Chapter 6

1. Measures of Query Cost:

- **Definition:**
    - Query cost refers to the resources and time required to execute a database query.
- **Measures:**
    - *Response Time:* The time taken to receive the first tuple of the query result.
    - *Processing Time:* The time taken to execute the query once it starts.
    - *Disk Accesses:* The number of times the system reads from or writes to the disk.
    - *CPU Time:* The time the CPU spends processing the query.
    - *Communication Cost:* The cost associated with transmitting data between different components.

2. Selection Operation, Sorting, and Join Operation:

- **Selection Operation:**
    - *Definition:* Choosing specific rows from a table based on a given condition.
    - *Cost Factors:* Depends on the selectivity of the condition and the size of the table.
- **Sorting Operation:**
    - *Definition:* Arranging data in a specified order.
    - *Cost Factors:* Depends on the number of records to be sorted and the available memory.
- **Join Operation:**
    - *Definition:* Combining rows from two or more tables based on a related column.
    - *Cost Factors:* Depends on the size of the tables and the efficiency of the join algorithm (e.g., nested loop join, hash join).

3. Transaction Concept:

- **Definition:**
    - A transaction is a sequence of one or more operations performed as a single logical unit of work.
- **ACID Properties:**

- *Atomicity:* All or nothing; a transaction is either fully completed or fully rolled back.
- *Consistency:* A transaction brings the database from one valid state to another.
- *Isolation:* The execution of one transaction is isolated from the execution of others.
- *Durability:* Once a transaction is committed, its effects persist.

4. Components of Transaction Management:

- *Transaction Manager:* Coordinates and manages the execution of transactions.
- *Transaction Log:* Records all changes made by transactions for recovery purposes.
- *Concurrency Control Manager:* Ensures proper isolation of transactions.
- *Recovery Manager:* Manages the database recovery process after a failure.
- *Lock Manager:* Manages locks to control access to data during transactions.

5. Concurrency and Recovery System:

- **Concurrency Control:**
  - *Definition:* Managing simultaneous execution of transactions to ensure database consistency.
  - *Methods:* Locking, timestamps, validation.
- **Recovery System:**
  - *Definition:* Restoring the database to a consistent state after a failure.
  - *Methods:* Log-based recovery, shadow-paging.

6. Different Concurrency Control Protocols:

- *Timestamps:* Assigns a unique timestamp to each transaction and uses them to determine order and concurrency.
- *Locking:* Controls access to data by acquiring and releasing locks.

7. Validation:

- *Definition:* Checking the correctness of a transaction before committing.
- *Use:* Ensures that a transaction, when committed, will not violate integrity constraints.

8. Multiple Granularity:

- *Definition:* Allowing different levels of locking, from fine to coarse, depending on the application needs.

- *Benefits:* Enhances concurrency by allowing multiple transactions to lock different parts of a resource simultaneously.

9. Deadlock Handling:

- *Definition:* A situation where two or more transactions are unable to proceed because each is waiting for the other to release a lock.
- *Methods:* Timeout, detection, prevention.

10. Different Crash Recovery Methods:

- *Log-Based Recovery:* Uses a transaction log to roll forward committed transactions and roll back uncommitted ones.
- *Shadow-Paging:* Maintains a shadow copy of the database that is updated and switched after a transaction commits.
- *Buffer Management:* Controls the transfer of data between the disk and memory buffer to optimize query performance.
- *Remote Backup System:* Regularly copies and stores database backups at a remote location for disaster recovery.