An Overview of RAID

András Petri
Technical University of Budapest
petri@mmt.bme.hu

August 31, 1997

TEMPUS S_JEP-8333-94

RAID Overview

Requirements for high-capacity mass storage devices:

- Higher performance
- Higher reliability

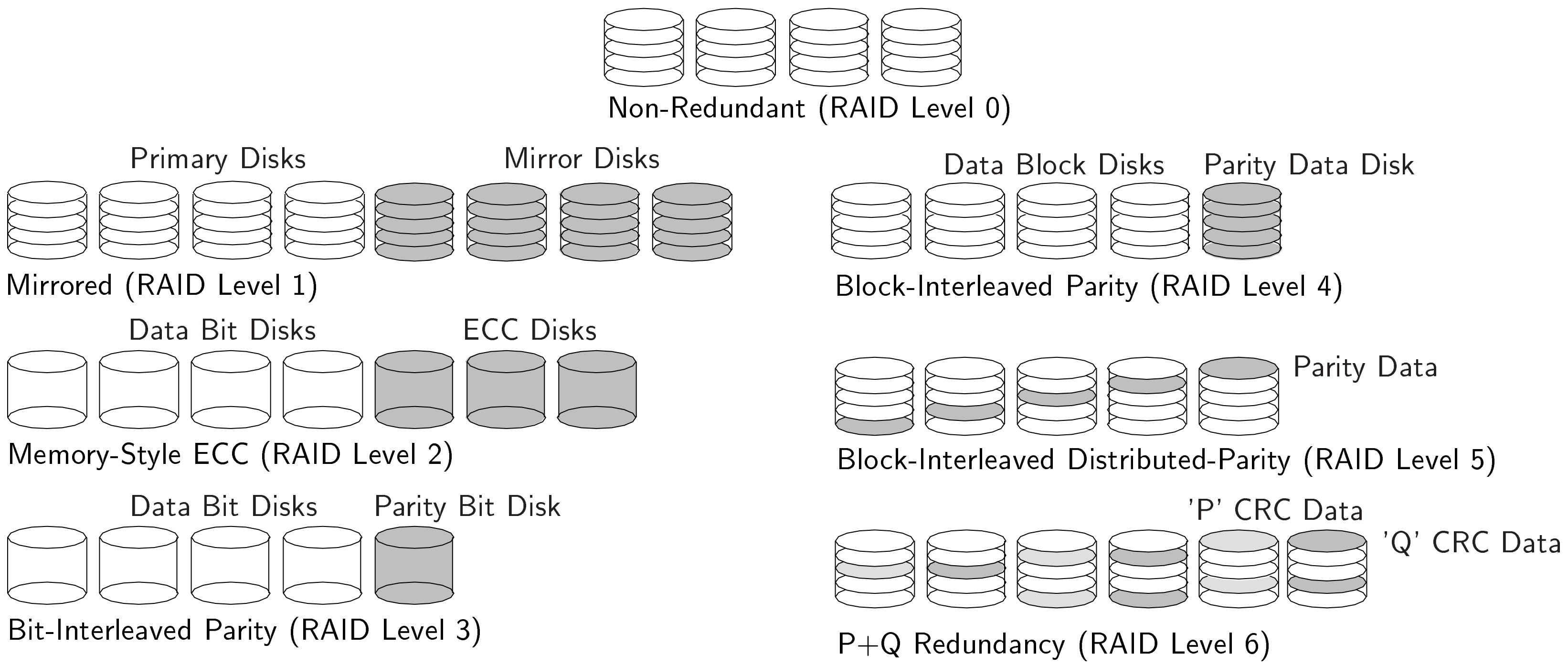
Most common solution: *disk arrays*

Reliability requires application of *redundancy*

RAID stands for **R**edundant **A**rray of **I**nexpensive **D**isks

(Inexpensive: disks of the array are not special, ultra-reliable ones)

