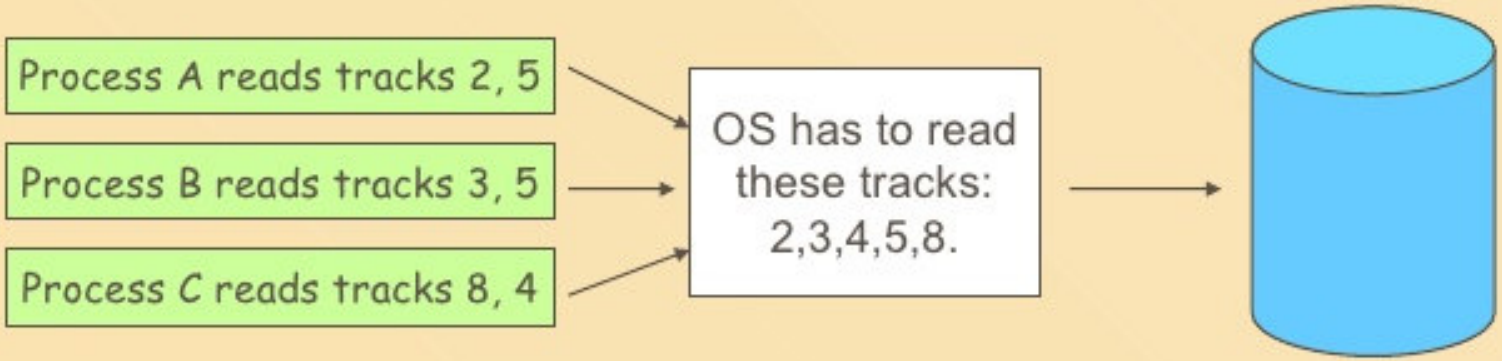
Disk Management

Disk Structure

- Disk drives are addressed as large 1-dimensional arrays of *logical blocks*, 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.

Disk Scheduling

- At runtime, I/O requests for disk tracks come from the processes
- OS has to choose an order to serve the requests



Access time

- Total access time = seek time + rotational delay + data transfer time
- Seek time – time required to move the disk arm to the required track
- Rotational delay – time required to rotate the disk to the required sector
- Data transfer time – time to read/write data from/to the disk

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.

Access time has two major components

Seek time is the time for the disk are to move the heads to the cylinder containing the desired sector.

Rotational latency is the additional time waiting for the disk to rotate the desired sector to the disk head.

Minimize seek time

Disk Scheduling

Seek time ≈ seek distance

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.

Several algorithms exist to schedule the servicing of disk I/O requests. Let us illustrate them with a request queue (0-199).

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

Head pointer 53

Disk Scheduling – FCFS (First Come First Serve)

