Sample Questions on Storage, RAID and Indexing (Solution)

Q1:

(a) RAID 5 Storage of 9 Blocks on 4 Disks

Disk 1	Disk 2	Disk 3	Disk 4
B1	B2	B3	P0
B4	B5	P1	B6
B7	P2	B8	B9

(b) RAID 1 Storage Requirement

- **Required effective storage** = 20TB
- **Each disk** = 4TB
- **RAID 1 stores a mirrored copy**, so effective storage = (Total Disks) / 2

Let **N** be the total disks required:

$N/2 \times 4TB = 20TB$
$N = (20TB / 4TB) \times 2 = 10$ Disks (Ans.)

(c) RAID Storage of 10 Blocks (Using 4TB Disks)

Disk 1	Disk 2	Disk 3	Disk 4
B1	B2	B3	P0
B4	B5	P1	B6
B7	P2	B8	B9
P3	B10	-	-

Q2:

(a) RAID 5 Storage Requirement

RAID 5 uses **(N-1) disks for storage** and **1 disk for parity**.

Effective Storage= $(N-1) \times 4TB$

Given **effective storage** = 20TB,

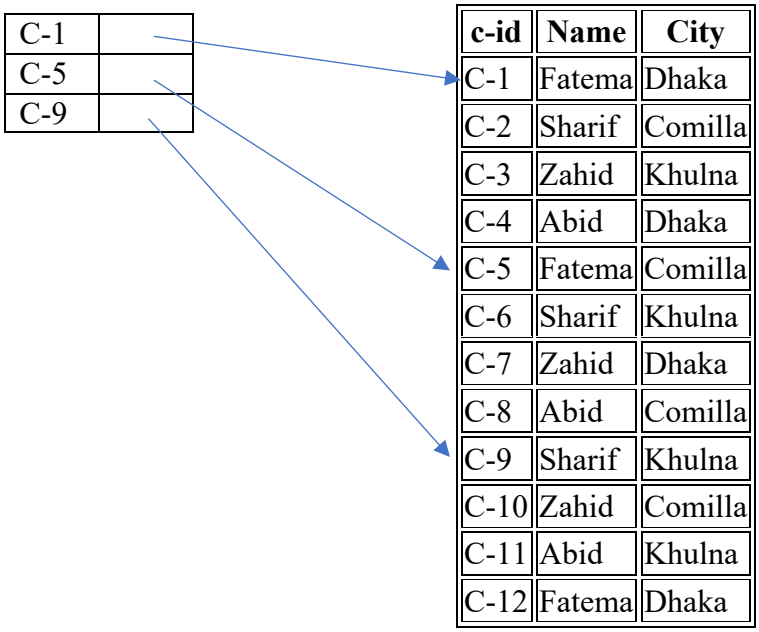
$(N-1) \times 4TB = 20TB$
$N - 1 = 20TB/4TB = 5$
$N = 6$ (Ans.)

(b) RAID 5 Storage of 10 Blocks (Using 4TB Disks)

Disk 1	Disk 2	Disk 3	Disk 4
B1	B2	B3	P0
B4	B5	P1	B6
B7	B8	P2	B9
P3	B10	-	-

Q3:

