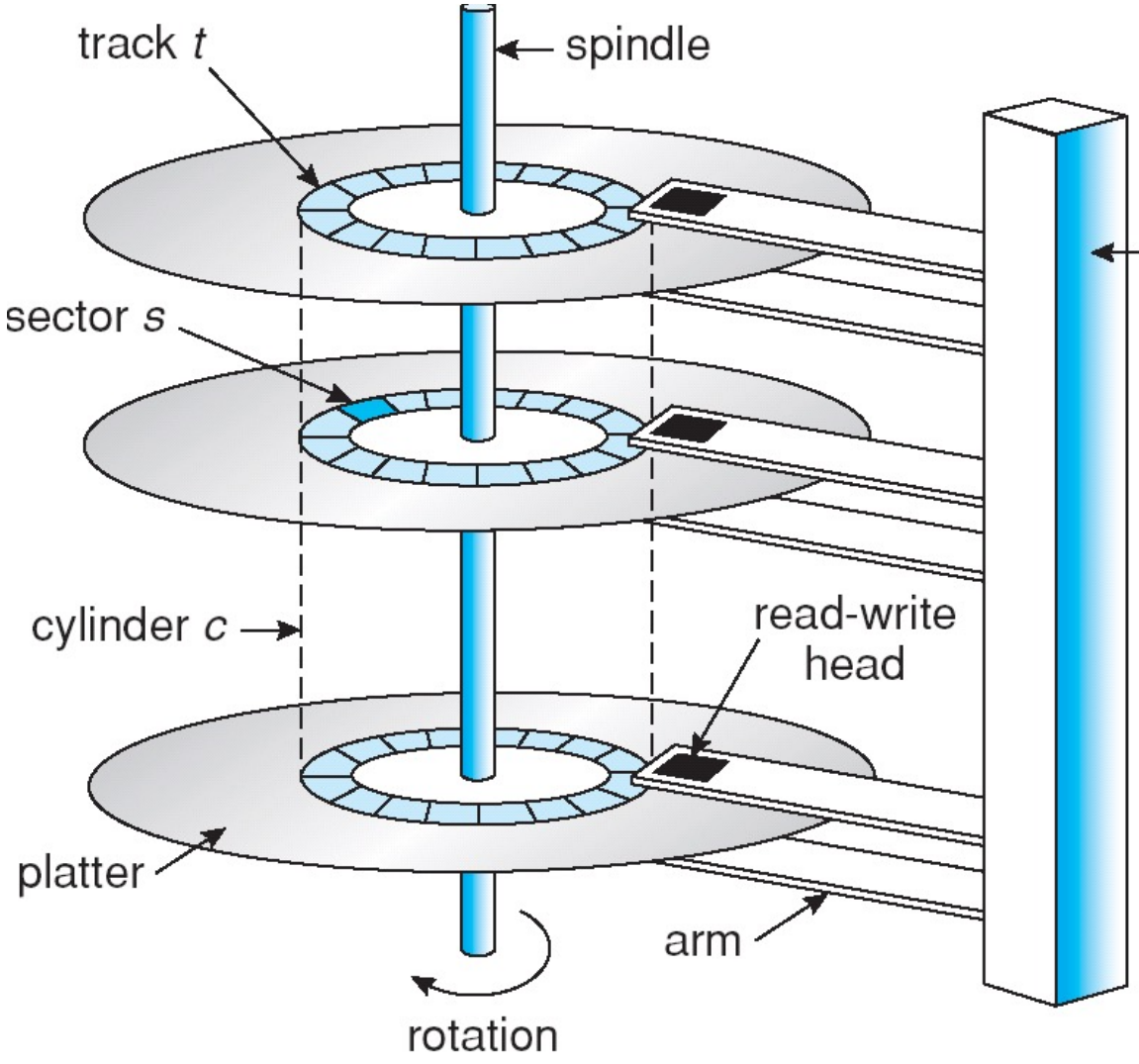
# HDD and SDD

COMS W4118

**References:** Operating Systems Concepts (9e), Linux Kernel Development, previous W4118s

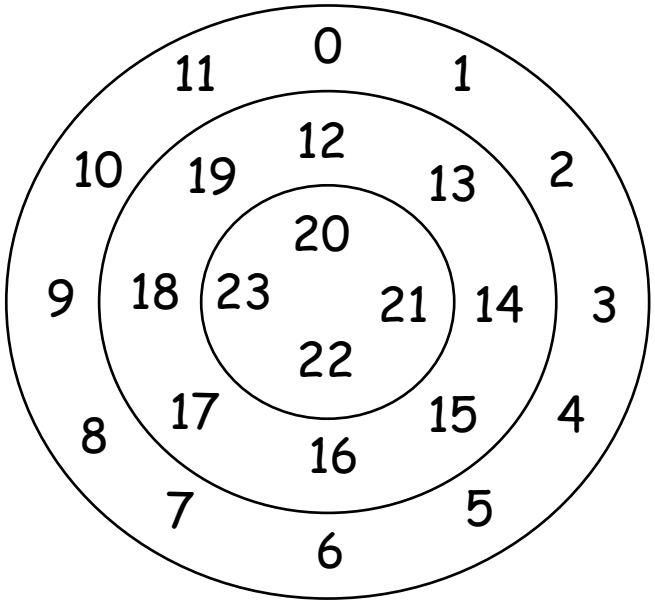
# Hard Disk Drive (HDD)




- Range from 30GB to 3TB per drive
- Aluminum platters with magnetic coating
- Commonly, 2-5 platters per drive
- Common platter sizes: 3.5”, 2.5”, and 1.8”
- Magnetic heads

