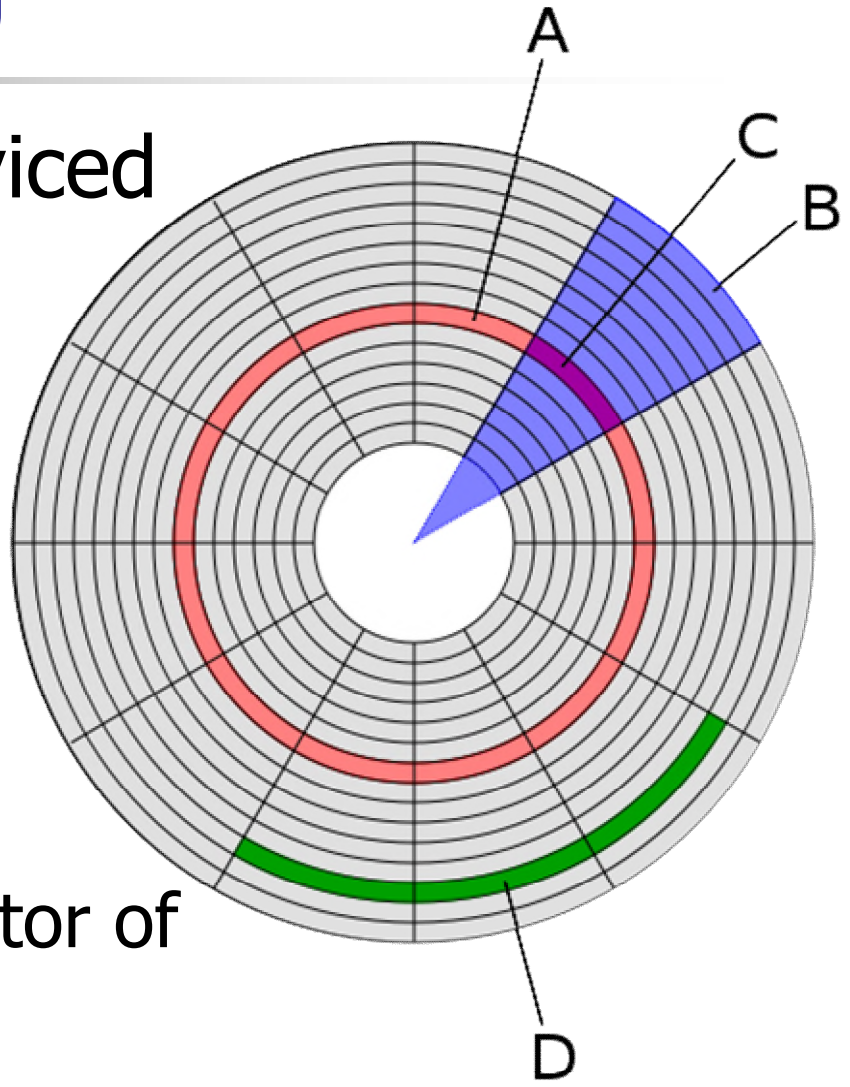


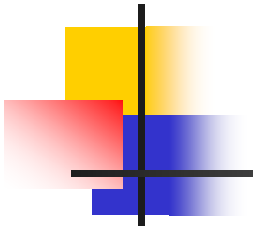
CS423 Operating Systems
File Systems

Disk Scheduling

- Which disk request is serviced first?
  - FCFS
  - Shortest seek time first
  - Elevator (SCAN)
  - C-SCAN (Circular SCAN)



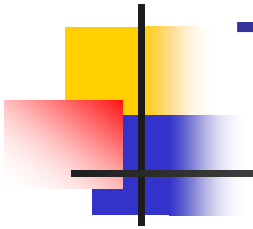
A: Track. B: Sector. C: Sector of Track. D: File



Scheduling Algorithms

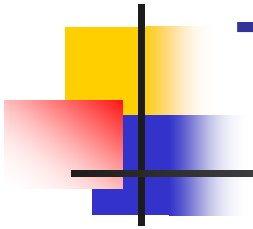
Algorithm Name	Description
FCFS	First-come first-served
SSTF	Shortest seek time first; process the request that reduces next seek time
SCAN (aka Elevator)	Move head from end to end (has a current direction)
C-SCAN	Only service requests in one direction (circular SCAN)
LOOK	Similar to SCAN, but donot go all the way to the end of the disk.
C-LOOK	Circular LOOK. Similar to C-SCAN, but donot go all the way to the end of the disk.

Estimate Sustained Average Transfer Rate



- Suppose that a disk drive spins at 7200 RPM (revolutions per minute), has a sector size of 512 bytes, and holds 160 sectors per track.
- What is sustained average transfer rate of this drive in megabytes per second?

