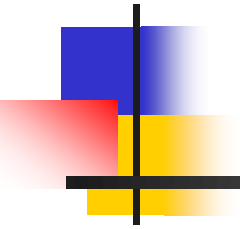
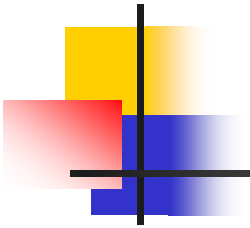
操作系统原理及应用

李　伟
xchlw@seu.edu.cn

计算机科学与工程学院、软件学院
江苏省网络与信息安全重点实验室



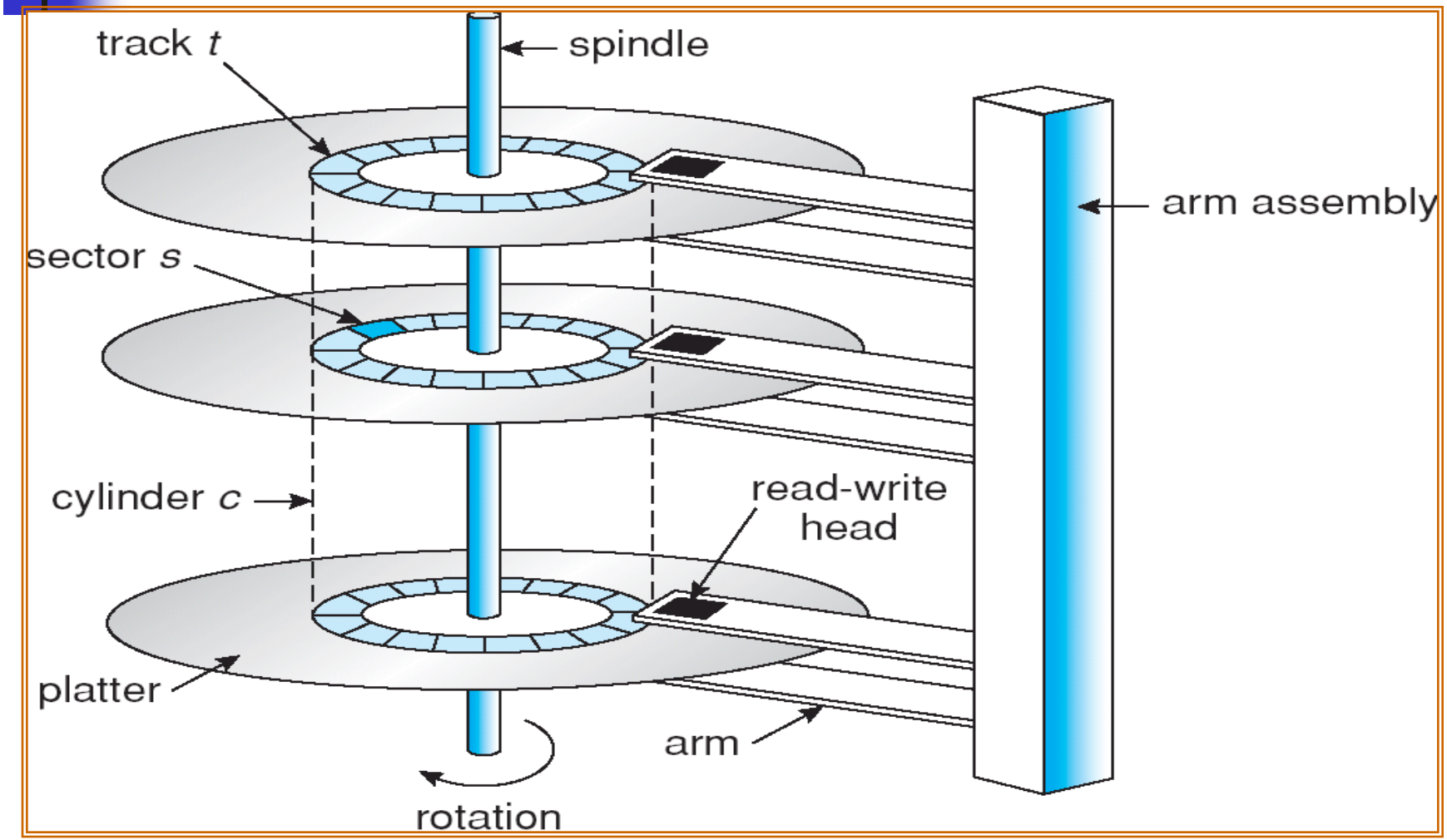
Chapter 12 Mass Storage Structure

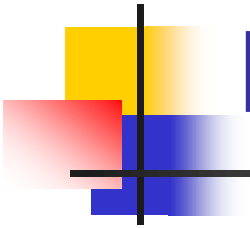


Outline

- Disk Structure
- Disk Attachment
- Disk Scheduling
- Disk Management
- Swap-Space Management
- RAID Structure
- Stable-Storage Implementation
- Tertiary Storage Devices

Disk Structure

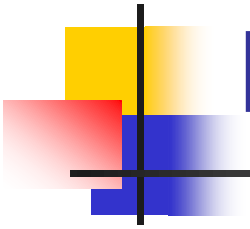




Disk Structure

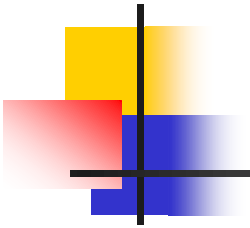
- Magnetic disks provide bulk of secondary storage of modern computers
  - Transfer rate is rate at which data flow between drive and computer
  - Positioning time (random-access time) is time to move disk arm to desired cylinder (seek time) and time for desired sector to rotate under the disk head (rotational latency)
  - Head crash results from disk head making contact with the disk surface

Parameter	IBM 360-KB floppy disk	WD 18300 hard disk
Number of cylinders	40	10601
Tracks per cylinder	2	12
Sectors per track	9	281 (avg)
Sectors per disk	720	35742000
Bytes per sector	512	512
Disk capacity	360 KB	18.3 GB
Seek time (adjacent cylinders)	6 msec	0.8 msec
Seek time (average case)	77 msec	6.9 msec
Rotation time	200 msec	8.33 msec
Motor stop/start time	250 msec	20 sec
Time to transfer 1 sector	22 msec	17 μsec



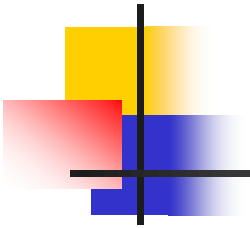
Disk Structure

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



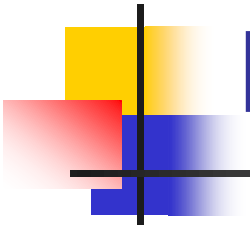
Disk Structure

- 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.



