

CS305

Computer Architecture

RAID: Redundant Array of Inexpensive Disks

Bhaskaran Raman
Room 406, KR Building
Department of CSE, IIT Bombay

<http://www.cse.iitb.ac.in/~br>

Hard Disk Failures

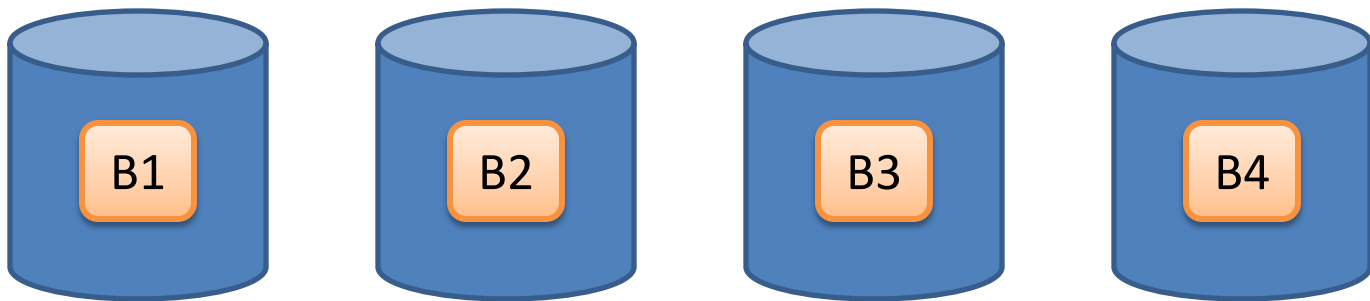
- Mechanical movements involved →
 - Failure probability high, increases with time
- Some terms:
 - **MTTF**: Mean Time To Failure
 - **MTTR**: Mean Time To Recovery (after failure)
- Example: MTTF = 1 year, MTTR = 2 days
 - Availability = $\text{MTTF} / (\text{MTTF} + \text{MTTR})$
 - Availability = $365 / 367 = 99.46\%$
 - Is this enough?

RAID:

Redundant Array of Inexpensive Disks

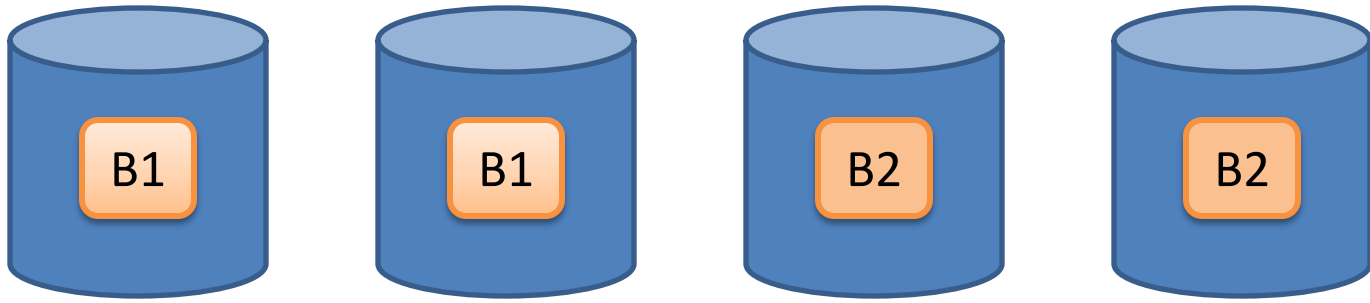
- Improving reliability of each disk: expensive, not a complete solution anyway
- Key idea in RAID: name states it all!
 - Reliability improved through redundancy
- Raid levels: 0-5 (also 6, 10)

