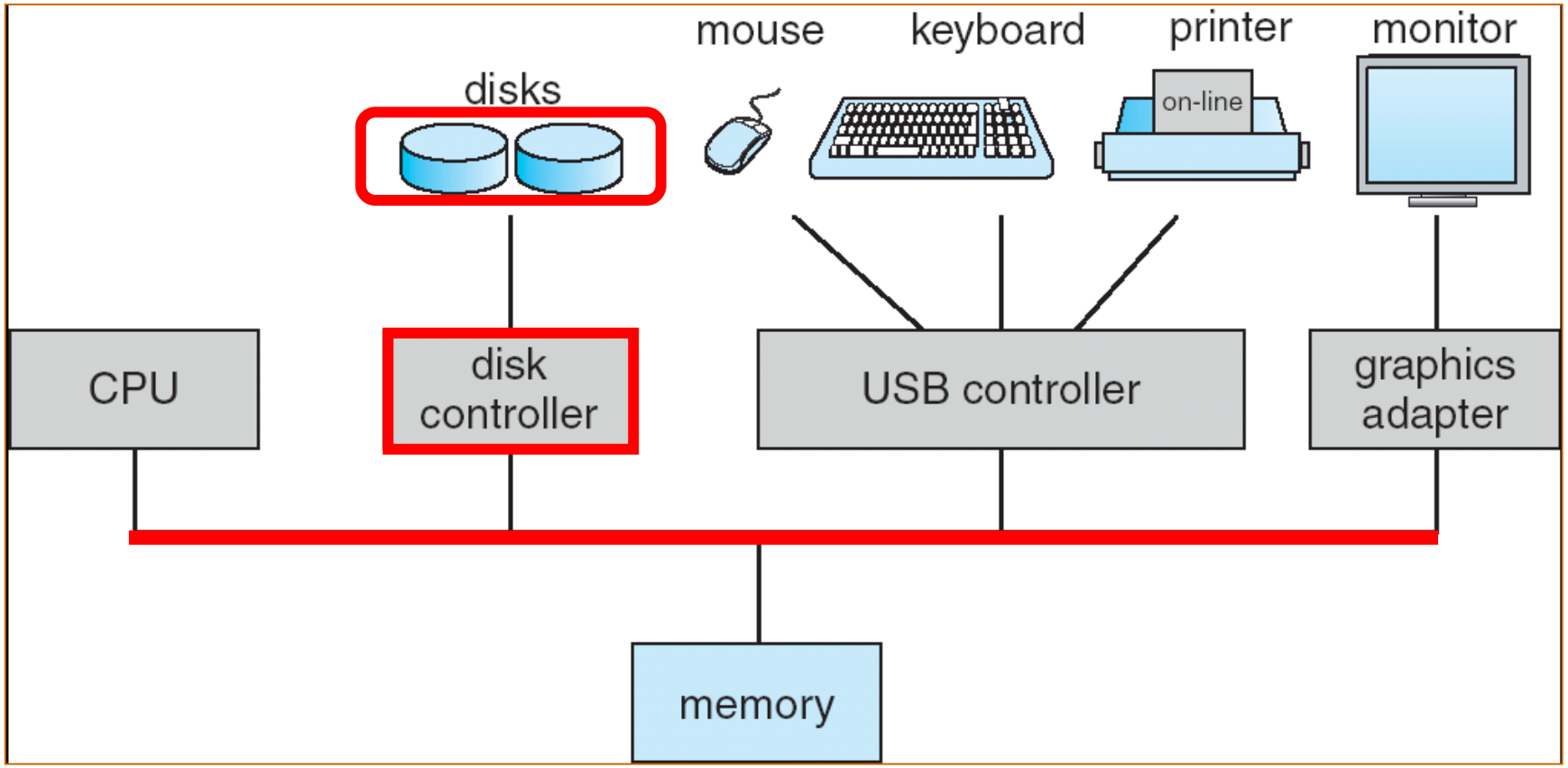
Mass-Storage Structure

Outline

- Overview of Mass-Storage
- Disk Structure
- Disk Scheduling
- Disk Management

Overview of Mass-Storage

Computer System Organization

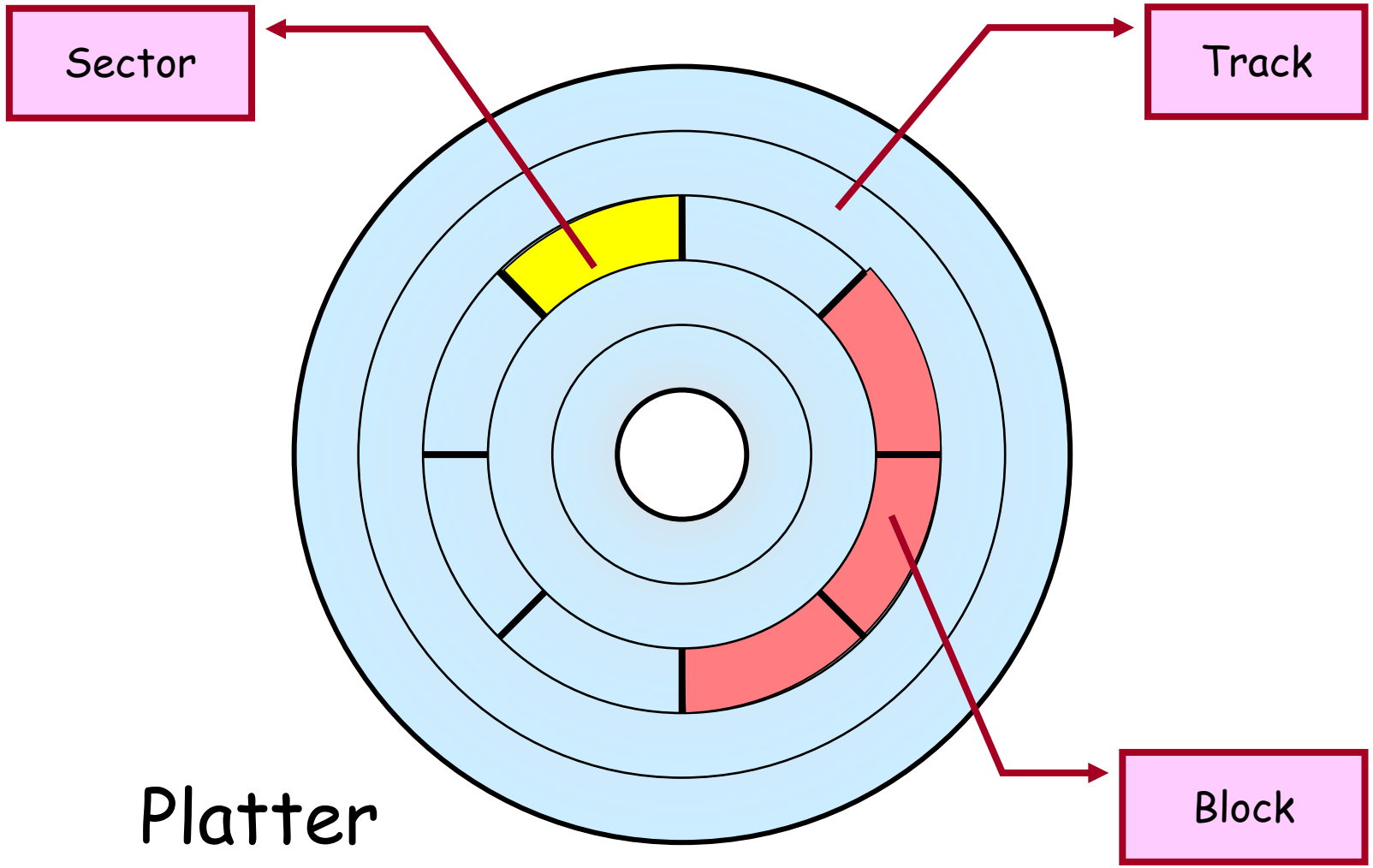


Overview of Mass Storage Structure

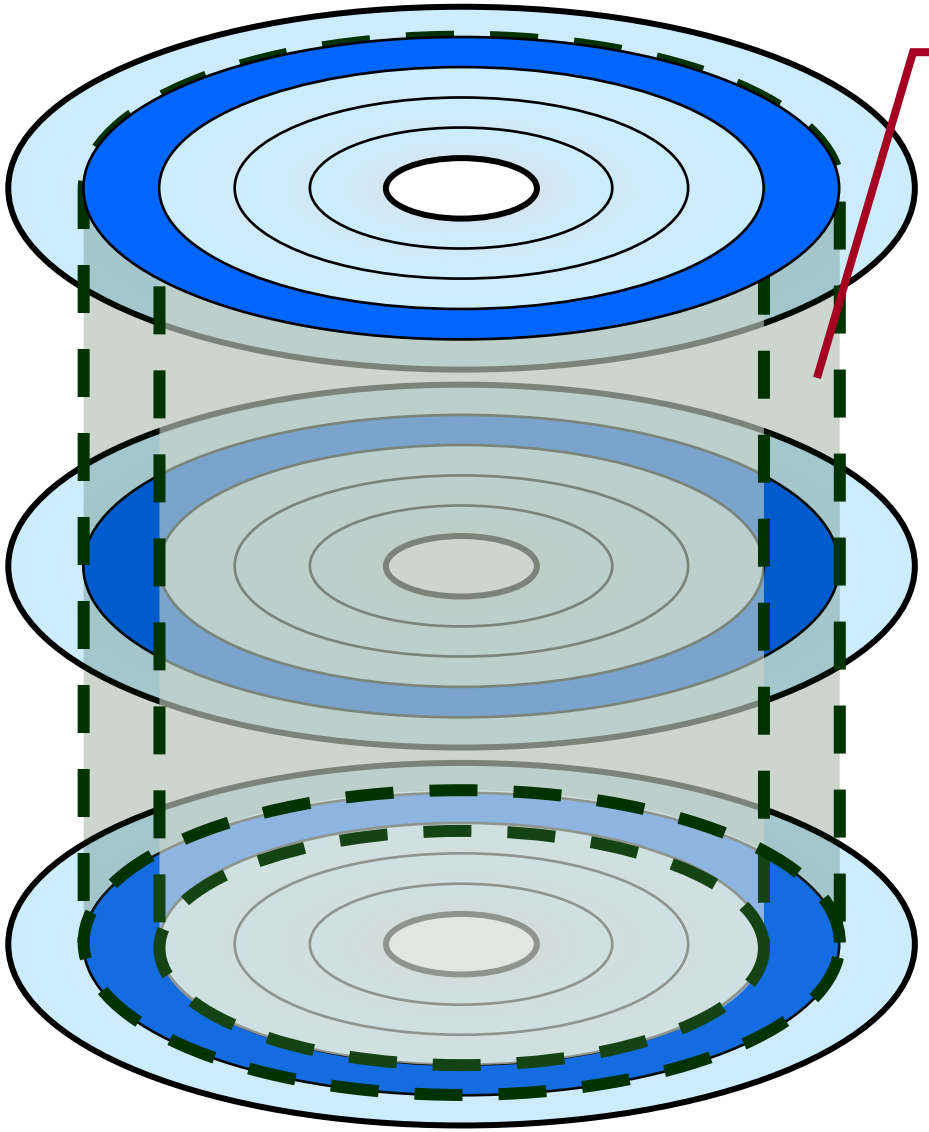
- Disks can be removable
- Drive attached to computer via **I/O bus**
  - Busses vary, including **EIDE, ATA, SATA, USB, Fibre Channel, SCSI**
  - **Host controller** in computer uses bus to talk to **disk controller** built into drive or storage array

