PRINCIPLES OF DATA ORGANISATION

RAID

MOTIVATION

- Data are *very* important
- Good hard drives (reliable, fast) are expensive
- Cheap hard drives lack certain qualities
- How we can get more from inexpensive hardware → RAID

RAID

- Redundant Arrays of Inexpensive (Independent) Disks
- Inexpensive
  - original motivation
  - alternative to high-capacity expensive disks
- Independent
  - present-day motivation
  - higher reliability – redundancy
  - higher bandwidth – parallelism

- Must be supported by the controller
  - HW/SW-based solution

RAID – MTTF

- Problem: Mean time to failure (MTTF) of a system is much lower than MTTF of an individual device
  - System with 100 disks each with MTTF 100,000 hours (11 years) will have system MTTF 1000 hours (41 days)
- Redundancy of information can help by storing multiple copies of data which are then used in case of a failure

RAID – MIRRORING

- Keeps copies of a disk → each write is carried out on multiple disks
- Data are read from one disk
  - if one goes awry, the backup disk can be utilized

RAID – PARITY

Example with 5 disks: 3 disks with data (D1, D2, D3), one parity disk (DP) and one hot spare disk (HS) used for recovery purposes:

To calculate parity XOR operation can be used

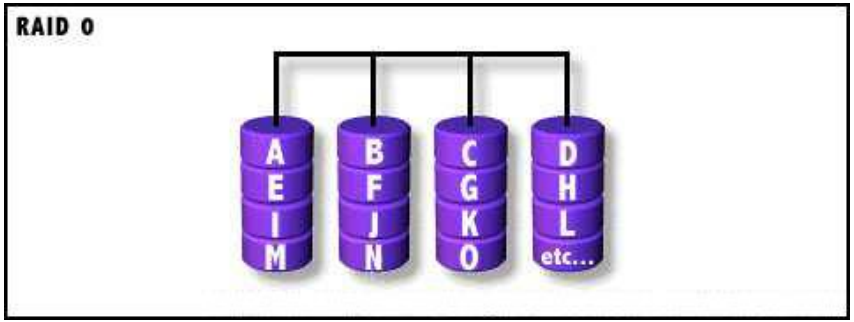
D1: 00100101
D2: 11101001
D3: 10101101
DP: 01100001 (= D1 XOR D2 XOR D3)

D2 breaks down

Now the original values of **D2** can be obtained from the parity information from **DP** (D1 XOR D3 XOR DP) and can be written to **HS** which can serve as a new **D2**

RAID 0

- No redundancy
- Stripping by blocks
  - Each block on one disk
  - Using mod to find out the position
- High performance
- Non-critical applications



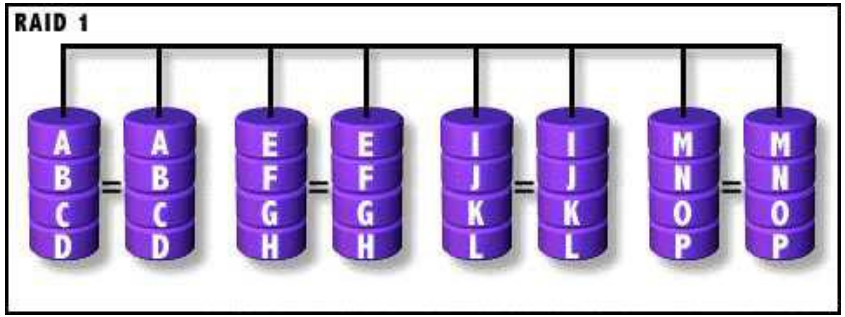
Advantages

- Easy to implement
- All storage capacity available
- Superior I/O performance

Disadvantages

- No fault-tolerance

RAID 1



- Disk mirroring
- Parallel writes
  - Optionally parallel reads (if supported by the controller)
- Data-critical applications (storing log files, accounting systems, …)