# Disk Interface

- From FS perspective: disk is addressed as a one dimension array of logical sectors
- Disk controller maps logical sector to physical sector identified by track #, surface #, and sector #



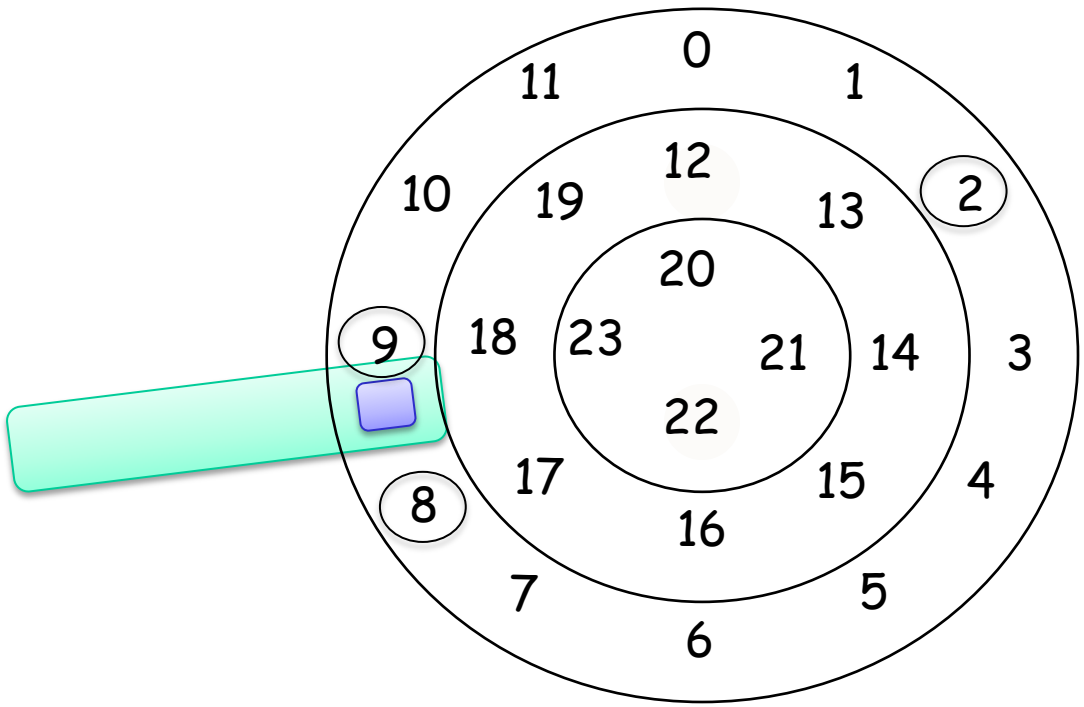

- Note: Old drives allowed direct C/H/S (cylinder/head/sector) addressing by OS. Modern drives export LBA (logical block address) and do the mapping to C/H/S internally.

