

A Case for Redundant Arrays Of Inexpensive Disks

Paper By

David A Patterson

Garth Gibson

Randy H Katz

University of California, Berkeley.

Background

- ❑ Computer Speed depends on
 - CPU Speed
 - Memory Speed
 - I/O Activity
 - ❑ The Pending I/O Crisis
-

Solution

❑ Array of Inexpensive Disks

- Have small size and consume low power
- Contains full track buffer and functions of Mainframe controllers
- Unreliable

MTTF of a disk array is given by

MTTF of a single disk/Number of disks in the
Array

A Better Solution: RAID

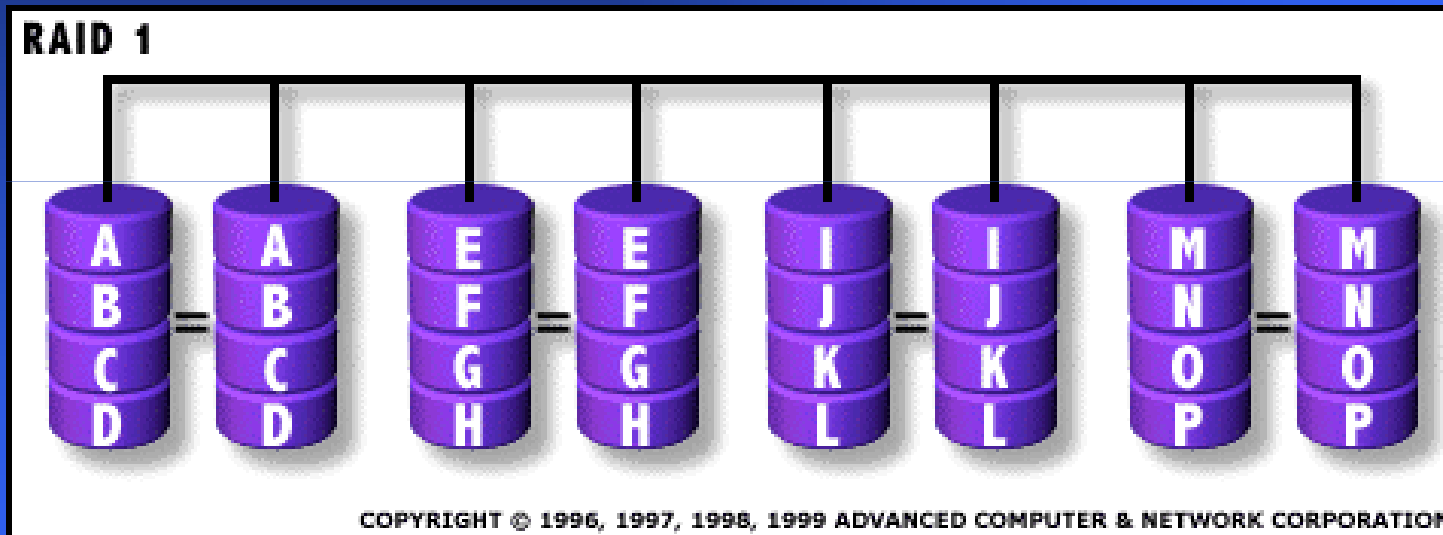
□ Reliability

- Arrays are divided into reliability groups
 - Each group has check disks
 - How to evaluate
 - Reliability Overhead Cost
 - Useable Storage Capacity Percentage
 - Performance
-

Different Levels of RAID

- RAID Level 1
 - RAID Level 2
 - RAID Level 3
 - RAID Level 4
 - RAID Level 5
- and others
-

RAID Level 1 (Mirroring)

