CS343: Operating System

Disk Management and RAID Structure

Lect36 : 7th Nov 2023

Dr. A. Sahu

Dept of Comp. Sc. & Engg.

Indian Institute of Technology Guwahati

Outline

- Mass Storage
  - Disk Structure
  - Disk Arm Scheduling
  - Disk Management
  - RAID Structure
- FS Basic
- FS Implementation
- I/O subsystem
- Device Drivers

Disk Management

- **Low-level formatting**, or **physical formatting** — Dividing a disk into sectors that the disk controller can read and write
  - Each sector can hold header information, plus data, plus error correction code (**ECC**)
  - Usually 512 bytes of data but can be selectable

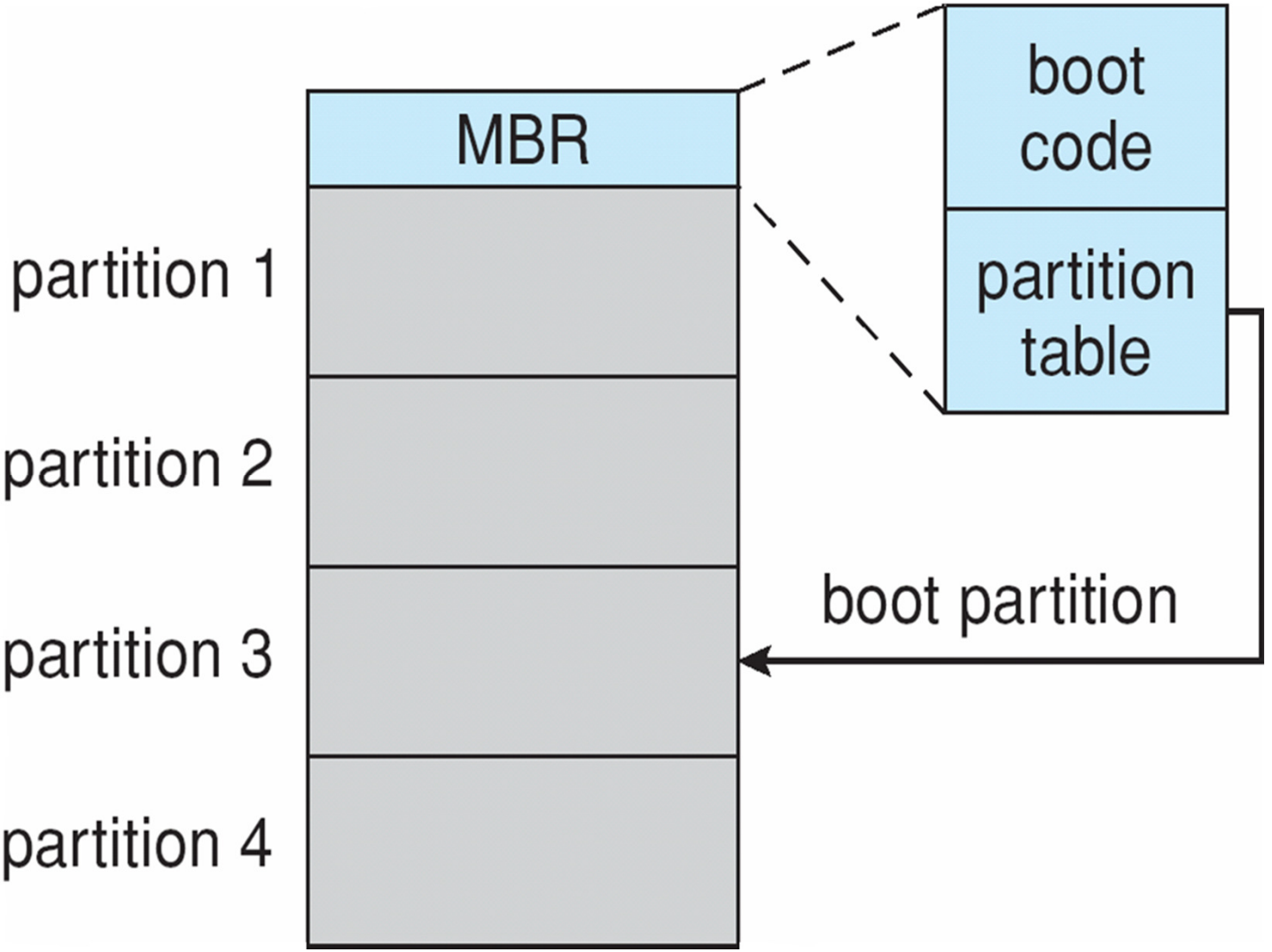
Disk Management

- To use a disk to hold files, OS still needs to record its own data structures on the disk
  - **Partition** the disk into one or more groups of cylinders, each treated as a logical disk
  - **Logical formatting** or "making a file system"
  - To increase efficiency most file systems **group blocks into clusters**
    - Disk I/O done in blocks
    - File I/O done in clusters

Disk Management (Cont.)

