CS420: Operating Systems

Mass Storage Structure

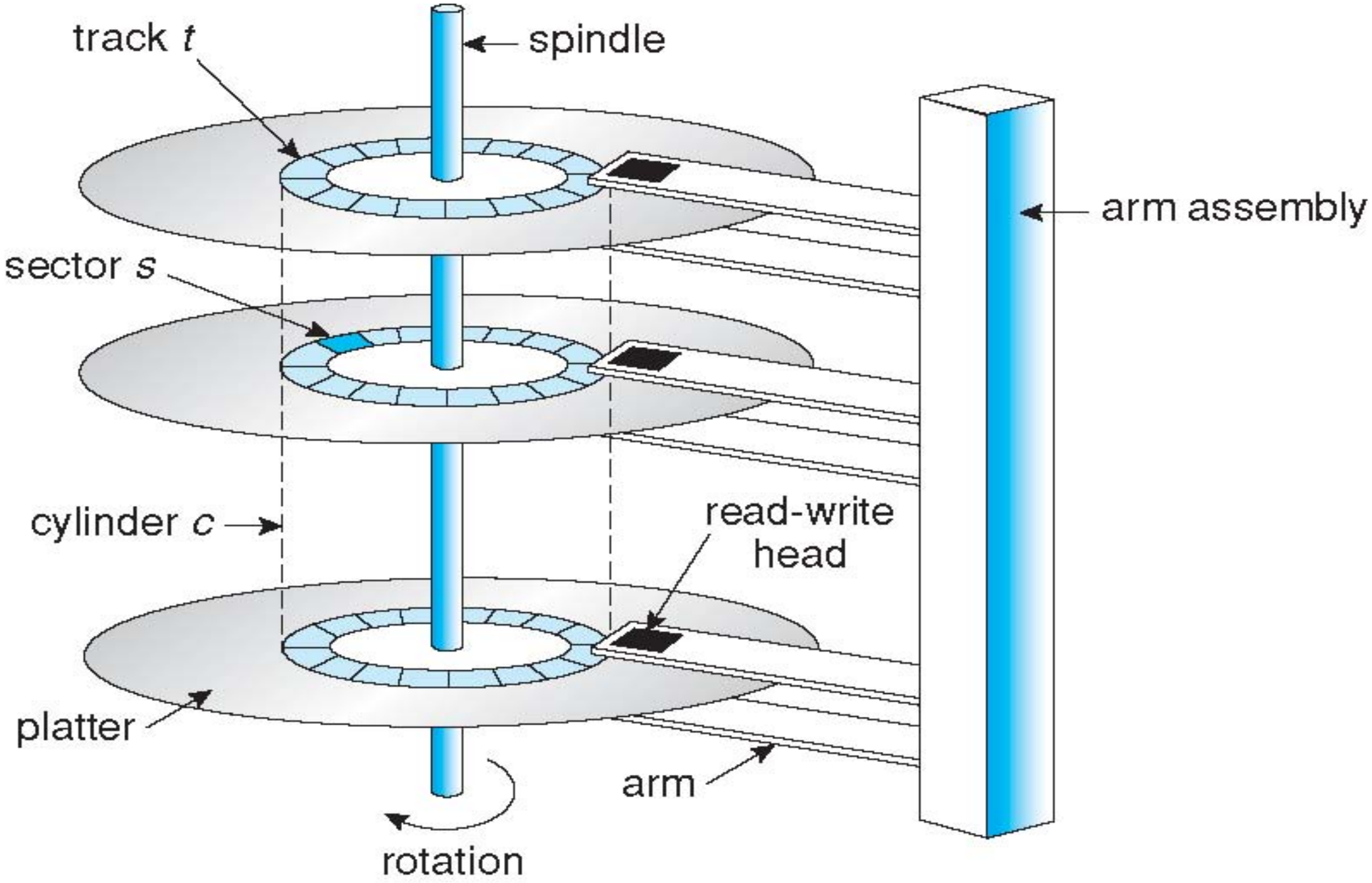
James Moscola
Department of Physical Sciences
York College of Pennsylvania

