ECE437/CS481

M06E: DISK SCHEDULING

CHAPTER 10.1-10.4

Xiang Sun

The University of New Mexico

I/O Scheduling for Block Devices

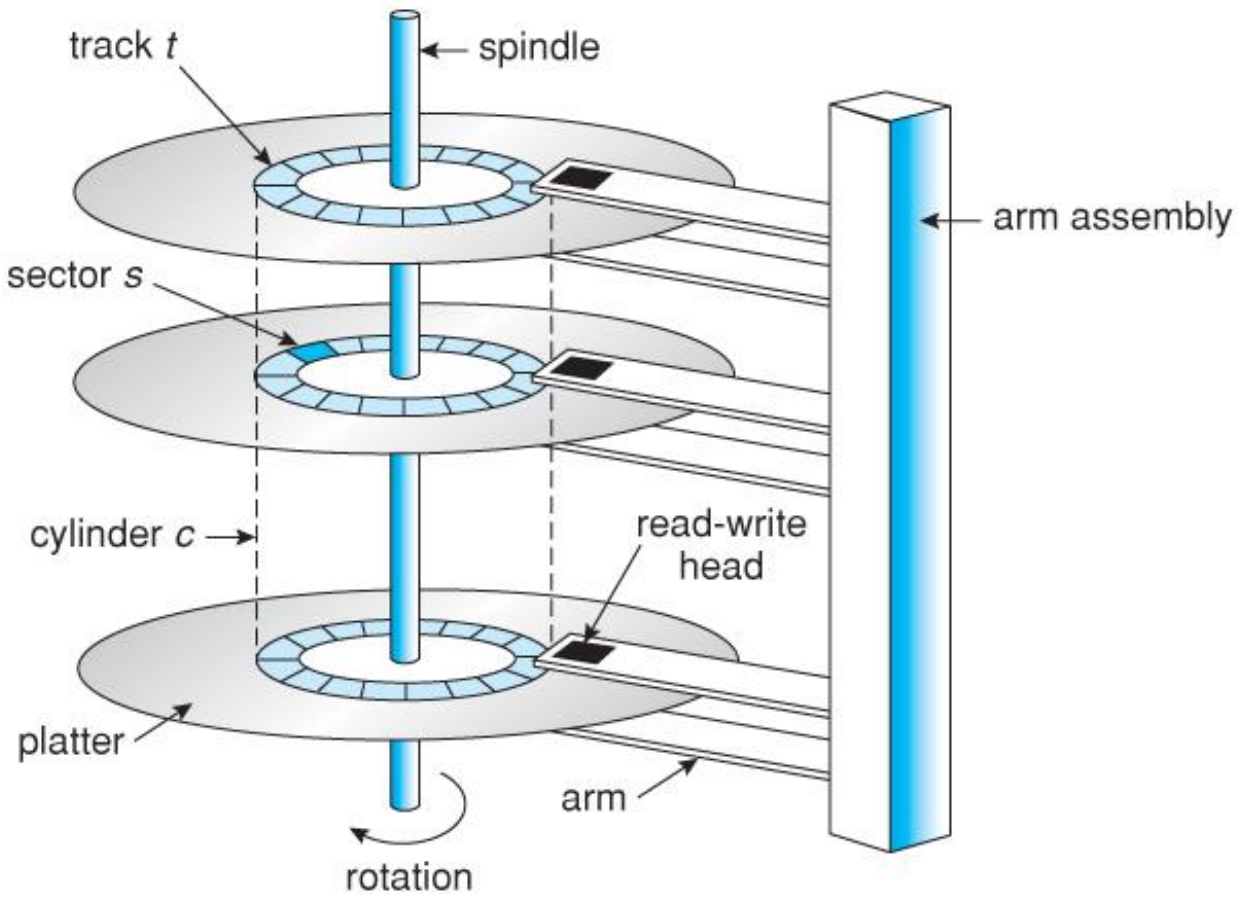
- Mass Storage structure
  - Magnetic disk provide bulk of secondary storage
  - Drivers rotate at 3,600 to 10,000 times per second

- Hard drive is attached to computer via I/O bus
  - ATA: Parallel ATA (PATA), Serial ATA (SATA)

  - SCSI (Small Computer System Interface)
  - USB

I/O Scheduling for Block Devices

❑ Recall Disk Structure



- **Sector**: Minimum unit (normally, 512 Byte for old hard disks; 4 KB for new hard disks) for data storage.
  - ✓ Use `cat /sys/block/sda/queue/physical_block_size` to see the size of sector for your hard disk.
- **Cluster**: Group of sectors
- **Track**: One cycle on a platter
- **Cylinder**: A semantic shape that consists of same level tracks on different platters.