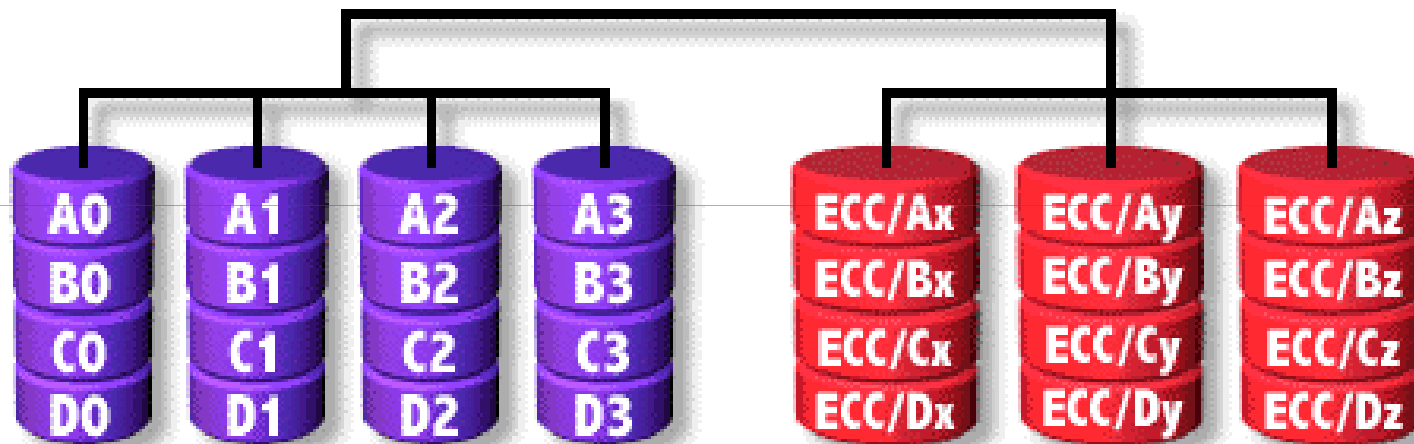


RAID Level 1

- ❑ Data duplicated, also the controller card
 - ❑ Requires only two drives to implement
 - ❑ Advantages
 - Random Read performance=better than single disk
 - Sequential Read Performance=fair, just as single disks
 - Sequential Write Performance=good
 - Random Write Performance=same as single disks
 - Simple to Implement
 - ❑ Disadvantage - high check disk overhead!
-

RAID Level 2

RAID 2



A0 to A3=Word A; B0 to B3 = Word B;
C0 to C3=Word C; D0 to D3 = Word D

ECC/Ax to Az=Word A ECC; ECC/Bx to Bz = Word B ECC;
ECC/Cx to Cz=Word C ECC; ECC/Dx to Dz = Word D ECC

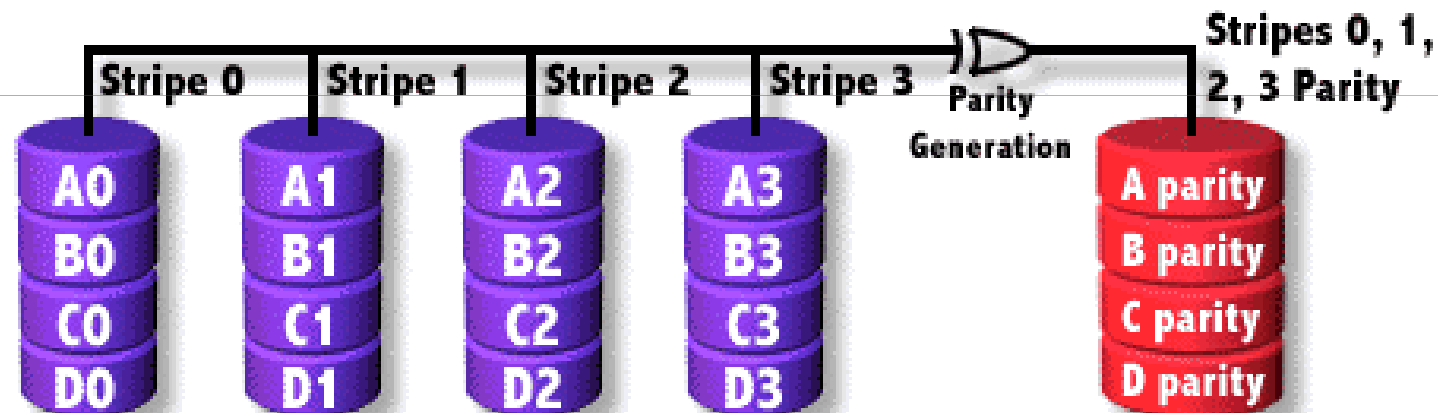
COPYRIGHT © 1996, 1997, 1998, 1999 ADVANCED COMPUTER & NETWORK CORPORATION

RAID Level 2

- ❑ Uses Bit-level striping with Hamming codes for ECC
- ❑ Number of disks required depends on exact implementation
- ❑ Only fair fault tolerance
- ❑ Advantages
 - Random Read performance=fair
 - Sequential Read Performance=very good
 - Sequential Write performance=good
- ❑ Disadvantages
 - Random Write Performance=poor
 - Requires a complex controller
 - High overhead for check disks
 - Not used in modern systems

