RAID

to achieve increased performance, data redundancy, and fault tolerance.

- RAID (Redundant Array of Independent Disks) is standardized scheme for multiple-disk database design.
  - scheme → Method
  - multiple → multiple hard disks
- The RAID scheme consists of seven levels, zero through six.
- These levels share three common characteristics:

(A) ➢ RAID is a set of physical disk drives viewed by the operating system as a single logical drive.

(B) ➢ Data are distributed across the physical drives of an array in a scheme known as striping.

(C) ➢ Redundant disk capacity is used to store parity information, which guarantees data recoverability in case of a disk failure.

drive → disk

RAID level 0

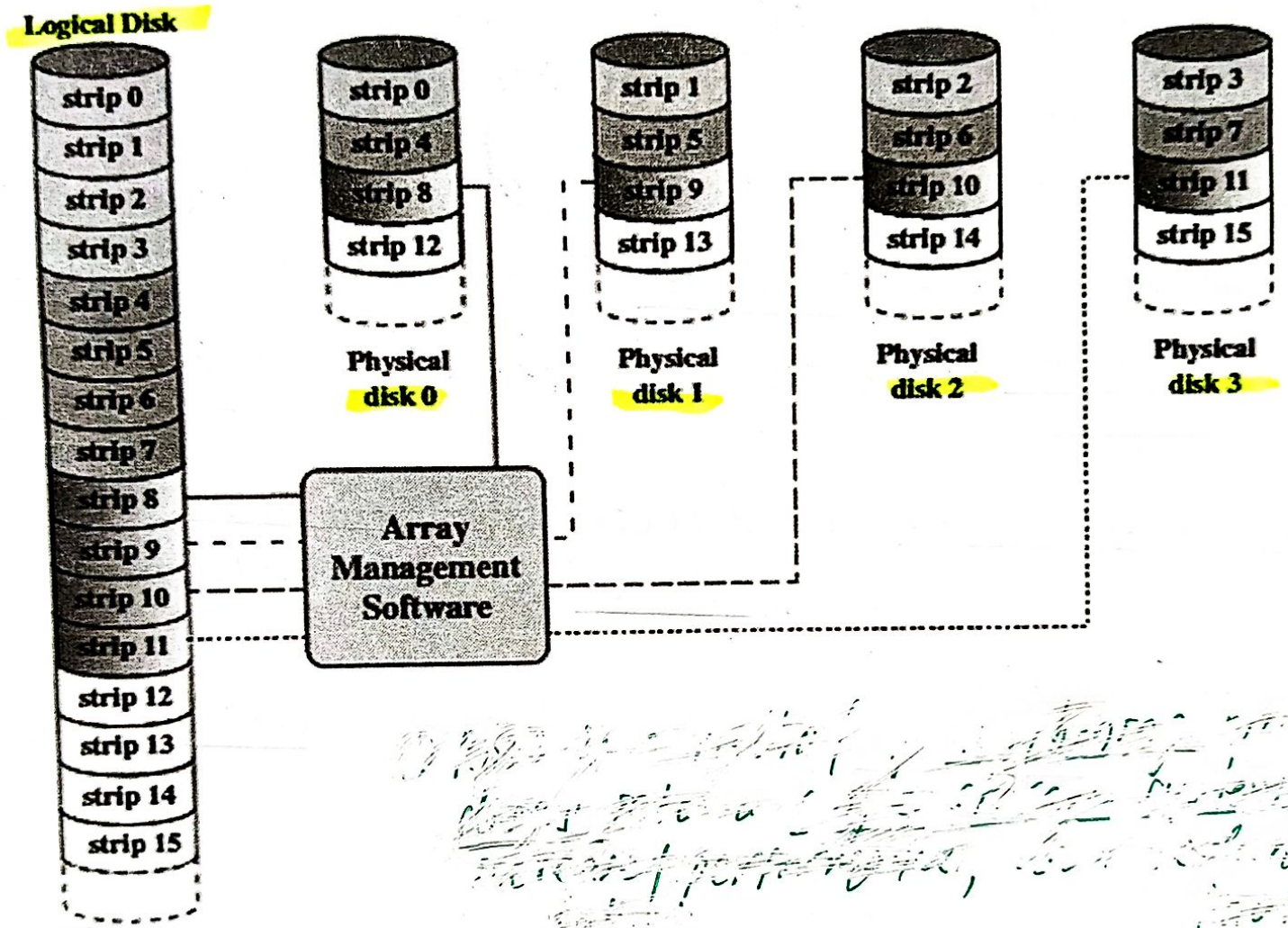


Figure 6.7 Data Mapping for a RAID Level 0 Array

(1) In RAID level 0, the data are striped across the available disks.