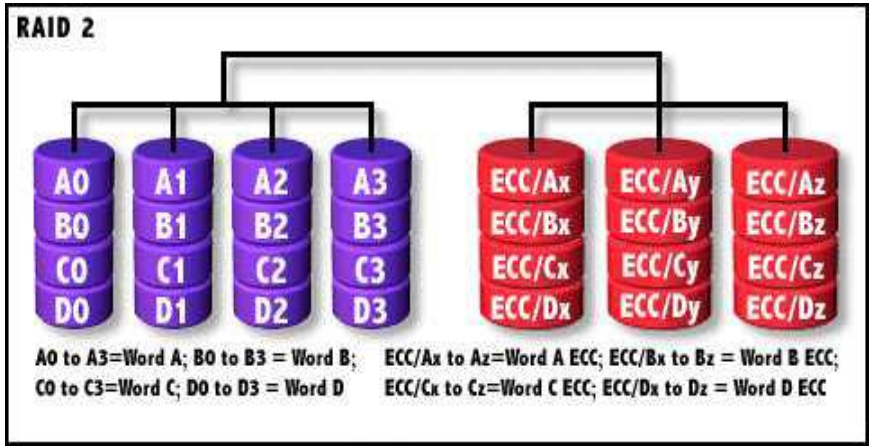
Advantages

- Easy to implement
- I/O speed comparable to a single disk
- In case of disk failure, data are only copied

Disadvantages

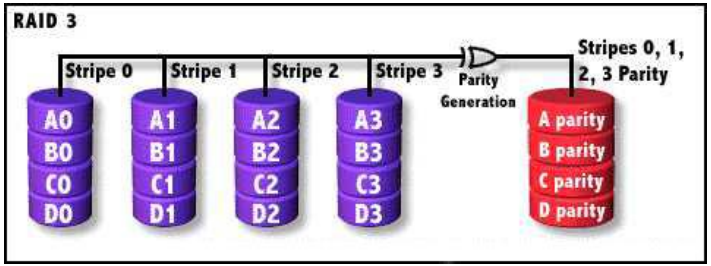
- 50% redundancy = 50% capacity used

RAID 2



- Bit-level striping
- Original idea: the same amount of discs as bits in a word
  - 32bit computer = 32 discs for data
- Hamming code parity (error correction)
  - Can detect up to two and correct up to one-bit errors
  - Redundant disks
    - The number of redundant disks is proportional to the log of the total number of disks on the system
- Synchronised rotations allow reading from all at once but not from different spaces
  - The controller is complex and expensive
- Rarely used

RAID 3



- Byte-level striping
- One parity disk (XOR) - performance bottleneck
  - Each write requires one more write (and computation) of the parity bit
  - Also parity disk checked on read
- Rarely used
  - Bottleneck

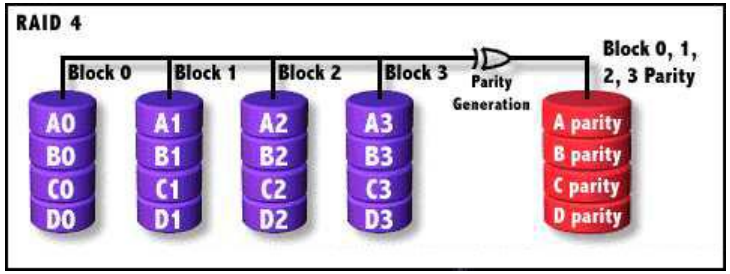
Advantages

- High-throughput for large I/O
  - Can utilize reading from multiple discs at once

Disadvantages

- At least 3 disks
- Slow for small I/O operations
- I/O requires activity on every disk

RAID 4



- Modification of RAID 3
- Block-level striping
- One parity disk (XOR)
  - Performance bottleneck
- Rarely used

Advantages

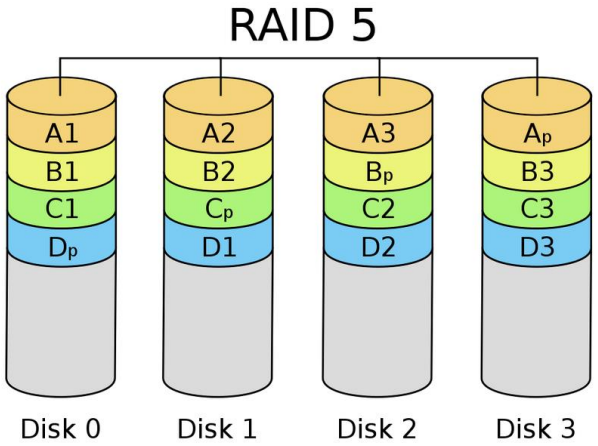
- I/O requests can be carried out in parallel