Outline

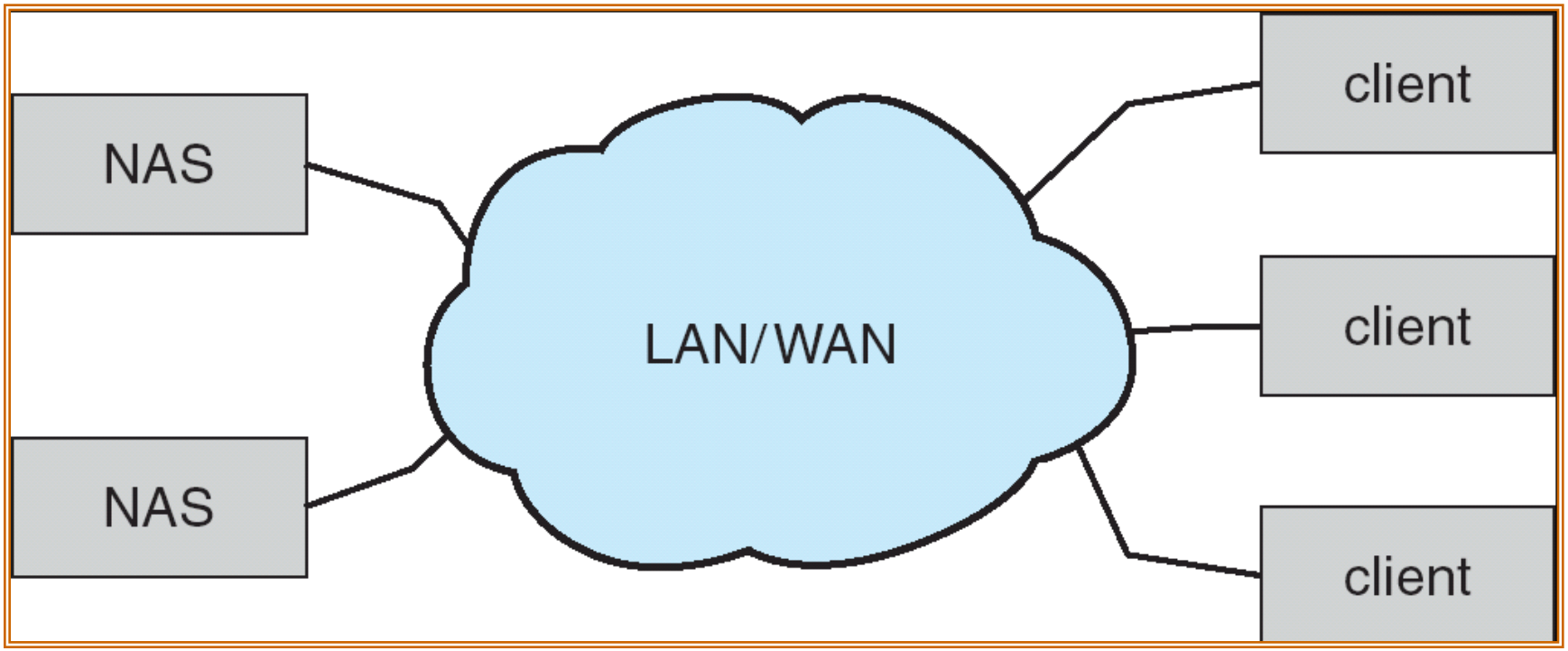
- Disk Structure
- Disk Attachment
- Disk Scheduling
- Disk Management
- Swap-Space Management
- RAID Structure
- Stable-Storage Implementation
- Tertiary Storage Devices