Estimate Sustained Average Transfer Rate



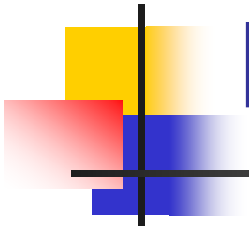
- Suppose that a disk drive spins at 7200 RPM (revolutions per minute), has a sector size of 512 bytes, and holds 160 sectors per track.
- What is sustained average transfer rate of this drive in megabytes per second?

Disk spins 120 times per second (7200 RPM/60)

Each spin transfers a track of 80 KB (160 sectors x0.5K)

Sustained average transfer rate is 120x80 = 9.6MB/s.

Average Performance of Random Access



- Suppose that a disk drive spins at 7200 RPM (revolutions per minute), has a sector size of 512 bytes, and holds 160 sectors per track.
- Average seek time for the drive is 8 milliseconds
- Estimate # of random sector I/Os per second that can be done and the effective average transfer rate for random-access of a sector?

Average Performance of Random Access

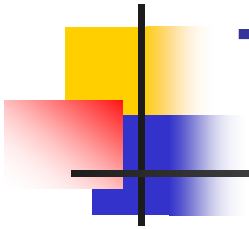


Disk spins 120 times per second
Average rotational cost is time to travel half track: 1/120 * 50%=4.167ms

Transfer time is 8ms to seek
+ 4.167 ms rotational latency
+ 0.052 ms (reading one sector takes 0.0005MB/ 9.6MB).
= 12.219 ms

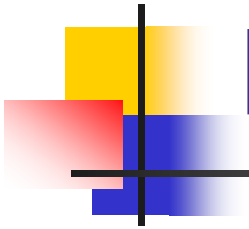
# of random sector access/second = 1/0.012219=81.8

Effective transferring rate: 0.5 KB * 81.8= 0.0409 MB/s.



Typical Modern Specifications

- Disk rotation speed: 5,000-15,000 RPM
- Number of sectors per track: 500-2000
- Number of tracks: 100,000-200,000
- Average seek latency: 2ms
- Block size: 0.5K-4K
- Multiple zones (layout is different for inner and outer tracks)
- Multiple bays (stacked drives)



Flash Memory

- Electronically Erasable Programmable Read Only Memory (EEPROM)
- Example specifications (NAND Flash):
  - Page size: 2KB (approx.)
  - Block size: 64 pages (128KB)
  - Device size: 16K blocks
- Random READ: 25μs, Sequential READ: 25ns
- WRITE performance
  - PROGRAM PAGE: 220μs, BLOCK ERASE: 1.5ms
- Endurance: 100,000 PROGRAM/ERASE cycles

RAID

(Redundant Array of Inexpensive Disks)

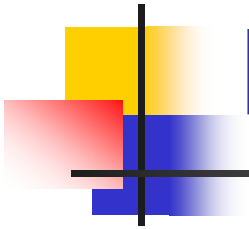


- Multiple disk drives provide reliability via **redundancy**. Increases the **mean time to failure**







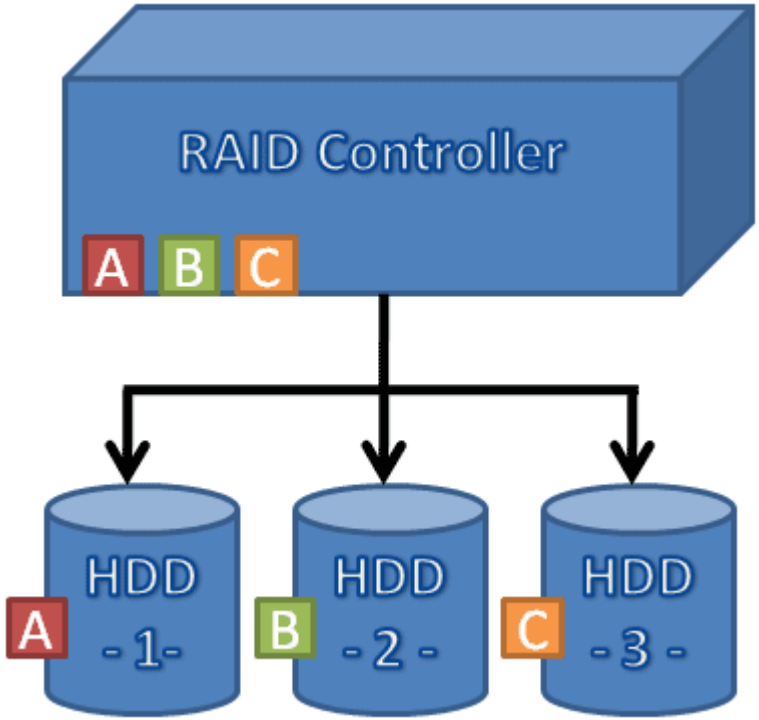


RAID (Cont.)