Raid-0: No Redundancy, Only Extra Storage



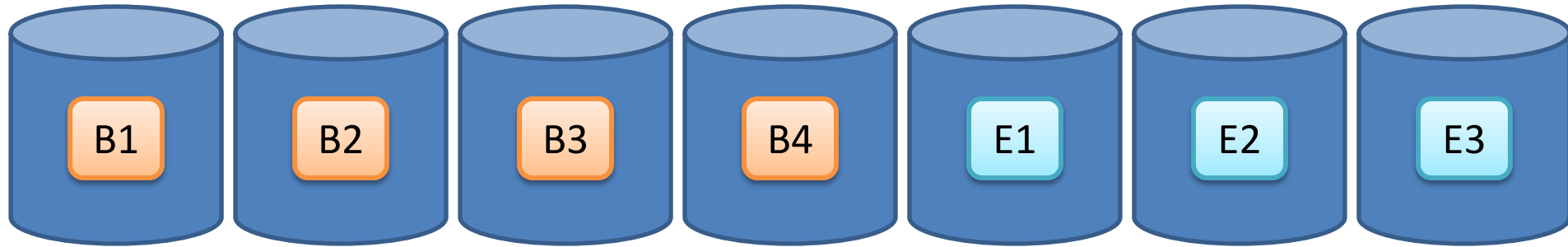
- (+) More storage
- (-) No fault tolerance

Raid-1: Mirroring



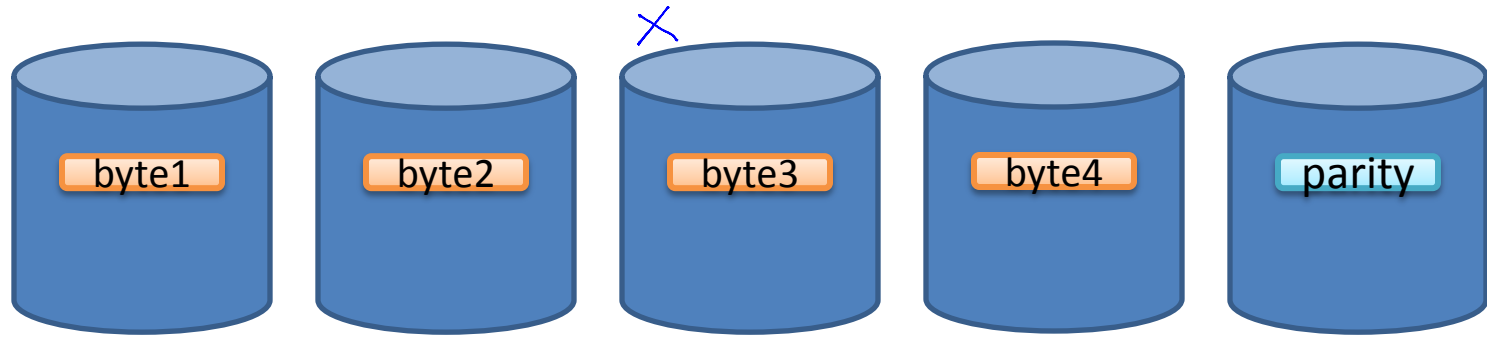
- (+) Single disk failure recovery: instant!
- (-) Twice the expense!

Raid-2: Error Correcting Codes



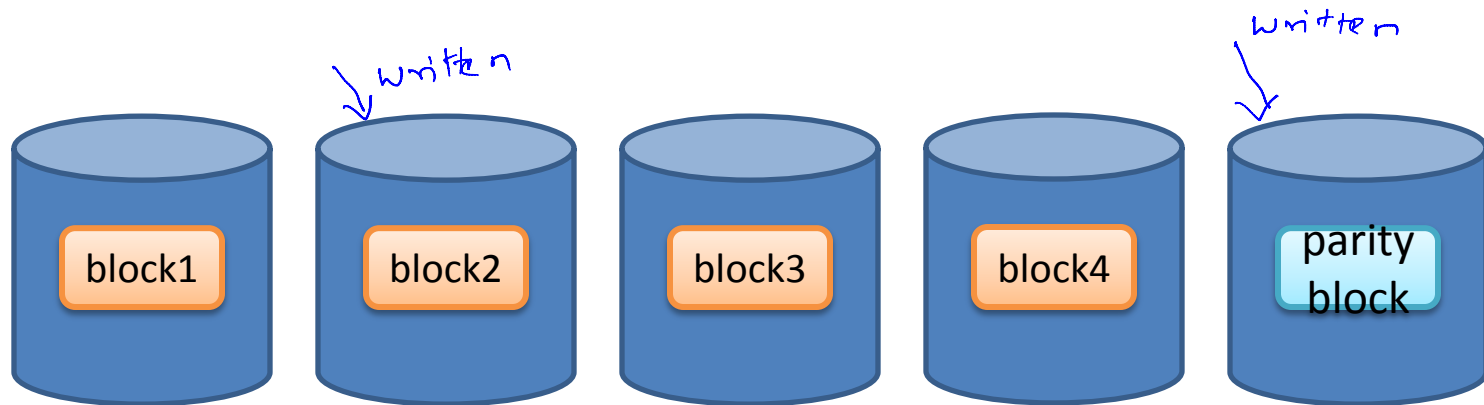
- (+) Single disk failure recovery: instant!
- (+) Lesser overhead than mirroring
- (-) Can do better: error correction not needed (hard disk failures are self-evident)