(2) RAID level 0 does not include redundancy to improve performance.

(3) suppose RAID 0 array consisting of 4 disks. One-forth of the data would reside on each of the 4 disks. The advantage of RAID 0 is that, because 4 disks are used, data can be written & read 4-times more quickly than with a single disk.

(4) The primary disadvantage of RAID 0 that, if any one of the disks is failed, then entire stripes in that disk will fail, because, it is not included with redundancy.

(4) RAID level 1

① On RAID level 1, data striping is used as in RAID level 0.

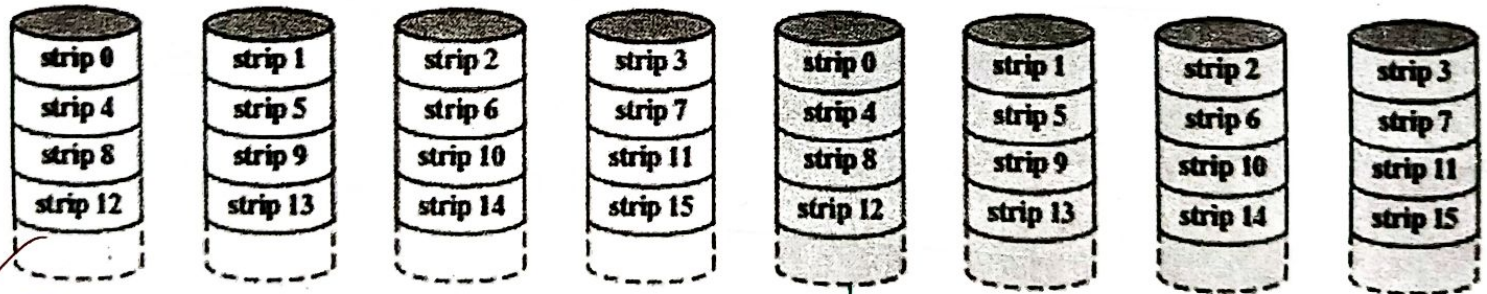
② On RAID level 1, every disk has a mirror disk that contains the same data to achieve redundancy. When a drive fails, the data may still be accessed from the 2nd drive.

③ All write operations are directed to both disks, so that the two disks are always in sync with one another.

④ All read operations are directed from either of the two disks.

RAID level 1:

- RAID 1 differs from RAID levels 2 through 6 in the way in which redundancy is achieved.
- In RAID 1, redundancy is achieved by the simple expedient of duplicating all the data.
- Data striping is used, as in RAID 0, but in this case, each logical strip is mapped to two separate physical disks so that every disk in the array has a mirror disk that contains the same data.
- RAID 1 can also be implemented without data striping, though this is less common.



(b) RAID 1 (Mirrored)

Mirror of disk 0

(6) RAID level 2 →

(1) In RAID level 2, data striping is used as level 1 & 0.

(2) RAID level 2 make use of a parallel access technique.