Disk Structure

Properties of a Hard Magnetic Disk



Properties of a Hard Magnetic Disk

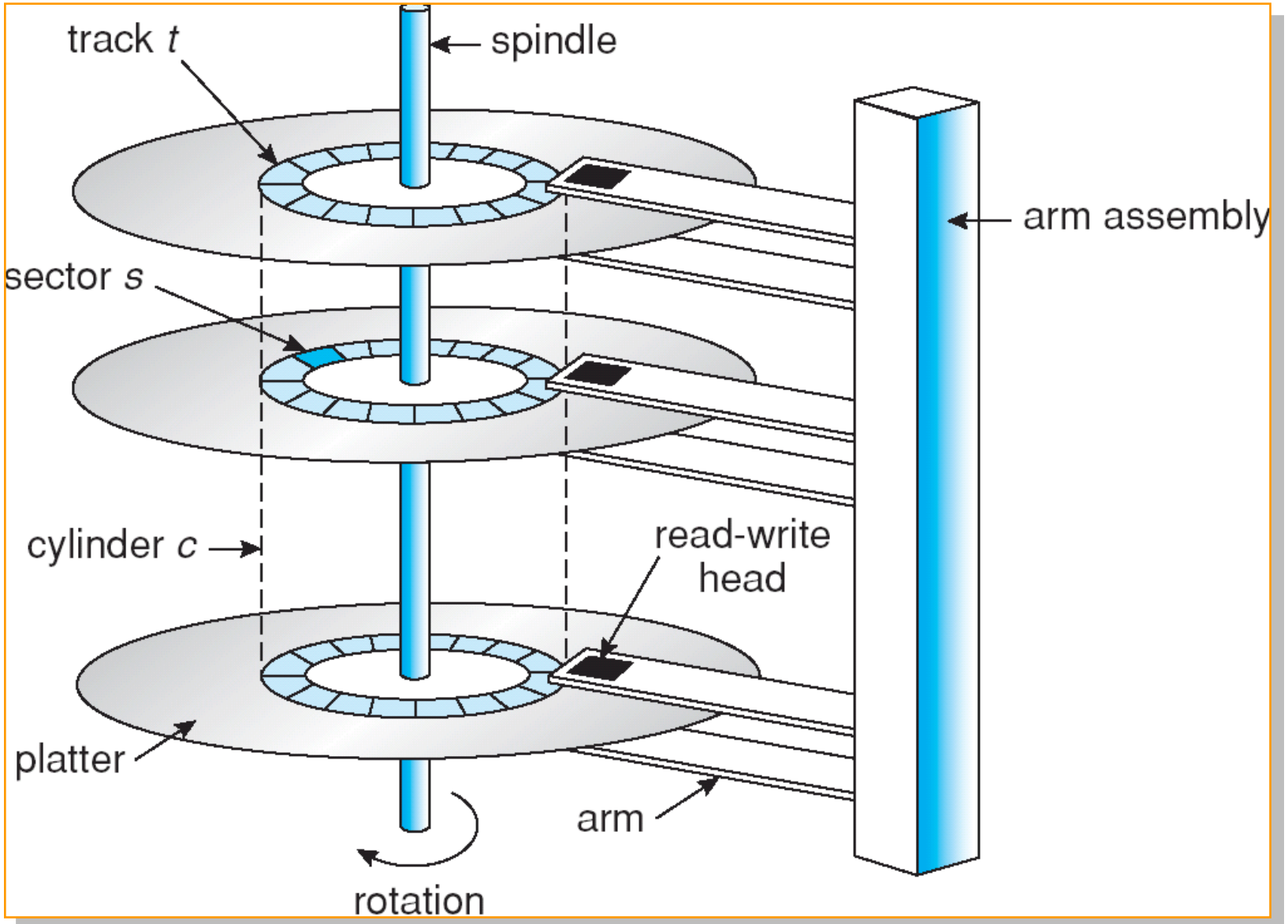


Platters

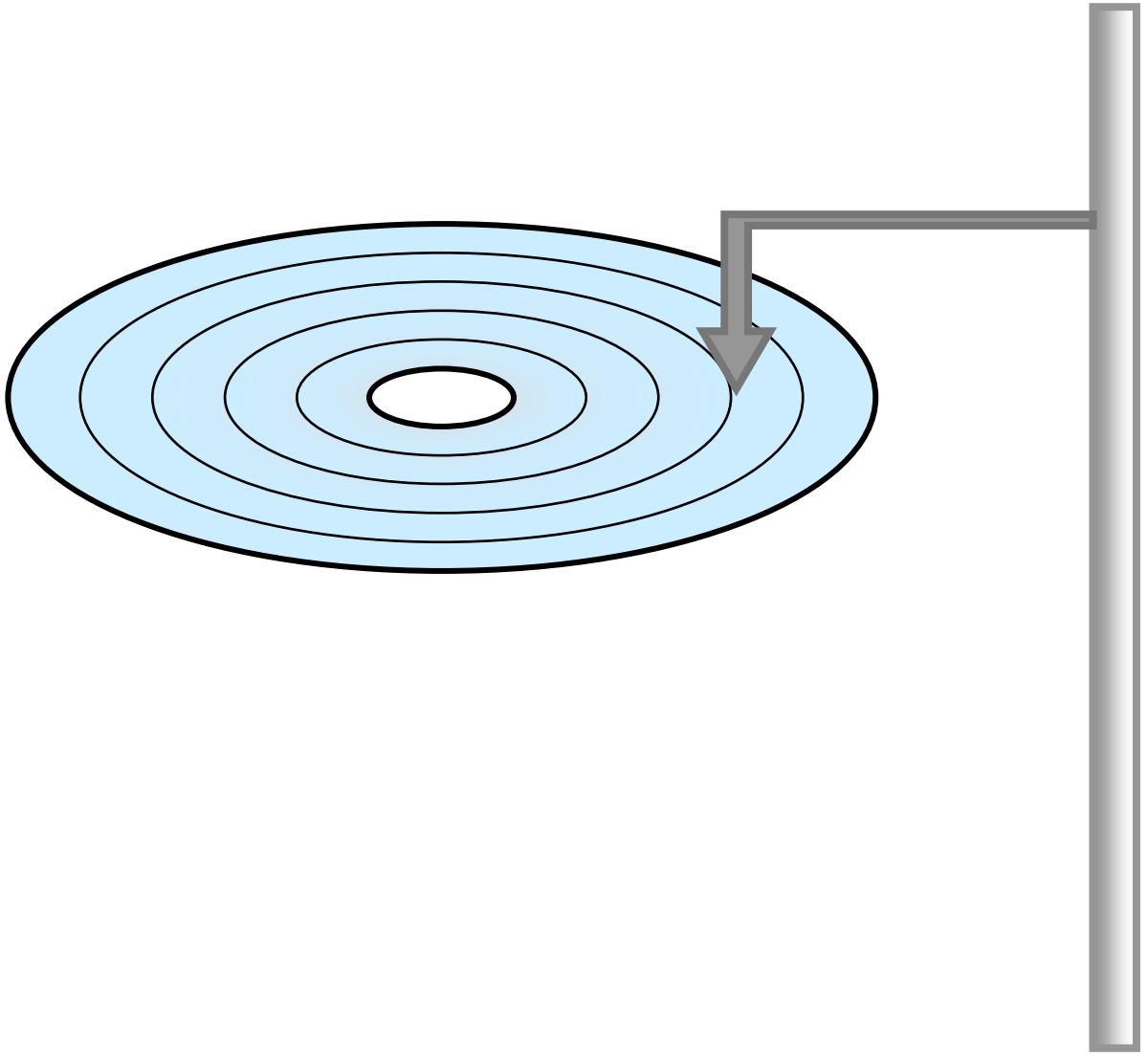
Cylinder

Cylinder: all the tracks under the head at a given point on all surface

Moving-head Disk Mechanism

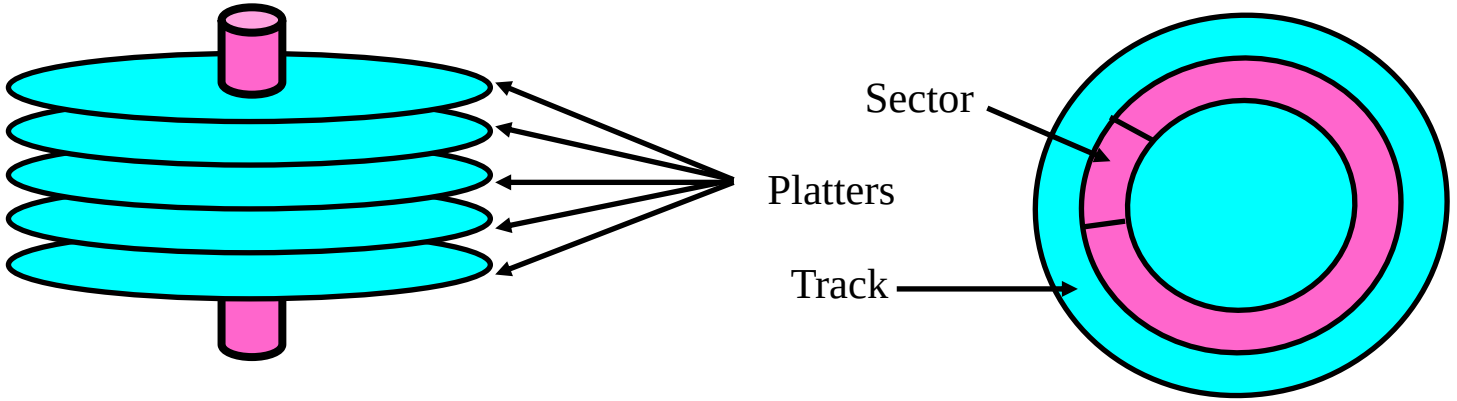


Moving-head Disk Mechanism



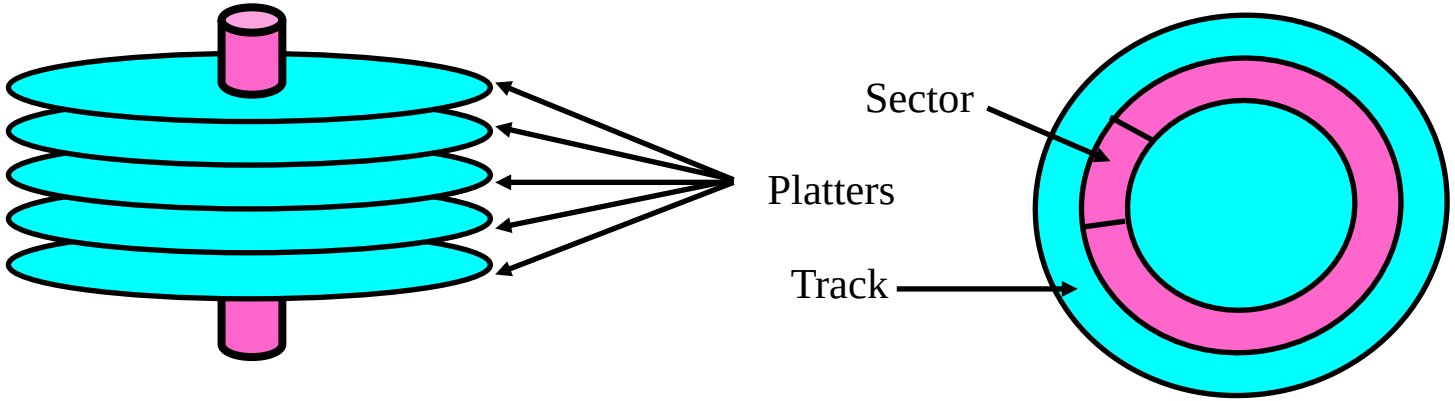
Properties of a Hard Magnetic Disk

- Properties
  - Independently addressable element: sector
    - OS always transfers groups of sectors together—"**blocks**"
  - A disk can access directly any given block of information it contains (random access).
  - A disk can be rewritten



Properties of a Hard Magnetic Disk

- Typical numbers (depending on the disk size):
  - 500 to more than 20,000 tracks per surface
  - 32 to 800 sectors per track
    - A sector is the smallest unit that can be read or written
- Constant bit density: more sectors on outer tracks



