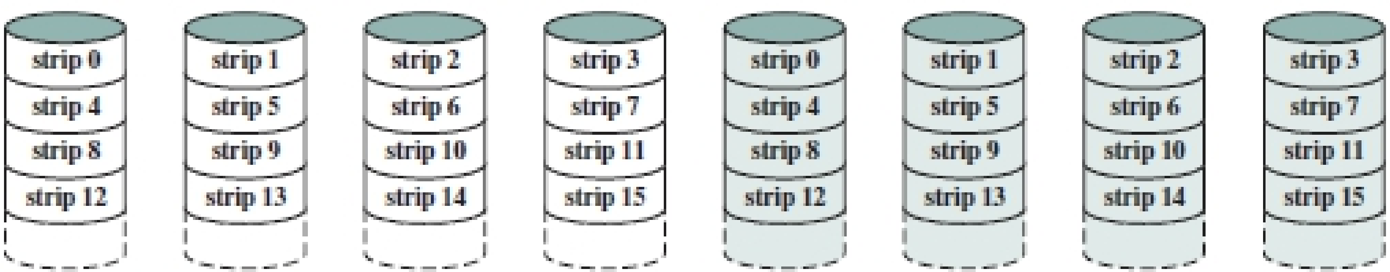
Computer Organization and Architecture (EET 2211)

Lecture 23

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)

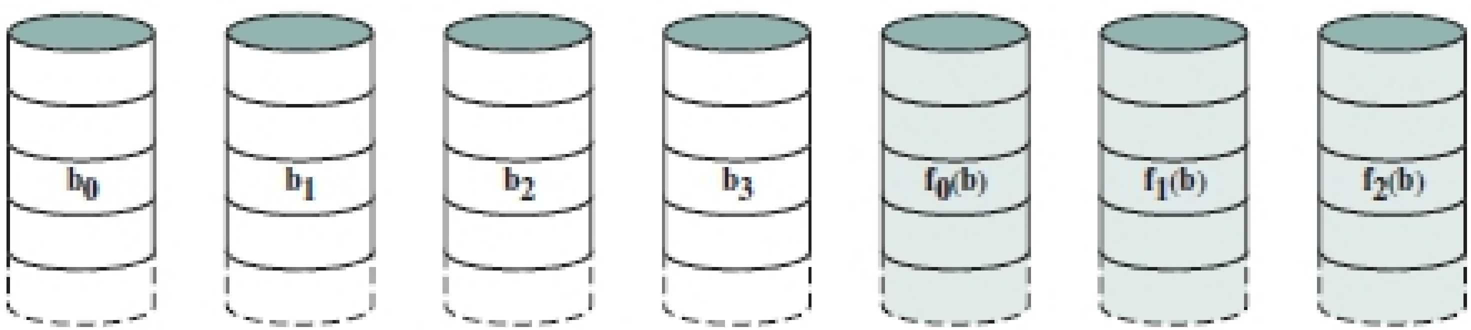
There are a number of positive aspects to the RAID 1 organization:

- A read request can be serviced by either of the two disks that contains the requested data, whichever one involves the minimum seek time plus rotational latency.
- A write request requires that both corresponding strips be updated, but this can be done in parallel. Thus, the write performance is dictated by the slower of the two writes.
- Recovery from a failure is simple. When a drive fails, the data may still be accessed from the second drive.

- The principal disadvantage of RAID 1 is the cost.
- In a transaction-oriented environment, RAID 1 can achieve high I/O request rates if the bulk of the requests are reads.
- If a substantial fraction of the I/O requests are write requests, then there may be no significant performance gain over RAID 0.
- RAID 1 may also provide improved performance over RAID 0 for data transfer intensive applications with a high percentage of reads.
- Improvement occurs if the application can spilt each read request so that both disk members participate.

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.

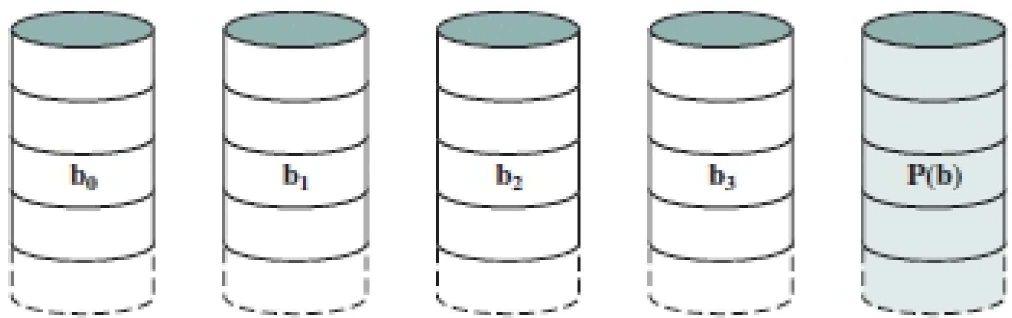


(c) RAID 2 (Redundancy through Hamming code)

- Although RAID 2 requires fewer disks than RAID 1, it is still rather costly.
- The number of redundant disks is proportional to the log of the number of data disks.
- On a single read, all disks are simultaneously accessed.
- On a single write, all data disks and parity disks must be accessed for the write operation.