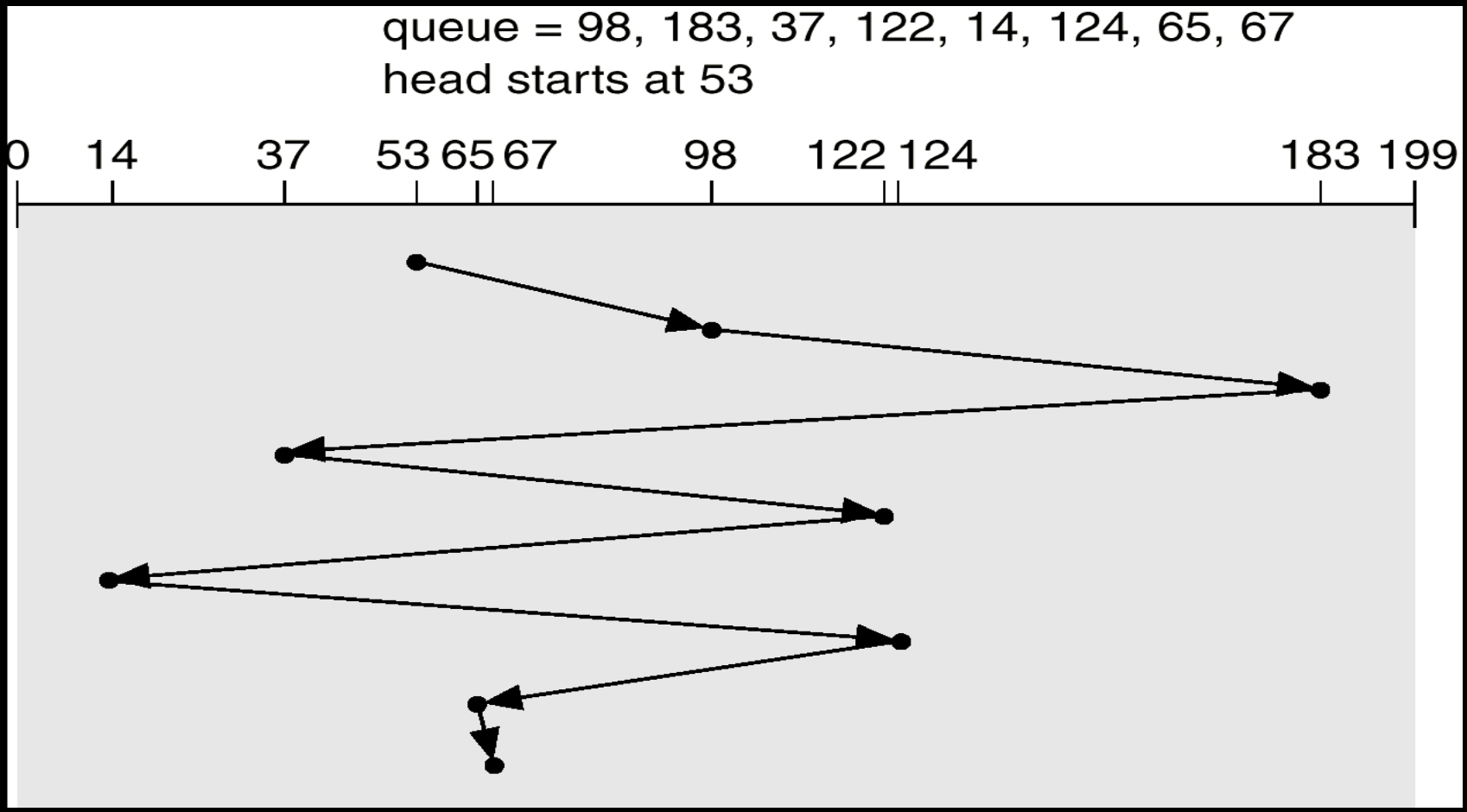
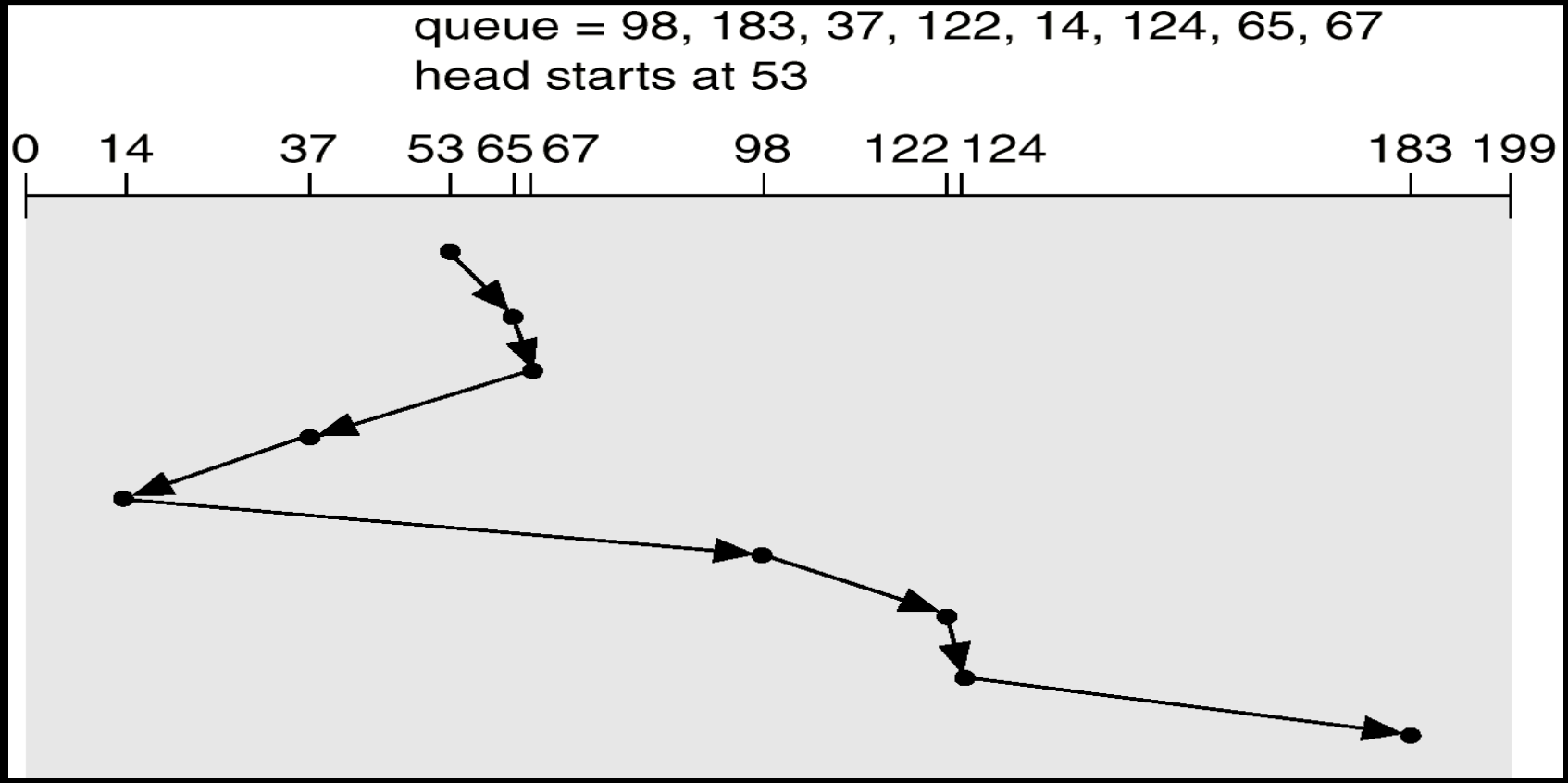


