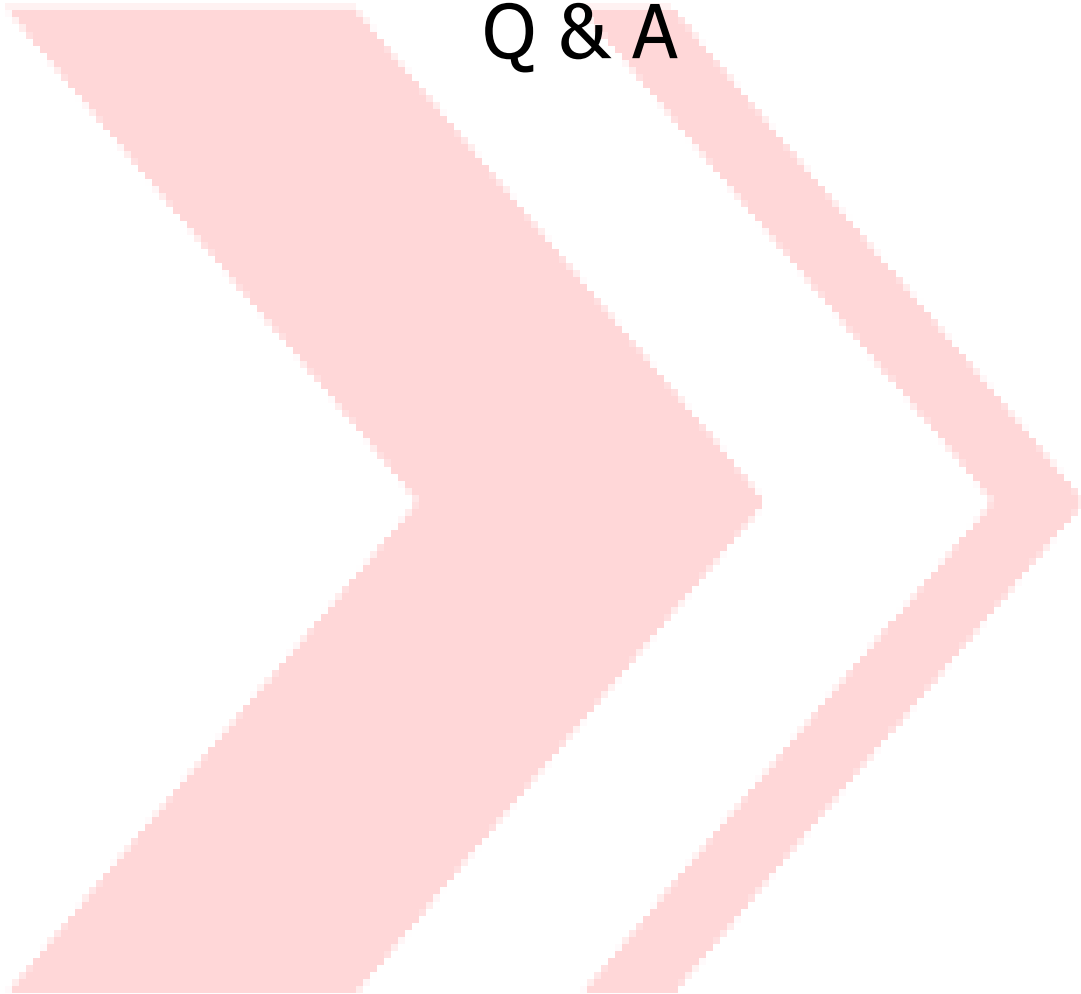


Nexwave DBMS Interview Q & A



What is a Database?

A database is a collection of related data .A database is a logically coherent collection of data with some inherent meaning.

What is DBMS?

Database Management system is a collection of programs that enables user to create and maintain a database. Thus a DBMS is a general purposed s/w system that facilitates the process of defining constructing and manipulating a database for various applications. (Defining a data base involves specifying the data types, structures and constraints for the data to be stored in the data database. Constructing a data base is the process of storing data itself on some storage medium that is controlled by DBMS. Manipulating a database includes such functions as querying the data base to retrieve specific data, updating the database to reflect the changes in the mini-world.

What are different DBMS languages?

1. DDL (Data definition language)
2. SDL (Storage definition language)
3. VDL (View definition language)
4. DML (Data manipulation language)

What are different types of DBMS?

1. DBMS
2. RDBMS (Relational)
3. ORDBMS (Object Relational)
4. DDBMS (Distributed)
5. FDBMS (Federated)
6. HDDBMS (Homogeneous)
7. HDBMS (Hierarchical)
8. NDBMS (Networked)

What is ORDBMS?

Object oriented RDBMS is a relational DBMS in which every thing is treated as objects. User can define operations on data as a part of the database definition.

What are the advantages of using a DBMS?

1. Controlling redundancy.
2. Restricting unauthorized access.
3. Providing persistent storage for program objects and data structures.
4. Permitting inferencing and actions using rules.
5. Providing multi-user interfaces.
6. Representing complex relationships among data.
7. Enforcing integrity constraints.
8. Providing backups and recovery.

What are the disadvantages of using a DBMS?