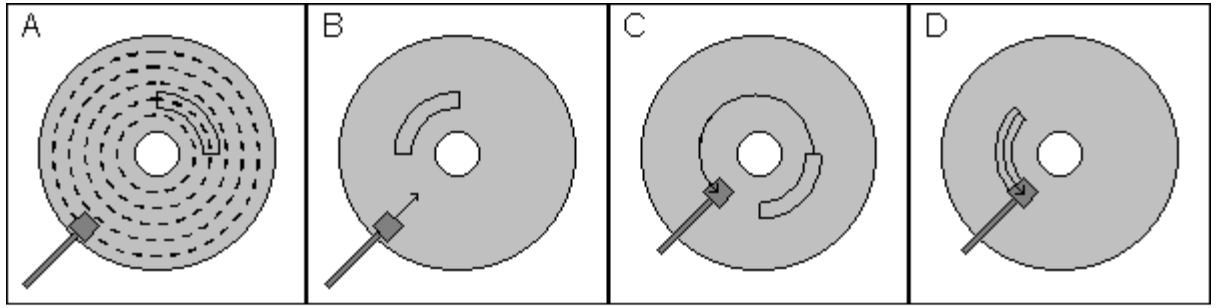
I/O Scheduling for Block Devices

❑ Response time

- ✓ Internal response time: the time for moving data between the disk surface and the controller on the drive.
- ✓ External response time : the time for moving data between the controller in the drive and the memory.

➢ Internal response time comprises

- ✓ Seek time: the time for moving disk arm to place the r/w head on the desired cylinder/track.
- ✓ Rotational latency: the time for desired sector to rotate under the disk head.
- ✓ Internal transfer time: the time for the disk platter to rotate until all the addressed sectors have passed under the head. The transfer time is directly proportional to the number of addressed sectors.
- ✓ Overhead: the time for searching the right platter, etc. The overhead is normally fixed.

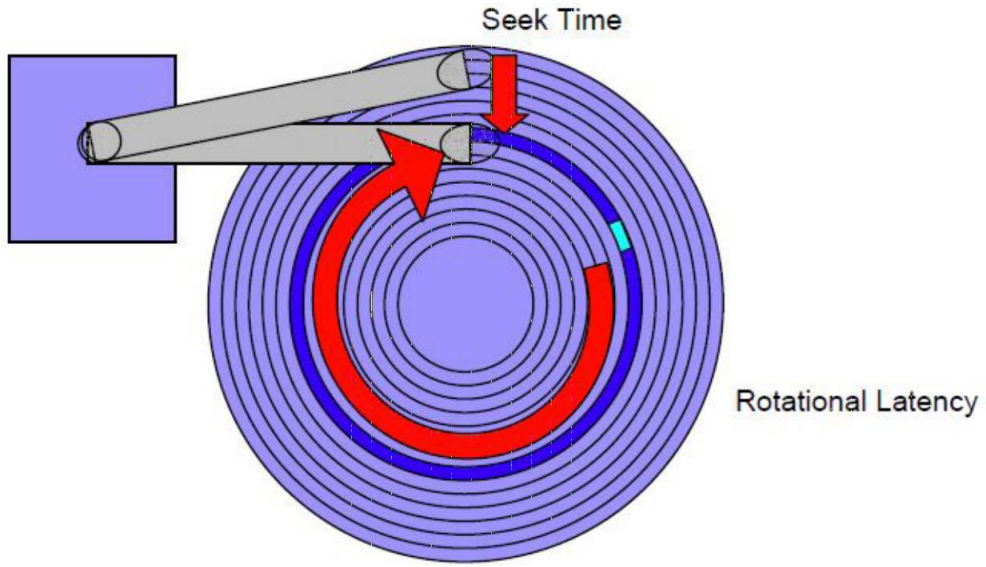


I/O Scheduling for Block Devices

❑ Seek time

- Manufacturers often report *average seek time* of 3-9 ms
  - ✓ These times average the time to seek from any track to any other track.

- In practice, the average seek time is often much lower.
  - ✓ For example, if the head is already on or near the desired track, then seek time is much smaller. In other words, locality is important!
  - ✓ Actual average seek times are often just 2-3ms.



I/O Scheduling for Block Devices

❑ Rotational latency

- Once the head is in place, we need to wait until the right sector is underneath the head.
  - ✓ This may require as little as no time (reading consecutive sectors) or as much as a full rotation (just missed it).
  - ✓ On average, for random reads/writes, we can assume that the disk spins halfway.
- Rotational delay is determined by how fast the disk platters spin.
  - ✓ Average rotational delay = 0.5 rotation ÷ rotational speed
  - ✓ For example, a 5400 RPM disk has an average rotational delay of:
    0.5 rotation / (5400 rotations/minute) x 60 secs/1 min ≈ 5.55ms

I/O Scheduling for Block Devices

- Transfer time
  - Time to read the bits in the target sector.
  - Average transfer time per sector= 1/RPM x 1/(avg # sectors/track) x 60 secs/1 min

- The overall internal response time is the sum of average seek time, rotational delay, transfer time, and overhead.