Overview of Mass Storage Structure

- **Magnetic disks provide bulk of secondary storage of modern computers**
  - Drives rotate at 60 to 250 times per second
  - Transfer rate is rate at which data flows between the drive and the computer
  - Positioning time (random-access time) consists of:
    (1) Seek time - time to move disk arm to desired cylinder
    (2) Rotational Latency - time for desired sector to rotate under the disk head
  - Head crash results from disk head making contact with the disk surface and typically results in a dead disk

- **Drive attached to computer via I/O bus**
  - Busses vary, including EIDE, ATA, SATA, USB, Fibre Channel, SCSI, SAS, Firewire
  - Host controller in computer uses bus to talk to disk controller built into drive or storage array

Moving-head Disk Mechanism



Disk Structure

- **Disk drives are addressed as large 1-dimensional arrays of logical blocks (typically 512 or 4096 bytes), where the logical block is the smallest unit of transfer**

- **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
    - Number of sectors per track decreases from outer tracks to inner tracks

Magnetic Disks

- **Performance**
  - Effective Transfer Rate – real – 1 Gb/s (~125 MB/s)
  - Seek time from 3 ms to 12 ms (9 ms common for desktop drives)
  - Rotational Latency based on spindle speed
    - 1/(RPM * 60)

Spindle [rpm]	Average latency [ms]
4200	7.14
5400	5.56
7200	4.17
10000	3
15000	2

(From Wikipedia)

Disk Scheduling

- **The operating system is responsible for using hardware efficiently — for the disk drives, this means having a fast access time and disk bandwidth**

- **Want to minimize seek time**

- **Disk bandwidth is the total number of bytes transferred, divided by the total time between the first request for service and the completion of the last transfer**
  - Affected by how efficiently data is read from the disk

Disk Scheduling (Cont.)

- **Each disk I/O request includes the following:**
  - Access Mode (i.e. read or write)
  - Disk address for the data transfer
  - A memory address
  - Number of sectors to transfer

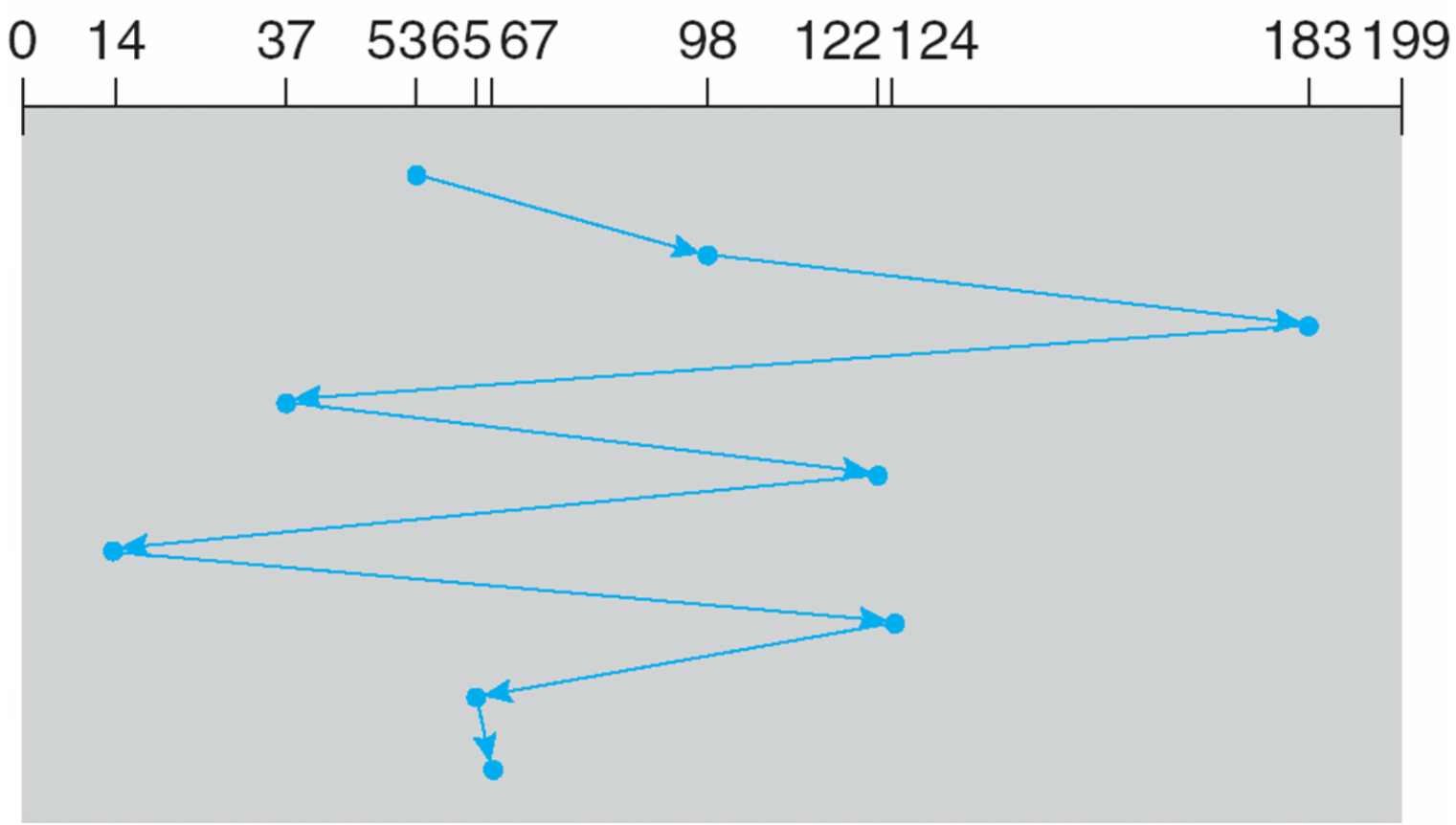
- **OS maintains queue of requests, per disk or device (remember wait-queues?)**