- Raw disk access for apps that want to do their own block management, keep OS out of the way (databases for example)
- Boot block initializes system
  - The bootstrap is stored in ROM
  - **Bootstrap loader** program stored in boot blocks of boot partition
- Methods such as **sector sparing** used to handle bad blocks

Booting from a Disk in Windows



Swap-Space Management

- Swap-space — Virtual memory uses disk space as an extension of main memory
  - Less common now due to memory capacity increases
- Swap-space can be carved out of the normal file system, or, more commonly
- It can be in a separate disk partition (raw)

RAID

- RAID – Redundant Array of Inexpensive Disks

- Creating a Disk System to provide
  - Reliability
  - Availability
  - Performance
  - Capacity

RAID

- RAID – redundant array of inexpensive disks
- Started by David Paterson University of Berkley
  - Author of Computer Architecture Book
- RAID
  - Multiple disk drive may provide better throughput and performance via interleaving /stripping
  - multiple disk drives provides reliability via **redundancy**
- Frequently combined with **NVRAM** to improve write performance
  - NVRAM: Faster then HDD, slower than RAM
  - Number write bound will not be a problem to use in between HDD and RAM

RAID

- Multiple disk provides reliability via **redundancy**
- Increases the **mean time to failure**
- **Mean time to repair –** exposure time when another failure could cause data loss
- **Mean time to data loss** based on above factors
- Several improvements in disk-use techniques involve the use of multiple disks working cooperatively

RAID (Cont.)

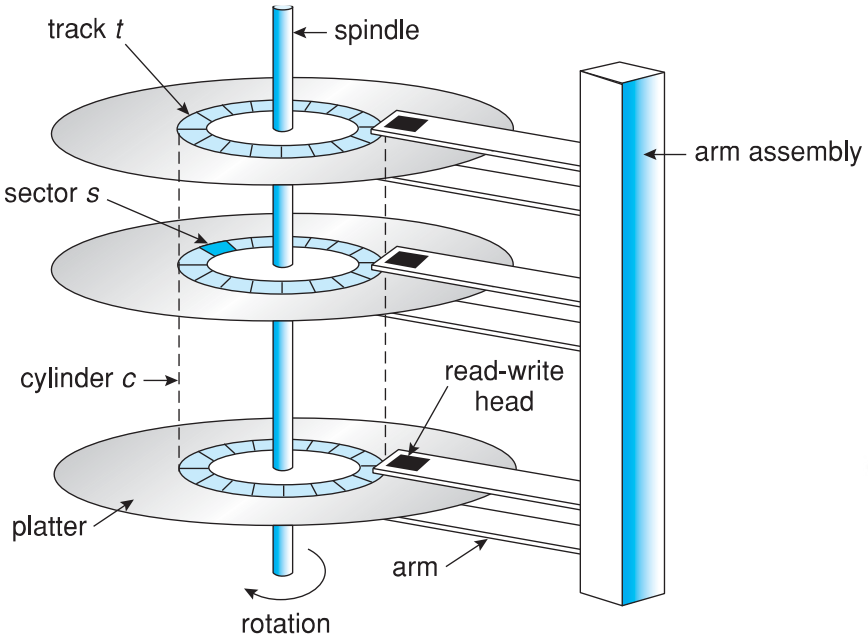
- Disk **striping** uses a group of disks as one storage unit
- RAID is arranged into six different levels
- RAID schemes improve performance and improve the reliability of the storage system by storing redundant data
  - **Mirroring** or **shadowing** (**RAID 1**) keeps duplicate of each disk
  - Striped 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

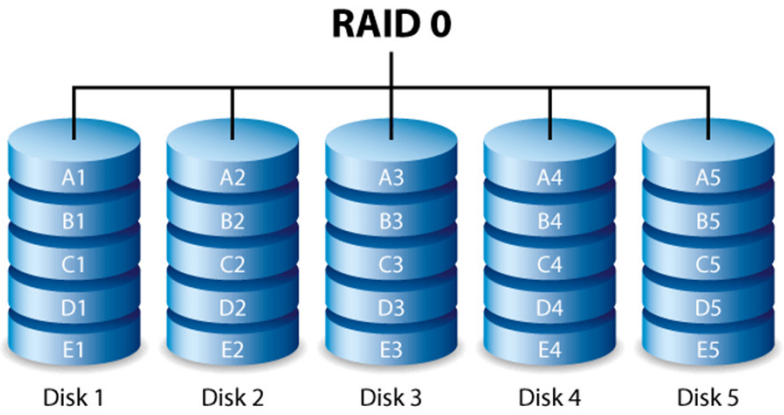
RAID (Cont.)

- RAID within a storage array can still fail
  - if the array fails, so automatic **replication** of the data between arrays is common
- Frequently, a small number of **hot-spare** disks are left unallocated
  - Automatically replacing a failed disk and having data rebuilt onto them

RAID 0 : David Paterson's

- RAID 0 : Main concentration on Performance
  - One HDD: Read/Write using multiple Head from many platter in parallel: Parallel Performance
  - Raid0: Read from multiple HDD in parallel
  - Assume 10 Platter/HDD and 10 HDDs: 100 blocks R/W in parallel

