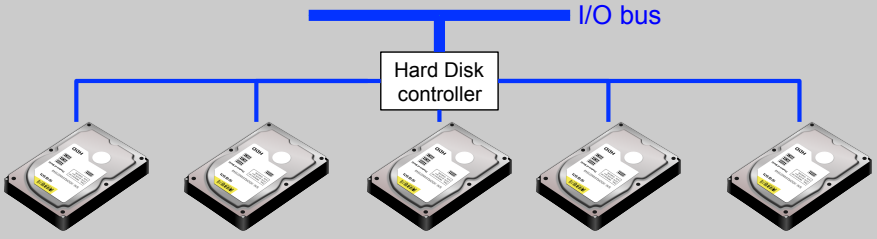
RAID

- Observations about Disk Performance
- RAID – Combining independent Disks
- RAID-0 to RAID-6
- Combining RAID Levels

RAID

Observation: Traditional secondary storage devices are slow!

Idea: Improve their performance by using multiple devices in parallel:



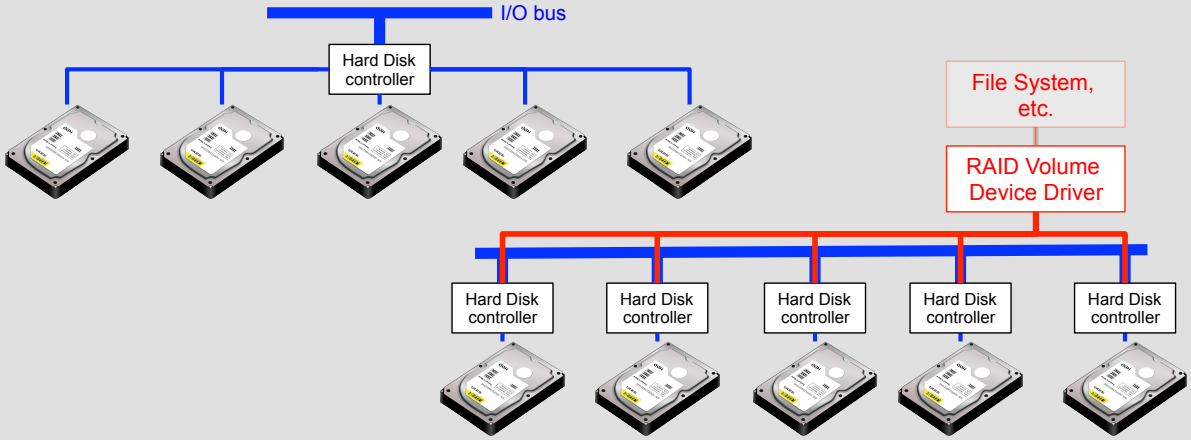
RAID

- **R**edundant **A**rrays of **I**ndependent **D**isks
- **R**edundant **A**rrays of **I**nexpensive **D**isks (Berkeley)

RAID

Common characteristics of RAID:

1. Array of physical disks that are visible as single device to OS.



RAID

Common characteristics of RAID:

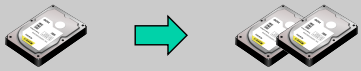
1. Array of physical disks that are visible as single device to OS.
2. Data is distributed across physical drives of array.
3. Redundant disk capacity is used for error detection/correction.

RAID (cont)

Approach: Replace single large-capacity disk with array of smaller-capacity disks.

Benefits:
- Improved I/O performance
- Enables incremental upgrade

Problem:
- **Reliability**: more devices increase the probability of failure.



failure rate doubles! → MTTF cut in half!