RAID Level 3

RAID 3

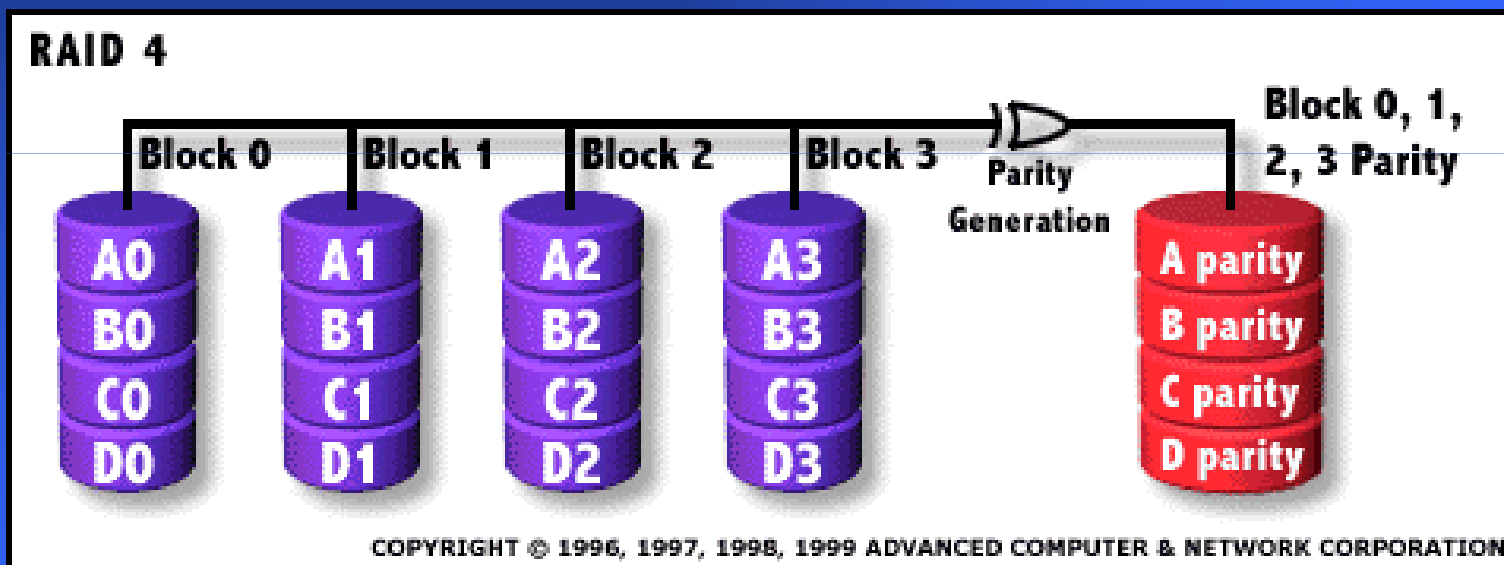


COPYRIGHT © 1996, 1997, 1998, 1999 ADVANCED COMPUTER & NETWORK CORPORATION

RAID Level 3

- ❑ Uses byte-level striping with dedicated parity
- ❑ Requires minimum three drives to implement
- ❑ Has good fault-tolerance
- ❑ Advantages
 - Random Read Performance=good
 - Sequential Read performance=very good
 - Sequential Write performance=fair to good
 - Lowest overhead for check disks
- ❑ Disadvantages
 - Random Write performance=poor
 - Complex controller design