Illustration shows total head movement of 640 cylinders.

Disk Scheduling – SSTF (Shortest Seek Time First)



Selects the request with the minimum seek time from the current head position. SSTF scheduling is a form of SJF scheduling; may cause starvation of some requests.

Illustration shows total head movement of 236 cylinders.

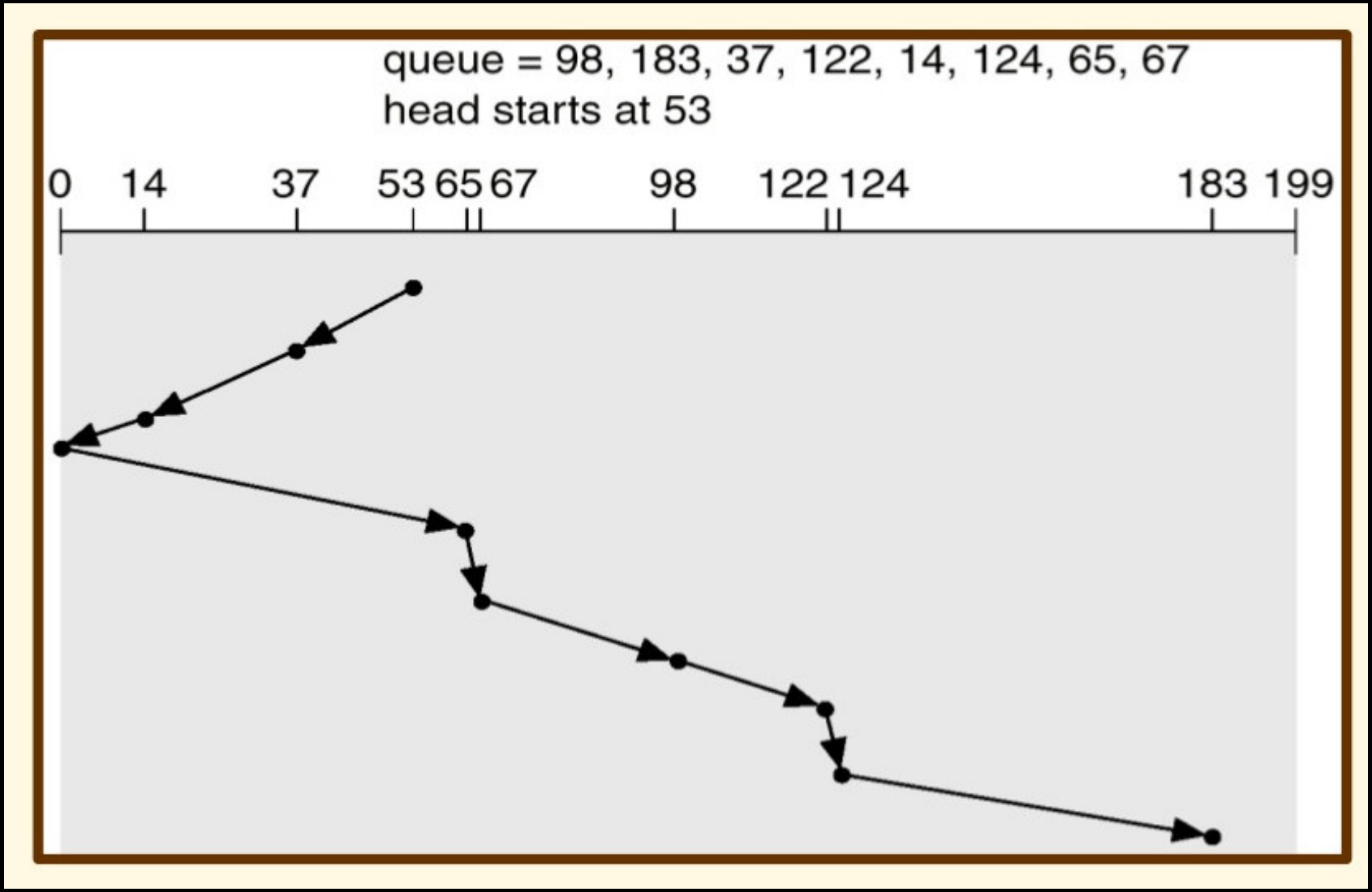
Disk Scheduling – Scan (Elevator)