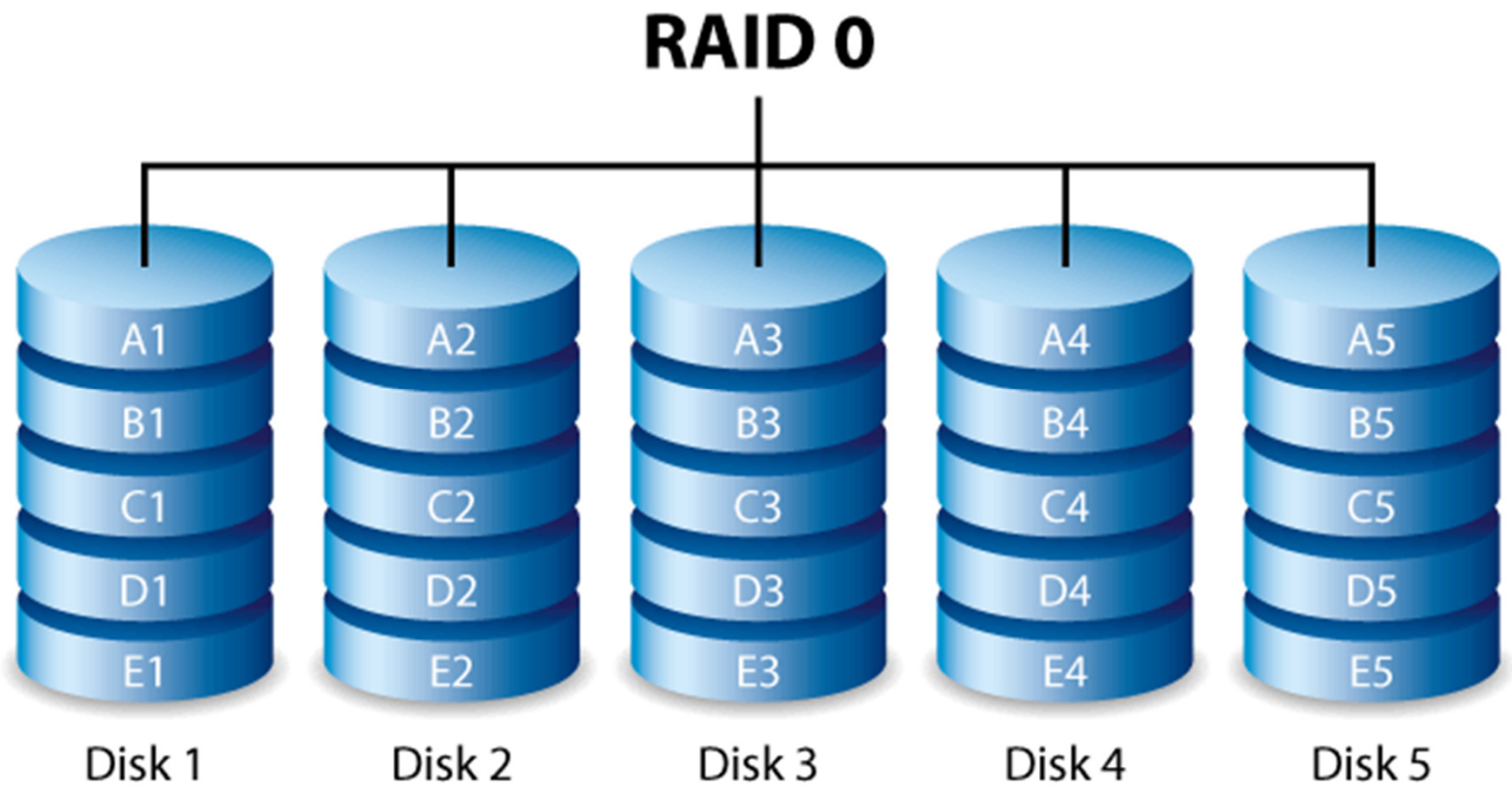




RAID 0

- Striping distributes contents of files roughly equally among all disks in the set
  - Which makes concurrent read or write operations on the multiple disks almost inevitable.
- Concurrent operations make the throughput of most read and write operations equal to the throughput of one disk multiplied by the number of disks.

RAID 0


RAID 0
A1
B1
C1
D1
E1
A2
B2
C2
D2
E2
A3
B3
C3
D3
E3
A4
B4
C4
D4
E4
A5
B5
C5
D5
E5
Disk 1
Disk 2
Disk 3
Disk 4
Disk 5