- Assume a disk has the following specifications.
  - An average seek time of 9ms
  - A 5400 RPM rotational speed
  - Avg number of sectors/track= 400
  - 0.2 ms of overheads

- How long does it take to read 2 continuous sectors?
  - The average rotational delay is 5.55 ms.
  - The transfer time will be about 2x1/5400x1/400x60≈ 0.05ms.
  - The internal response time is then 9ms + 5.55ms + 0.05ms + 0.2ms = 14.8ms. That's 14,800,000 CPU cycles for a 1GHz processor!

I/O Scheduling for Block Devices

- Disk scheduling
  - Disks are several orders of magnitude slower than main memory
    - ✓ The performance of disk I/O is vital for system
    - ✓ Internal response time >> external response time for reading/writing a sector
    - ✓ How to reduce the internal response time?
      - Rotational latency
      - Internal transfer time
      - Overhead
      - Seek time
  - Disk scheduling, minimizing seek distance (seek time)
    - ✓ Possibly reorder stream of read/write requests to improve performance.

I/O Scheduling for Block Devices

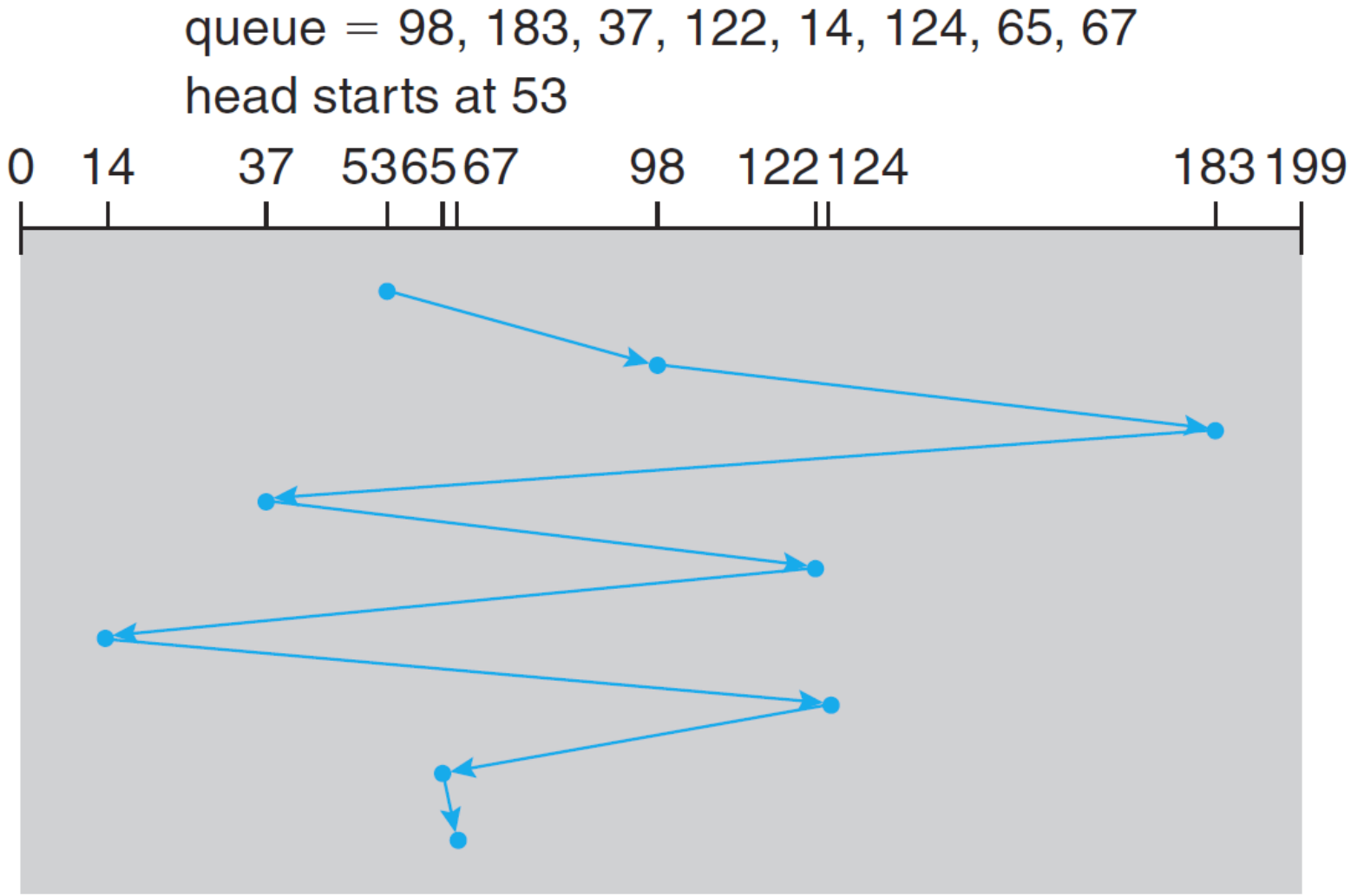
- Algorithms to schedule disk I/O requests
  - FCFS
  - SSTF
    - ✓ (Shortest-service-time-first): pick the request that requires the least movement of the head
  - SCAN
    - ✓ (back and forth over disk): good service distribution
  - C-SCAN
    - ✓ (Circular Scan): one way with fast return
  - Look
  - C-Look