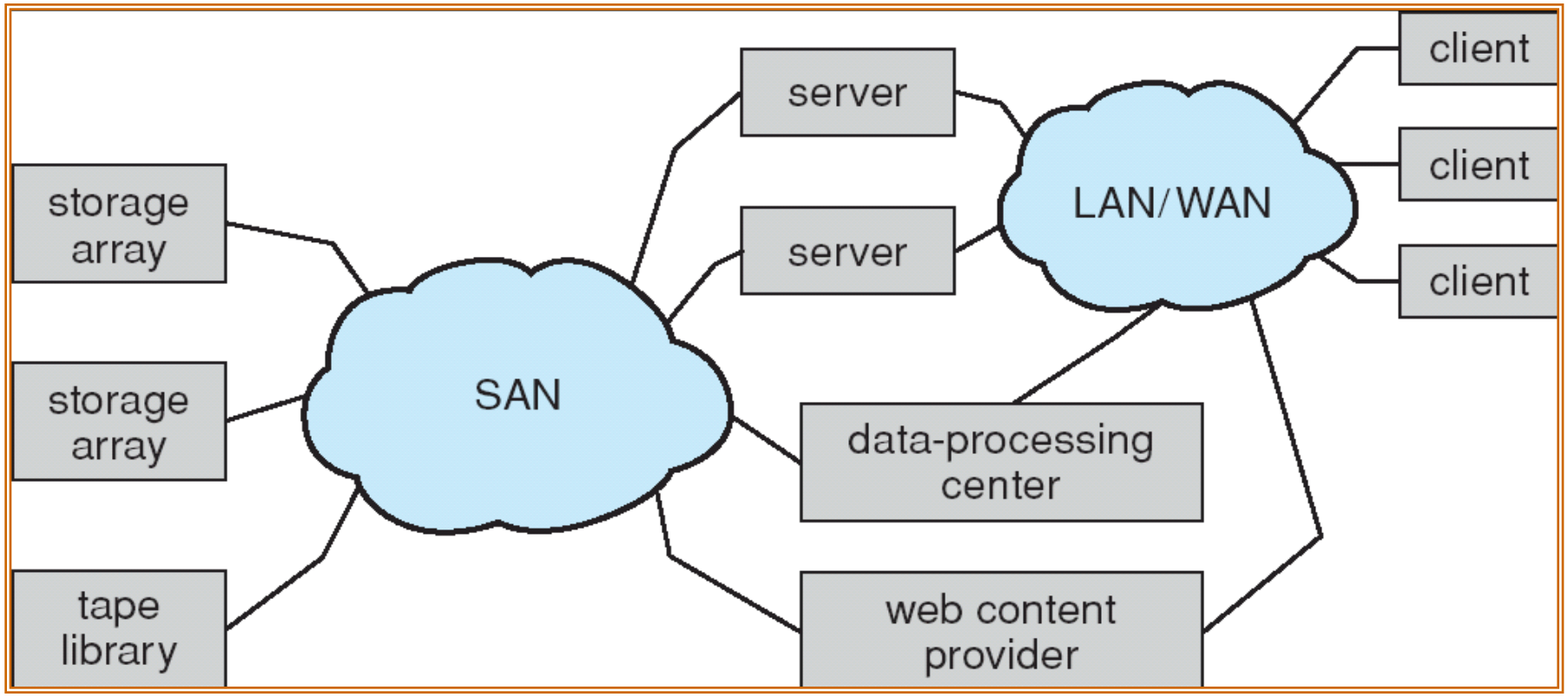
Disk Attachment

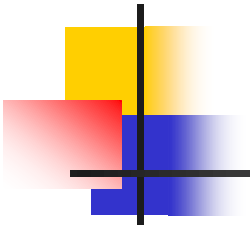
- **Disks may be attached one of two ways**
  - **Host attached via an I/O port**
  - **Network attached via a network connection**

Network-Attached Storage



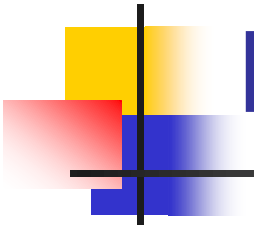
Storage-Area Network





Outline

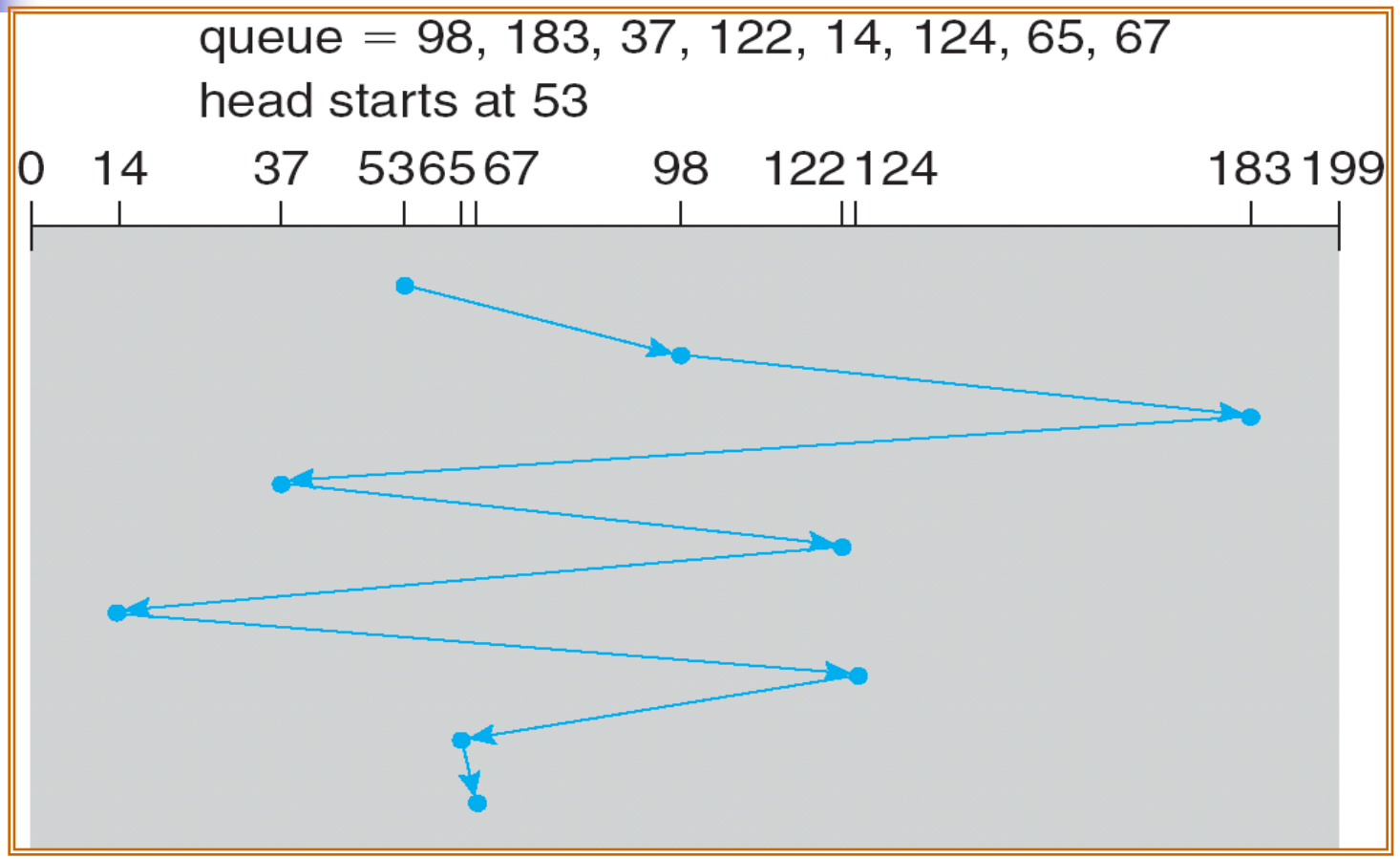
- Disk Structure
- Disk Attachment
- Disk Scheduling
- Disk Management
- Swap-Space Management
- RAID Structure
- Stable-Storage Implementation
- Tertiary Storage Devices