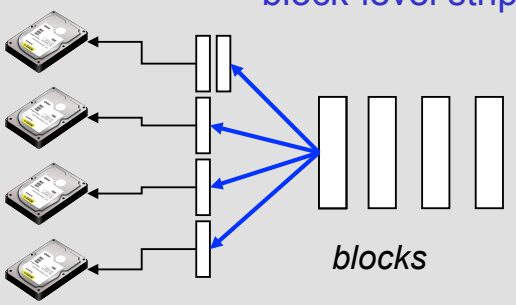
Solution: redundancy

RAID: Basic Principles

RAID = Striping + Redundancy

Striping: block-level, bit-level

block-level striping

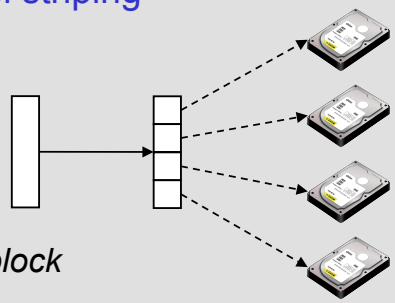


RAID: Basic Principles

RAID = Striping + Redundancy

- block-level
- bit-level

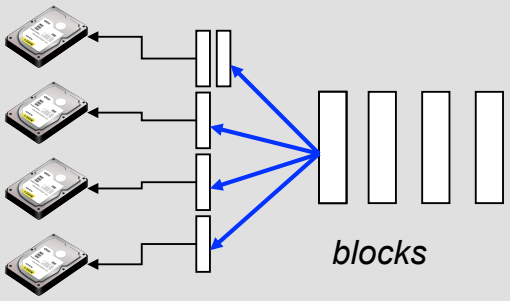
bit-level striping



RAID Level 0

"Block-level Striped Set without Parity"

RAID-0



Cons:
Tolerates no disk failures!

Pros:
- Increased Capacity
- Read/Write Performance

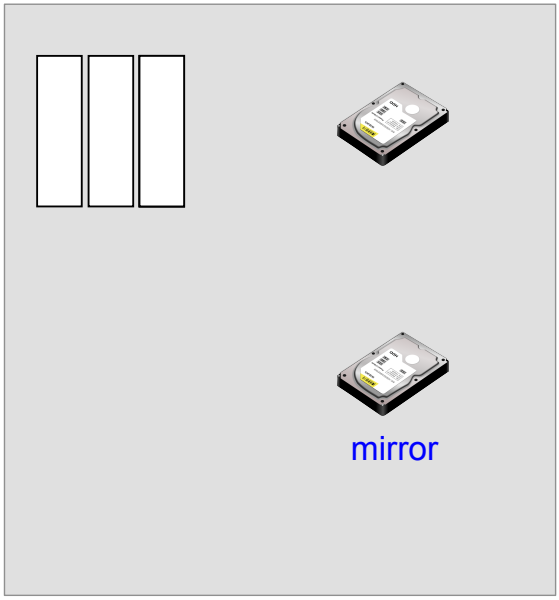
Placement:

Disk Id = B / N
Offset = B % N

where
B = Block number
N = Number of disks

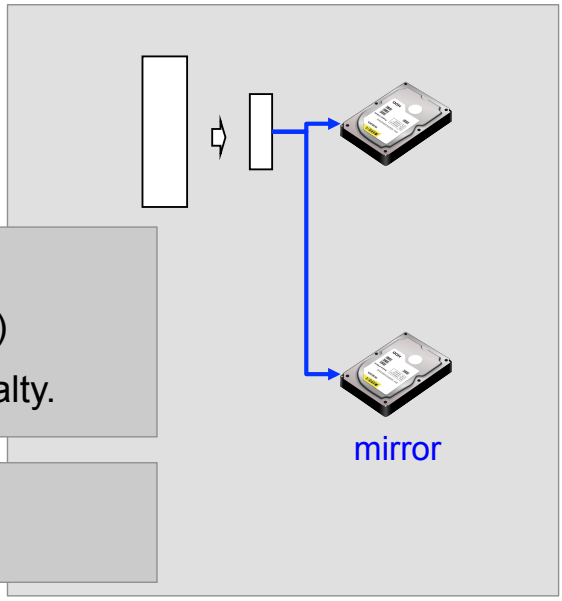
RAID Level 1

"Mirrored Set without Parity"



RAID Level 1

"Mirrored Set without Parity"



Performance:

- READs : good (with "split reads")
- WRITEs: small performance penalty.