  - Illustrate with a platter containing 200 tracks (#0-#199) and a request queue that buffers 8 read/write requests on the #98, #183, #37, #122, #14, #124, #65, #67 track.
  - Current head location: #53 track.

I/O Scheduling for Block Devices

FCFS

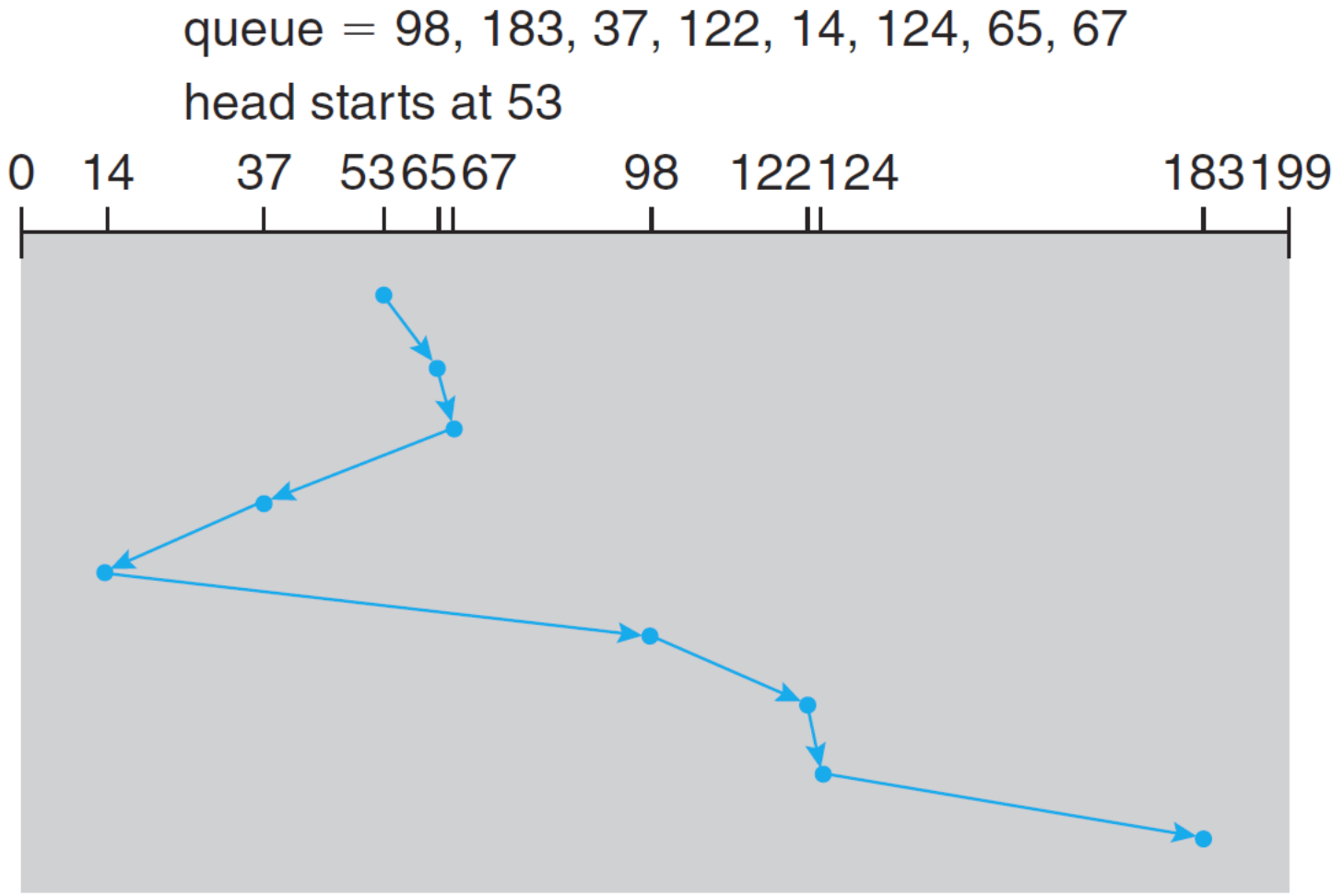
- Most common baseline policy
- Total head movement of 640 tracks