Sector Format and Structure

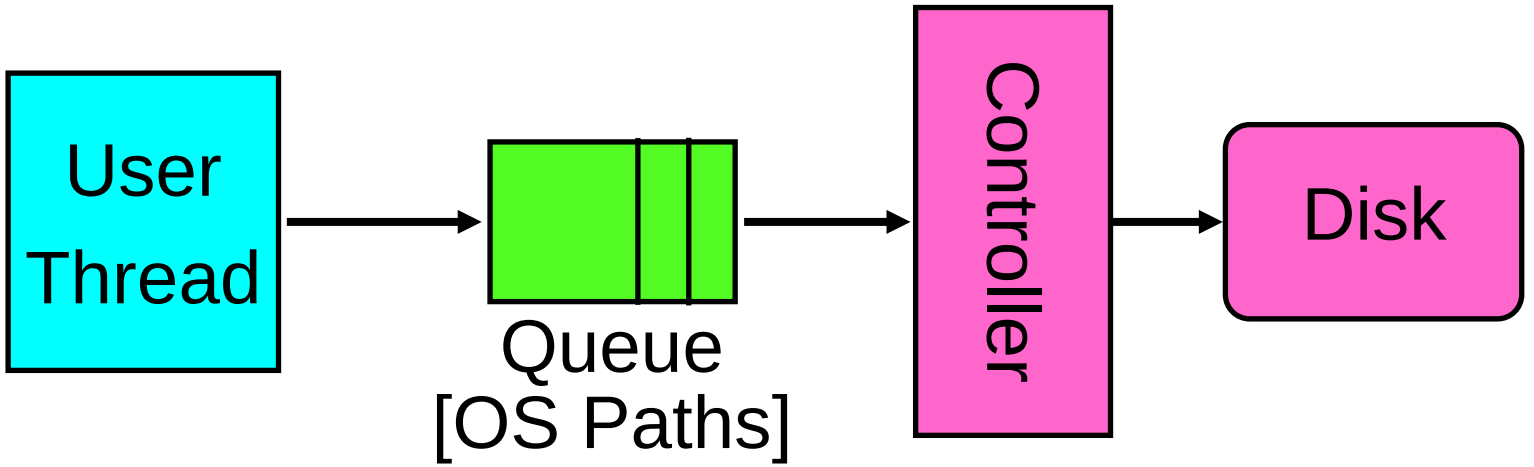
- The basic unit of data storage on a hard disk is the **sector**.
- One kind of sector format

Preamble 16B	**Data** 512B	**ECC** 12-16B

- Where is the desired Byte?

Equipment#	Head#	Cylinder#	Sector #	Offset

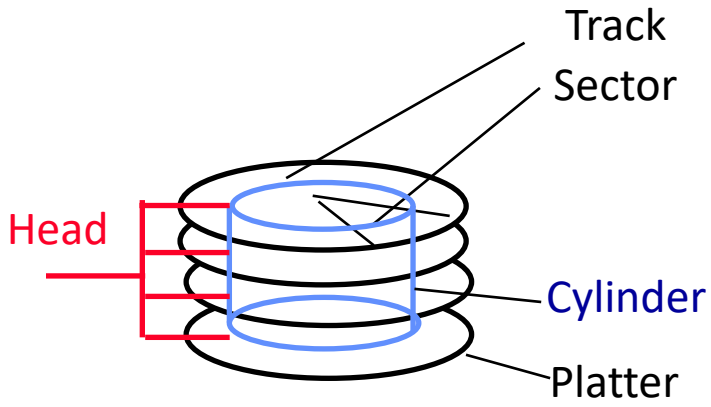
Disk I/O Performance



Response Time = Queue + Disk Service Time

Disk Performance

- Read/write data is a three-stage process:
  - **Seek time**: position the head/arm over the proper track (into proper cylinder)
  - **Rotational latency**: wait for the desired sector to rotate under the read/write head
  - **Transfer time**: transfer a block of bits (sector) under the read-write head



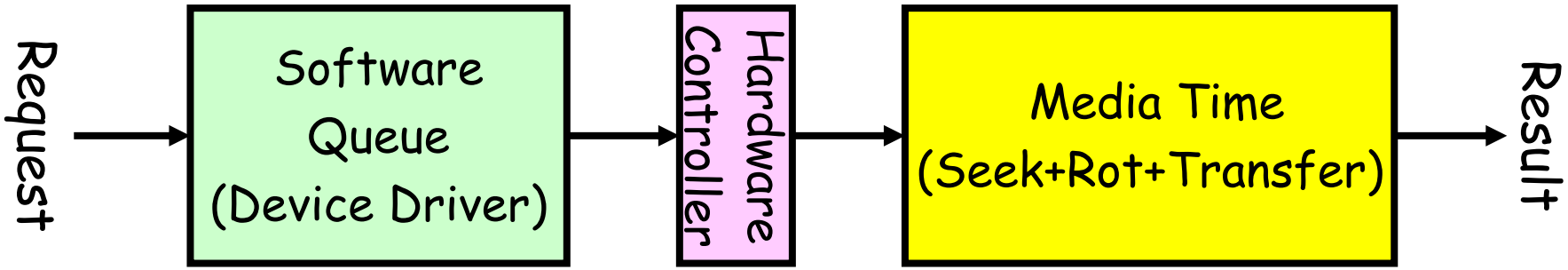
Positioning time (random-access time)

Disk Performance

- Disk Latency

  = Queueing Time + Controller time +

  Seek Time + Rotation Time + Transfer Time



Disk Performance

- 8 platter surfaces; 1024 tracks per surface; 64 sectors per track; 1KByte per sector. Average seek time is 8ms; rotate at 7200 rpm.
- What is the size of this hard disk?

  8×1024×64×1K = 512 MB

- What is the average positioning time?

  8+(60×1000)/7200/2 = 12.17 ms

Typical Numbers of a Magnetic Disk

- Average seek time as reported by the industry:
  - Typically in the range of 8 ms to 12 ms
- Rotational Latency:
  - **Most** disks rotate at 3,600 to 7200 RPM (Up to 15,000RPM or more)
  - Approximately 16 ms to 8 ms per revolution, respectively
  - An average latency to the desired information is halfway around the disk: 8 ms at 3600 RPM, 4 ms at 7200 RPM
- Transfer Time is a function of:
  - Transfer size (usually a sector): 512B – 1KB per sector
  - Rotation speed: 3600 RPM to 15000 RPM
  - Recording density: bits per inch on a track
  - Diameter: ranges from 1 in to 5.25 in
  - Typical values: 2 to 50 MB per second
- Controller time depends on controller hardware
- Cost drops by factor of two per year (since 1991)

Disk Scheduling

Disk Scheduling

- Disk can do only one request at a time; What order do you choose to do queued requests?

a request queue (0-31).
29, 11, 8, 15, 2

- Minimize seek time

Seek time ≈ seek distance

Disk Scheduling

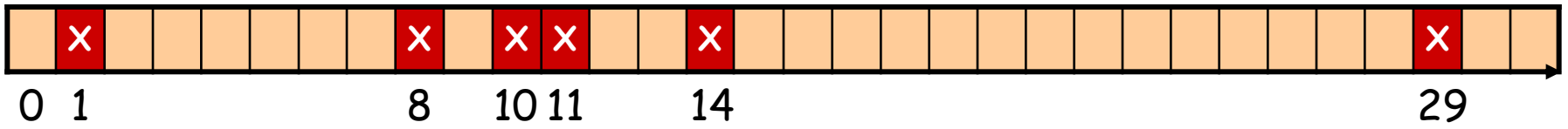
- Several algorithms exist to schedule the servicing of disk I/O requests.
  - FCFS (first-come, first-served)
  - SSTF (shortest-seek-time-first)
  - SCAN
  - C-SCAN (circular SCAN)
  - LOOK and C-LOOK

Disk Scheduling

- We illustrate them with a request queue (0-31).