RAID Level 4

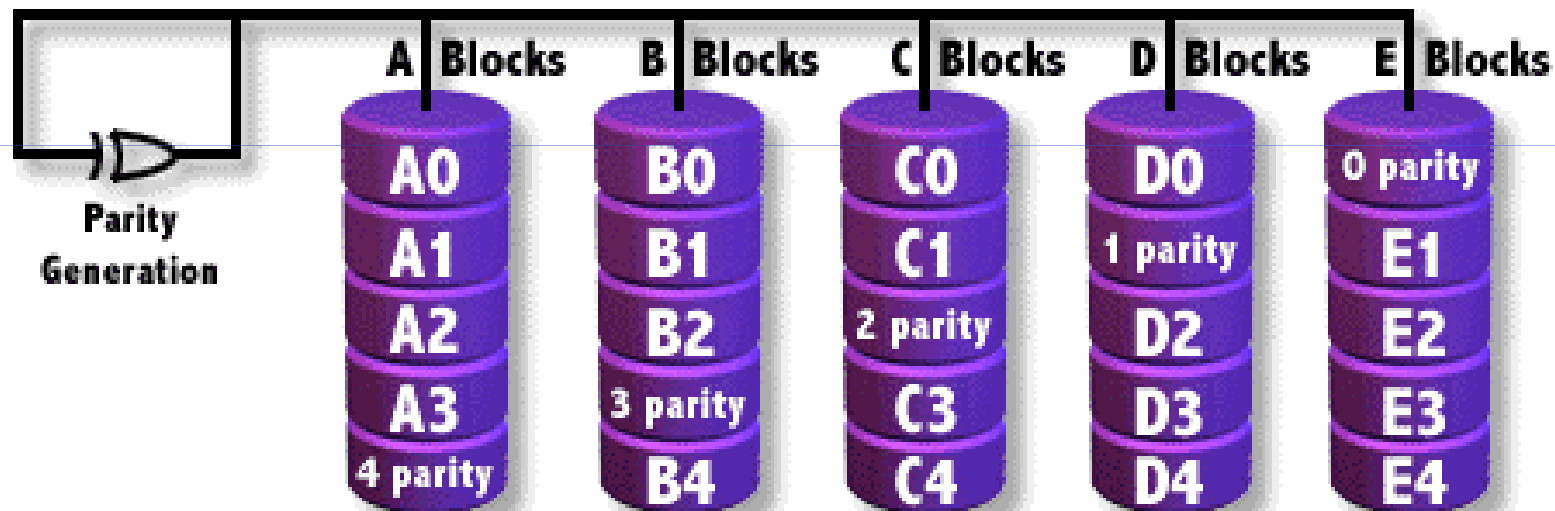


RAID Level 4

- ❑ Uses Block-level striping with dedicated parity
 - ❑ Requires minimum of 3 drives to implement
 - ❑ Has good fault-tolerance
 - ❑ Advantages
 - Random Read Performance=very good
 - Sequential Read and Write performance=good
 - Lowest overhead of check disks
 - ❑ Disadvantages
 - Quite complex controller design
 - Random write performance=poor
 - Not commonly used
-

RAID Level 5

RAID 5

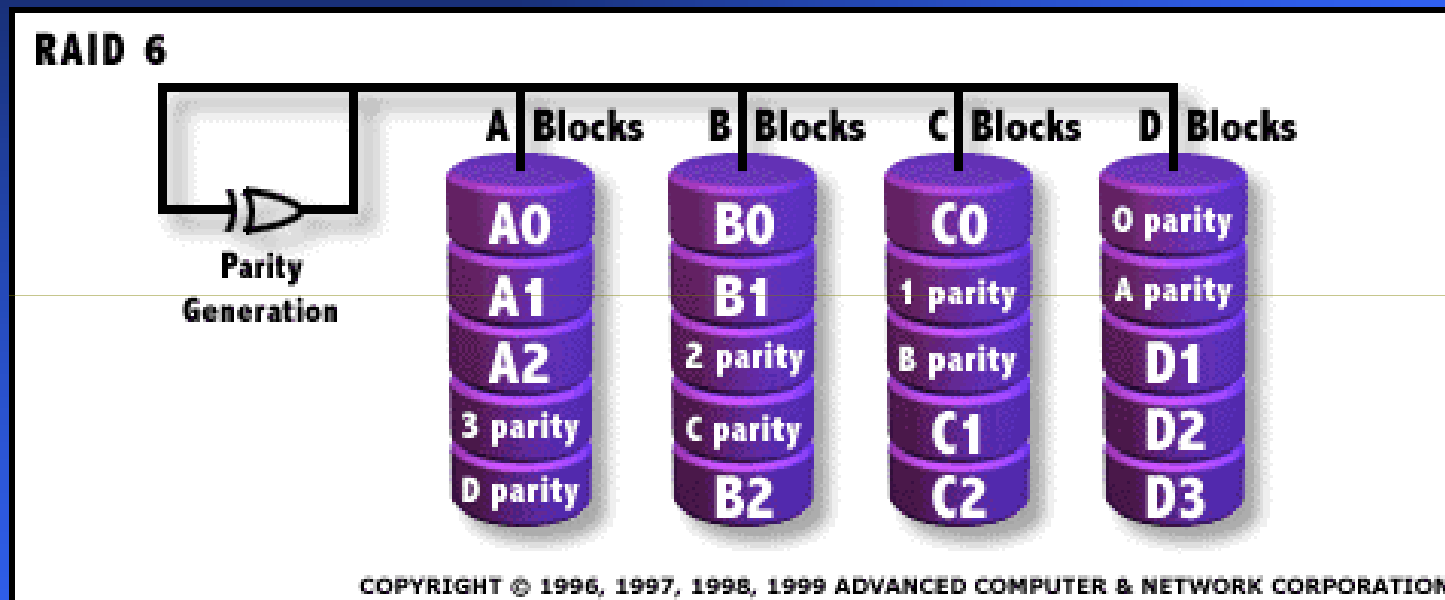


COPYRIGHT © 1996, 1997, 1998, 1999 ADVANCED COMPUTER & NETWORK CORPORATION

RAID Level 5

- ❑ Uses Block-level striping with distributed parity
 - ❑ Requires a minimum of 3 drives to implement
 - ❑ Advantages
 - Random Read performance=very good
 - Random Write performance=fair
 - Sequential Read and Write performance=good
 - Lowest overhead of check disks
 - ❑ Disadvantages
 - Most complex controller design
 - Difficult to rebuild in the event of a disk failure
-