- **Several algorithms exist to schedule the servicing of disk I/O requests**
  - First-Come, First Served (FCFS) scheduling
  - Shortest-Seek Time First (SSTF) scheduling
  - SCAN / C-SCAN scheduling
  - LOOK / C-LOOK scheduling

First-Come, First-Served (FCFS) Scheduling

- **Serve I/O requests in the order that they are received**
  - Fair, but poor performance

queue = 98, 183, 37, 122, 14, 124, 65, 67
head starts at 53

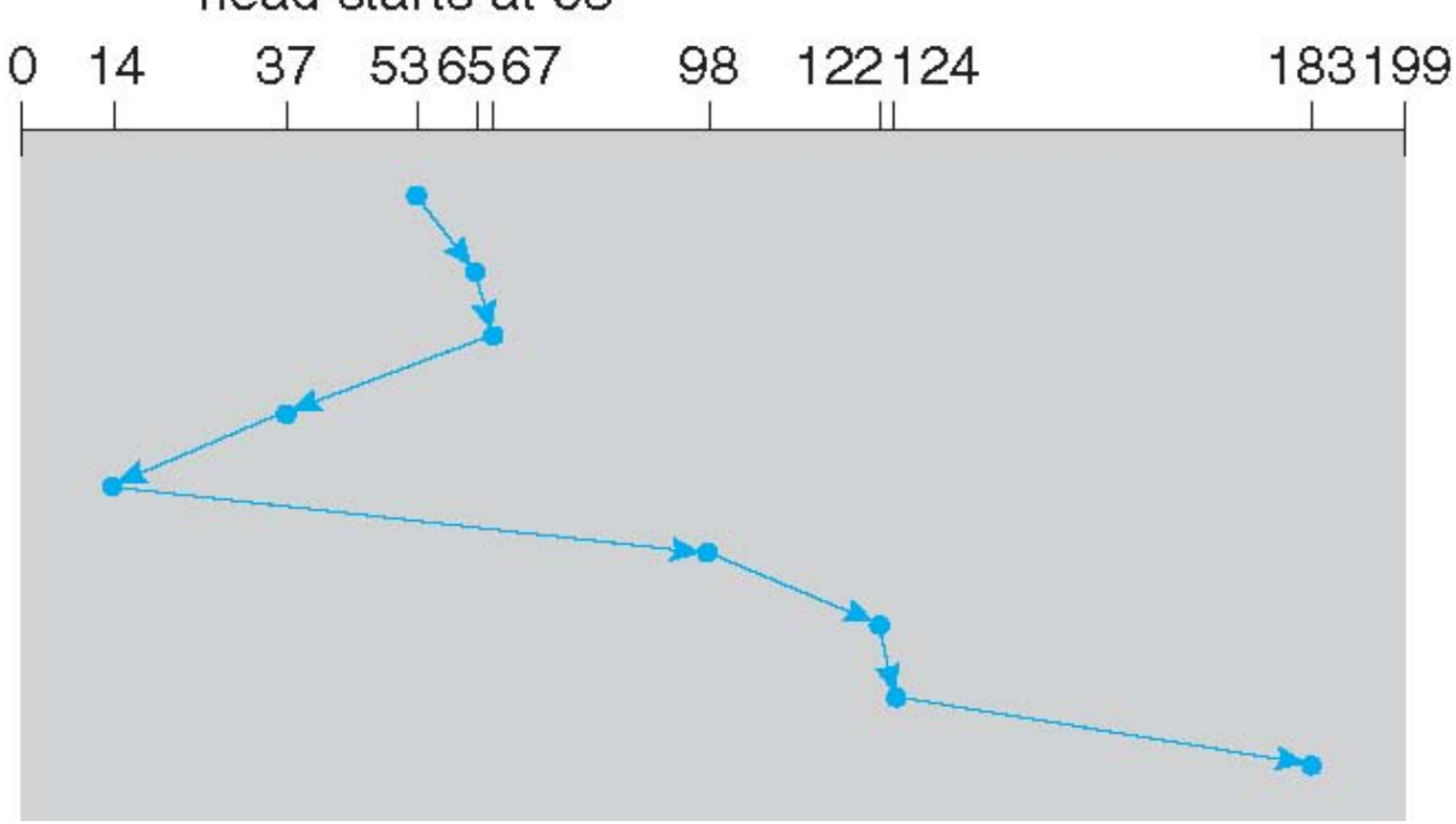


Total head movement of 640 cylinders

Shortest-Seek-Time-First (SSTF) Scheduling

- **Selects the request with the minimum seek time from the current head position**
  - May cause starvation of some requests if requests keep coming in close to current location while a request is waiting on the far side of the disk