Raid-3: Byte-Level Parity



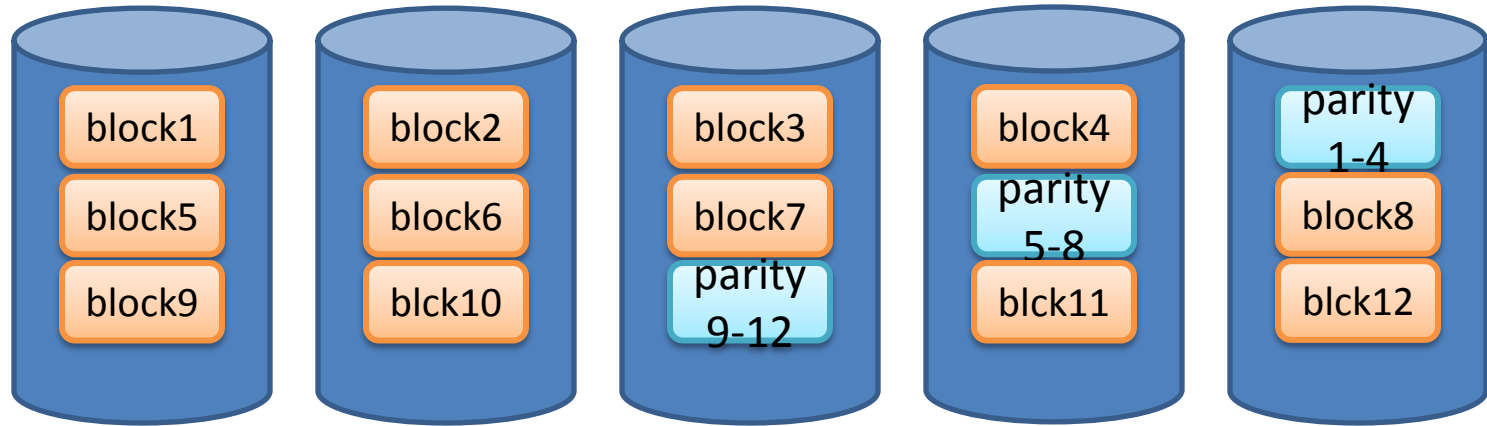
- (+) Single disk failure can be tolerated
- (+) Low overhead
- (-) All disks have to be synchronized in reading/writing

Raid-4: Block-Level Parity



- (+) Single disk failure can be tolerated
- (+) Low overhead
- (+) Disks do not have to be synchronized in reading/writing
- (-) Write throughput bottleneck: single block at a time

Raid-5: Block-Level Distributed Parity



- (+) Single disk failure can be tolerated, low overhead
- (+) Disks do not have to be synchronized in reading/writing
- (+) No write throughput bottleneck: parallel writes possible
- (-) Only one disk failure at a time can be tolerated

Summary

- Hard disks can fail: not good for high end systems
- RAID: use more number of inexpensive disks to get overall good fault tolerance
 - Raid-0: Only additional storage
 - Raid-1: Mirroring
 - Raid-2: ECC
 - Raid-3: Byte-level parity
 - Raid-4: Block-level parity
 - Raid-5: Block level distributed parity