Disk Scheduling

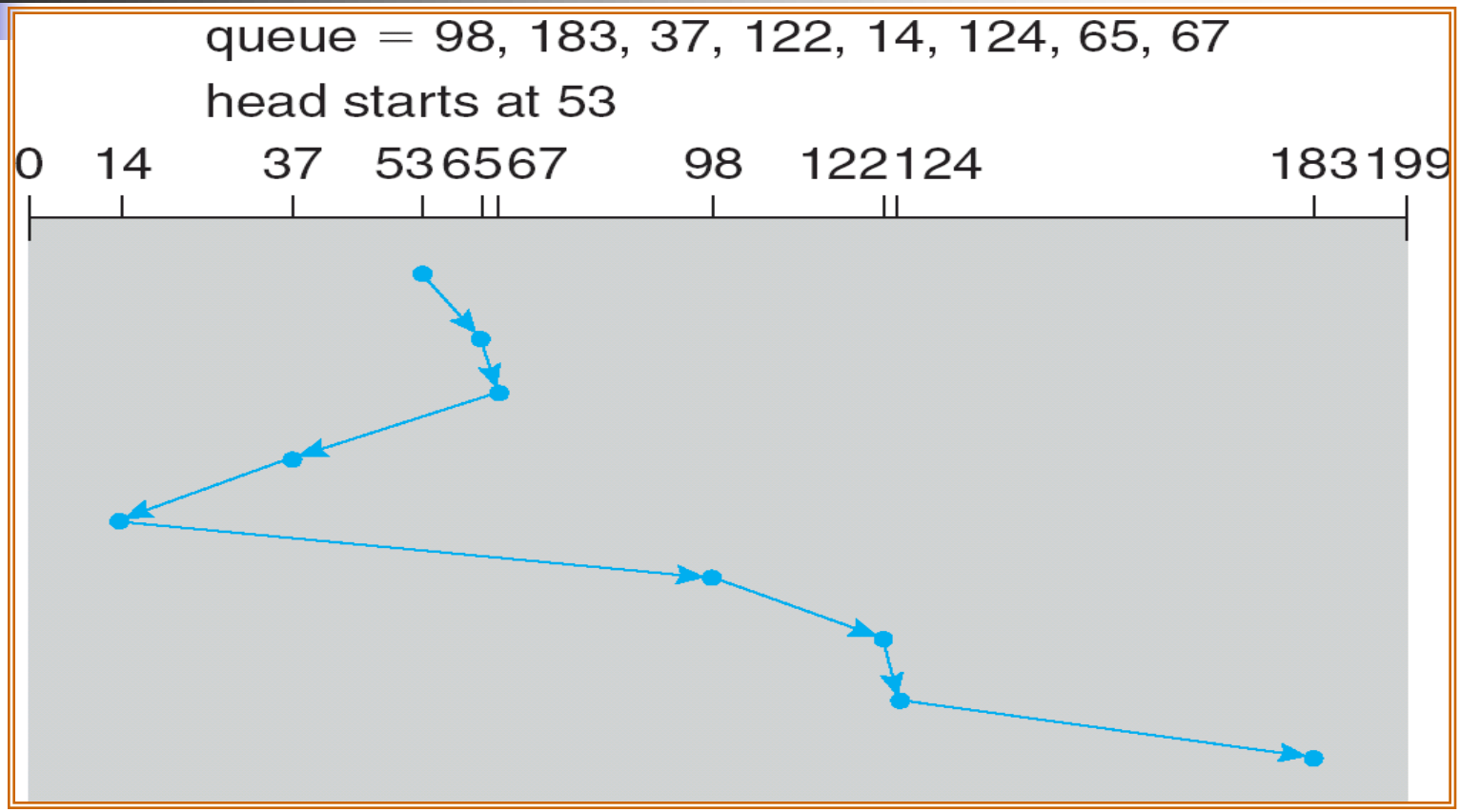
- **Access time has two major components**
  - ***Seek time*** **is the time for the disk are to move the heads to the cylinder containing the desired sector.**
  - ***Rotational latency*** **is the additional time waiting for the disk to rotate the desired sector to the disk head.**
- **Disk bandwidth**
  - **the total number of bytes transferred, divided by the total time between the first request for service and the completion of the last transfer.**