Disadvantages

- Lot of small write operations can be problematic
- Complex controller design

RAID 5

- Block-level striping
- Distributed parity
- Reads do not check the parity block (too expensive)
- Can handle single disk failure
- Most common



Advantages

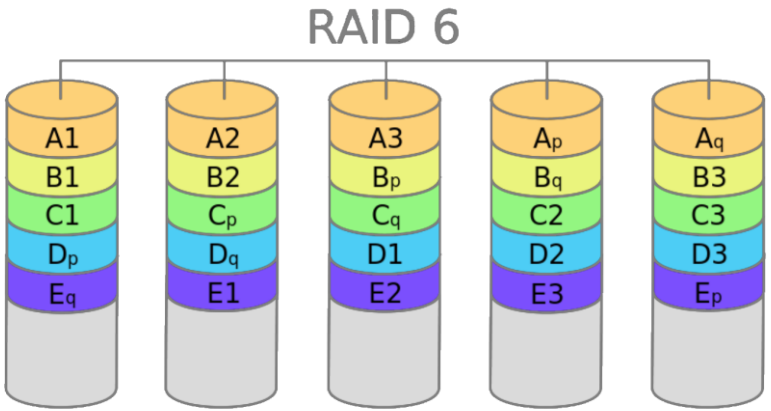
- High-throughput read operation
  - But not as good as for mirroring
- Good aggregate transfer rate

Disadvantages

- Write operation is slower (parity computation)
- At least 3 disks

RAID 6

- Enterprise version of RAID 5
- Block-level striping
- 2 distributed parity blocks
  - Each counted in a different way
- RAID 5 for enterprise



Advantages

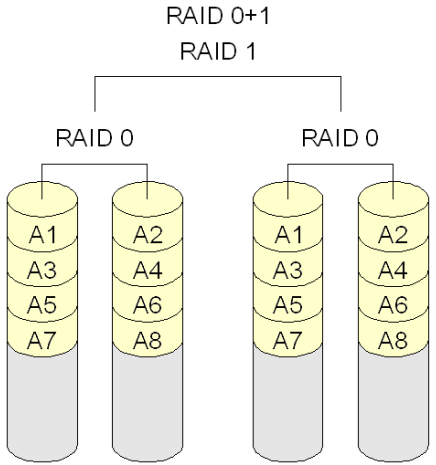
- Can handle failure of 2 drives

Disadvantages

- Expensive write

RAID 0 + 1

- Alternative to enterprise RAID 6
- RAID 1 of RAID 0
  - Mirroring on top of stripping
- 6 disks (2 sets of 3 disks)



Advantages

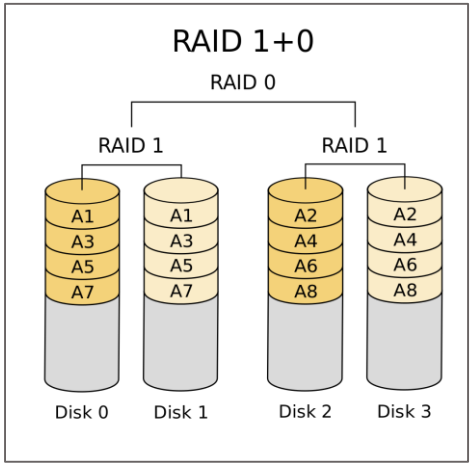
- Additional fault tolerance (as RAID 5)
- High data transfer rate
  - Data can be read from multiple place at once

Disadvantages

- High overhead
  - Half of the capacity is utilized
- Limited scalability for more disks
- Requires at least 4 disks

RAID 1 + 0

- RAID 0 of RAID 1
  - Stripping on top of mirroring
- 6 disks (3 sets of 2 disks)
- Database server requiring high performance and fault tolerance



Advantages

- High reliability (in each RAID 1 array, 1 disk can fail)

Disadvantages

- High overhead
- Limited scalability
- Requires at least 4 disks