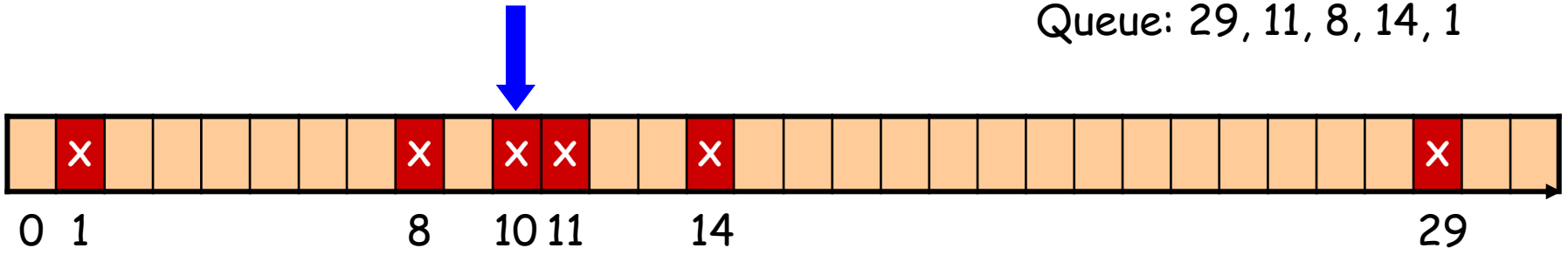
29, 11, 8, 14, 1

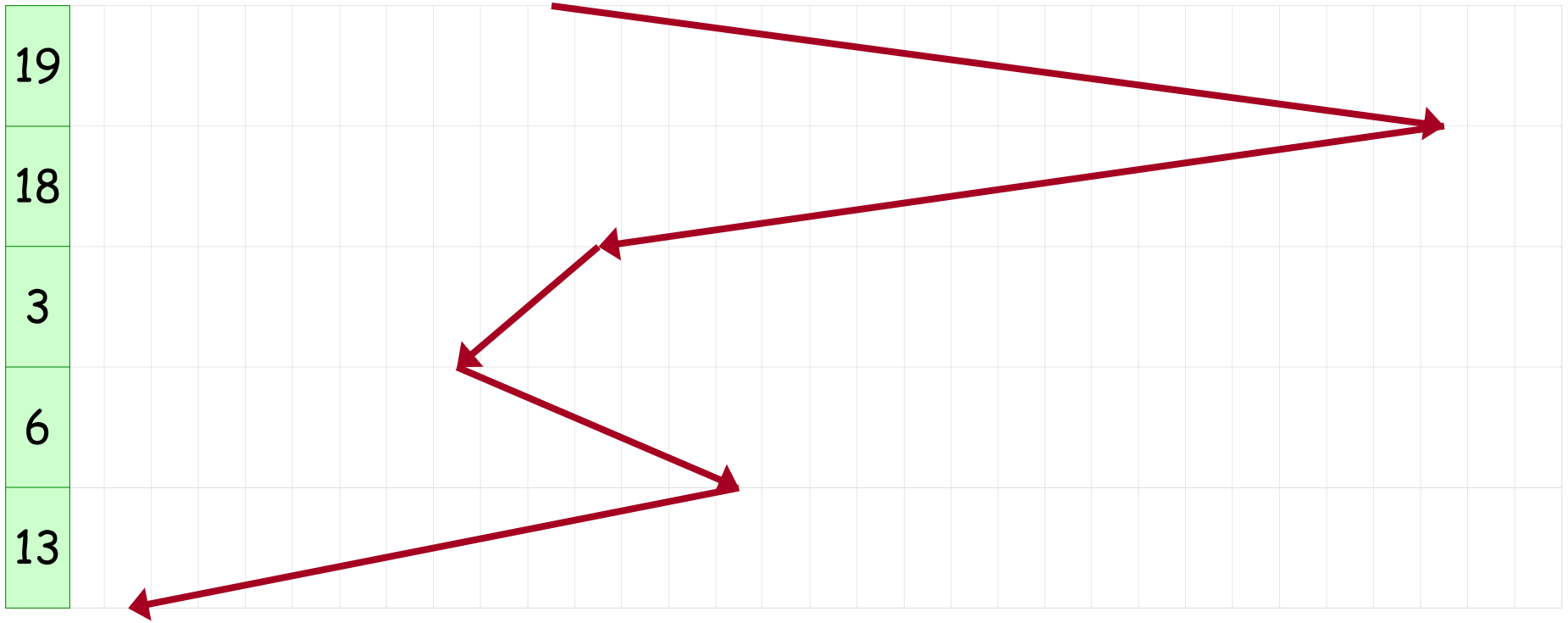
Head pointer 10



FCFS

Queue: 29, 11, 8, 14, 1





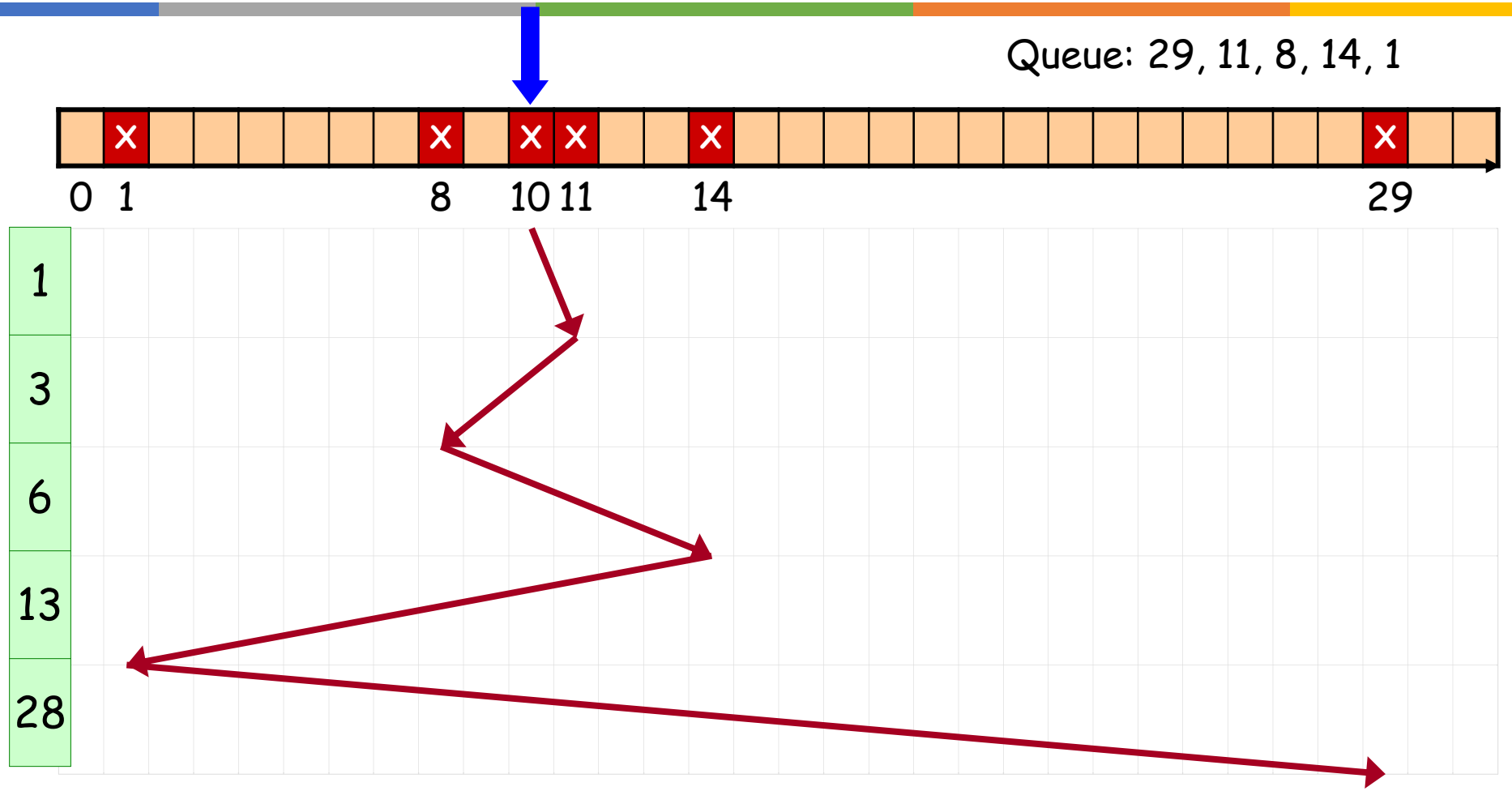
total head movement of 59 cylinders

SSTF

- 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.