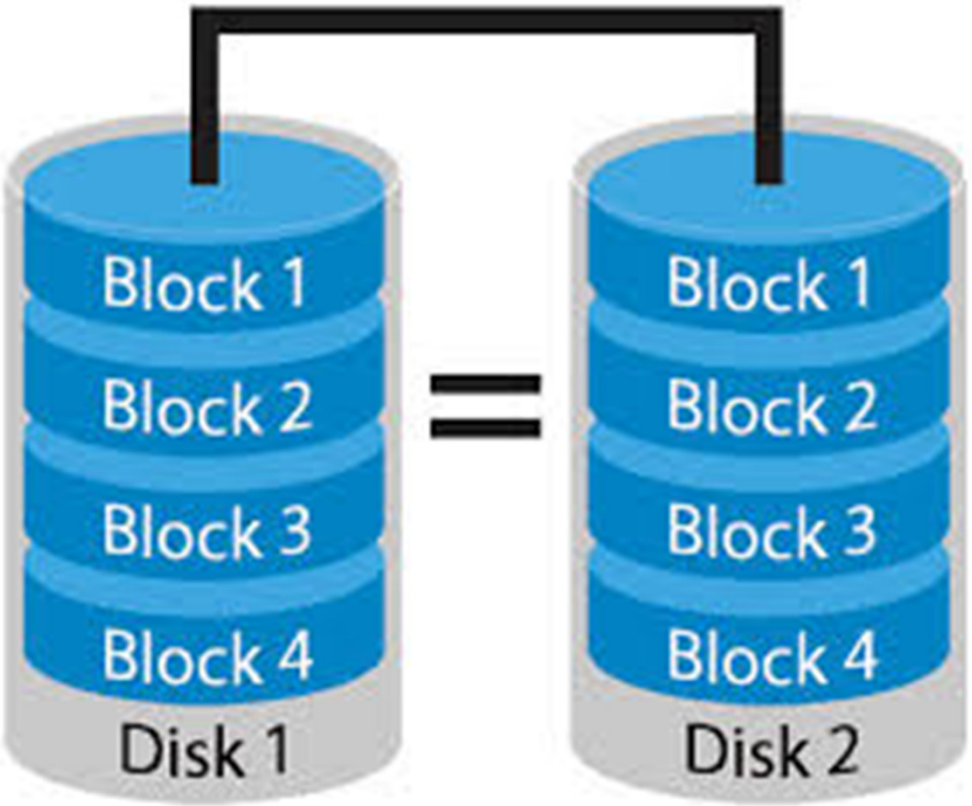
RAID 0

- **Increased throughput is the big benefit of RAID 0** versus spanned volume
- Capacity of a RAID 0 volume
  - **Sum of the capacities of the disks in the set**
  - Same as with a spanned volume

RAID 0

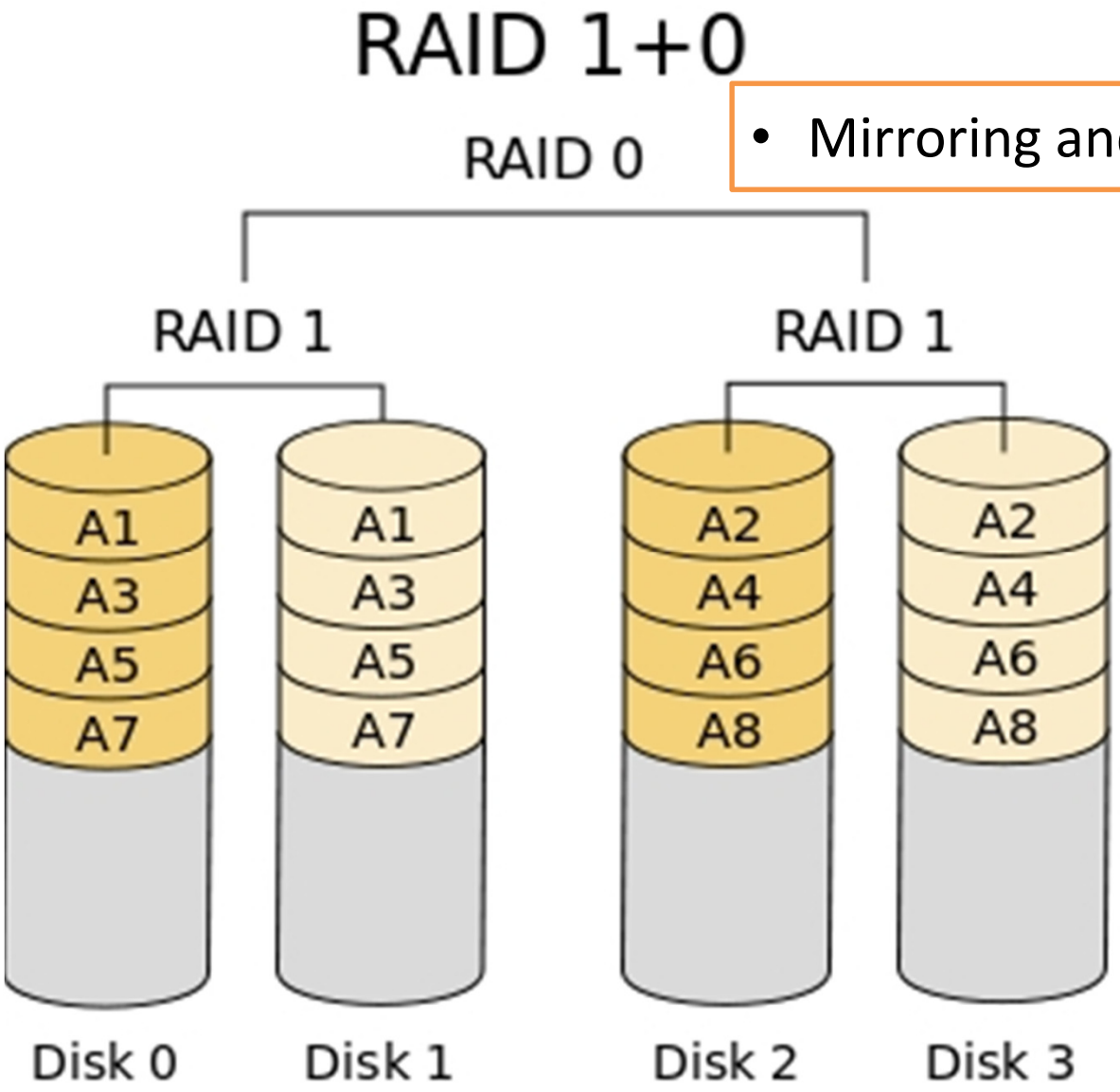
- RAID 0 consists of striping, without mirroring or parity.
- There is no added redundancy for handling disk failures, just as with a spanned volume
- Thus, failure of one disk causes the loss of the entire RAID 0 volume, with reduced possibilities of data recovery when compared to a broken spanned volume.