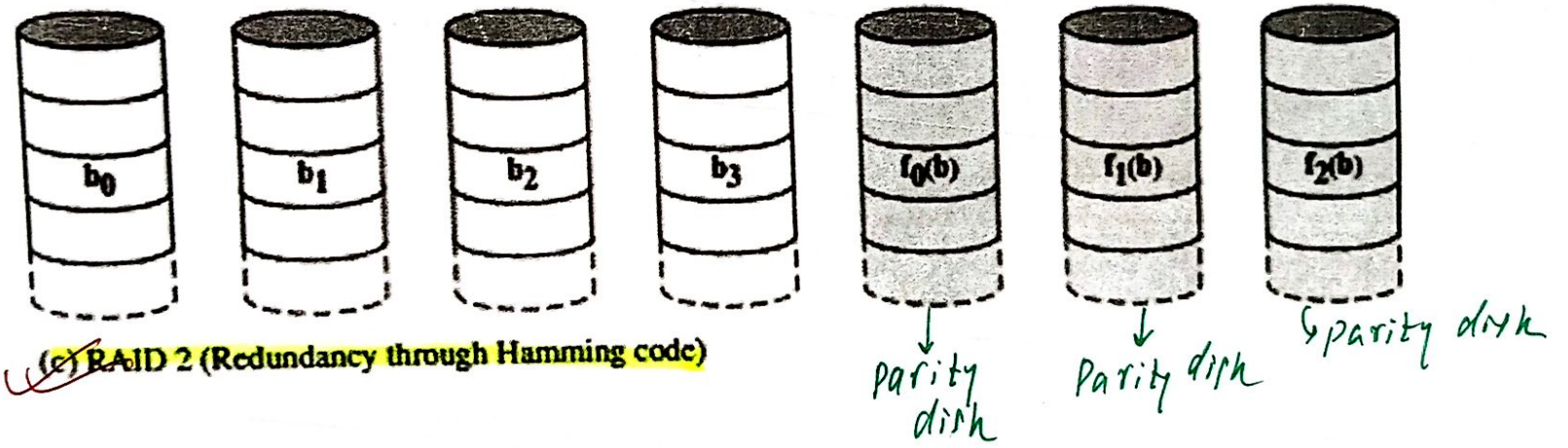
(3) In RAID level 2, every disk has a redundant disk that contain the error-correcting codes for corresponding bits on each data disk to achieve redundancy.

(4) It uses a synchronized spinning of the disk drives, which means that each disk is in sync with the others, so they all read & write simultaneously.

(5) RAID 2 employs Hamming code for error correction, which means it can correct single-bit errors & detect two-bit errors.

RAID level 2:

- RAID levels 2 and 3 make use of a parallel access technique.
- As in other RAID schemes, data striping is used.
- With RAID 2, an error-correcting code is calculated across corresponding bits on each data disk, and the bits of the code are stored in the corresponding bit position on multiple parity disks.



(c) RAID 2 (Redundancy through Hamming code)

(8) RAID level 3 →

(1) In RAID level 3, data striping is used as level 0, 1, 2 but it is very small.

(2) RAID level 3 make use of a parallel access technique.

(3) In RAID level 3, a single redundant disk is used that contain bit-interleaved parity for all physical disks.

(4) This redundant disk contain the bit-interleaved Parity of the each data bit of the physical disk.

(5) In the event of a drive failure, the redundant disk accessed, & data is reconstructed from the remaining drives (disks)

Performance →

(1) RAID level 3 can achieve very high data transfer rates, because of data striping is very small.

(2) Any I/O request will involve the parallel transfer of data from all of the data disks.

(9)

Data Reconstruction (RAID level 3):

Consider an array of 5-drives, X0 to X3 contain data & X4 is parity disk.

The parity for the ith bit is calculated by:

$$X4(i) = X3(i) \oplus X2(i) \oplus X1(i) \oplus X0(i)$$

- $X4(i)$ → Parity bit of ith data bit
- $X3(i)$ → ith data bit in X3 drive
- $X2(i)$ → ith data bit in X2 drive
- $X1(i)$ → ith data bit in X1 drive
- $X0(i)$ → ith data bit in X0 drive

