Disk Arrays

COMP375 Computer Architecture and Organization

Goals

- Understand the different options for improving disk reliability and performance.
- Be able to determine the appropriate RAID solution for a given situation.
- Be able to calculate the overhead for different RAID configurations.

Large Storage Needs

- Some applications need more storage than a single disk can hold.
- Multiple disks need to be used.
- Such application may wish to have quick access to large files.

RAID

- **R**edundant **A**rray of **I**ndependent **D**isks
- A collection of disks are used as one large unit of mass storage.
- Multiple disks operating simultaneously can increase the data transfer rate.
- Extra data stored on the disks can recover the information should a disk fail.

Historical Note

- In the mid 1980's mainframe disk drives were physically large and expensive, about $50,000.
- PC disk drives were about 2/3 as fast, held about 1/5 as much and cost under $1,000.
- People began to wonder if they couldn't use a collection of small drives.
- <u>R</u>edundant <u>A</u>rray of **<u>I</u>nexpensive** <u>D</u>isks

Hardware or Software

- RAID can be implemented by the I/O controller or by software in the OS.
- A RAID controller hides the RAID functions from the OS and makes a set of disk drives look like a single large drive.
- The OS can implement RAID on a set of regular disks. The user is unaware of the RAID features.

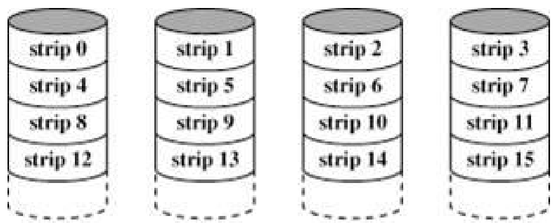
Failure Rates

- As the number of parts in a system increases, so does the probability of failure
- A device's reliability can be expressed as its probability of failure, **P**.
- If a system is composed of **N** components each with reliability **P**, the reliability of the system is $\mathbf{P}^{\mathbf{N}}$

RAID Types

- RAID 0 - Striping
- RAID 1 - Mirroring
- RAID 2 - Hamming code error recovery
- RAID 3 - Bit-interleaved parity
- RAID 4 - Block-level parity
- RAID 5 - Block-level distributed parity
- RAID 6 - Dual redundancy

RAID 0 (non-redundant)



(a) RAID 0 (non-redundant)

- Improved transfer rate
- Decreased reliability

http://www.acnc.com/04_01_00.html

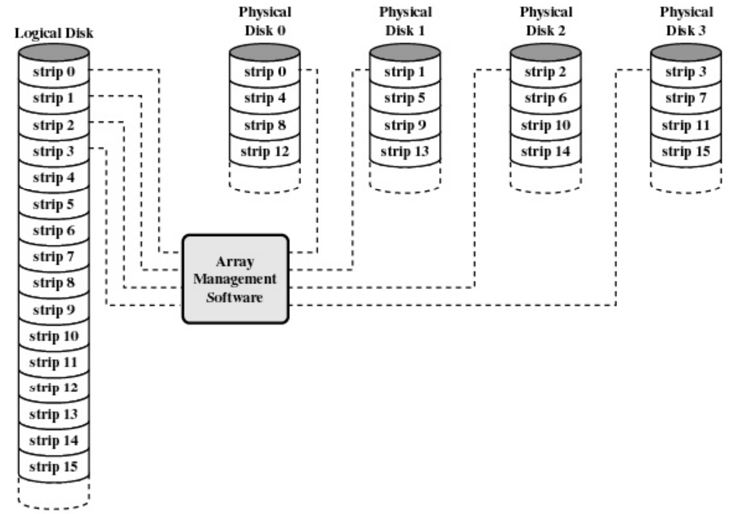
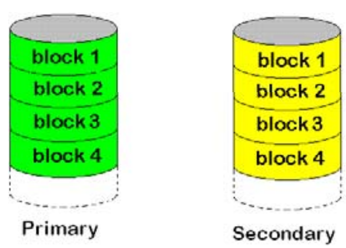


Figure 11.10 Data Mapping for a RAID Level 0 Array [MASS97]

RAID 1 (mirrored)



- Improved Reliability
- Slightly slower writes.
- Possibly faster reads
- Twice the disk space required

XOR Parity

- Consider the exclusive OR of several values

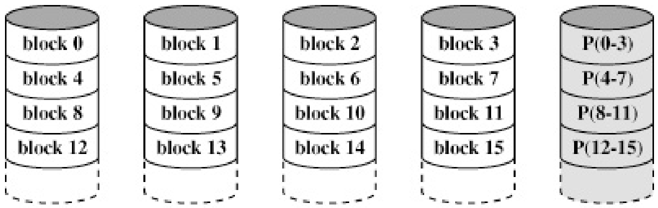
$$X = A \oplus B \oplus C$$

- If you XOR any of the three values, you will get the fourth.

$$B = X \oplus A \oplus C$$

- RAID 3, 4, 5 & 6 write the XOR of data to an additional disk to provide recovery in the event a disk fails.

RAID 4 (block-level parity)

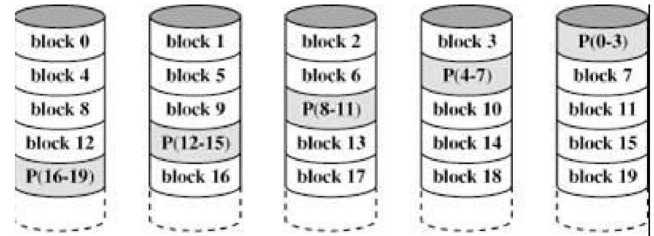


(e) RAID 4 (block-level parity)

Figure 11.9 RAID Levels (page 2 of 2)

Rarely used

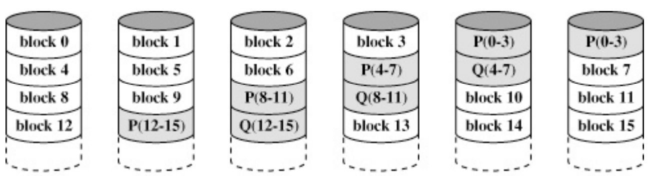
RAID 5 (distributed block parity)



RAID 5 (block-level distributed parity)

- Striping improves read performance
- Parity improves reliability
- N+1 disks are required

RAID 6 (dual redundancy)



(g) RAID 6 (dual redundancy)

Figure 11.9 RAID Levels (page 2 of 2)

- Like RAID 5 but with two parity blocks for each data block
- Slow writes
- N+2 disks required

Synchronized Rotation

- All of the drives in an disk array spin together
- Most disk drives have a feature allowing synchronization.
- Raid 0 simultaneously reads a block from all drives. If they were not synchronized, the system would have to wait for the last drive.
- Synchronized rotational delay = spin/2
- Unsynchronized rotational delay = spin*(n-1)/n

RAID Comparison

RAID	Disks	Reads	Writes	Survives failures
0	N	faster	faster	0
1	2N	slightly faster	slightly slower	1
5	N+1	faster	slightly slower	1
6	N+2	faster	slightly slower	2