SSTF

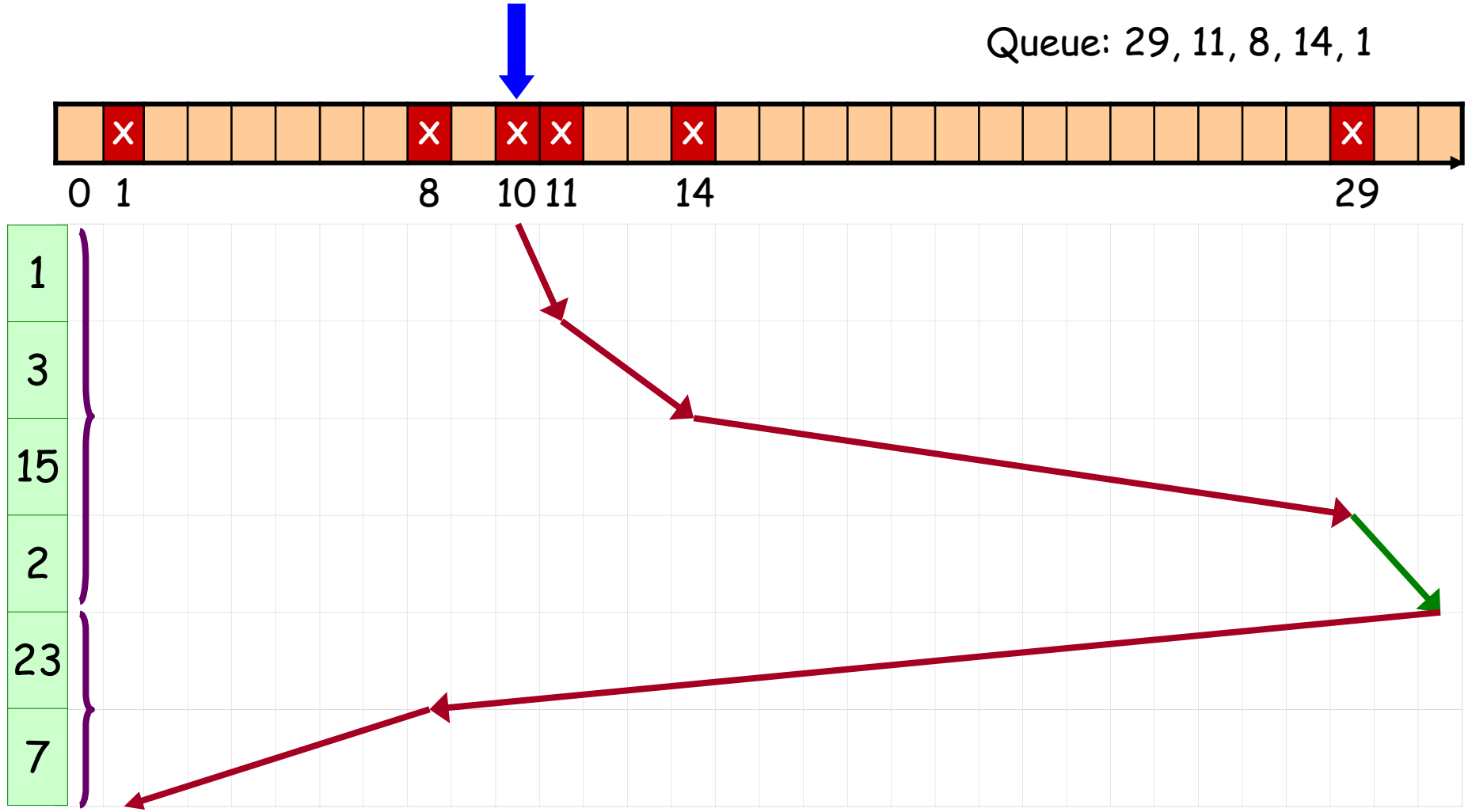


total head movement of 51 cylinders

SCAN

- 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.
- SCAN algorithm sometimes called the elevator algorithm

SCAN

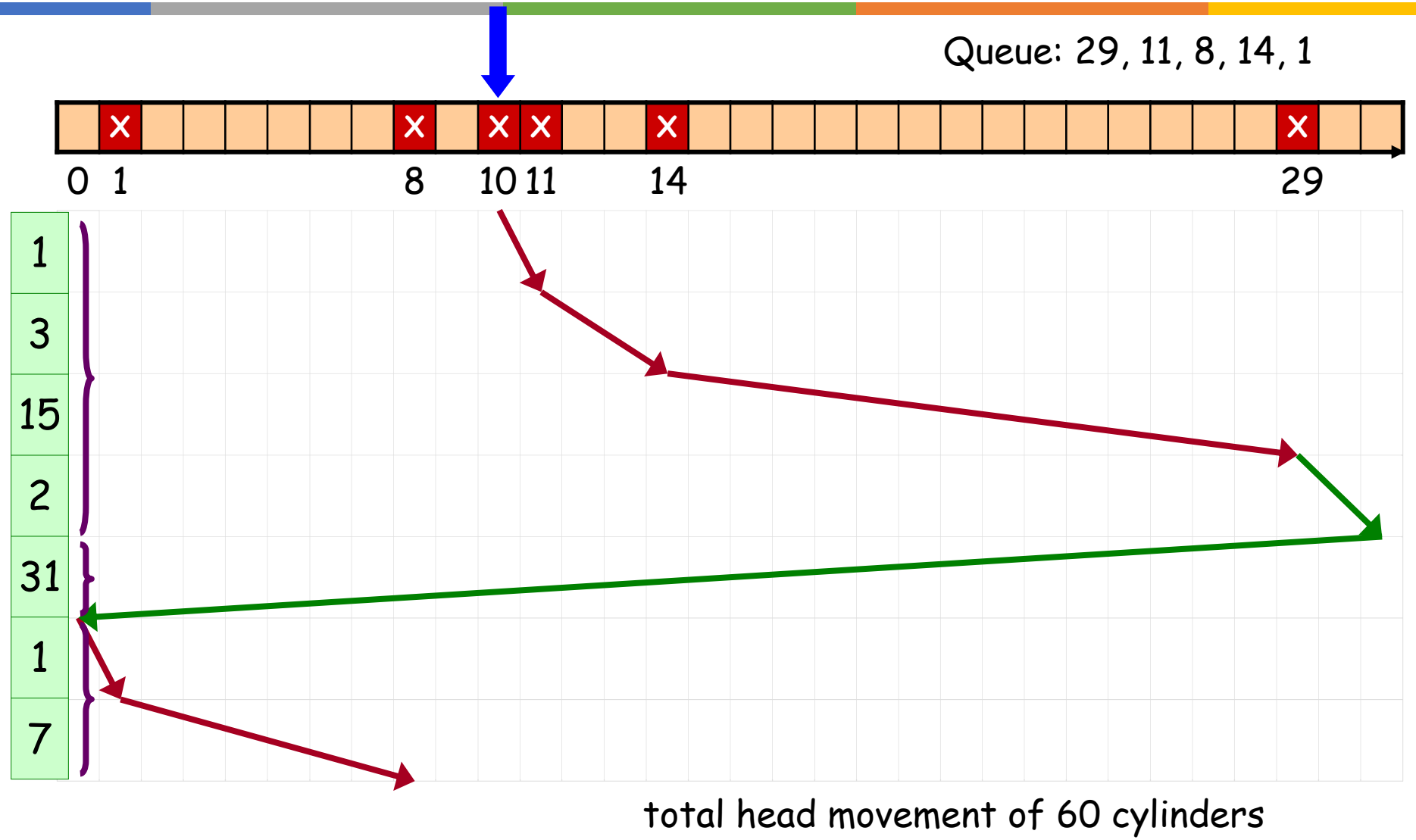


total head movement of 51 cylinders

C-SCAN

- Provides a more uniform wait time than SCAN.
- The head moves from one end of the disk to the other and serves 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.