queue = 98, 183, 37, 122, 14, 124, 65, 67
head starts at 53

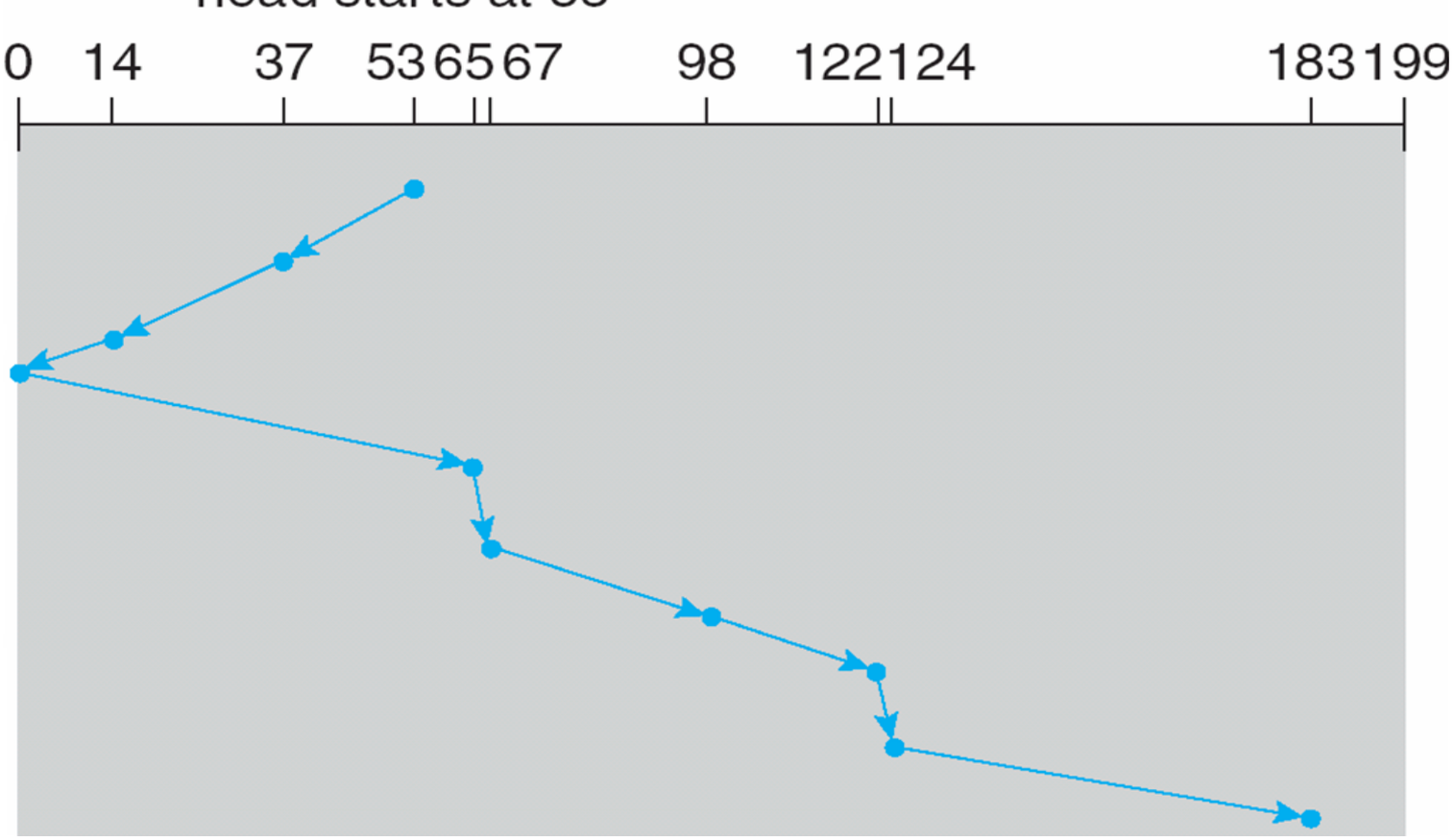


Total head movement of 236 cylinders

SCAN Scheduling

- **Disk arm starts at one end of the disk, and moves toward the other end, servicing requests until it gets to the other end of the disk, where the head movement is reversed and servicing continues**
  - Requests at far end of disk may have to wait long time