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)

Redundancy:

- In the event of a drive failure, the parity drive is accessed [redacted]ructed from the remaining devices.

Data reconstruction:

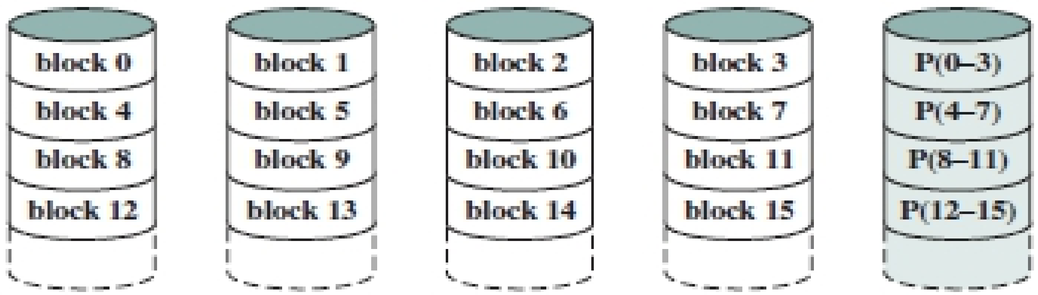
- Consider an array of five drives in which X0 through X3 contain data and X4 is the parity disk.
- The parity for the *i*th bit is calculated as follows:
- $X4(i) = X3(i) \oplus X2(i) \oplus X1(i) \oplus X0(i)$
- where $\oplus$ is exclusive-OR function.
- Suppose that drive X1 has failed.
- If we add $X4(i) \oplus X1(i)$ to both sides of the preceding equation, we get
- $X1(i) = X4(i) \oplus X3(i) \oplus X2(i) \oplus X0(i)$
- In the event of a disk failure, all of the data are still available in what is referred to as *reduced mode*.

Performance:

- Because data are striped in very small strips, RAID 3 can achieve very high data transfer rates.
- Any I/O request will involve the parallel transfer of data from all of the data disks.
- Only one I/O request can be executed at a time.

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.
- Consider an array of five drives in which X0 through X3 contain data and X4 is the parity disk.
- Suppose that a write is performed that only involves a strip on [redacted]

- Suppose that a write is performed that only involves a strip on disk X1.
- Initially, for each bit *i*, we have the following relationship:

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

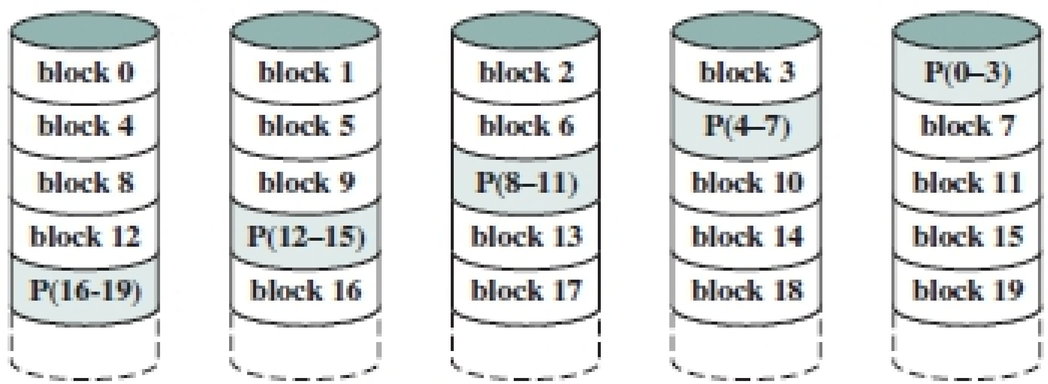
- After the update, with potentially altered bits indicated by a prime symbol:

$$\begin{aligned} X4'(i) &= X3(i) \oplus X2(i) \oplus X1'(i) \oplus X0(i) \\ &= X3(i) \oplus X2(i) \oplus X1'(i) \oplus X0(i) \oplus X1(i) \oplus X1(i) \\ &= X3(i) \oplus X2(i) \oplus X1(i) \oplus X0(i) \oplus X1(i) \oplus X1'(i) \\ &= X4(i) \oplus X1(i) \oplus X1'(i) \end{aligned}$$

- To calculate the new parity, the array management software must read the old user strip and the old parity strip.
- Then it can update these two strips with the new data and the newly calculated parity.
- [redacted] write operation must involve the parity [redacted]ore can become a bottleneck.