C-SCAN

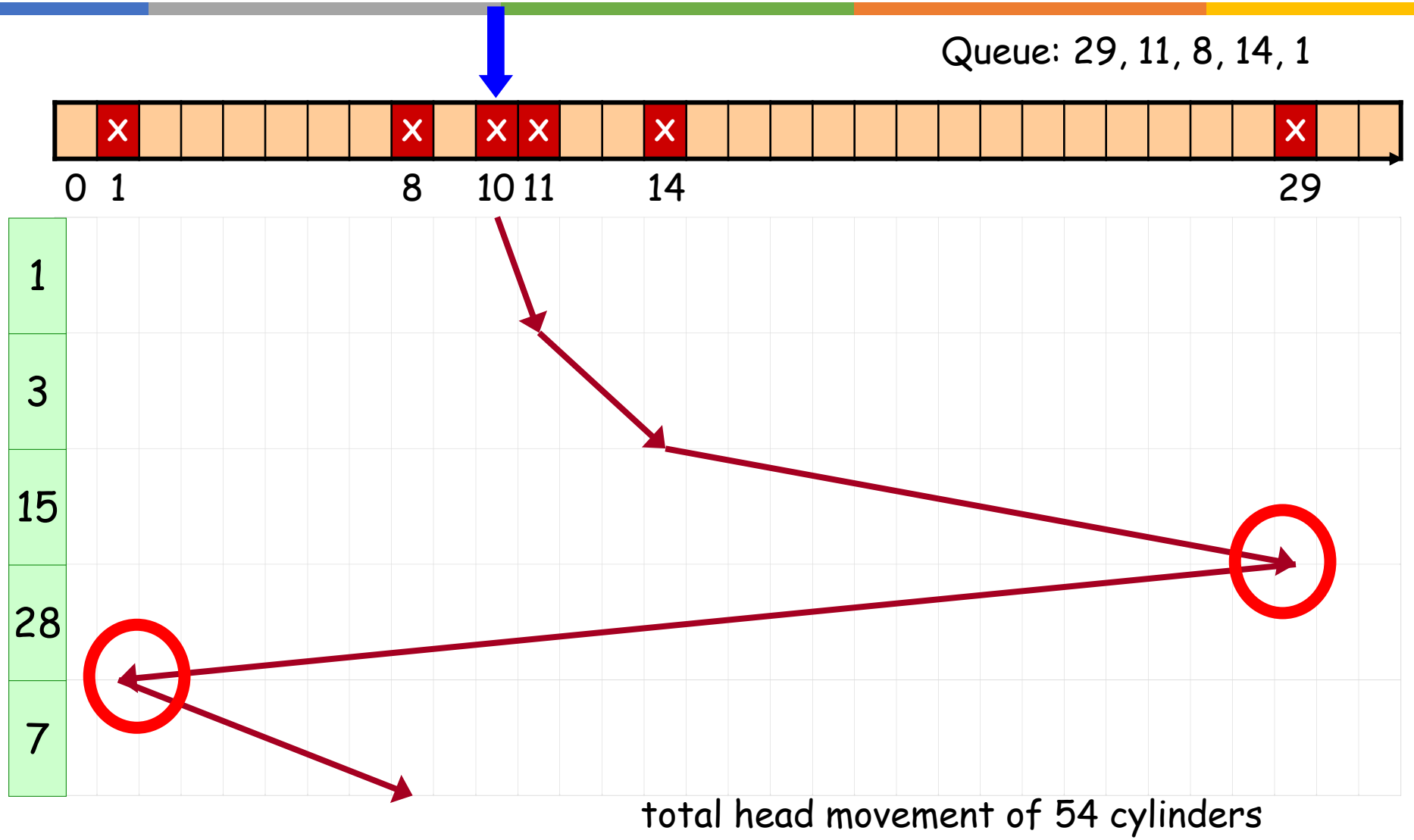


total head movement of 60 cylinders

C-LOOK

- Version of C-SCAN
- 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.

C-LOOK


Queue: 29, 11, 8, 14, 1
0 1 8 10 11 14 29
1
3
15
28
7
total head movement of 54 cylinders

Selecting Disk-Scheduling Algorithm

- SSTF is common
- SCAN and C-SCAN
  - perform better for systems that place a heavy load on the disk.
  - no starvation problem
- 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.

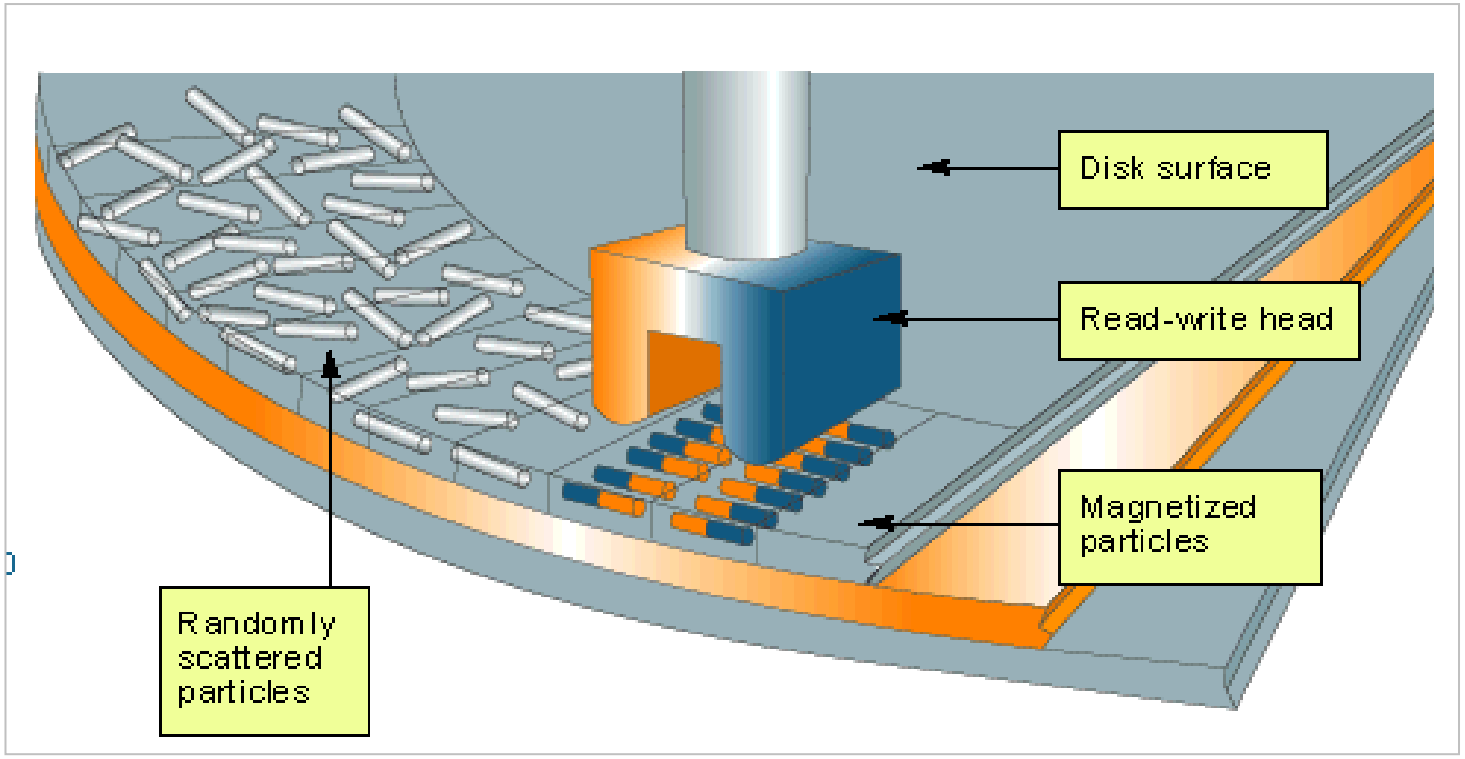
Selecting Disk-Scheduling Algorithm

- For modern disks
  - rotational latency is nearly as large as seek time
  - manufacturers implement disk-scheduling in the controller hardware
  - OS do not need to schedule disk requests

Disk Management

Disk Management

- Low-level formatting, or physical formatting — Dividing a disk into sectors that the disk controller can read and write.