(a) Sparse Index Structure on c-id

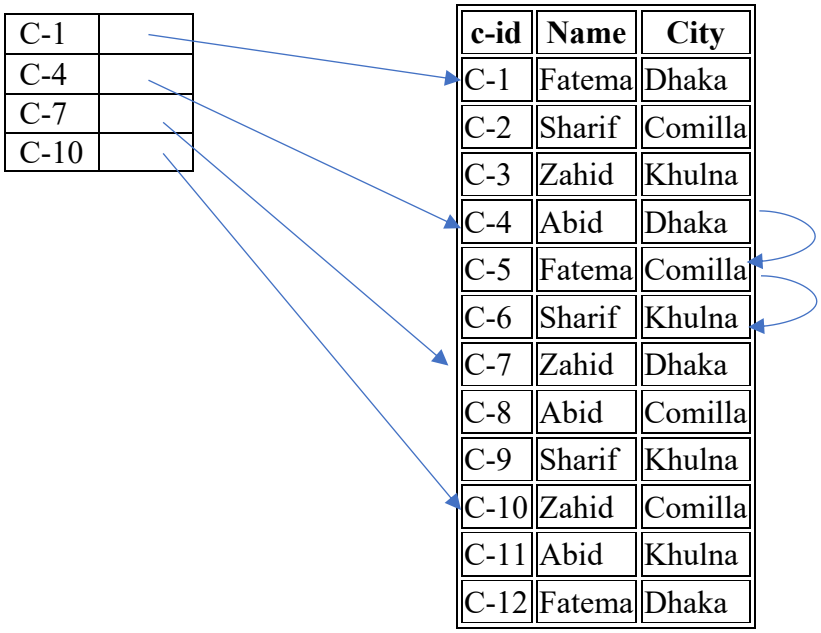


c-id	Name	City
C-1	Fatema	Dhaka
C-2	Sharif	Comilla
C-3	Zahid	Khulna
C-4	Abid	Dhaka
C-5	Fatema	Comilla
C-6	Sharif	Khulna
C-7	Zahid	Dhaka
C-8	Abid	Comilla
C-9	Sharif	Khulna
C-10	Zahid	Comilla
C-11	Abid	Khulna
C-12	Fatema	Dhaka

(b) Index for city

Indexing **City** is needed if queries frequently **search, filter, group, or sort** by it. If City has **many unique values (high cardinality)**, indexing improves performance. However, if City has **few unique values (low cardinality)** or is rarely used in queries, indexing is unnecessary.

(c)



(d)

SQL:

CREATE INDEX idx_customer_name ON Customer(Name);

Type of Index:

1. Single-column index → Built on a single attribute (`Name`).
2. Non-clustered index → It does not define the physical order of table records.
3. B-Tree index → Default indexing method in most databases.
4. Secondary index → Because it is not on the primary key (`c-id`).

Q4: RAID Level 1 vs. RAID Level 5

Feature	RAID 1 (Mirroring)	RAID 5 (Striping with Parity)
Data Storage	Each disk has an exact mirror copy.	Data is striped across disks with parity for redundancy.
Redundancy	High (100% redundancy).	Moderate (parity provides redundancy).
Fault Tolerance	Can tolerate failure of **half** the disks.	Can tolerate failure of **one** disk.

Feature	RAID 1 (Mirroring)	RAID 5 (Striping with Parity)
Performance (Read/Write)	**Read**: Fast (parallel reads).**Write**: Slower (writes occur on both disks).	**Read**: Fast (parallel reads).**Write**: Slower (parity calculation required).
Storage Efficiency	50% (half the total capacity is used for mirroring).	**(N-1)/N** efficiency (one disk used for parity).
Best Use Case	High-reliability systems requiring fast reads (e.g., **databases, critical servers**).	Cost-effective **storage with redundancy** (e.g., **file servers, backup storage**).

Example:

- **RAID 1:** If you use two 2TB disks, you only get **2TB usable storage** since data is mirrored.
- **RAID 5:** If you use four 2TB disks, you get **(4-1) × 2TB = 6TB usable storage**.

Q5: Seek Time and Block Transfer Time

(a) Seek Time

The **seek time** is the time taken for the disk's read/write head to move to the desired track.

- **Best-Case Seek Time:** When the read/write head is already at the required track (seek time = 0 ms).
- **Worst-Case Seek Time:** When the read/write head has to move across the entire disk from the outermost to the innermost track.

(b) Block Transfer Time

The **block transfer time** is the time taken to read/write a block once the head is positioned.

- **Best-Case Block Transfer Time:** When the requested block is **immediately** available under the read/write head.
- **Worst-Case Block Transfer Time:** When the disk must rotate a full revolution before the requested block reaches the head.

Q6: Reliable 10TB Storage with Frequent Updates

Best Choice: RAID 1+0 (RAID 10)

- RAID 10 combines **mirroring (RAID 1) and striping (RAID 0)**.
- Provides **high reliability** and **fast writes** (since updates occur in parallel).
- Requires **2 × the required storage** (because of mirroring).

Total Disks Required:

(10TB/2TB) × 2 =10 disks

Storage Layout:

Disk 1	Disk 2	Disk 3	Disk 4	Disk 5
B1	B1'	B2	B2'	B3
B3'	B4	B4'	B5	B5'

Why RAID 10?

- **Reliability:** Even if **half the disks** fail, data remains intact.
- **Frequent Updates:** RAID 10 provides fast write speeds.
- **Redundancy:** Ensures data protection.