RAID - Redundant Arrays of Inexpensive Disks

RAID: High-Performance, Reliable Secondary Storage by P. Chen, E. Lee, G. Gibson, R. Katz and D. Patterson in *ACM Computing Surveys 26(2)*, June 1994, pp. 145 - 185

Outline of Lecture

- Motivation
- Disk Arrays
- Data Striping
- Redundancy
- Basic Organizations

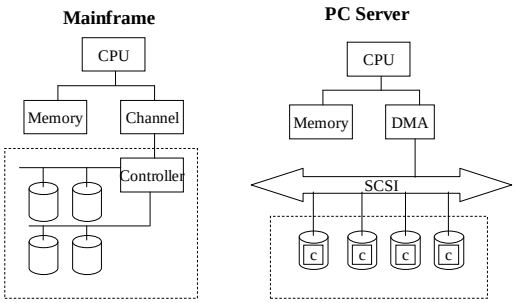
Motivation

- Driving force behind interest in RAID is the sustained exponential improvements in semiconductor technology
  - processor performance improving ≥ 50% / year
  - disk performance improving < 10% / year
- Amdahl's Law predicts only marginal overall improvements

Motivation (cont.)

- Faster CPUs make possible new applications and expand range of existing ones
  - new problems typically require fast access to larger data sets
  - trend towards large, shared, high-performance, network-based storage systems
- Disks are potential bottlenecks for system performance!

Disk Organizations



Disk Arrays

- Organize multiple independent disks into a single large high-performance *logical* disk
  - How does this increase performance?
- Large disk arrays highly vulnerable to failure
  - How do we increase reliability?

Data Striping

- Distributes data transparently over multiple disks
- Disks are accessed in parallel
- Array of disks appears a single large fast disk

Data Striping (cont.)

- Data segmented into equal-sized partitions distributed over multiple disks
  - Size of partition called the striping unit
  - Striping unit characterizes how disks are utilized
  - Units usually assigned in a round robin fashion
- Improves aggregate I/O performance by allowing multiple I/Os to be serviced in parallel

Redundancy

- Reliability of disk array decreases as number of disks increases

$$MTTF_{Array\ of\ Disks} = \frac{MTTF_{Single\ Disk}}{Number\ of\ Disks}$$

- Redundancy necessary to tolerate disk failures and allow continuous operation without data loss

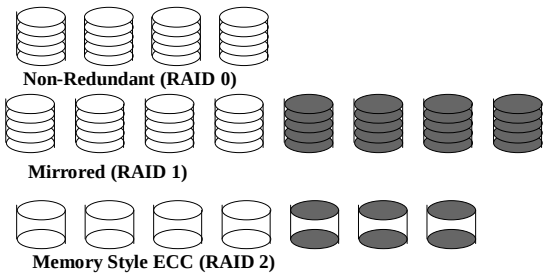
Redundancy (cont.)

- Introduces 2 problems:
  - Method for distributing redundant info across the disks?
    - Concentrate redundant info on a small number of check disks
    - Spread redundant info uniformly across all disks
    - Advantages? Disadvantages?
  - Method for computing the redundant info?
    - Most disk arrays store parity info

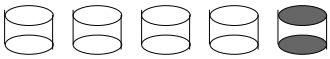
Redundancy (Cont.)

- In RAID disk array partitioned into **reliability groups**
  - Reliability group consists of a set of data disks and a set of check disks
  - Common redundancy scheme applied to each group

Basic RAID Organizations



Basic RAID Organizations (Cont.)



Bit Interleaved Parity (RAID 3)

Block Interleaved Parity (RAID 4)

Block Interleaved Distributed Parity (RAID 5)

Basic RAID Organizations (cont.)

P + Q Redundancy (RAID 6)