I/O Scheduling for Block Devices

❑ SSTF

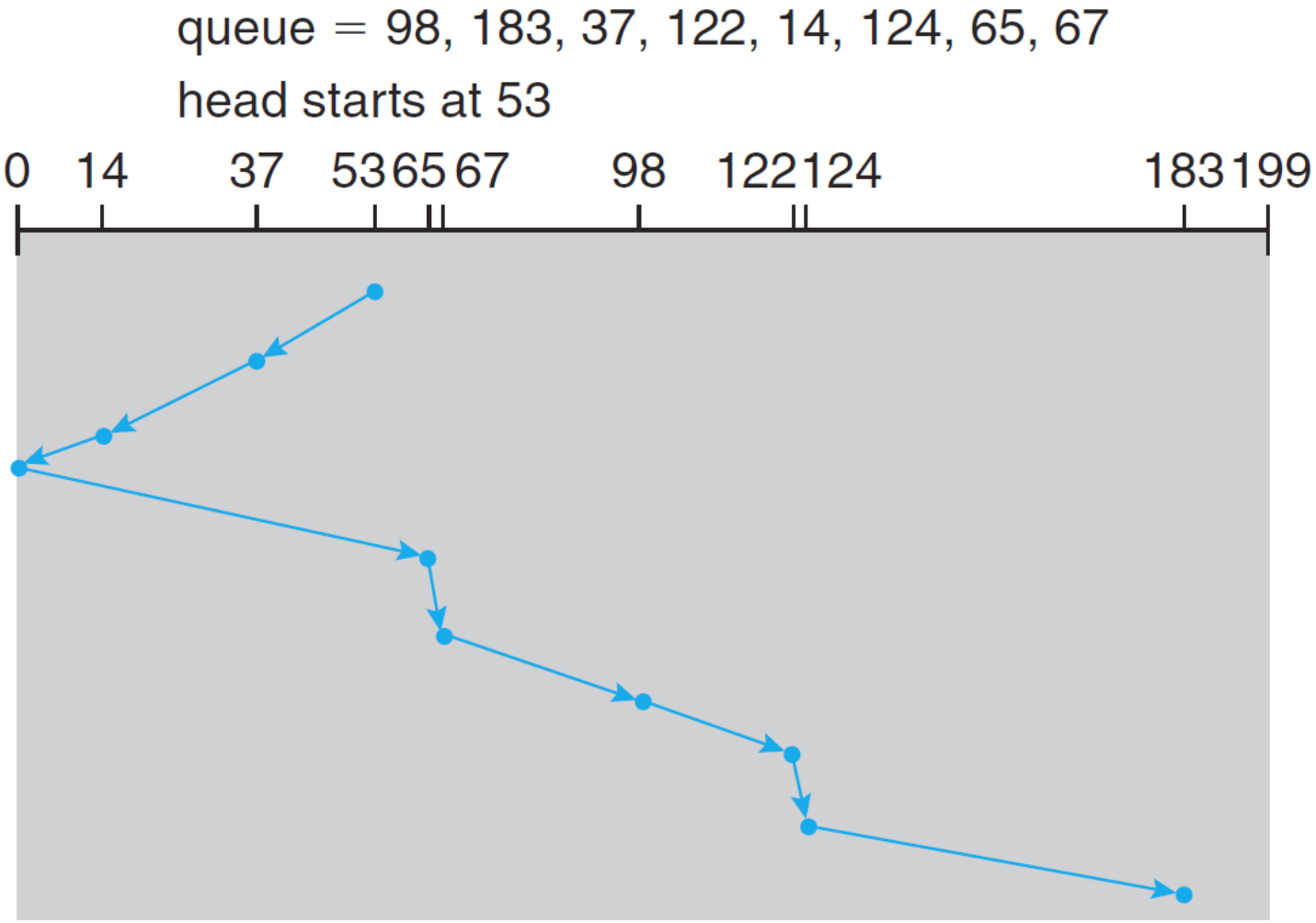
- Selects the request with the minimum seek time from the current head position.
- May cause starvation of some requests (e.g., 183 is scheduled last).
- Total head movement of 236 tracks.