RAID 1: Mirroring, Full redundancy



- But Fully Redundancy (Always have a safe copy if one fail)
- No Capacity Increase
- No Performance Increase

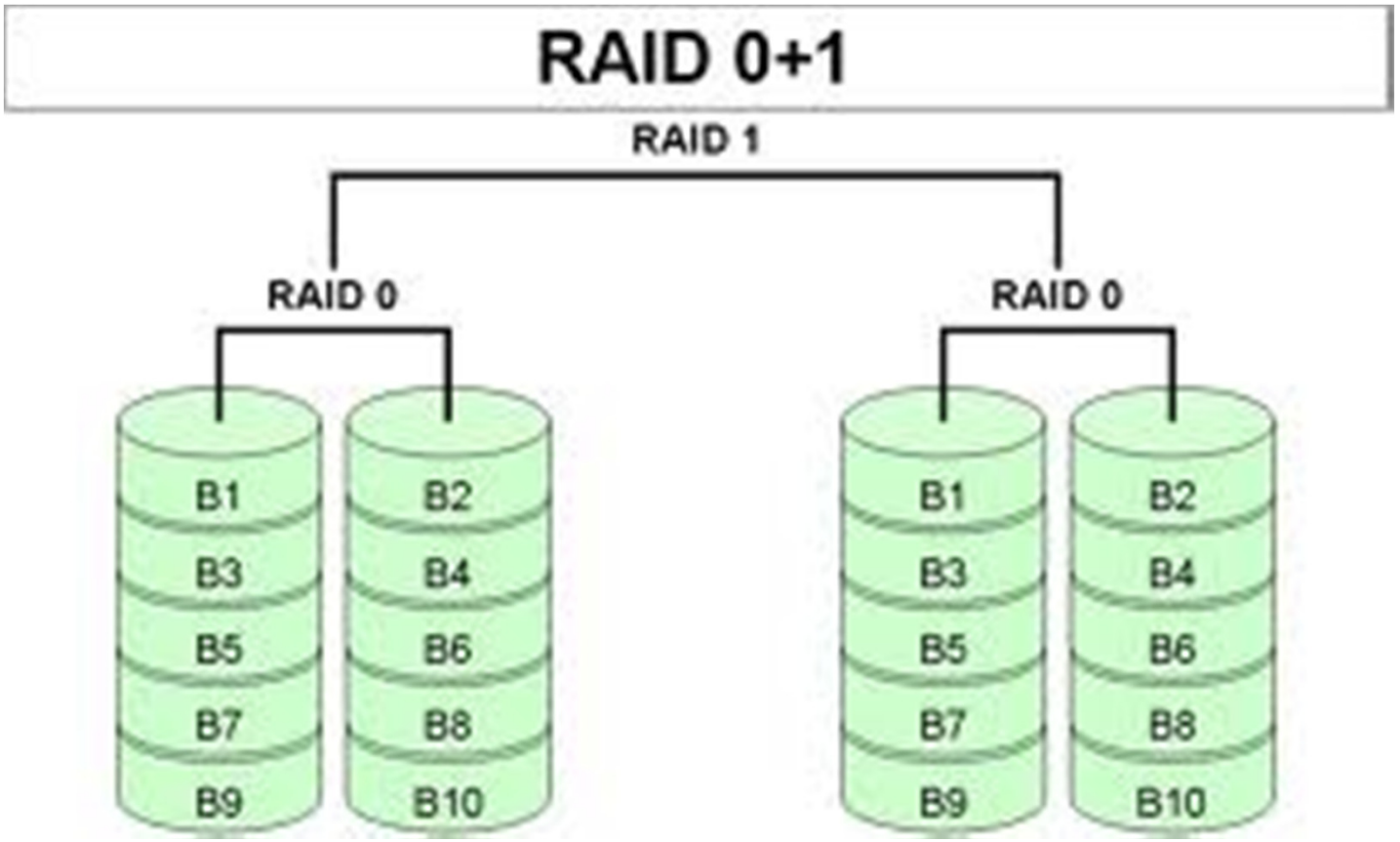
RAID 1+0 : Mirroring + Striping



- Mirroring and Striping (2)

