B.Com (Computers) II Year
RELATIONAL DATABASE MANAGEMENT SYSTEM
UNIT-IV

1. **What is Disk performance Factors?**

A. There are four factors that directly affect the speed with which data are transferred to and from disk storage: access motion time, head activation time, head activation, rotational, delay and data transfer rate.

1) Access motion Time: - The time required to move the read/write heads of the disk drive over the desired cylinder. Sometimes termed seek time, is the time required to move the read/write heads from their current position to a new cylinder address. Obviously, a move to an adjacent cylinder will not take the same amount of time as a move across the entire disk surface (inner most track to outer most track, or vice versa). As a compromise in calculation, the average access motion time may be used-roughly the time required to move across one-half of the cylinders, although more sophisticated methods may be used. A standard assumption is that the likelihood of access for every record is the same giving a uniform probability distribution. The average for a uniform distribution is halfway between the extreme values. For access motion time, the extreme would be (1) Stay positioned over the current cylinder, or (2) Move from the inner most cylinders to the outer most or (vice versa). Given the uniform distribution assumption, the average will be the time to move across one-half of the cylinders. 12-20 million seconds are typical average access motion time, varying with the make and model of the disk drive.

2) Head Activation Time: - The time required to activate a read /write head. It is the time required to electronically activate the head i.e., over the surface the data transfer is to take place relatively other performance factors, this time generally regarded as being negligible. Consequently, head activation time is seldom used in performance calculations.

3) Rotational Delay: - The time required for the disk to rotate the sought-for record under the read/write head. Rotational delay, or latency, is the third timing factor. It denotes the amount of time required for the desired block to rotate to the head, so that the data transfer may commence. Rotational delay depends upon two factors: How fast the disk is rotating and the location of the block being sought in relationship to the read/write head at time of its activation physically, this time reach from zero to the time required to complete one complete revolution of the disk(R). Performance computations usually assume an average rotational delay of R/2.

4) Data Transfer Rate: - The rate at which data can be read from the disk from the main, memory, or equivalently, the rate at which data are written from main memory to disk. Data Transfer Rate refers to the amount of time required to transfer data from the disk to primary memory. It is a function of rotational speed and the density recorded data. Data Transfer Time is usually expressed in thousands of bytes per second.

5) Data Transfer Time: - The expected time (T) to access a disk address and transfer a block of data is estimated as

T = A+R/2+L/2

Where A is the Access motion time, R is the Rotational delay, L is the length of the block in bytes, and D is the Data Transfer Rate.

2Q) What is Secondary Storage Data?

A) Secondary storage for Database system is usually comprised of disk storage and magnetic tape storage. Typically, the entire database is stored on disk, and portions are transferred from disk to primary memory as needed disk storage is the principal form direct-access storage, since individual records can be accessed directly or sequentially although magnetic tape storage is less expensive than disk storage records can only be accessed sequentially. It role in the database system basically limited to archiving data.

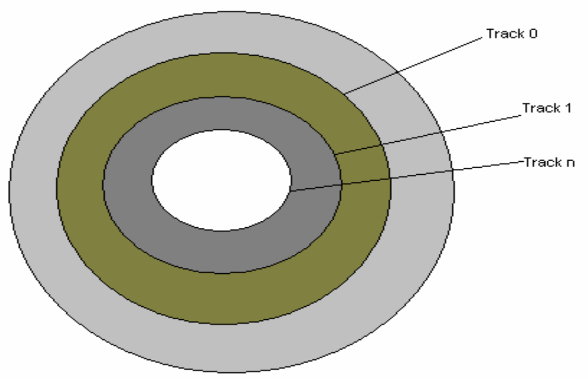
Disk Drive: - Physical unit that contains the disk storage unit each disk drive contain one disk pack or volume. Figure shows the principal components of the disk pack and the read/write mechanism required for the data transmission. The disk pack is made of a set of recording surfaces (disk) mounted on a shaft. In operation, the shaft and the disk rotate a high rate of speed. The data are recorded on tracks, which are circular recording positions found each surface. There may be several hundred tracks on a single surface a common metaphor for the disk pack is stack of phonograph records on a spindle: except that here the tracks are concentric and therefore do not spiral inward to the center.

As shown in the figure a set of read/write head roughly like the teeth of comb moves as a group so that the read/write heads at the end of an arm can be positioned overall those tracks having the same radius. The rate of such track is termed as cylinder.

Cylinder: - The same track extending through all surfaces of the disk storage unit.

That is a set of tracks of the same diameter rotating at a high speed forms a conceptual cylinder. This is a useful definition since any positioning.