I/O Scheduling for Block Devices

SCAN- (or Elevator algorithm)

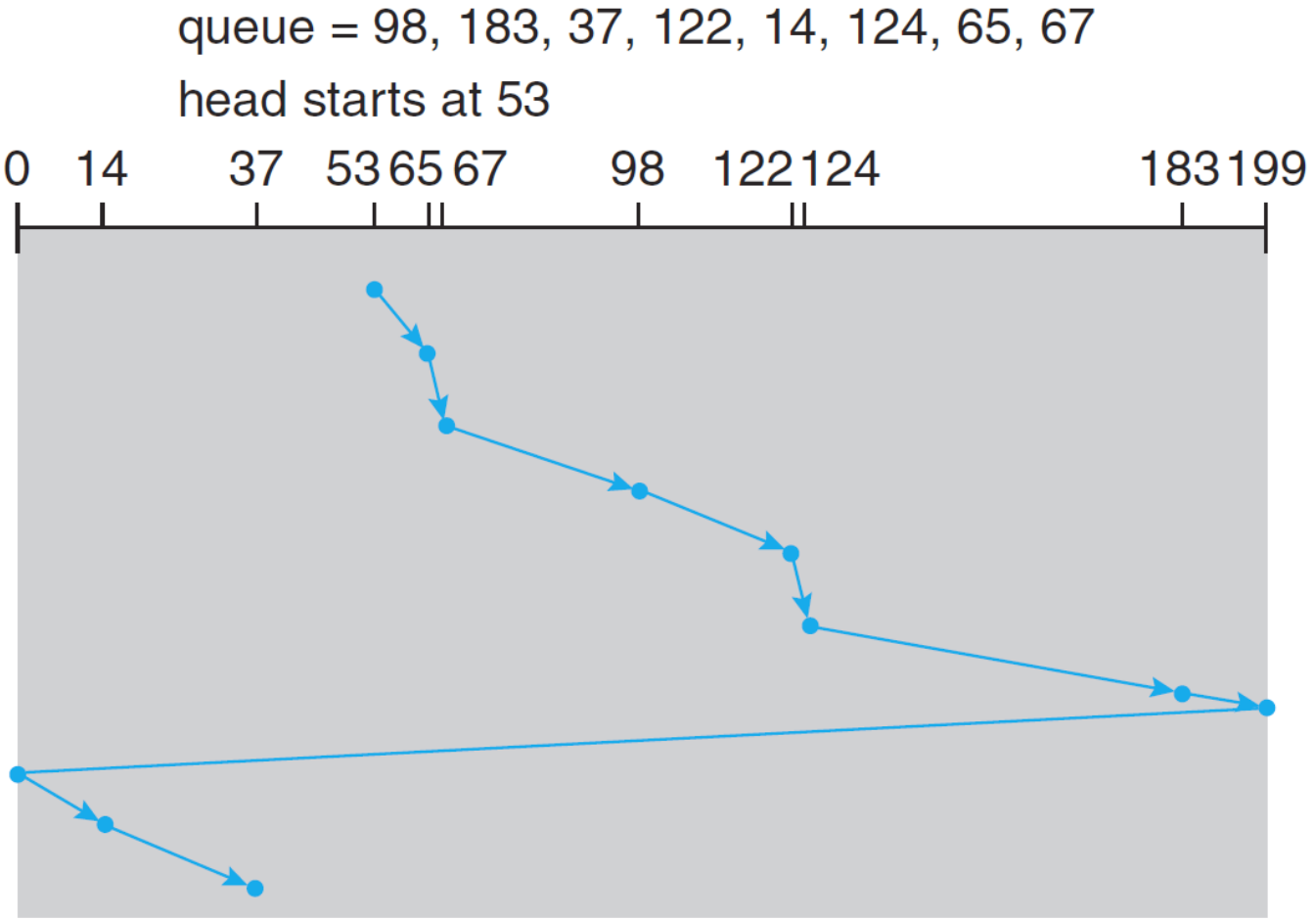
- The disk arm continues to move to one end of the disk, and then move toward the other end, serving requests during the movement.
- Assume initially moving backwards, total head movement of 236 tracks.