FCFS



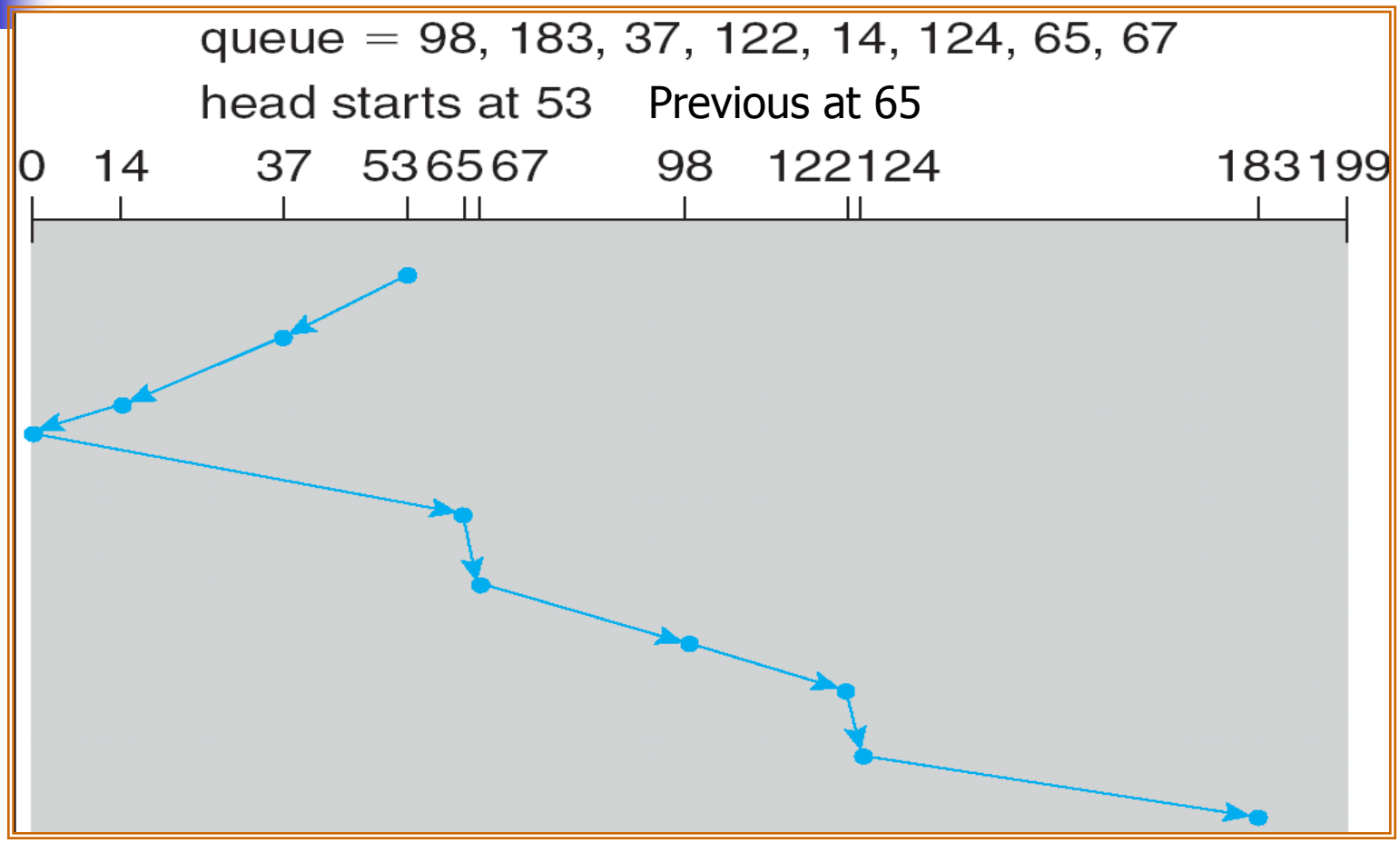
Seek Distance: total head movement of 640 cylinders

SSTF (Shortest Seek Time First)



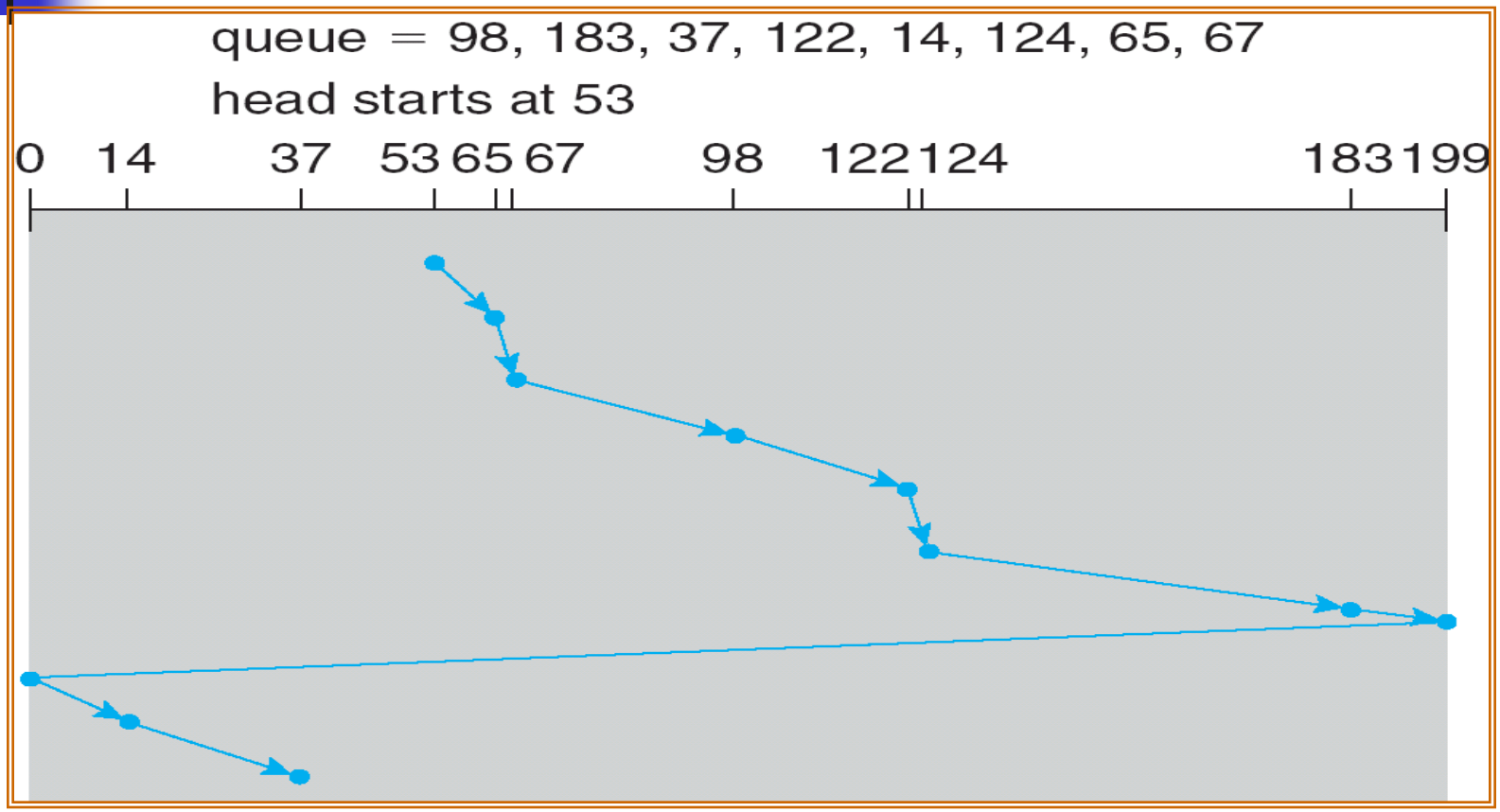
- **may cause starvation of some requests.**
- **total head movement of 236 cylinders**