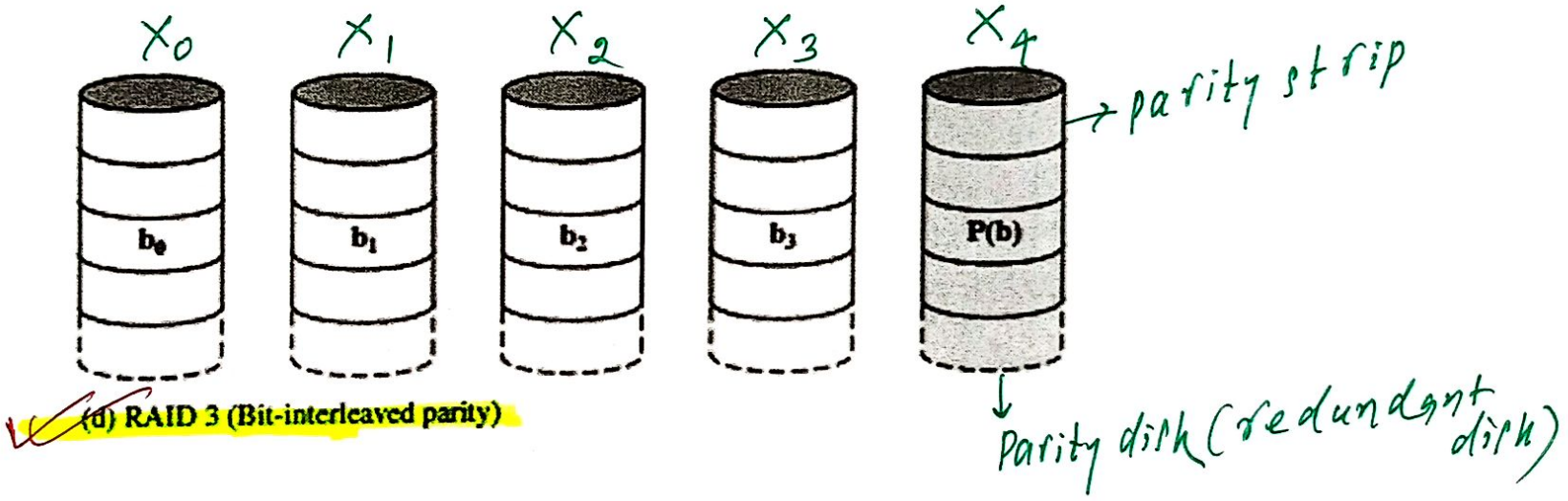
Suppose drive X1 has failed, then data of X1 drive can be reconstructed as follows:

$$X1(i) = X4(i) \oplus X3(i) \oplus X2(i) \oplus X0(i)$$

- $X1(i)$ → ith data bit in X1 drive

RAID level 3:

- RAID 3 is organized in a similar fashion to RAID 2.
- The difference is that RAID 3 requires only a single redundant disk, no matter how large the disk array.
- RAID 3 employs parallel access, with data distributed in small strips.



(d) RAID 3 (Bit-interleaved parity)

- In the event of a drive failure, the parity drive is accessed and data is reconstructed from the remaining devices.

(11) RAID level 4 →

(1) In RAID level 4, data striping is used, but strips are relatively large.

(2) RAID 4 make use of an independent access technique.

(3) In RAID level 4, a single redundant disk is used that contain block-level parity of all the physical disks.

* Updation of parity bit on redundant disk after writing of ith bit on any disk

consider an array of 5-drives in which X0 to X3 are data & X4 is parity disk.

suppose that a write is performed that only involves a strip on disk X1

Initially, for each bit i, the following relation:

$$X4(i) = X3(i) \oplus X2(i) \oplus X1(i) \oplus X0(i)$$

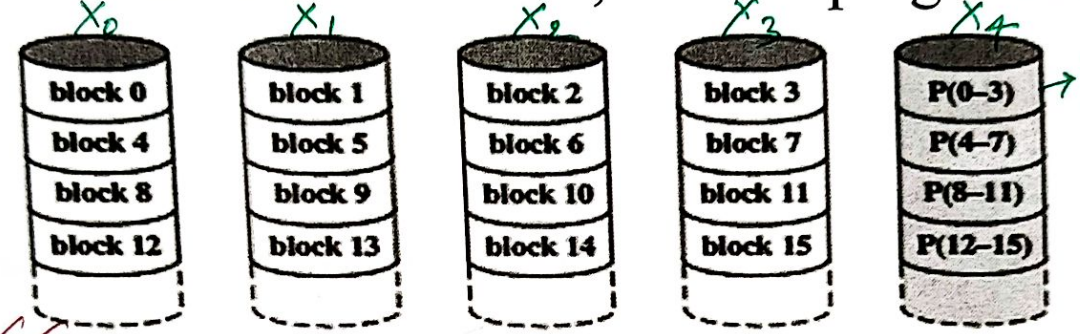
After the update, new parity bit:

$$X4'(i) = X4(i) \oplus X1(i) \oplus X1'(i)$$

- $X4'(i)$ ↓ new parity ith-bit
- $X4(i)$ ↓ old parity-ith bit
- $X1(i)$ ↓ old ith-data bit
- $X1'(i)$ ↓ new ith-data bit

RAID level 4:

- RAID levels 4 through 6 make use of an independent access technique.
- As in other RAID schemes, data striping is used.

