Google Paper:
Design Project 3?

PageRank

Redundant Array of Disks

6.033 Quiz 2 Review

April 18, 2007

Why RAID?

- Economics
- Aggregate I/O of multiple cheap disks can surpass individual SLED
- Lower power
- Belief in unreliable components

RAID Level 0

- Interleave or stripe data (bit, bytes, sectors)
  - Block A is spread across multiple disks

RAID Level 1

- Mirroring; duplicate data on second disk
- Advantages:
  - Could double read throughput by reading in parallel
- Disadvantages:
  - Write to multiple disks; subject to slowdown
  - 50% of disk capacity wasted

RAID Level 2

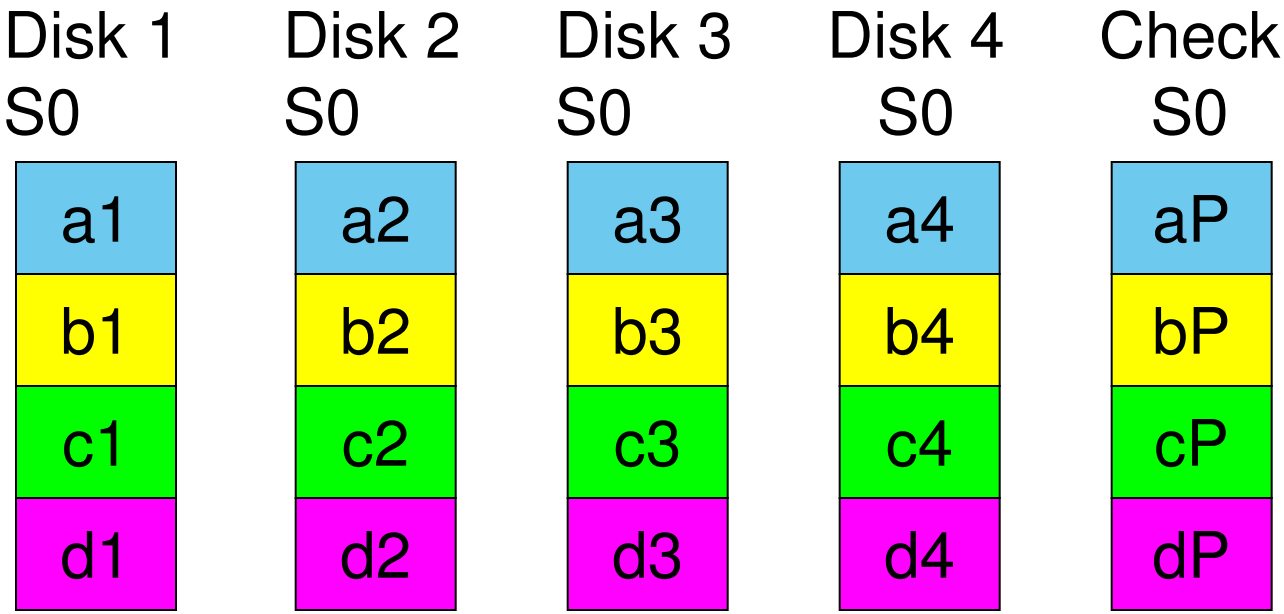
- Reduce number of check disks using ECC
  - e.g. 10 data disks, need 4 check disks

D3	D2	D1	P3	D0	P2	P1
7	6	5	4	3	2	1

$$P3 = D3 \oplus D1 \oplus D0$$
$$P2 = D3 \oplus D2 \oplus D0$$
$$P1 = D3 \oplus D2 \oplus D1$$

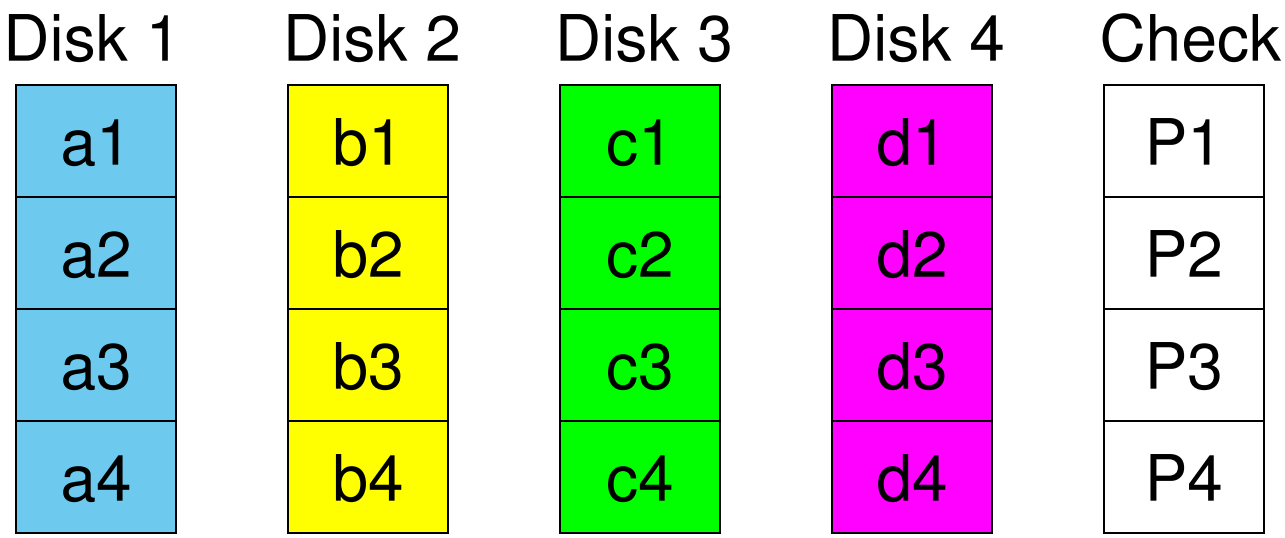
RAID Level 3

- Rely on hardware error-detection (fail-fast)
- Bit/Byte-level interleaving + parity
  - Block A is broken up into individual bytes



RAID Level 4

- Keep files together; group blocks into single sector



RAID Level 5

- Spread parity sector across disks