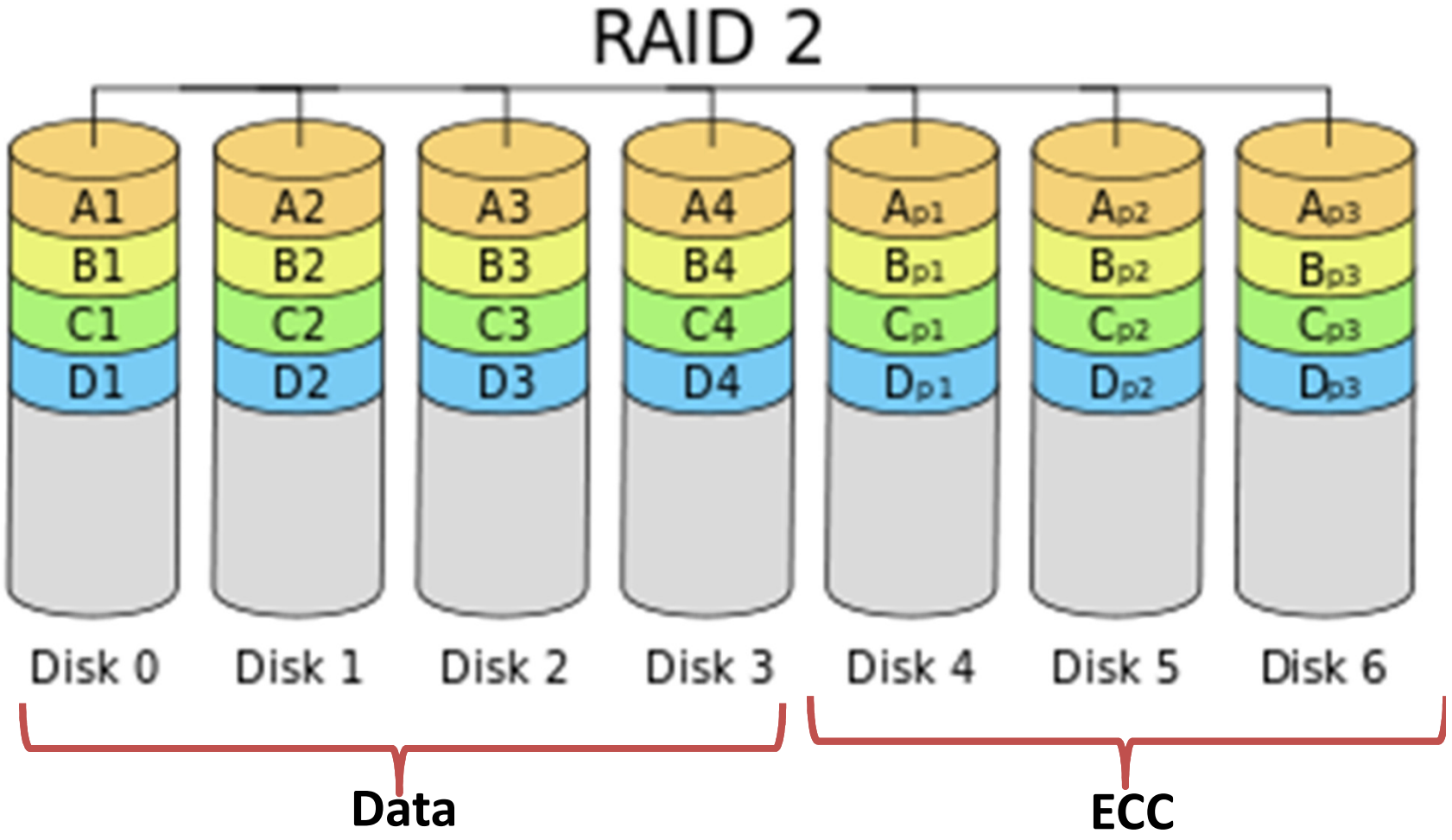
RAID 0+1

- Mirroring and Striping (2)