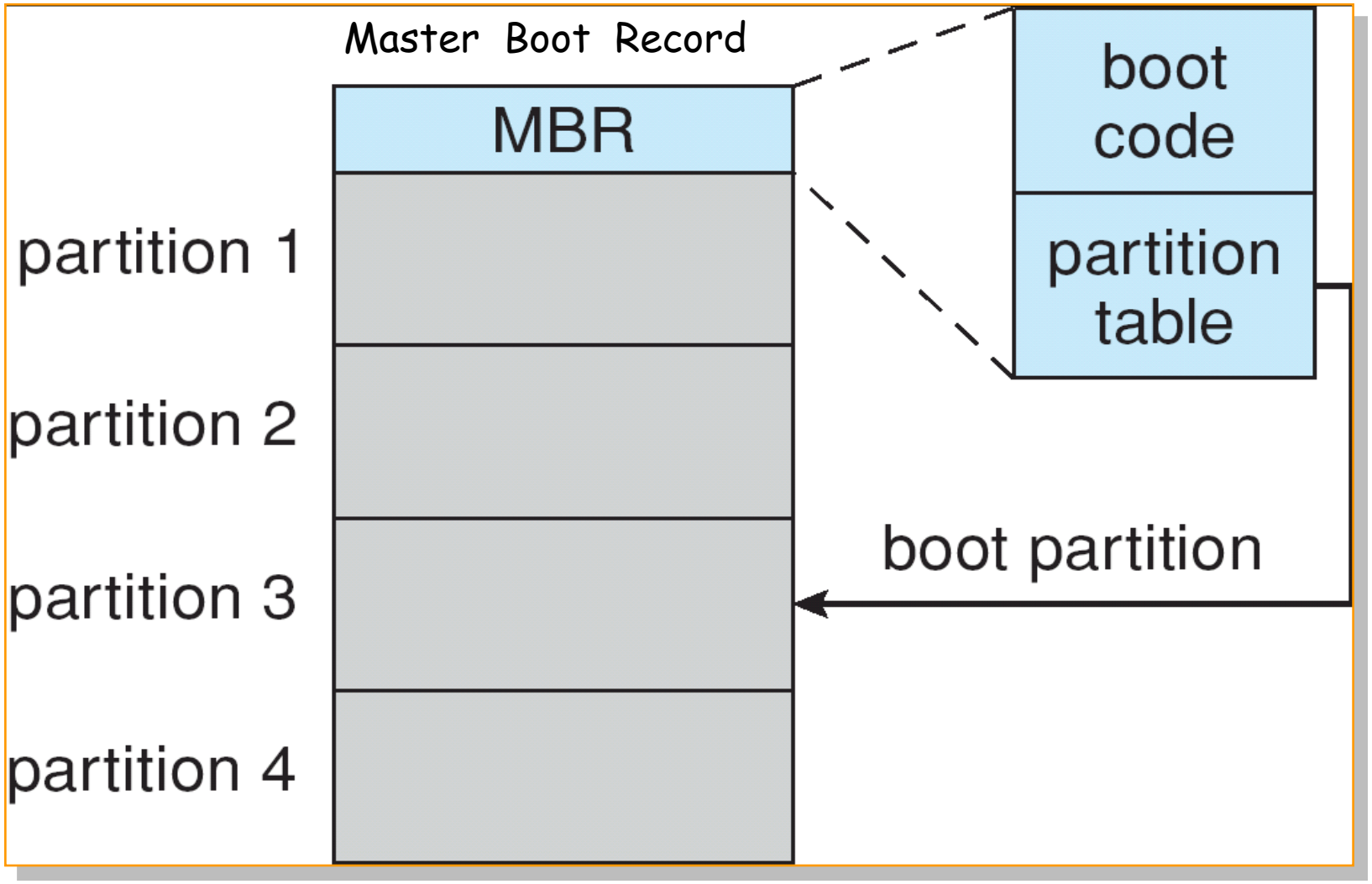
Operating System Support

- Major OS jobs are to manage physical devices and to present a virtual machine abstraction to applications
- For hard disks, OS provides two abstraction:
  - Raw device – an array of data blocks
  - File system

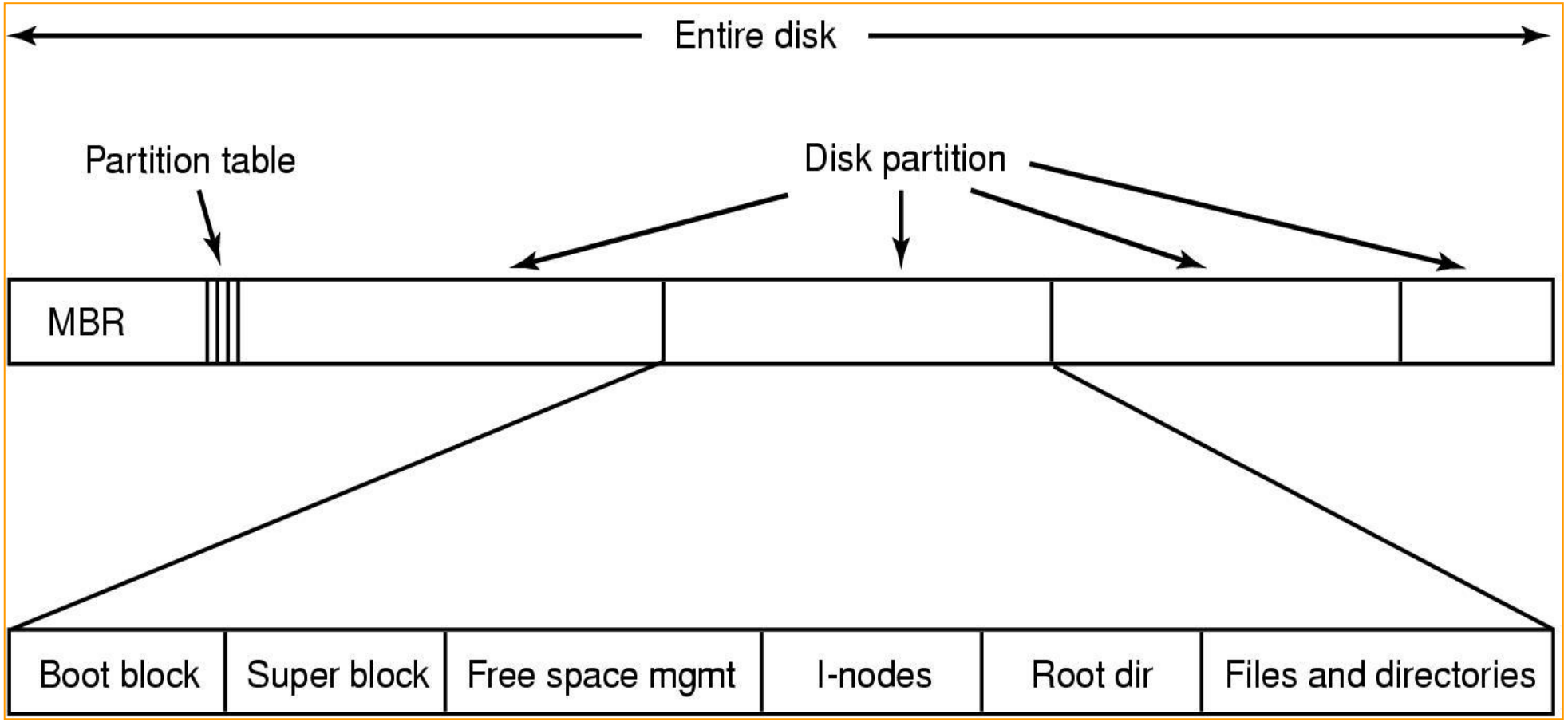
Disk Management

- 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.
  - Logical formatting or “making a file system”.

Booting from a Disk in Windows



One Possible File System Location



Summary

Summary

- Disk Structure:
  - Platter
  - Track
  - Sector
  - Cylinder
  - Block
- Disk Performance:
  - Queuing time + Controller + Seek + Rotational + Transfer
  - Rotational latency: on average ½ rotation

Summary

- FCFS Order
  - Fair among requesters, but order of arrival may be to random spots on the disk ⇒ Very long seeks
- SSTF: Shortest Seek Time First
  - Pick the request that's closest on the disk
  - Although called SSTF, today must include rotational delay in calculation, since rotation can be as long as seek
  - Con: SSTF good at reducing seeks, but may lead to starvation
- SCAN: Implements an Elevator Algorithm, take the closest request in the direction of travel
  - No starvation, but retains flavor of SSTF
- C-SCAN: Circular-Scan, only goes in one direction
  - Skips any requests on the way back
  - Fairer than SCAN, not biased towards pages in middle

北京交通大学

Thank you!

Q & A