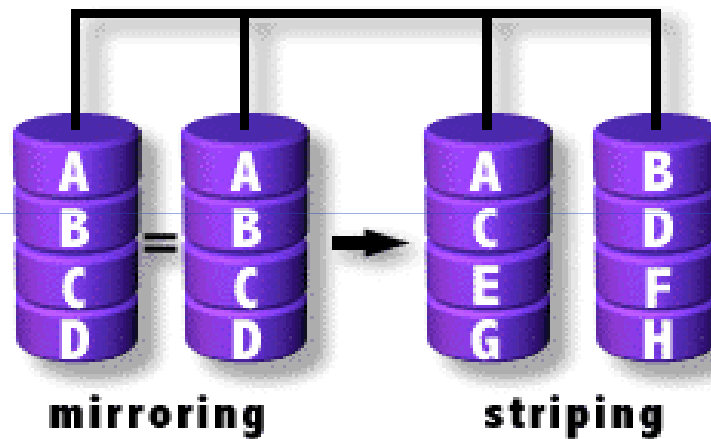
RAID Level 6



Raid Level 6 uses Block-level striping with *dual* distributed parity

RAID Level 8

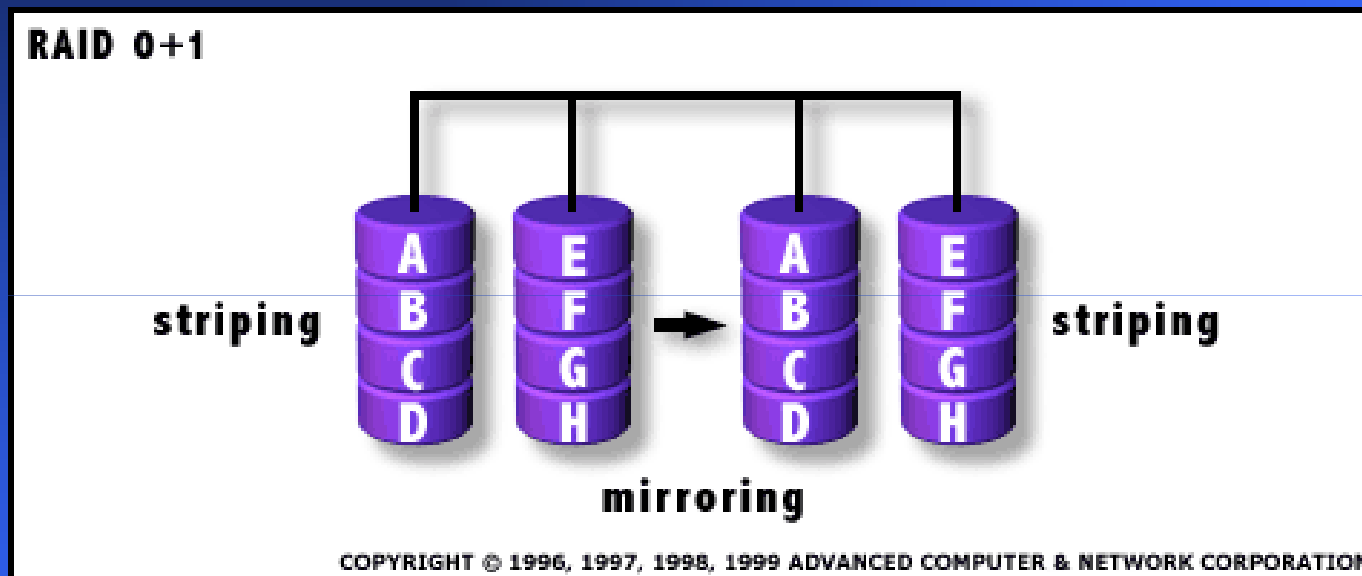
RAID 10



COPYRIGHT © 1996, 1997, 1998, 1999 ADVANCED COMPUTER & NETWORK CORPORATION

Uses Mirroring and striping without parity

RAID Level 0+1



Uses Mirroring and striping without Parity