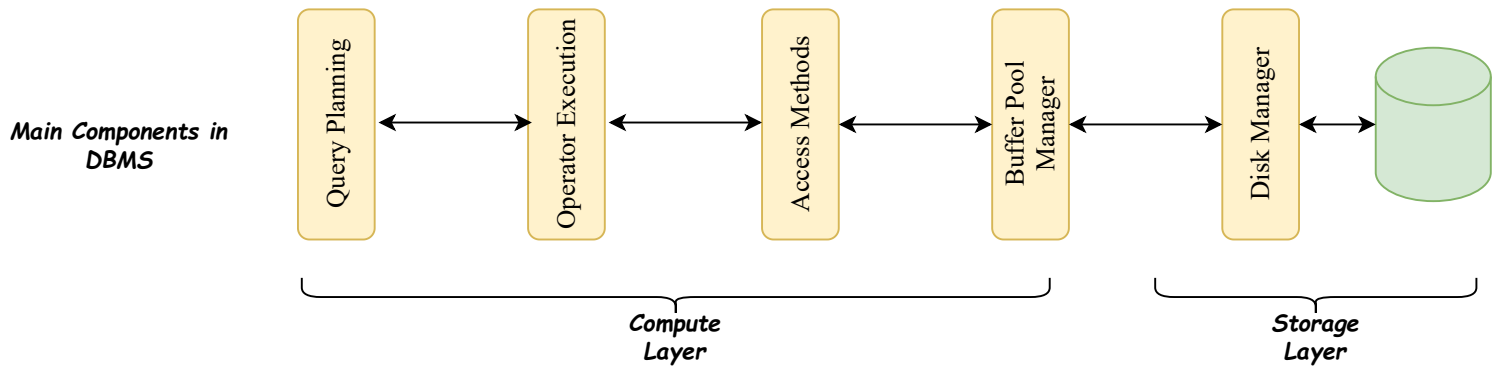


Storage Layer

How do we represent the database in files on Disk?

- File Storage
- Page Layout
- Tuple Layout

Storage Manager is responsible for maintaining a database files. It organizes the files as a collection of pages.

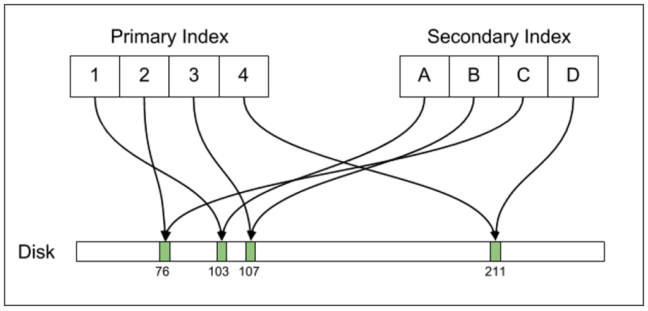
A page is a fixed size of block on Disk.

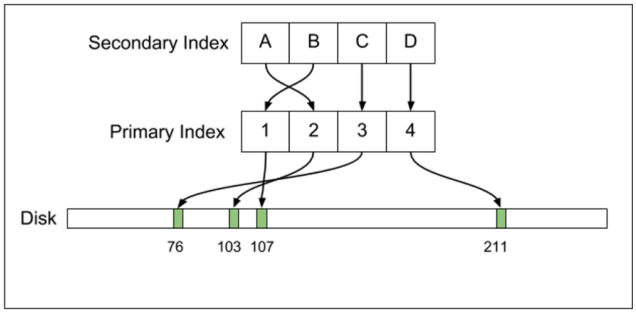
B-Tree Based Engine Vs Log Structured Merge (LSM) Tree:

Most RDBMS systems uses B-tree based storage engine. Provides low latency reads. However, the need to maintain a well-ordered data structure with random writes usually leads to poor write performance.

Using B-Tree, Range queries possible on Primary Keys.

Indexing:





If the tuple is immutable, Then this approach will have Write amplification.

write amplification refers to a problem with writing data to SSD disks: a small logical update (say, writing a few bytes) becomes a much larger, costlier update when translated to the physical layer.

Join with parent table for accessing data.

If Tuple is immutable, How update works? Configure Fill Factor optimization parameter, to keep changes in same block.

Full Page Writes/Double Write Buffer.

Linux filesystems typically use 4kB pages, but the page size DB uses is 8 to 16 KB. To prevent partial writes, databases uses different techniques.

Distributed Systems

Why we need distributed Systems:

- Fault Tolerance (High Availability)
- Scalability
- Latency (Geo-Graphic data centers)

Replication

Keeping a copy of the same data on several different nodes potentially in different locations. (Data Redundancy)

Replication Strategy: Sync vs Async

Most RDBMS Uses Below Architectures:

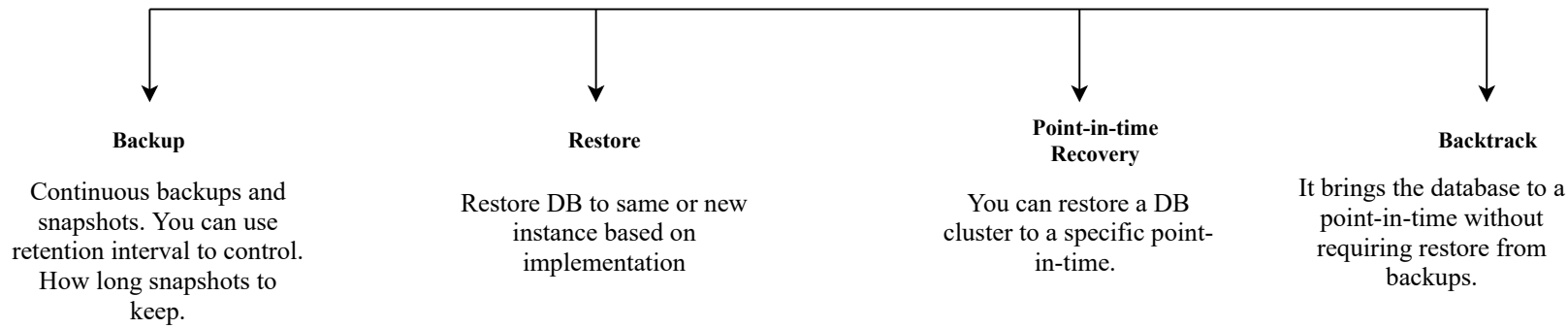
Leaders and Followers

- *Single Leader*
- *Multi Leader (Handle Write Conflicts)*

Partitioning

Range Partitioning
Hash Partitioning
Consistent Hashing

Backups



Monitoring

- What are your monitoring goals?

- Which resources will you monitor?
- How often will you monitor these resources?
- Which monitoring tools will you use?
- Who will perform the monitoring tasks?
- Whom should be notified when something goes wrong?

Performance

- Read Latency
- Throughput
- Transactions per second