Problem:

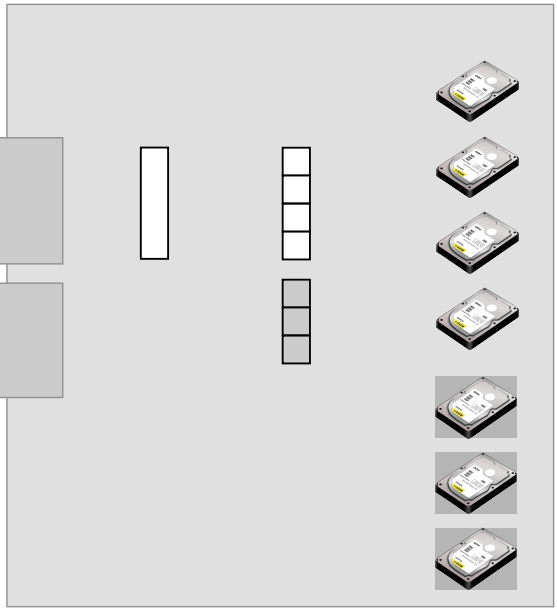
- cost (100% redundancy)

RAID Level 2

"Memory-Style Error-Correcting Parity"

- Small strips
- Head and spindles synchronized

- Error correction code calculated over bits of data disks. (Hamming Code)



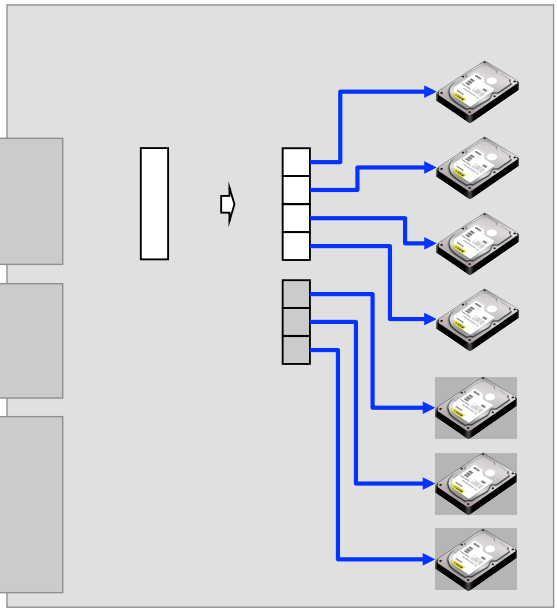
RAID Level 2

"Memory-Style Error-Correcting Parity"

- Small strips
- Head and spindles synchronized

- Error correction code calculated over bits of data disks. (Hamming Code)

- Appropriate for systems with many failures.
- Typically not implemented.

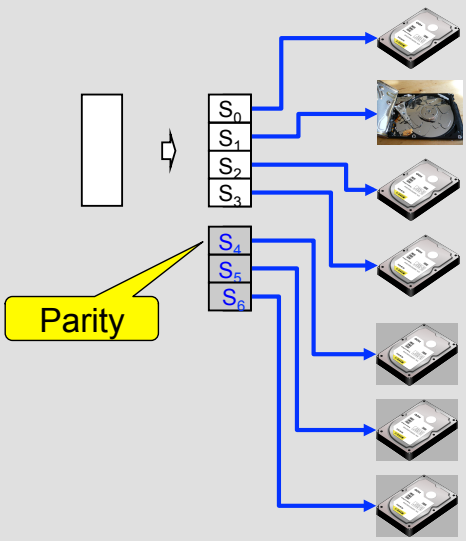


RAID Level 3

"Bit-Interleaved Parity"

- Small strips.
- Heads and spindles synchronized.
- Simple parity bits instead of ECC.

e.g.
$P(S) = S_4 = S_3$ xor S_2 xor S_1 xor S_0
Disk 1 fails:
$S_1 = S_4$ xor S_3 xor S_2 xor S_0

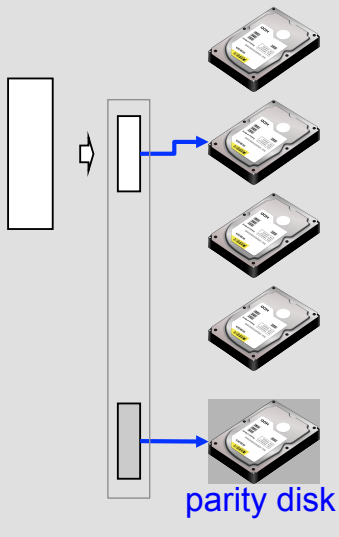


RAID Level 4

"Block level Parity"

- Same as RAID 3, but with block-level striping.
- Large strips.
- No synchronization across disks.

- Each block on parity disk contains parity information for all corresponding blocks.