Structure Of A Disk Srface



Of the set of read/write heads can be described by the cylinder location. Thus, all tracks in a given cylinder can be written to, or read from, without further movement of the read/write heads.

3) What is File Organization system?

A) There are three basic ways of physically organizing files on storage devices. Sequential organization, indexed-sequential organization and direct organization. This is not an entire set of all organization options available but those that are omitted are modifications of these basic organization types. Therefore it is not necessary to be exhaustive in order to cover the essential concepts.

In discussing the topic at hand, the terms organization and access are often used loosely if not interchangeably. The reason is that the way in which data are stored is closely intertwined with the method access.

Sequential File Organization: - Sequential file organizational means that records are stored adjacent to one another according to a key such as employee number, account number, and so forth. A conventional implementation arranges the records in acesending order of key values. This is efficient method of organizing records when an application. Such as a payroll program, will be updating a significant number of the stored records.

If a sequential file is maintained on magnetic tape, its records can only be accessed in a sequential manner. That is, if access to the tenth record in sequence is desired generally the preceding nine records must be read. Direct access of a particular record is impossible. Consequently magnetic tapes are not well suited for database operations and are usually relegated log files and recording archival information.

Indexed- Sequential File Organization: - When files are sequentially organized on a disk pack, however, direct access of records is possible. Indexed-sequential file organization provides facilities for accessing records both sequentially and directly. Records are stored in the usual physical sequence by primary key. In addition an index of record locations is stored on the disk. This allows records to be accessed sequentially for applications requiring the updating of large numbers of records, as well as providing the ability to access records directly in response to user queries.

Direct File Organization: - We have studied two forms of file organization: Sequential and indexed sequential. We have concurrently outlined the two associated methods of file access: sequential access and direct access. Records in a simple sequential file organization can be accessed directly, as well as sequentially. We now turn to a discussion of a third type of file organization called direct or hashed. Only direct access methods are applicable to this type of file organization.

4) What are the functions of DBA? What are the Establishing and Procedures?

A) DBA functions may generally fall into the areas of communicating with database users; planning, designing, and implementing database systems; and establishing standards and procedures. The planning, designing, and implementing of database systems.

Communicating with Users: - Database systems often have three components: a central, widely used database containing much of the firm's data; several functional database (e.g., for accounting) used by a more limited set of programs; and perhaps a few dedicated database, used for a single application (e.g., a bill-of-materials database). The important organizational issue here is that the general impact of implementing a database system is the centralization of a significant portion of the firm's data.

Centralizing data through a database system tends to eliminate local ownership of data and to reduce redundancy. Ownership and control are transferred to the central data dictionary, which maintains a record of the ownership and use of each data element. Such a shifting of control over data may generate resistance from some users. This resistance can be mitigated by actively educating users as to the advantages of learning database technology: how it can make them more effective and efficient at their jobs. The DBA, in cooperation with top management, should provide this education

Establishing standards and procedures: - organizations having few standards and procedures may encounter difficulty in converting to the database environment, since the record shows that the integrated data management facilitated by database systems requires good, comprehensive standards and procedures. An organization that is beginning to implement a database system may find it useful to examine the standards in use at other organizations that are already using database systems.

1. **Analysis and routing of trouble reports: -** A formal trouble-reporting system was established in order to report all errors to the DBA. Trouble reports are analyzed to determine the likely cause of each reported problem. The reports are then routed to the appropriate manager or user group for disposition. Each trouble report contains a complete log and descriptive information. Each report requires a formal response to the report's initiator specifying how the problem has been resolved.

2. **Monitoring of hardware and software**: The status of all hardware and software is regularly monitored, and reports of failures and consequent action are made to appropriate mangers and user groups. Periodic analysis of hardware and software requirements is made, forming the basis for decisions on replacement and upgrading, including needs for additional database storage media.

3. **Testing:** Performance acceptance testing is conducted to evaluate all new procedures, software, and hardware. Structural and consistency checks of the database are conducted on a regular basis.

4. **Security:** security classifications are implemented that identify which user groups are authorized to access specific data elements sin the database and what actions may be performed thereon. Computer operations area frequently monitored to assure that these access controls are functioning in the intended way.

5. **Backup and recovery:** Backup and recovery procedures are tested regularly to assure their effectiveness in restoring the database after any disruption of service/ a disaster plan has been drawn up and is tested periodically to make sure it works.

6. **Performance evaluation:** Priorities have assigned to activities that compete for database resources, such as processing transactions, generating reports, and processing queries, system performance is monitored by collecting statistics on transaction volume, response time, error rates, and hardware utilization. Input is elicited from system users to monitor their satisfaction with the system's performance. Database size and growth is also tracked. File expansion programs are run and database reorganizations are performed as necessary. Activity logs and abnormal termination logs are reviewed and summaries prepared for management evaluation.