The 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.

Sometimes called the *elevator algorithm*.

Illustration shows total head movement of 208 cylinders.

Disk Scheduling – Scan (Elevator)



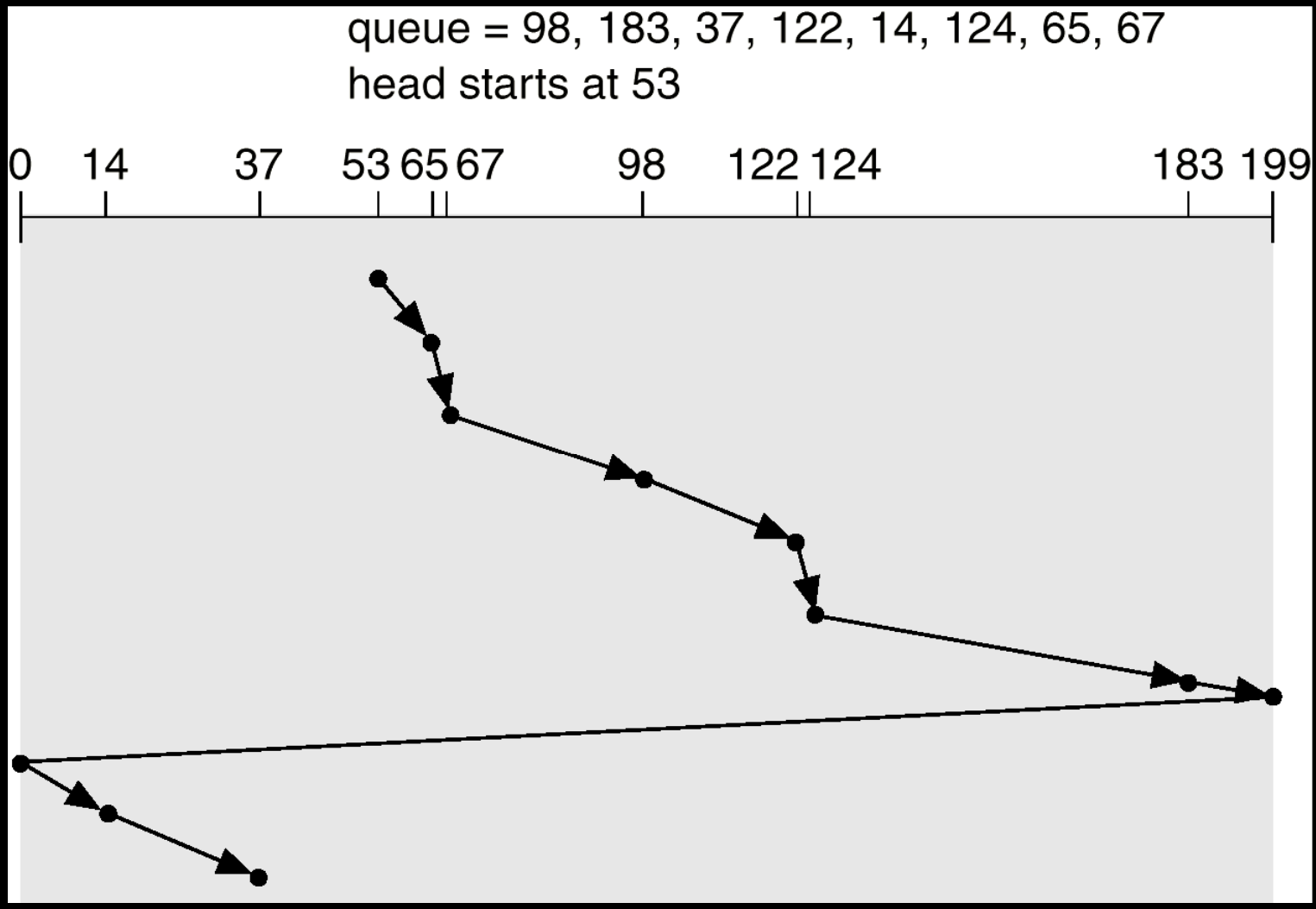
Disk Scheduling – C-Scan

Provides a more uniform wait time than SCAN.