RAID Level 4

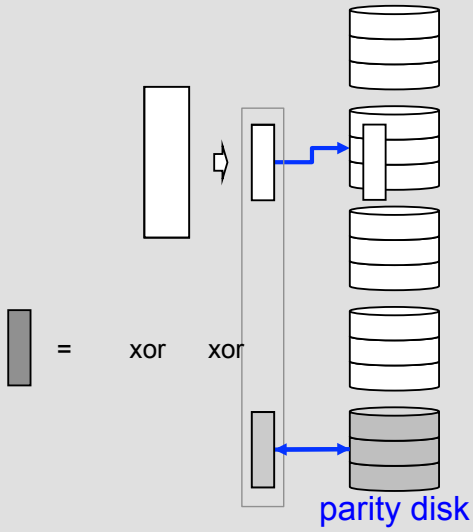
"Block-level Parity"

- Same as RAID 3, but with block-level striping.
- Large strips.
- No synchronization across disks.

- Each block on parity disk contains parity information for all corresponding blocks.

- Upon WRITE, update parity bits:

$X_4(i)' = X_4(i)$ xor $S_1(i)$ xor $S_1(i)'$



RAID Level 4

"Block-level Parity"

Note:

We do not have to read from other disks to update parity disk!

$X_4(i)' = X_4(i)$ xor $S_1(i)$ xor $S_1(i)'$

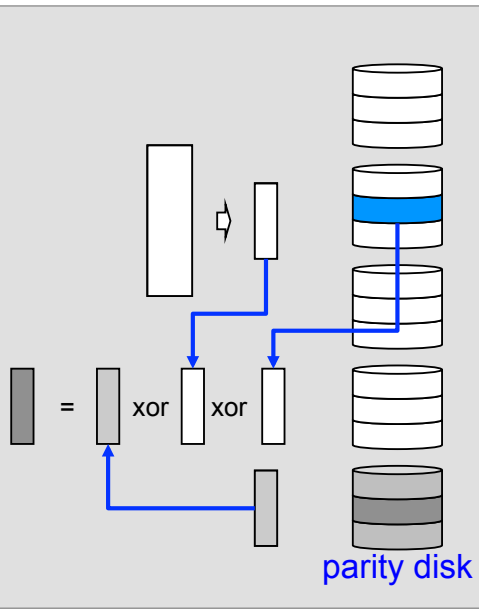
$X_4 = X_0$ xor X_1 xor X_2 xor X_3

$X_4' = X_0$ xor X_1' xor X_2 xor X_3

$X_4' = X_0$ xor X_1' xor X_2 xor X_3 xor X_1 xor X_1

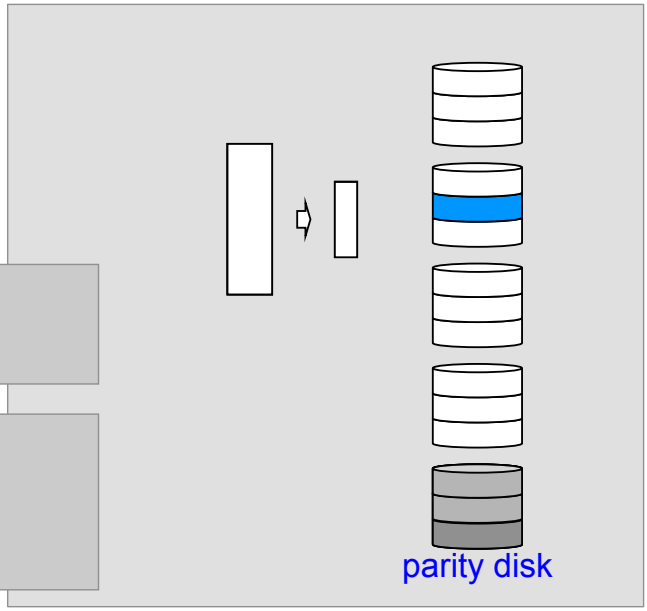
$X_4' = X_0$ xor X_1 xor X_2 xor X_3 xor X_1' xor X_1

$X_4' = X_4$ xor X_1 xor X_1'



RAID Level 4

"Block-level Parity"