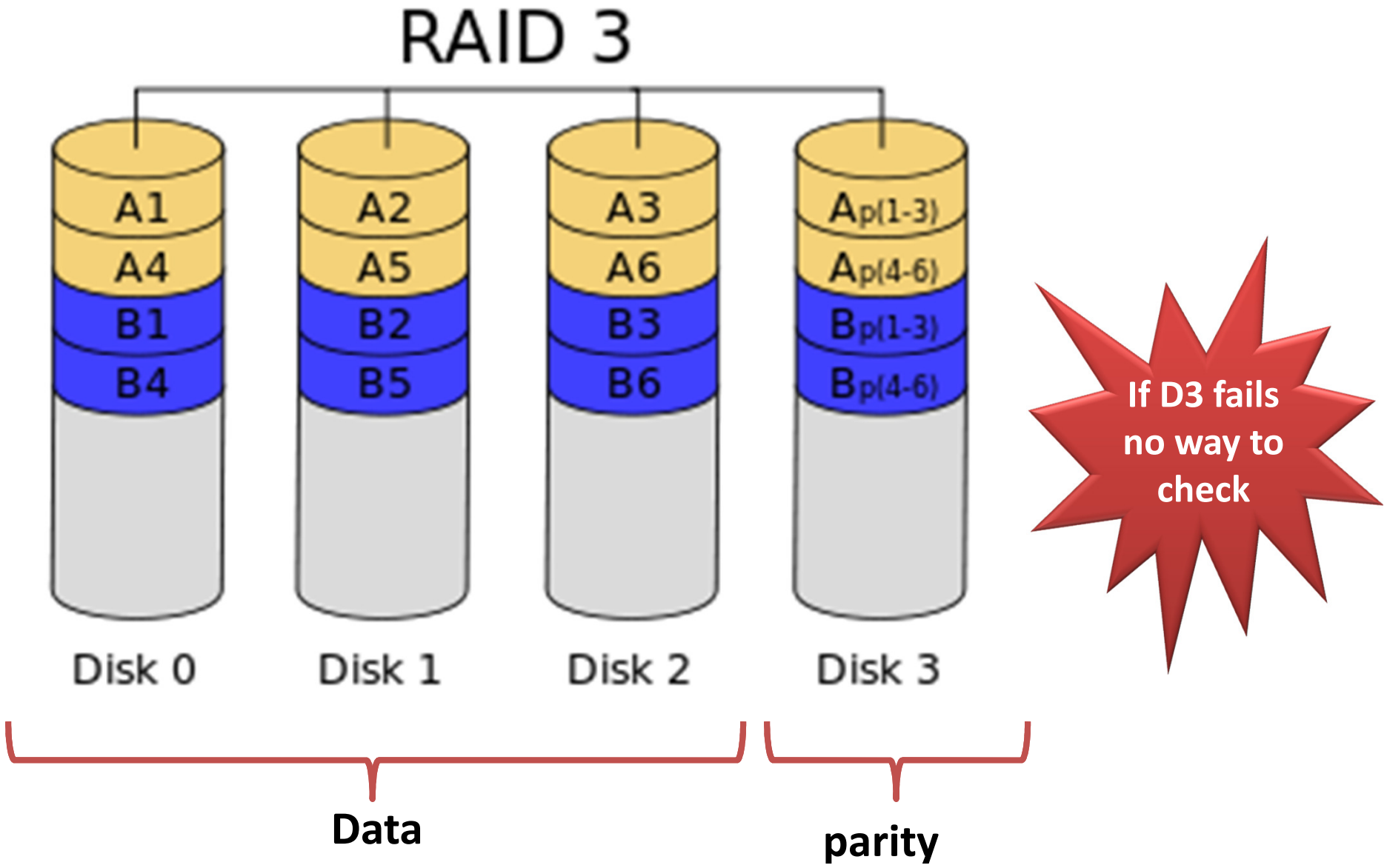


RAID 2: Bit level Stripping + ECC

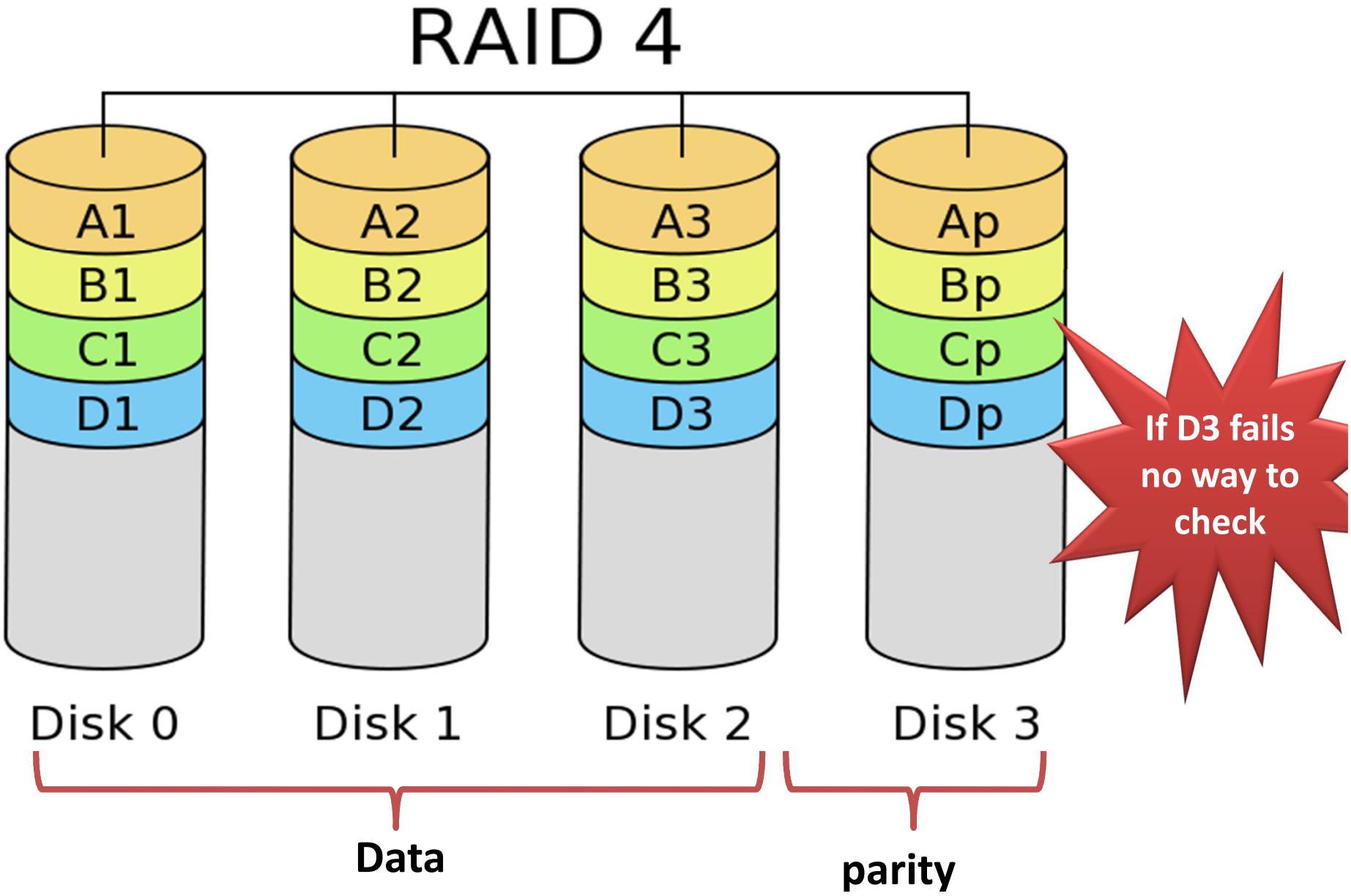
- Example **4 for Data** and **3 for ECC parity**: reduced capacity and performance but reliability