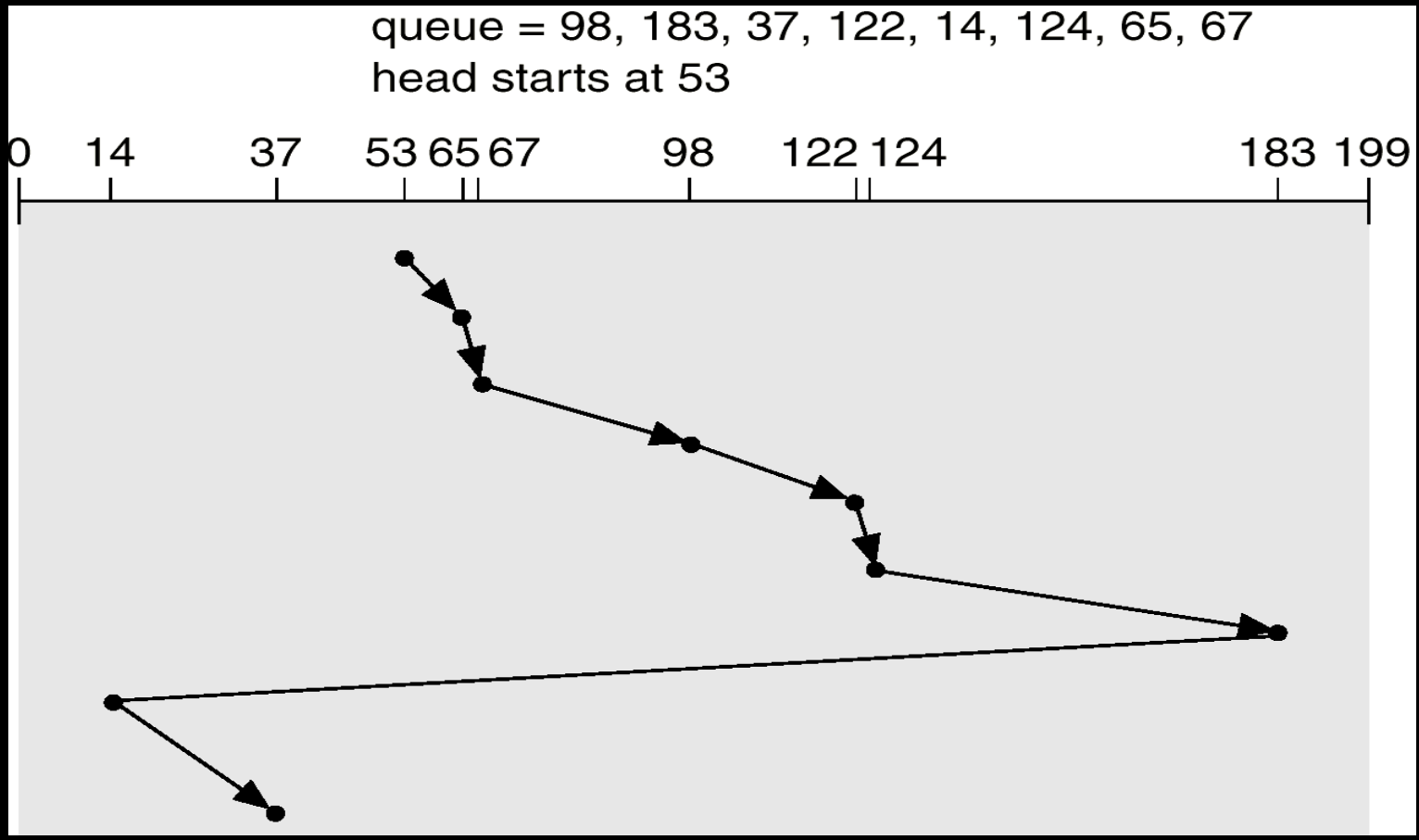
The head moves from one end of the disk to the other. servicing requests as it goes. When it reaches the other end, however, it immediately returns to the beginning of the disk, without servicing any requests on the return trip.