I/O Scheduling for Block Devices

C-SCAN

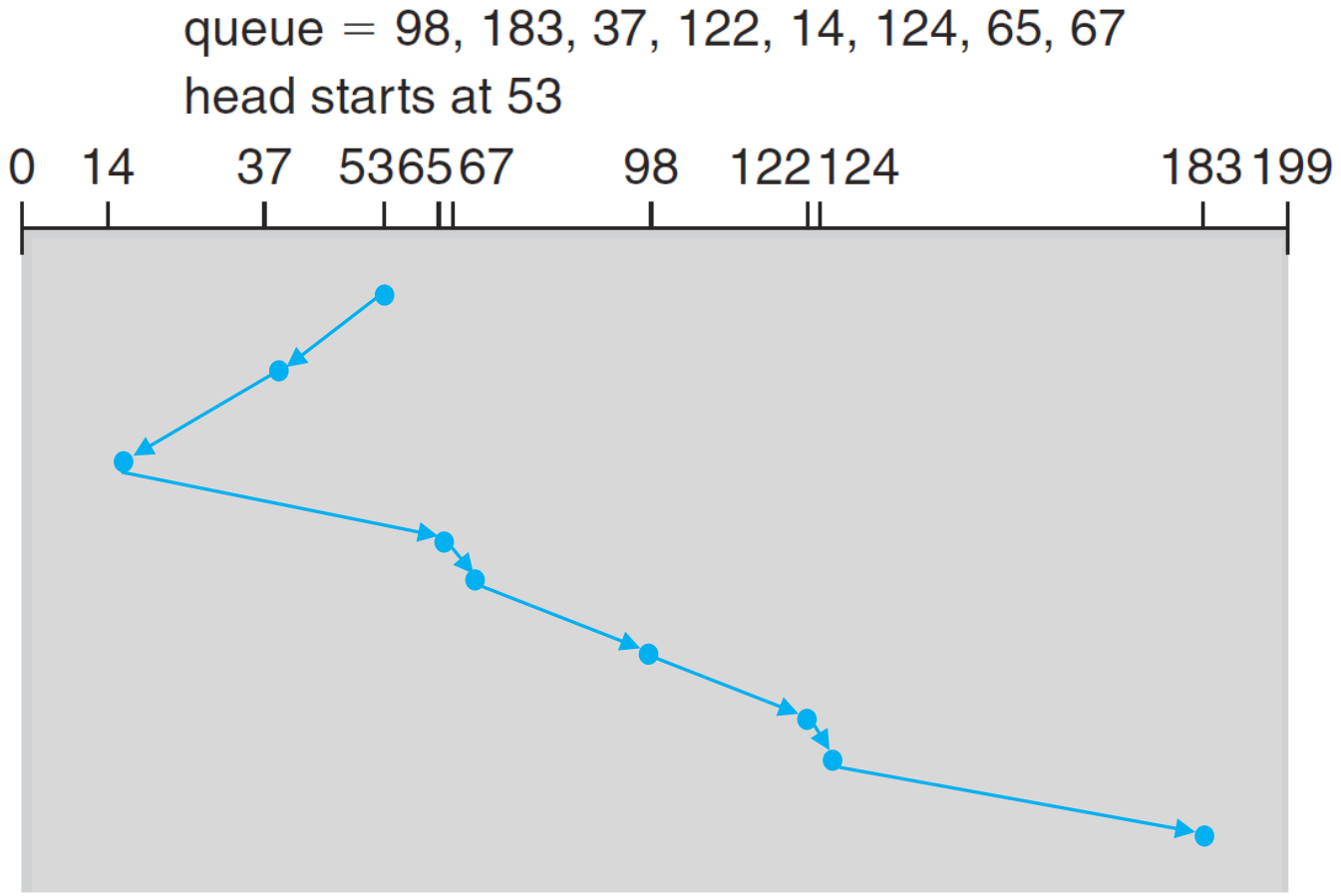
- Provides a more uniform wait time than SCAN. The head moves to one end of the disk, serving requests as it goes.
- When it reaches the end, it immediately returns to the other end of the disk, without servicing any requests on the return trip.
- Treats the tracks as a circular list that wraps around from the last track to the first one.
- Assume initially moving forwards, total head movement of 183(+199) tracks.