Basic RAID organizations



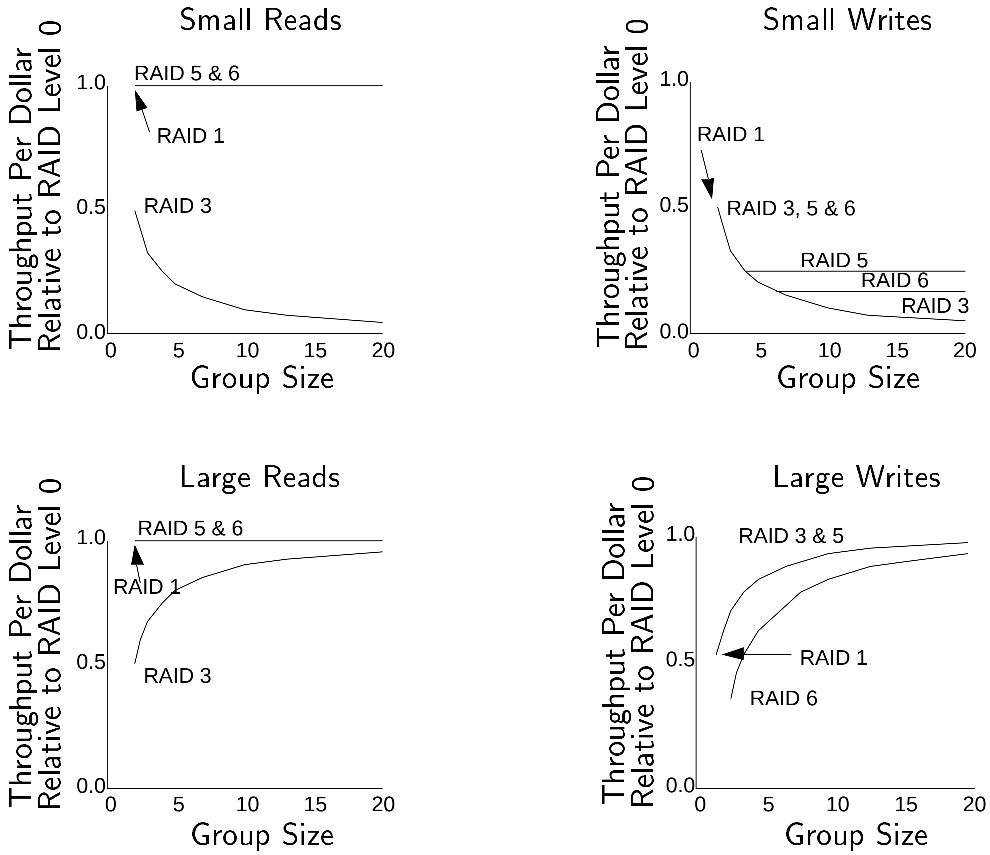
Comparisons of the different RAID classes

RAID Level	Small Reads	Small Writes	Large Reads	Large Writes	Storage Efficiency
1	1	1/2	1	1/2	1/2
3	1/G	1/G	(G-1)/G	(G-1)/G	(G-1)/G
5	1	max(1/G,1/4)	1	(G-1)/G	(G-1)/G
6	1	max(1/G,1/6)	1	(G-2)/G	(G-2)/G

Throughput per dollar metrics of different RAID levels relative to RAID level 0

- G is the size of the parity group

How the parity group size affects performance



Reliability

Common problems in RAID structures:

- System crashes $\Rightarrow$ inconsistent parity data
- Uncorrectable disk read errors $\Rightarrow$ unreadable parity data
- Correlated disk failures $\Rightarrow$ insufficient parity data

Reliability measurements must be made with respect to these factors

Example: Reliability measures of a RAID level 5 array

Type of Failure	Mean Time To Data Loss	MTTDL (Years)	10 Year Failure Rate
Double Disk Failure	$\frac{MTTF(disk) \times MTTF(disk2)}{N \times (G-1) \times MTTR(disk)}$	285	3.4%
System Crash + Disk Failure	$\frac{MTTF(system) \times MTTF(disk)}{N \times MTTR(system)}$	154	6.3%
Disk Failure + Bit Error	$\frac{MTTF(disk)}{N \times (1-(p(disk)^{G-1}))}$	36	24.4%
Software RAID	(harmonic sum of above)	26	31.6%
Hardware RAID (with NVRAM)	(harmonic sum excluding system crash + disk failure)	32	26.8 %

Note that the total number of disks in the system, N, is equal to the number of data disks times $G/(G-1)$.

Implementation considerations

- Mechanisms for dealing special conditions:
  - avoidance of reading stale (invalid) data
  - reconstruction of parity data after system crash
  - operation with failed disk(s)
- Considering failures of other system components:
  - orthogonal RAID
- Performance tuning:
  - buffering (caching)
  - service of multiple I/O requests
  - optimal arrangement of redundant data