RAID Level 5

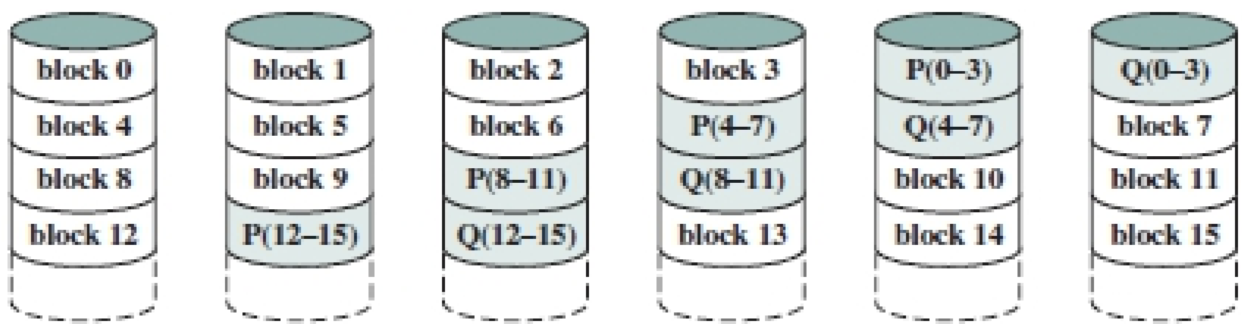
- RAID 5 is organized in a similar fashion to RAID 4.
- The difference is that RAID 5 distributes the parity strips across all disks.
- A typical allocation is a round-robin scheme.
- For an n-disk array, the parity strip is on a different disk for the first n stripes and the pattern then repeats.
-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



(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.

- On the other hand, RAID 6 incurs a substantial write penalty, because each write affects two parity blocks.
- Performance benchmarks [EISC07] show a RAID 6 controller can suffer more than a 30% drop in overall write performance compared with a RAID 5 implementation.
- RAID 5 and RAID 6 read performance is comparable.

Table 6.4 RAID Comparison

Level	Advantages	Disadvantages	Applications
0	I/O performance is greatly improved by spreading the I/O load across many channels and drives No parity calculation overhead is involved Very simple design Easy to implement	The failure of just one drive will result in all data in an array being lost	Video production and editing Image Editing Pre-press applications Any application requiring high bandwidth
1	100% redundancy of data means no rebuild is necessary in case of a disk failure, just a copy to the replacement disk Under certain circumstances, RAID 1 can sustain multiple simultaneous drive failures Simplest RAID storage subsystem design	Highest disk overhead of all RAID types (100%)—inefficient	Accounting Payroll Financial Any application requiring very high availability
2	Extremely high data transfer rates possible The higher the data transfer rate required, the better the ratio of data disks to ECC disks Relatively simple controller design compared to RAID levels 3, 4, & 5	Very high ratio of ECC disks to data disks with smaller word sizes—inefficient Entry level cost very high—requires very high transfer rate requirement to justify	No commercial implementations exist/not commercially viable
3	Very high read data transfer rate Very high write data transfer rate Disk failure has an insignificant impact on throughput Low ratio of ECC (parity) disks to data disks means high efficiency	Transaction rate equal to that of a single disk drive at best (if spindles are synchronized) Controller design is fairly complex	Video production and live streaming Image editing Video editing Prepress applications Any application requiring high throughput
4	Very high Read data transaction rate Low ratio of ECC (parity) disks to data disks means high efficiency	Quite complex controller design Worst write transaction rate and Write aggregate transfer rate Difficult and inefficient data rebuild in the event of disk failure	No commercial implementations exist/not commercially viable
5	Highest Read data transaction rate Low ratio of ECC (parity) disks to data disks means high efficiency Good aggregate transfer rate	Most complex controller design Difficult to rebuild in the event of a disk failure (as compared to RAID level 1)	File and application servers Database servers Web, e-mail, and news servers Intranet servers Most versatile RAID level
6	Provides for an extremely high data fault tolerance and can sustain multiple simultaneous drive failures	More complex controller design Controller overhead to compute parity addresses is extremely high	Perfect solution for mission critical applications