# Disk Latencies

- Rotational delay: rotate disk to get to the right sector
- Seek time: move disk arm to get to the right track
- Transfer time: get bits off the disk

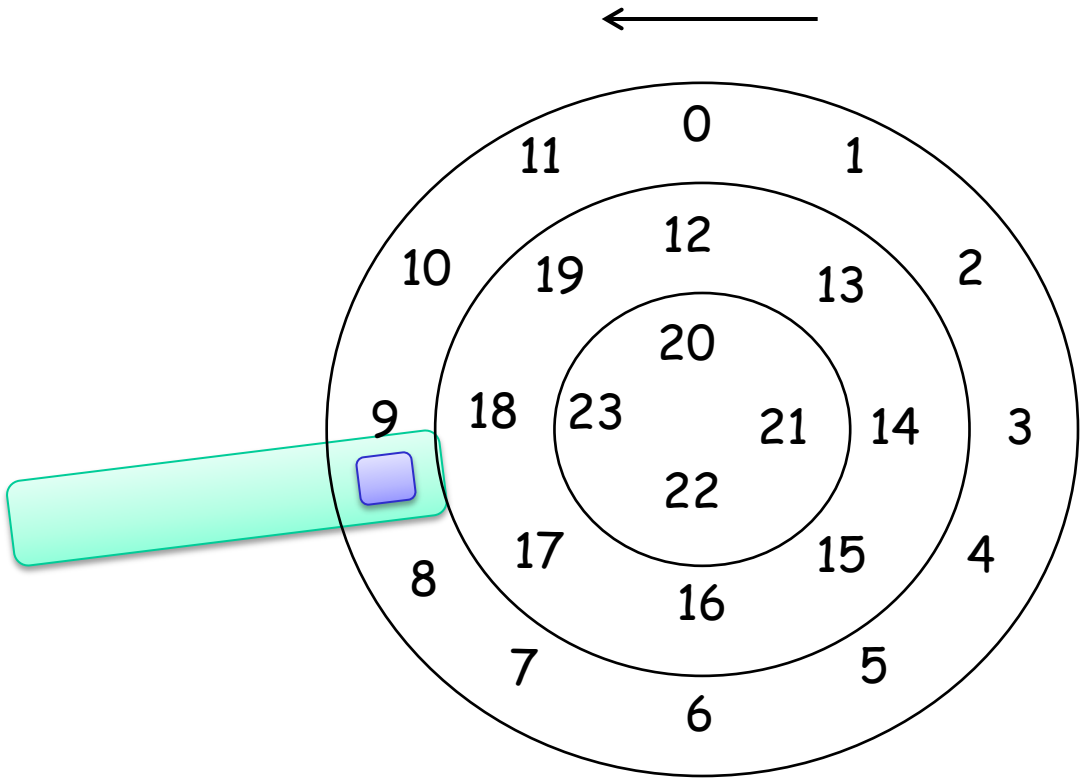
# Seek Time

- Must move arm to the right track
- Can take a while (e.g., 5– 10ms)
  - Acceleration, coasting, settling (can be significant, e.g., 2ms)



# Transfer Time

- Transfer bits out of disk
- Actually pretty fast (e.g., 125MB/s)



# I/O Time (T) and Rate (R)

- T = Rotational delay + seek time + txfer time
- R = Size of transfer / T
- Workload 1: large sequential accesses?
- Workload 2: small random accesses?

# Design tip: Use Disks Sequentially!

- Disk performance differs by a factor of 200 or 300 for random v.s. sequential accesses
- When possible, access disks sequentially

# Solid-state Storage Device (SSD)

- Pros:
  - No moving parts – less fragile, less power hungry
  - Faster sequential read/write
  - Orders of magnitude faster random read/write
- Cons:
  - Expensive
  - Slow erase
  - Limited life span

# Flash Translation Layer (FTL)

- Goal
  - Provide LBA on top of flash weirdness
- Challenges
  - Flash device is written in pages (1KB-8KB)
  - But erased in erase blocks (128KB-2MB)
  - Updating a page requires erasing the whole block!
  - Limited number of erase & write cycles
- Solutions
  - Log-structured I/O
    - Instead of updating, write new version somewhere else
    - Needs garbage collection, which leads to write amplification
  - Wear leveling to spread writes
  - OS informs deleted blocks to SSD using TRIM command