I/O Scheduling for Block Devices

❑ Look

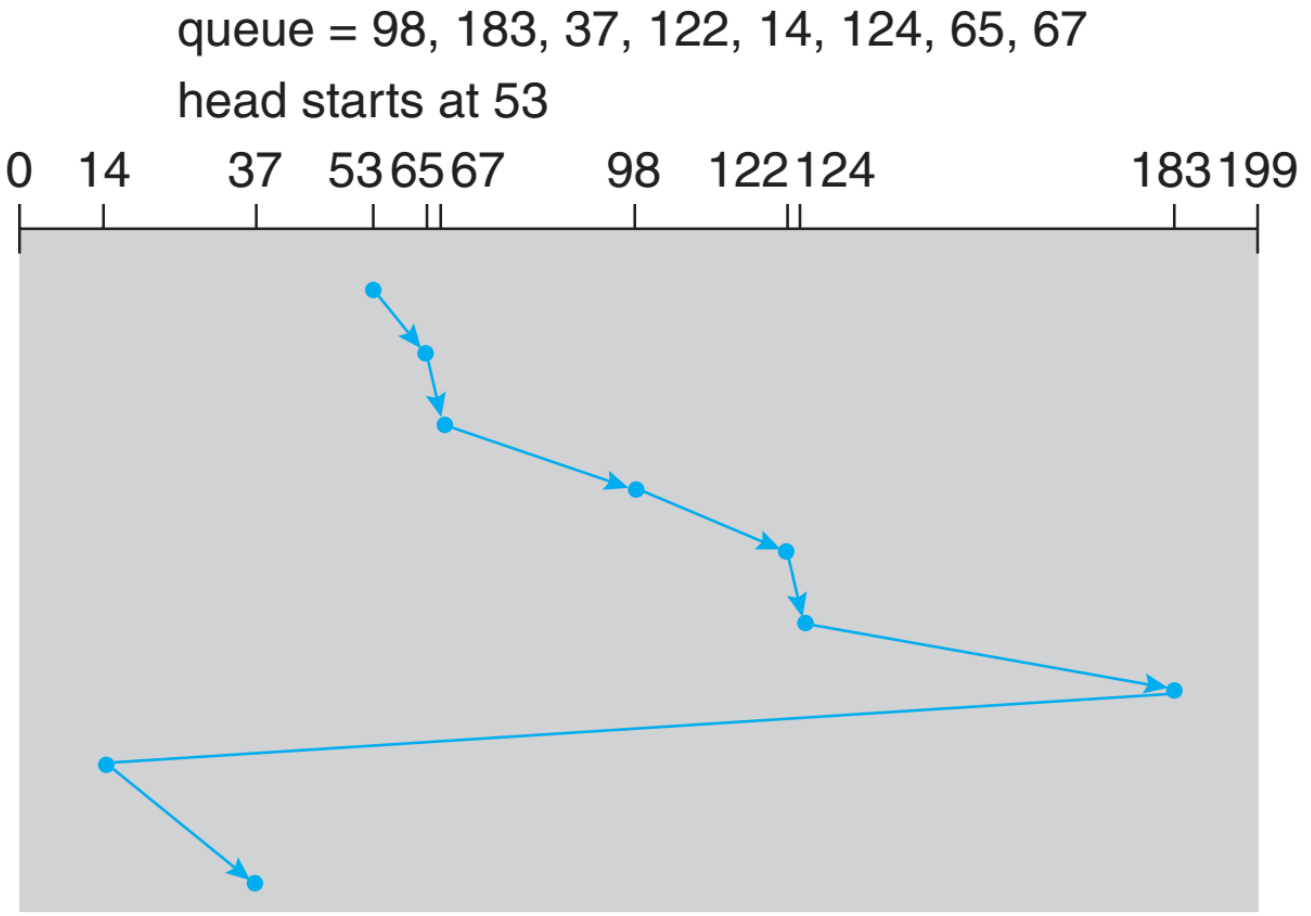
- Similar to SCAN.
- The disk arm only goes as far as the last request in each direction, without going all the way to the end of the disk.
- Assume initially moving backwards, sequence will be (53), 37, 14, 65, 67, 98, 122, 124, 183.
- Total head movement of 53-14+183-14=208 tracks.



I/O Scheduling for Block Devices

C-Look

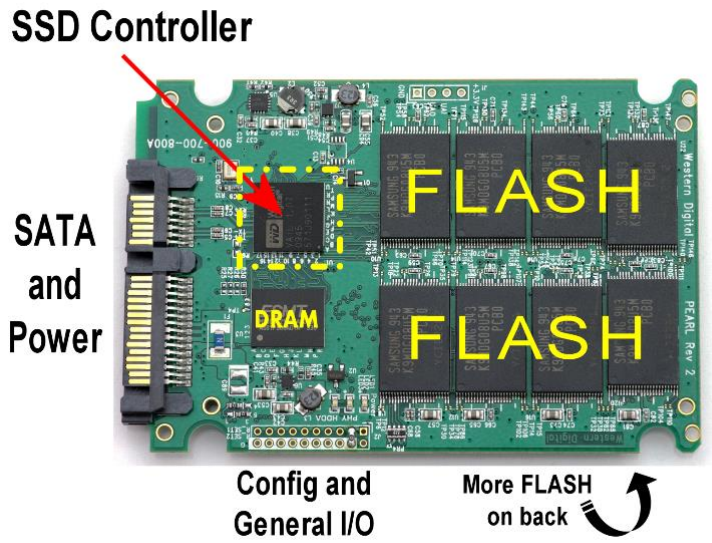
- Similar to C-SCAN.
- The disk arm only goes as far as the last request in each direction, then reverses direction immediately, without first going all the way to the end of the disk.
- Assume initially moving forwards, total head movement of 153(+169) tracks.



I/O Scheduling for Block Devices

❑ Solid State Drive (SSD)

- ➢ No moving mechanical components
- ➢ Have lower internal response time.
- ➢ Use flash memory
  - ✓ retaining data without power
  - ✓ No seek latency
  - ✓ Efficient random I/O, 100ns compared to 1-10ms with HDD
  - ✓ Limited life time by # of rewrites
  - ✓ Data loss from SSD may not be recovered
- ➢ Small and light in weight
- ➢ Generally require $\frac{1}{2}$ of the power of HDDs.



❑ In order to provision high reliability and performance, hybrid drives, which containing a large hard disk drive and an SSD, are applied.