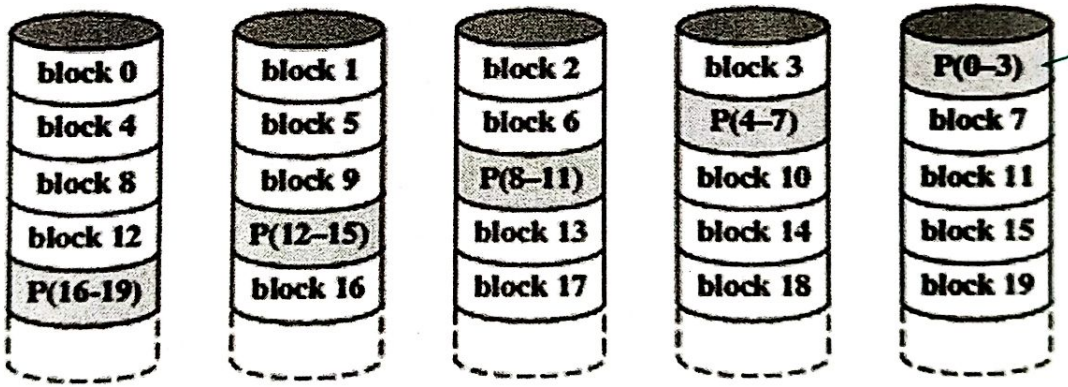


(c) RAID 4 (Block-level parity)

- In the case of RAID 4 through 6, the strips are relatively large.
- RAID 4 involves a write penalty when an I/O write request of small size is performed.

RAID level 5:

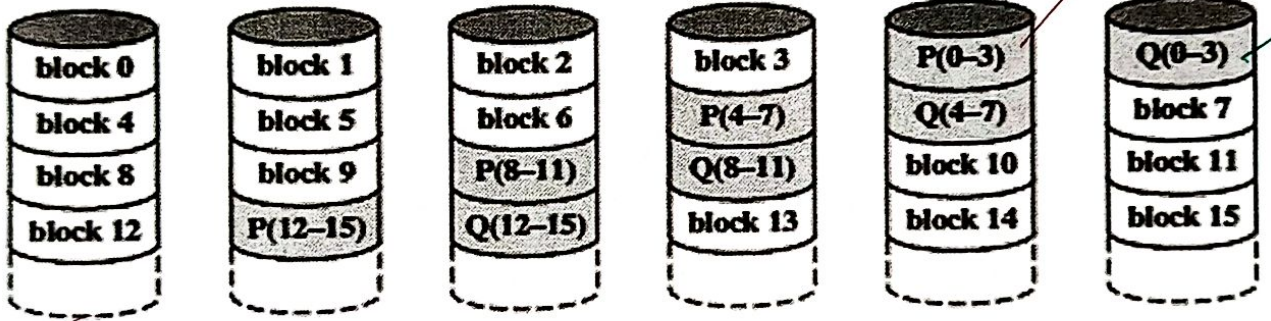
- RAID 5 is organized in a similar fashion to RAID 4.
- The difference is that RAID 5 distributes the parity strips across all disks.
- A typical allocation is a round-robin scheme.
- For an n-disk array, the parity strip is on a different disk for the first n stripes and the pattern then repeats.



(f) RAID 5 (Block-level distributed parity)

RAID level 6:

In RAID 6 scheme, two different parity calculations are carried out and stored in separate blocks on different disks. Thus, a RAID 6 array whose user data require N disks consists of N+2 disks.



(g) RAID 6 (Dual redundancy)

The advantage of RAID 6 is that it provides extremely high data availability.

- Three disks would have to fail within the MTTR (mean time to repair) interval to cause data to be lost.