queue = 98, 183, 37, 122, 14, 124, 65, 67
head starts at 53

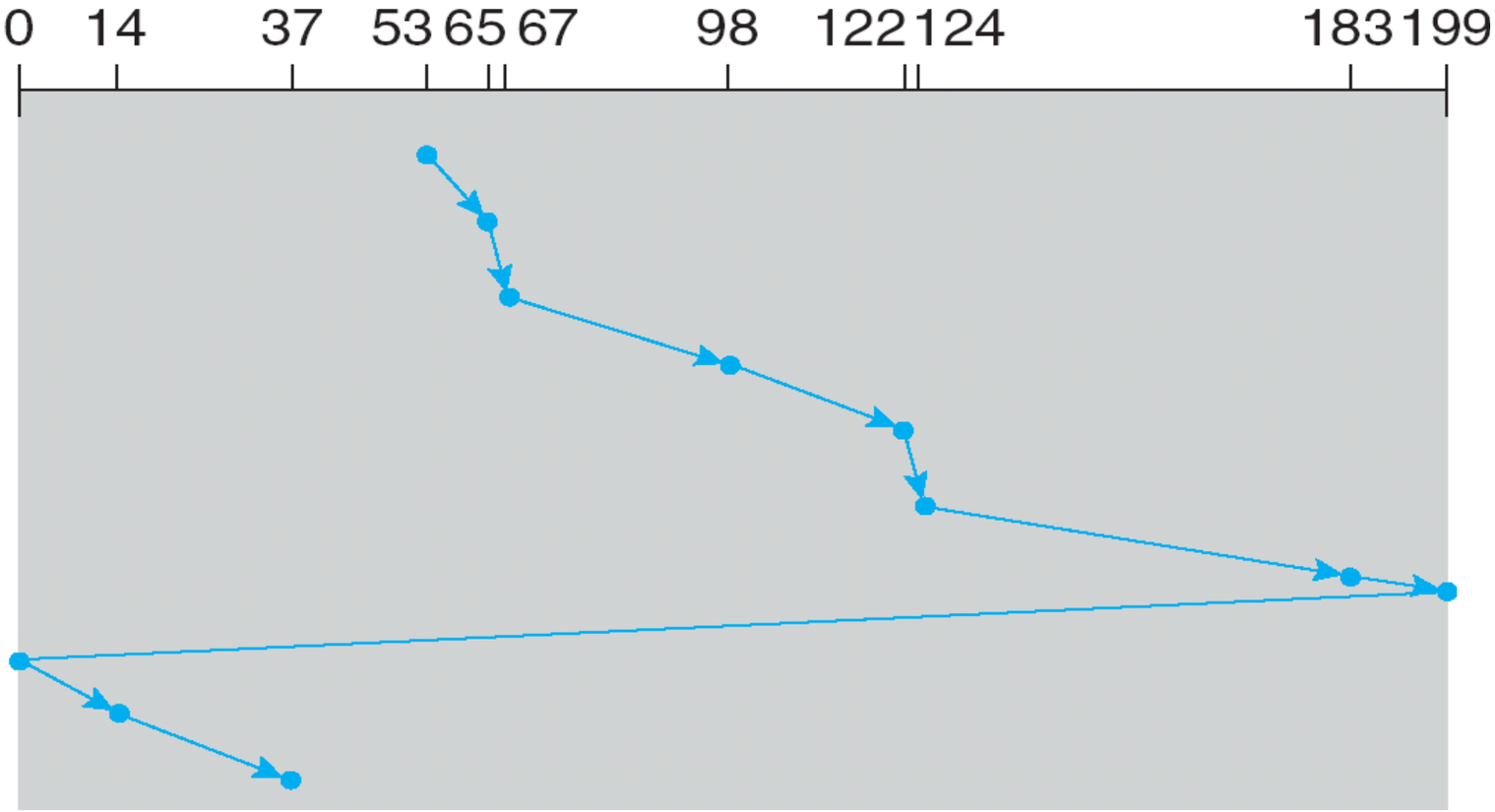


Total head movement of 208 cylinders

C-SCAN (circular scan) Scheduling

- **Disk head moves from one end of the disk to the other, servicing requests as it goes. When it reaches the other end, it immediately returns to the beginning of the disk, without servicing any requests on the return trip**
  - Provides a more uniform wait time than SCAN