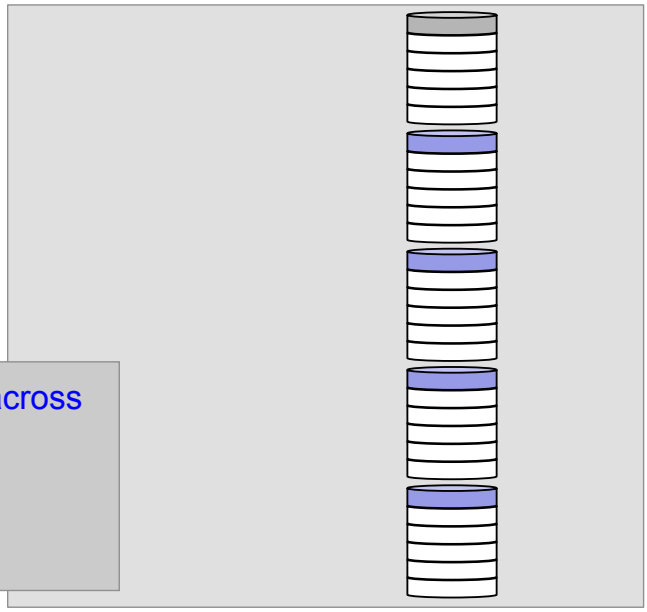


Problem:
- Parity disk gets lots of load!

Solution:
- Spread parity across disks!
- Raid 5

RAID Level 5

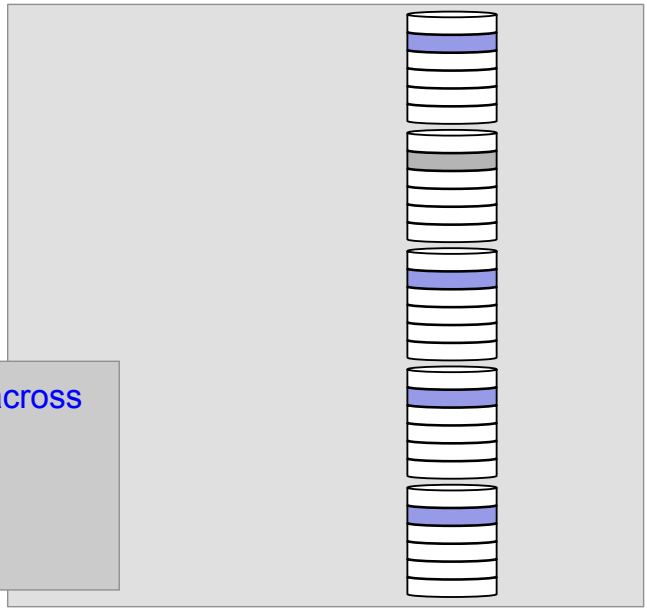
"Striped Set with Interleaved Parity"



- Same as RAID 4, but parity spread across all disks.
- No synchronization across disks.
- Large strips.

RAID Level 5

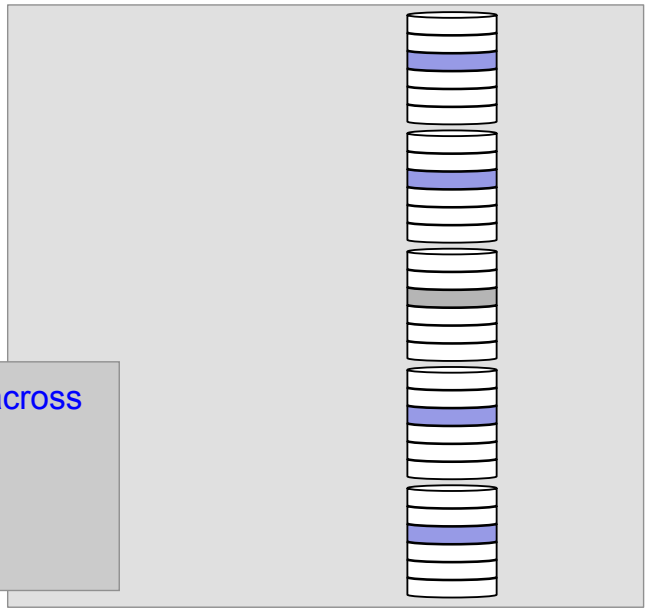
"Striped Set with Interleaved Parity"



- Same as RAID 4, but parity spread across all disks.
- No synchronization across disks.
- Large strips.