SCAN — elevator algorithm



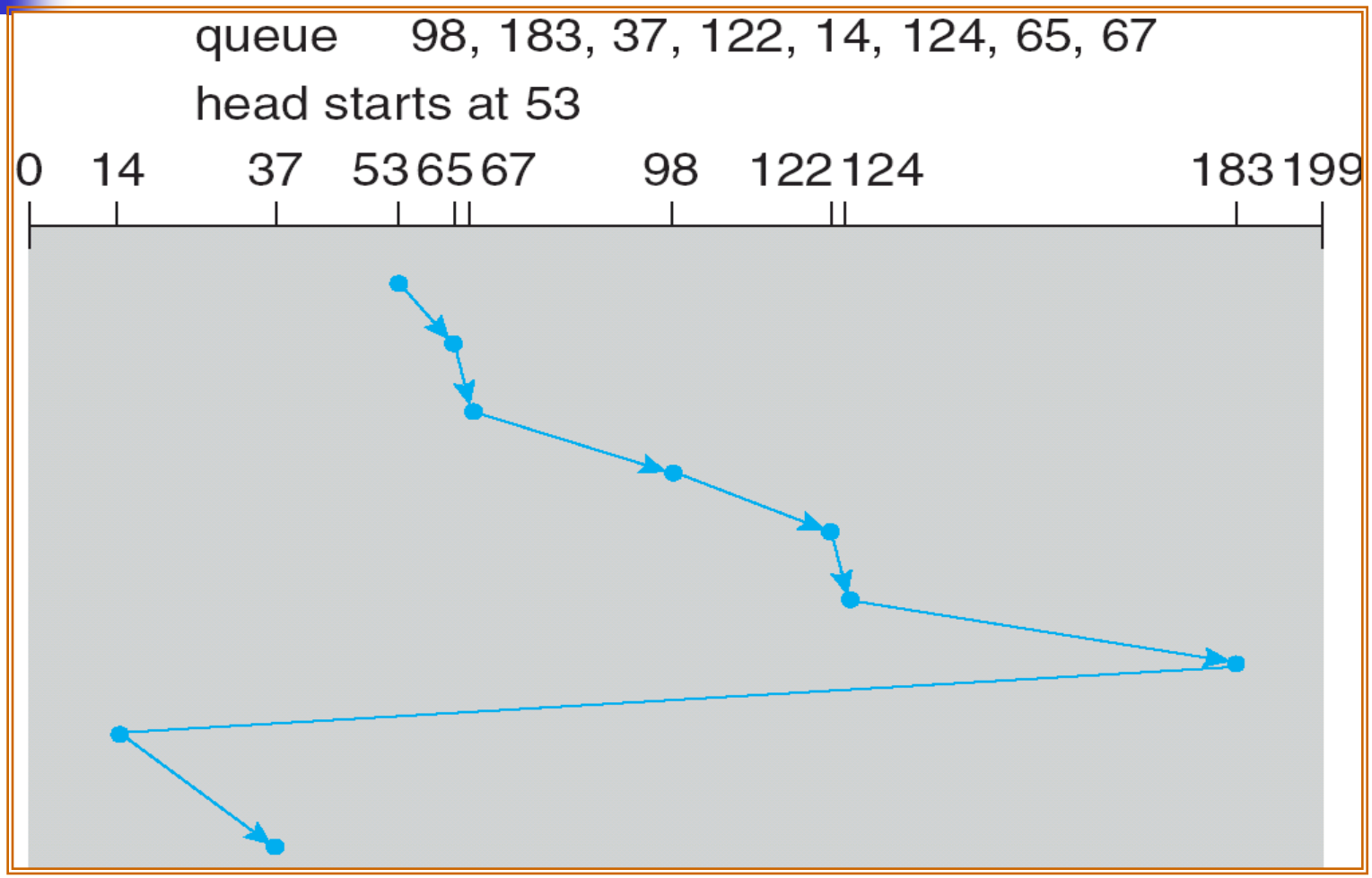
- **total head movement of 236 cylinders**

C-SCAN

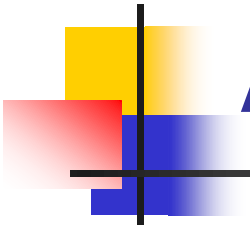


- **Provides a more uniform wait time than SCAN.**

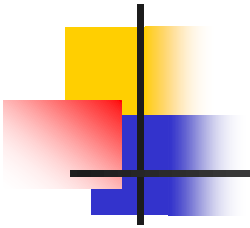
C-LOOK


queue 98, 183, 37, 122, 14, 124, 65, 67
head starts at 53
0
14
37
53
65
67
98
122
124
183
199

Selecting a Disk-Scheduling 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.**
- **Either SSTF or LOOK is a reasonable choice for the default algorithm.**
- **Performance depends on the number and types of requests.**
- **Requests for disk service can be influenced by the file-allocation method.**



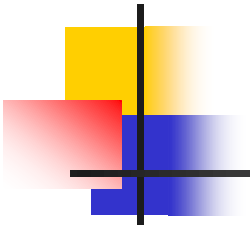
Outline

- Disk Structure
- Disk Attachment
- Disk Scheduling
- Disk Management
- Swap-Space Management
- RAID Structure
- Stable-Storage Implementation
- Tertiary Storage Devices