- RAID
  - multiple disks work cooperatively
  - Improve reliability by storing redundant data
  - **Striping** (**RAID 0**) improves performance with disk **striping** (use a group of disks as one storage unit)
  - **Mirroring** (**RAID 1**) keeps duplicate of each disk
  - Striped mirrors (**RAID 1+0**) or mirrored stripes (**RAID 0+1**) provides high performance and high reliability
  - **Block interleaved parity** (**RAID 4, 5, 6**) uses much less redundancy

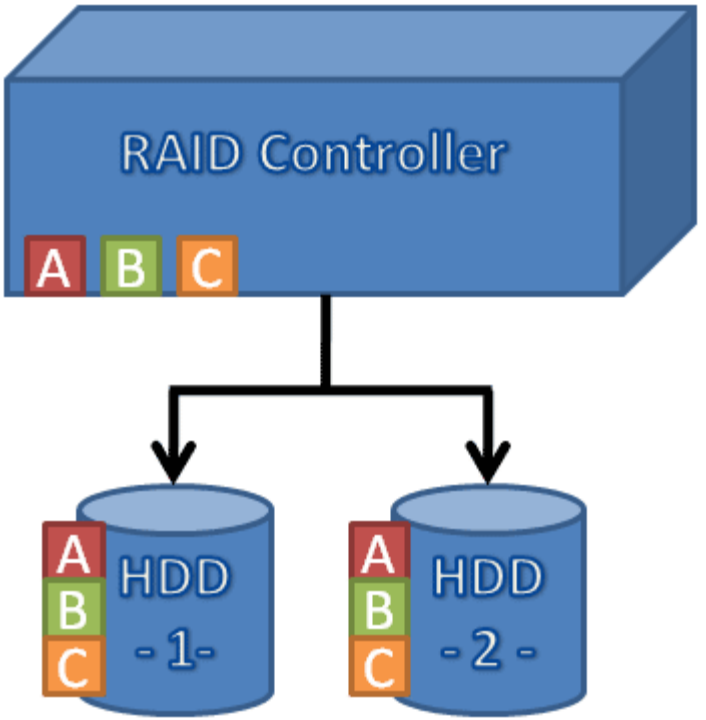
Raid Level 0

- Level 0 is nonredundant disk array
- Files are striped across disks, no redundant info
- High read throughput
- Best write throughput (no redundant ir
- Any disk failure results in data loss



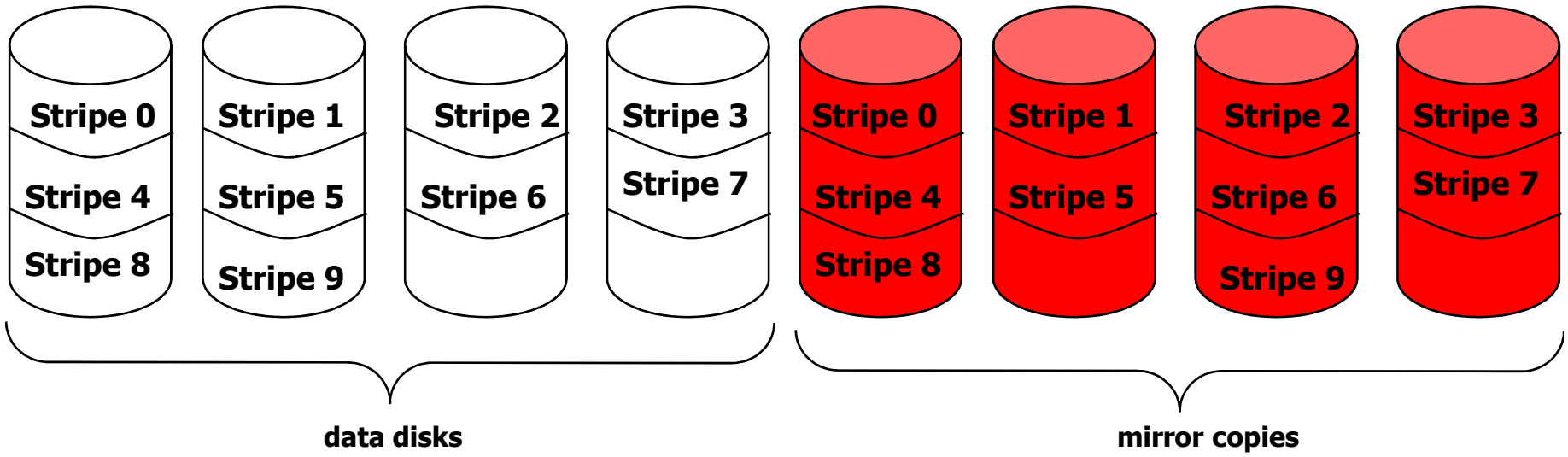
Raid Level 1

- Mirrored Disks
- Data is written to two places
  - On failure, just use surviving disk (easy to rebuild)
- On read, choose fastest to read
  - Write performance is same as single drive, read performance is 2x better
- Expensive (high space overhead)



Raid Level 0+1

- Stripe on a set of disks
- Then mirror of data blocks is striped on the second set.



Raid Level 1+0

- Pair mirrors first.
- Then stripe on a set of paired mirrors