Treats the cylinders as a circular list that wraps around from the last cylinder to the first one.

Disk Scheduling – C-Scan



Disk Scheduling – C-Look



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.

Disk Scheduling – Choice of Algorithm

SSTF is common and has a natural appeal

SCAN and C-SCAN perform better for systems that place a heavy load on the disk.

Performance depends on the number and types of requests.

Requests for disk service can be influenced by the file-allocation method.

The disk-scheduling algorithm should be written as a separate module of the operating system, allowing it to be replaced with a different algorithm if necessary.

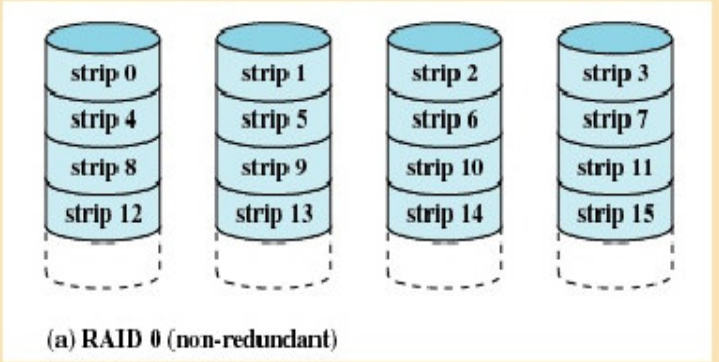
Either SSTF or LOOK is a reasonable choice for the default algorithm.

RAID

- Redundant Array of Independent Disks
- A set of physical disk drives viewed by the OS as a single logical drive
- Data are distributed across the physical drives. May improve performance.
- Redundant disk stores parity information. Recoverability, reliability.

RAID 0 (Non-redundant)

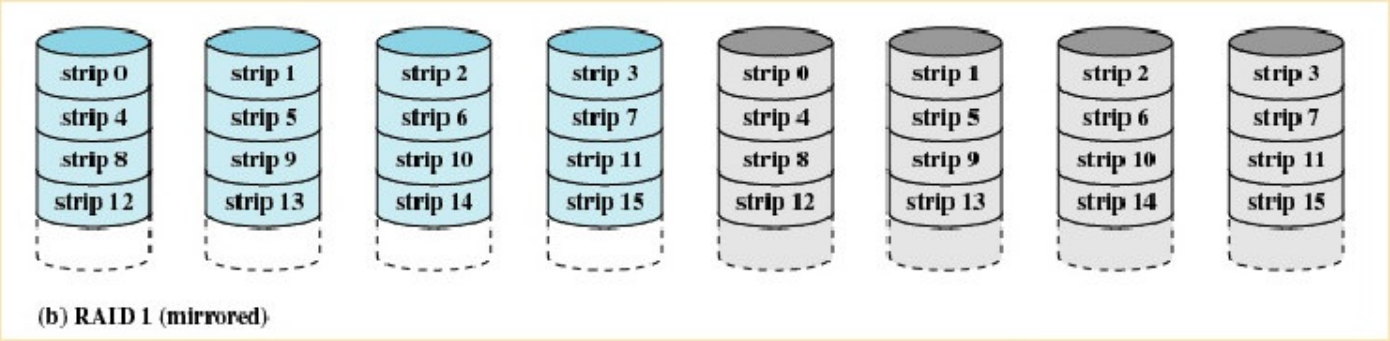
- The logical disk is divided into strips, mapped round robin to consecutive physical disks
- Improve performance in disk read/write
- Not fault tolerant



(a) RAID 0 (non-redundant)

RAID 1 (Mirrored)

- Each disk is mirrored by another disk
- Good performance if the hardware supports concurrent read/write to the mirrored pair
- Reliable, but expensive



(b) RAID 1 (mirrored)