RAID Level 5

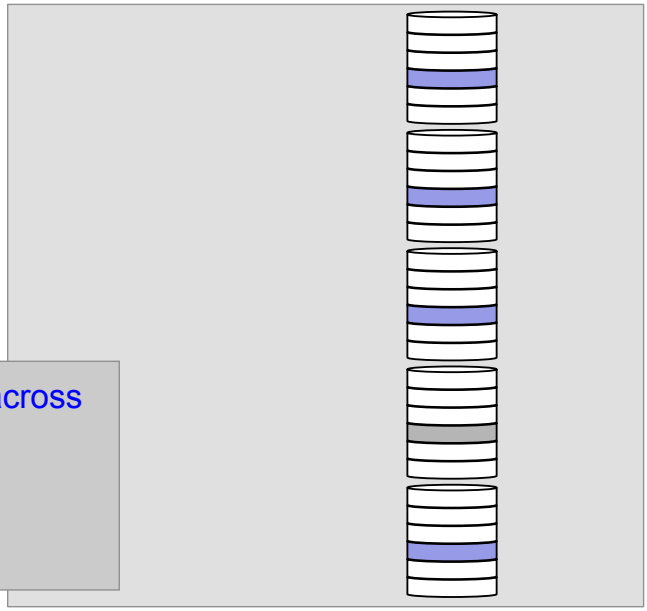
"Striped Set with Interleaved Parity"



- Same as RAID 4, but parity spread across all disks.
- No synchronization across disks.
- Large strips.

RAID Level 5

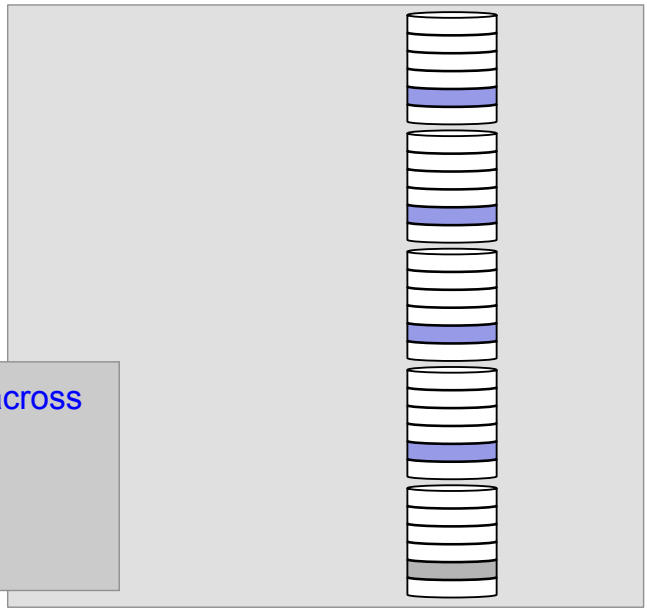
"Striped Set with Interleaved Parity"



- Same as RAID 4, but parity spread across all disks.
- No synchronization across disks.
- Large strips.

RAID Level 5

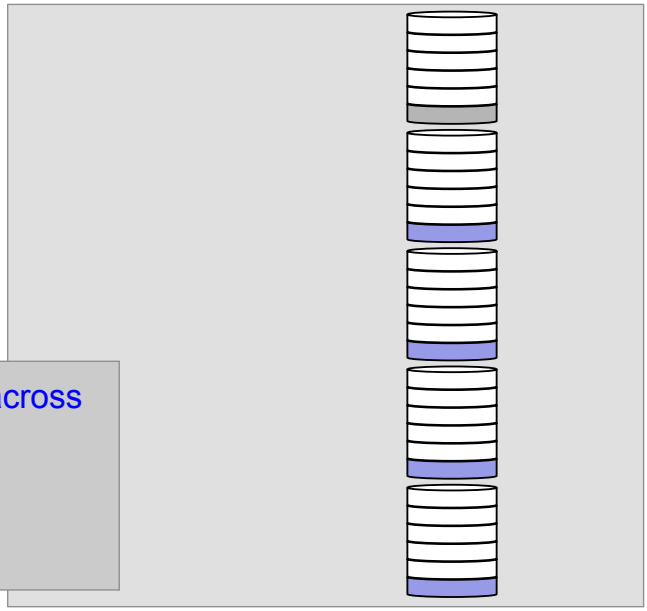
"Striped Set with Interleaved Parity"



- Same as RAID 4, but parity spread across all disks.
- No synchronization across disks.
- Large strips.