Disk Formatting

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

A disk sector

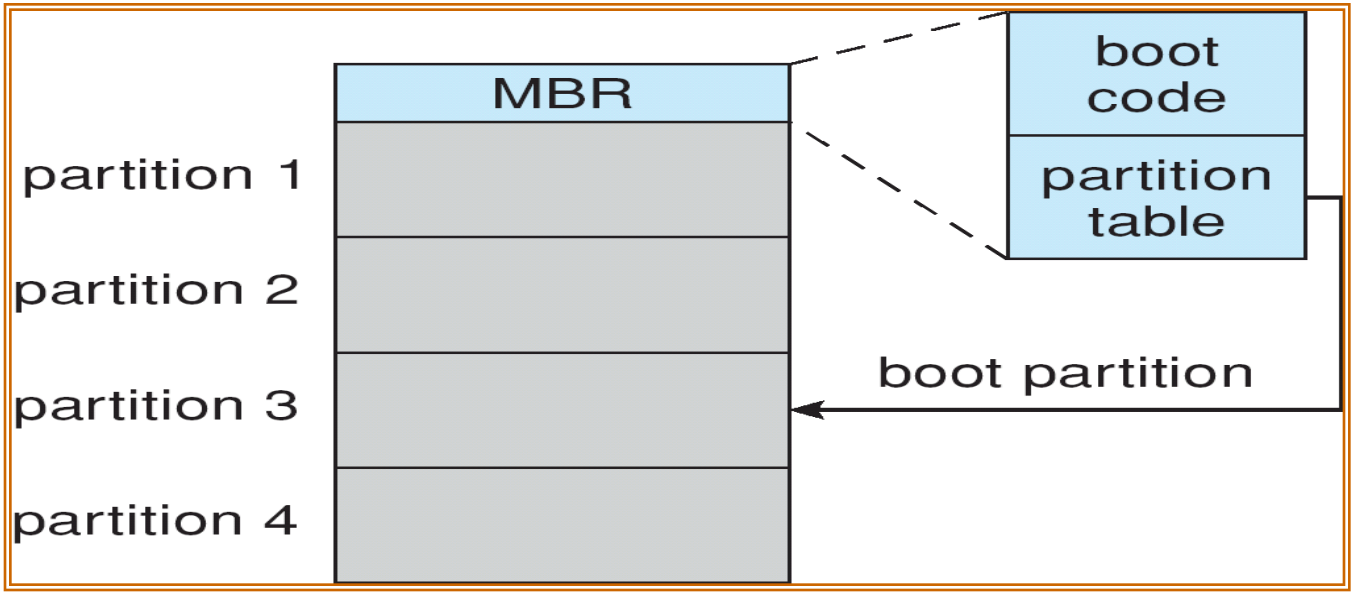


Disk Formatting

- 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”.

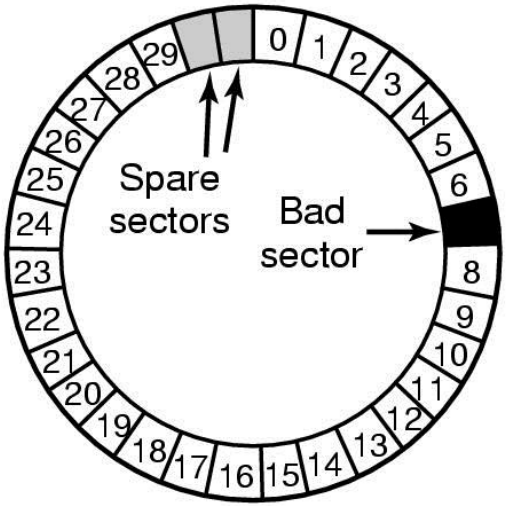
Boot Block

- **Boot block initializes system**
  - **The bootstrap is stored in ROM.**
  - ***Bootstrap loader* program.**
- **Booting from disk in Windows 2000**

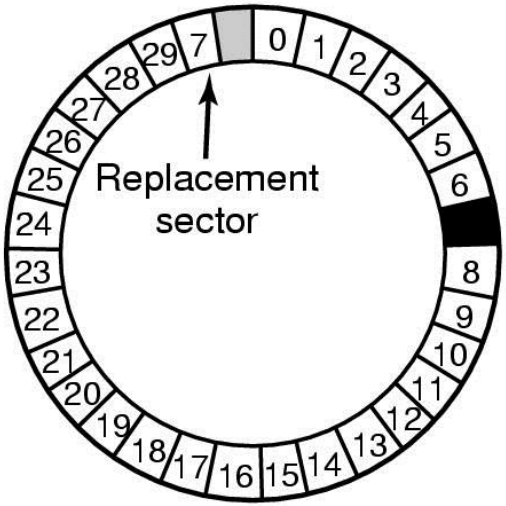


Error Handling

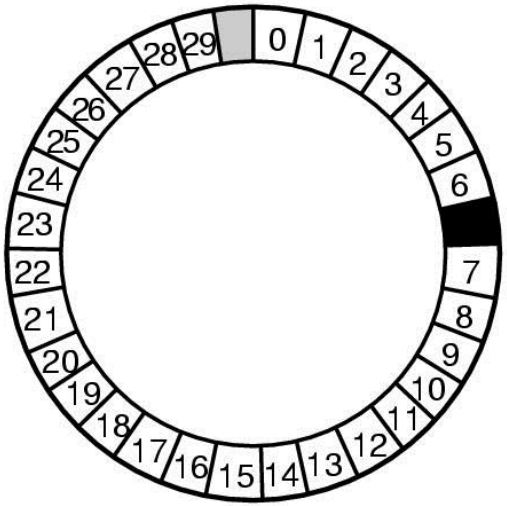
- A disk track with a bad sector
- Substituting a spare for the bad sector
- Shifting all the sectors to bypass the bad one



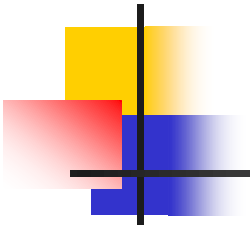
(a)