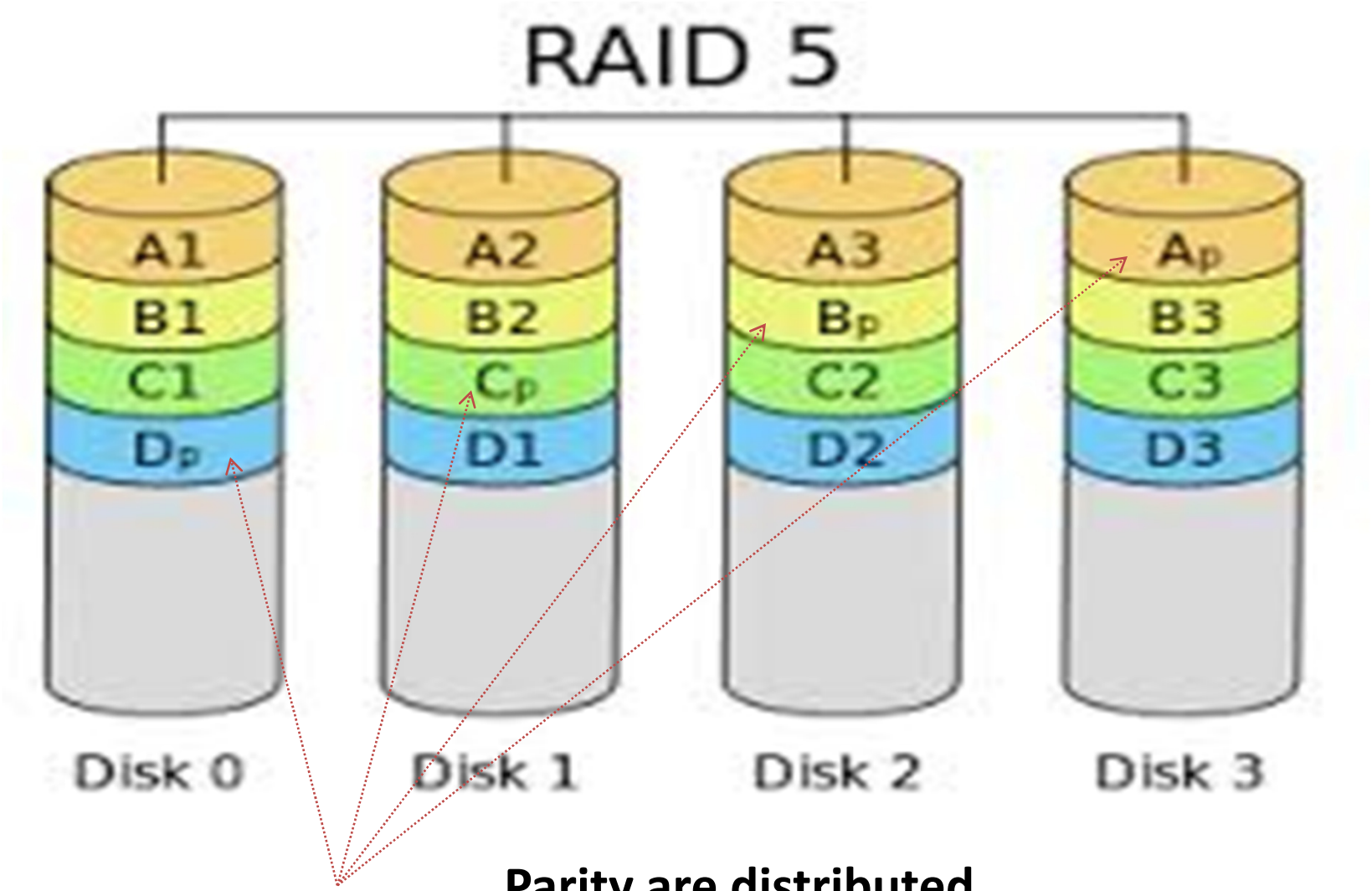
RAID 3: Byte level Stripping +Parity



RAID 4 : Block level Stripping +Parity



RAID 5: Stripping + Distributed Parity



Parity are distributed