7. **Integrity checking:** Schedules have been developed for testing the integrity of the date stored in the database.

5) What are the Goals of DBA?

A) A database must be protected from accidents, such as input or programming errors, from malicious use of the database, and from hardware of software failures that corrupt data. Protection from accidents that cause data inaccuracies is part of the goal of maintaining data integrity. These accidents include failures during transaction processing. Logical errors that violate the assumption that transactions preserve database consistency constraints, and anomalies due to concurrent access to the database (concurrent processing).

Protecting the database from unauthorized or malicious use is termed data security. Although the dividing line between data integrity and data security is not precise, a working definition is as follows:

1. Integrity is concerned with making certain that operations performed by users are correct and maintain database consistency.
2. Security is concerned with limiting users to performing only those operations that are allowed.

The possibility of hardware or software failure requires that database recovery procedures be implemented as well. That is, means must be provided to restore databases that have been corrupted by system malfunctions to a consistent state.

6) What is database Integrity?

A) A condition or integrity that is applied to a particular set of data is commonly termed **Integrity Control** or **Constraint.** In relational model terminology, integrity controls may apply to (1) Individual attributes, (2) the relationship between two different attributes (perhaps in different relations) or, (3) the relationship between tuples of one or more tables. Ideally, the enforcement of integrity constraints would be carried out by the DBMS currently as each new data item is entered.

7) What is Database Security?

A) Database integrity problems can be challenging, but they are generally easier to cope with than malicious access to the database, which includes the following:

1. Theft of information
2. Unauthorized modification of data
3. Unauthorized destruction of data

Thus, database security methods focus on preventing unauthorized users from accessing the database. Because DBMS features that make the database easy to access and manipulate also open doors to intruders, most DBMS include security features that allow only authorized persons or processing that can be accompanied once access is made.

Authentication: - Database access usually requires user authentication and authorization. For user authentication, the first level of security establishes that the person seeking system the user knows, such as log-on number and password, (2) something the user possesses, such as plastic ID card, or (3) a physical representation of the user, such as fingerprint or voiceprint.

Authorization and views: - A view is a means of providing a user with a personalized model of the database. It is also a useful way of limiting a user’s access to various positions of the database: Data a user does not need to se are simply hidden from view. This simplifies system usage while promoting security. Executing selects, projections, and joins on existing relations can represent views. The user might also be restricted from seeing any part of the existing relation or from executing joins on certain relations.

Types of Views: - Different types of access authorization may be allowed for a particular view, such as the following:

1. Read authorization: allows reading, but not modification of data.
2. Insert authorization: allows insertion of new data, but no modification of existing data.
3. Update authorization: allows modification of data, but not deletion.
4. Delete authorization: allows deletion of data.

Views and security in SQL: -

```
  CREATE VIEW viewname (list of attributes desired, if different from base table)
 As query.
```

Encryption: - The various authentication and authorization measures that are standard for protection access to database may not be adequate for highly sensitive data. In such instances, it may be desirable to encrypt the data. Encrypted data cannot be read by an intruder unless that party knows the method of encryption. Considerable research has been devoted to developing encryption methods.

8) What is Database Recovery?

A) Information stored on computer media is subject to loss or corruption caused by a wide range of events, it is important to provide means for resorting correct data to the database. Resorting the database to precisely the same state that existed at the time of system failure is not always possible, but database recovery procedures can restore the database to the state that existed shortly before the failure and identify the status of transaction processing at the time of the failure. With this capability, unprocessed transactions can be processed against the restored database to bring it back to a fully current status.

Sources of Failure:

A useful classification of failure types includes the following:

1. **System errors:** the system has entered an undesirable state, such as deadlock, which prevents the program from continuing with normal processing. This type of failure may or may not result in corruption of data files.
2. **Hardware failures:** Two of the most common types of hardware failure and loss of transmission capability over a transmission link. In the former case, the cause usually results from the disk read/write head coming in physical contact with the disk surface.
3. **Logical errors:** Bad data or missing data are common conditions that may preclude a program's continuing with normal execution.

Recovery Procedures: -

To maintain data integrity, a transaction must be in one of the two following states:

1. **Aborted:** A transaction may not always complete its process successfully. To be sure the incomplete transaction will not affect the consistent state of the database, such transactions must be aborted, and resorting the database to the state it was in before the transaction in question began execution. Such restoration is achieved by rollback.
2. **Committed:** A transaction that successfully completes its processing is said to be committed. A committed transaction always leaves the database in a new consistent state.