(b)

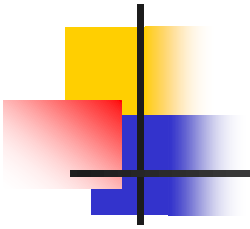


(c)



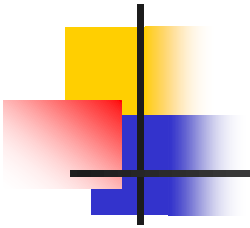
Outline

- Disk Structure
- Disk Attachment
- Disk Scheduling
- Disk Management
- Swap-Space Management
- RAID Structure
- Stable-Storage Implementation
- Tertiary Storage Devices



Swap-Space Management

- **Swap-space — Virtual memory uses disk space as an extension of main memory.**
- **Swap-space can be carved out of the normal file system,or, more commonly, it can be in a separate disk partition.**



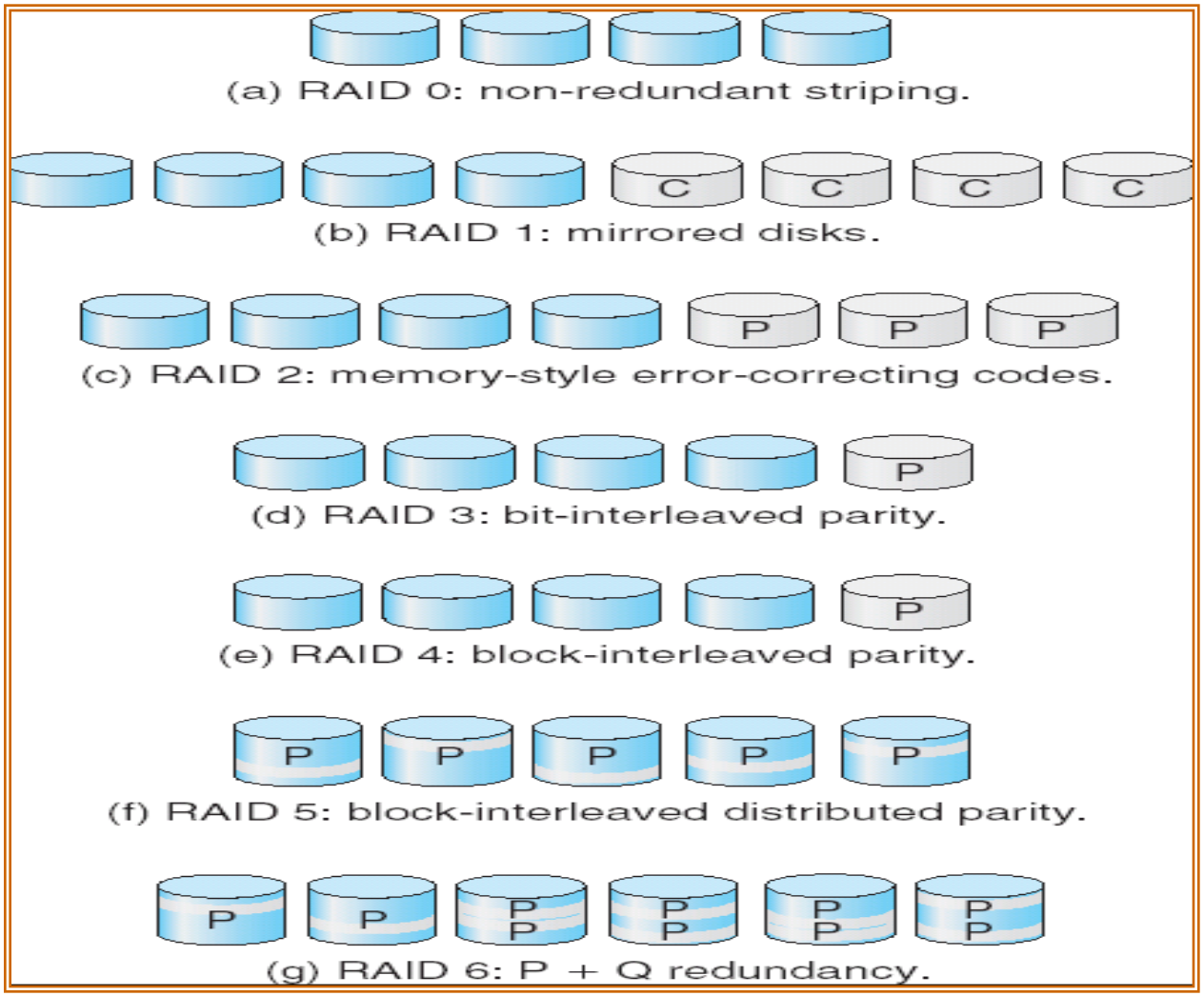
Outline

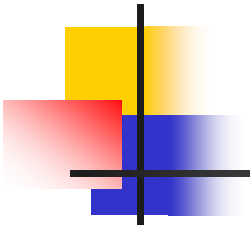
- Disk Structure
- Disk Attachment
- Disk Scheduling
- Disk Management
- Swap-Space Management
- RAID Structure
- Stable-Storage Implementation
- Tertiary Storage Devices

RAID Structure

- RAID – Redundant Array of Independent Disk（磁盘冗余阵列）
  - multiple disk drives improves reliability via redundancy and performance via parallelism.
- RAID is arranged into six different levels.

RAID Levels


C
P
(a) RAID 0: non-redundant striping.
(b) RAID 1: mirrored disks.
(c) RAID 2: memory-style error-correcting codes.
(d) RAID 3: bit-interleaved parity.
(e) RAID 4: block-interleaved parity.
(f) RAID 5: block-interleaved distributed parity.
(g) RAID 6: P + Q redundancy.



Outline

- Disk Structure
- Disk Attachment
- Disk Scheduling
- Disk Management
- Swap-Space Management
- RAID Structure
- Stable-Storage Implementation
- Tertiary Storage Devices