RAID Level 5

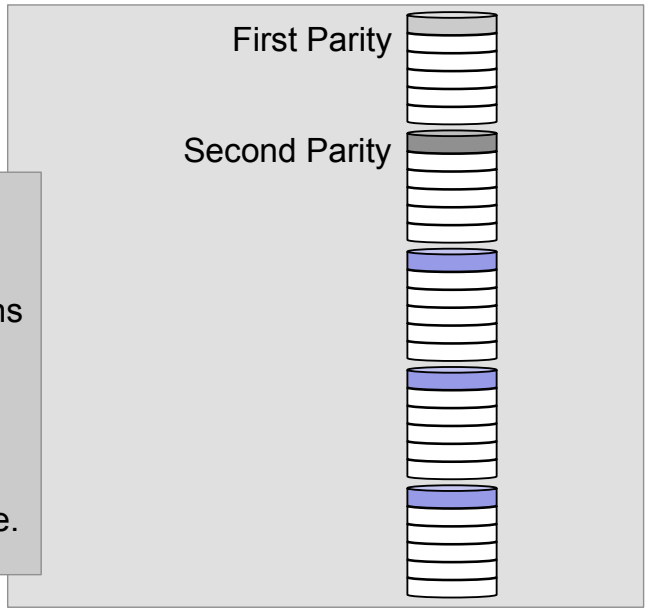
"Striped Set with Interleaved Parity"



- Same as RAID 4, but parity spread across all disks.
- No synchronization across disks.
- Large strips.

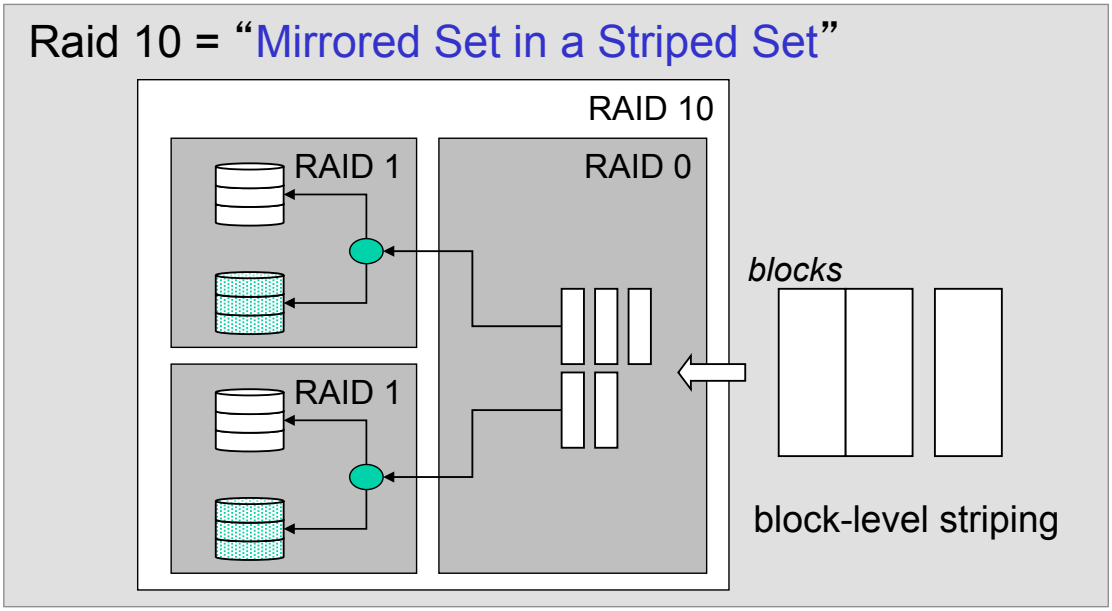
RAID Level 6

"Striped Set with Dual Interleaved Parity"



- Same as RAID 5, but uses 2 bits to store "parity".
- Uses two independent parity functions and writes to two disks
- Tolerates two failures.
- In practice, a second drive can fail during recovery from first drive failure.

Nested Levels: e.g. RAID Level 1+0 = RAID 10



RAID

- Observations about Disk Performance
- RAID – Combining independent
- RAID-0 to RAID-6
- Combining RAID Levels