queue = 98, 183, 37, 122, 14, 124, 65, 67
head starts at 53



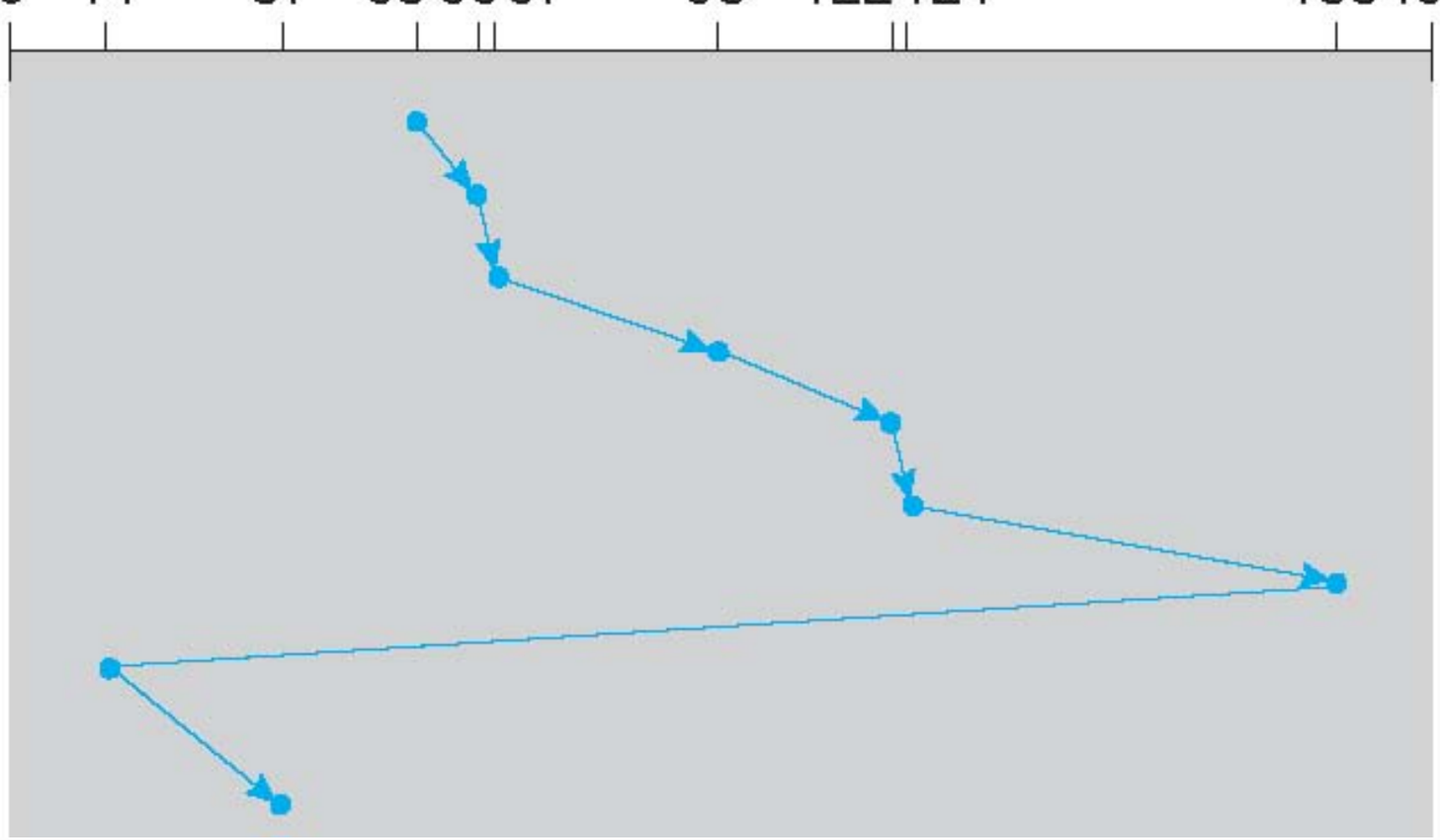
Total head movement of 382 cylinders

C-LOOK Scheduling

- **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**
  - Eliminates unnecessarily traversing to the edge of the disk before turning around

queue = 98, 183, 37, 122, 14, 124, 65, 67

head starts at 53



LOOK is a version of SCAN
C-LOOK a version of C-SCAN

Total head movement of 322 cylinders

Selecting a Disk-Scheduling Algorithm

- **Dependent on the system -- performance depends on the number and types of requests**
  - SSTF is common
  - SCAN and C-SCAN perform better for systems that place a heavy load on disk
    - Less starvation
- **Requests for disk service can be influenced by the file-allocation method**
  - And metadata layout

Disk Management

- **Low-level formatting, or physical formatting — Dividing a disk into sectors that the disk controller can read and write**
  - Each sector can hold header information, plus data, plus error correction code (ECC)
  - Usually 512 bytes of data but can be selectable

- **To use a disk to hold files, the operating system still needs to record its own data structures on the disk**
  - Partition the disk into one or more groups of cylinders, each treated as a logical disk
  - Logical formatting or “making a file system”
  - To increase efficiency most file systems group blocks into clusters
    - Disk I/O done in blocks
    - File I/O done in clusters