1. High initial investments in h/w, s/w, and training.
2. Generality that a DBMS provides for defining and processing data.
3. Overhead for providing security, concurrency control, recovery, and integrity functions.

What is RDBMS?

Relational Data Base Management Systems (RDBMS) are database management systems that maintain data records and indices in tables. Relationships may be created and maintained across and among the data and tables. In a relational database, relationships between data items are expressed by means of tables. Interdependencies among these tables are expressed by data values rather than by pointers. This allows a high degree of data independence. An RDBMS has the capability to recombine the data items from different files, providing powerful tools for data usage.

List the advantages of DBMS?

Here is the list of advantages of using DBMS

1. Redundancy is controlled. (Duplication is controlled)
2. Unauthorized access is restricted. (Restriction others to access data base information)
3. Providing multiple user interfaces. (Providing multi user access also limiting their permission)
4. Enforcing integrity constraints.
5. Providing backup and recovery. (Back memory provided when data base crash occurs)

What are the properties of the Relational tables?

Relational tables have six properties:

- Values are atomic.
- Column values are of the same kind.
- Each row is unique.
- The sequence of columns is insignificant.
- The sequence of rows is insignificant.
- Each column must have a unique name.

What is SQL and where does it come from?

Structured Query Language (SQL) is a language that provides an interface to relational database systems. SQL was developed by IBM in the 1970s for use in System R, and is a de facto standard, as well as an ISO and ANSI standard. SQL is often pronounced SEQUEL. In common usage SQL also encompasses DML (Data Manipulation Language), for INSERTs, UPDATEs, DELETEs and DDL (Data Definition Language), used for creating and modifying tables and other database structures. The development of SQL is governed by standards. A major revision to the SQL standard was completed in 1992, called SQL2. SQL3 support object extensions and are (partially?) implemented in Oracle8 and 9.

Describe the level of data abstraction?

There are three levels for data abstraction

- Physical level: The lowest level of abstraction describes how data are stored
- Logical level: Next higher level in abstraction, describes which data is stored and relation between them
- View Level: This describes the part of entire database

What are the difference between DDL, DML and DCL commands?

DDL is Data Definition Language statements. Some examples:

- CREATE - to create objects in the database
- ALTER - alters the structure of the database
- DROP - delete objects from the database
- TRUNCATE - remove all records from a table, including all spaces allocated for the records are removed
- COMMENT - add comments to the data dictionary
- GRANT - gives user's access privileges to database
- REVOKE - withdraw access privileges given with the GRANT command

DML is Data Manipulation Language statements. Some examples:

- SELECT - retrieve data from the a database
- INSERT - insert data into a table
- UPDATE - updates existing data within a table
- DELETE - deletes all records from a table, the space for the records remain
- CALL - call a PL/SQL or Java subprogram
- EXPLAIN PLAN - explain access path to data
- LOCK TABLE - control concurrency

DCL is Data Control Language statements. Some examples:

- COMMIT - save work done
- SAVEPOINT - identify a point in a transaction to which you can later roll back
- ROLLBACK - restore database to original since the last COMMIT
- SET TRANSACTION - Change transaction options like what rollback segment to use

What are primary keys and foreign keys?

Primary keys are the unique identifiers for each row. They must contain unique values and cannot be null. Due to their importance in relational databases, Primary keys are the most fundamental of all keys and constraints. A table can have only one Primary key. Foreign keys are both a method of ensuring data integrity and a manifestation of the relationship between tables.

- A **PRIMARY KEY** constraint is a unique identifier for a row within a database table. Every table should have a primary key constraint to uniquely identify each row and only one primary key constraint can be created for each table. The primary key constraints are used to enforce entity integrity.
- A **UNIQUE** constraint enforces the uniqueness of the values in a set of columns, so no duplicate values are entered. The unique key constraints are used to enforce entity integrity as the primary key constraints.
- A **FOREIGN KEY** constraint prevents any actions that would destroy links between tables with the corresponding data values. A foreign key in one table points to a primary key in another table. Foreign keys prevent actions that would leave rows with foreign key values when there are no primary keys with that value. The foreign key constraints are used to enforce referential integrity.
- A **CHECK** constraint is used to limit the values that can be placed in a column. The check constraints are used to enforce domain integrity.
- A **NOT NULL** constraint enforces that the column will not accept null values. The not null constraints are used to enforce domain integrity, as the check constraints.

What is difference between SQL and SQL SERVER?

SQL (sequel) is a language that provides an interface to RDBMS, developed by IBM. SQL SERVER is a RDBMS just like Oracle, DB2.

What is Difference between DBMS and File System?

DBMS and File system are two ways of saving data, including managing data.

File System: It is a collection of raw data files stored in hard drive.

DBMS: It is a bundle of application that is dedicated for managing data stored in databases

In File System, files are used to store data while, collections of databases are utilized for the storage of data in DBMS. In file system storing altering data done manually. DBMS provide automated methods to complete managing data.

What is a data model?

It is a collection of concepts that can be used to describe the structure of a database. It provides necessary means to achieve this abstraction. By structure of a database we mean the data types, relations, and constraints that should hold on the data.

What are different categories of data models?

1. High-level or conceptual data models.
2. Representational data models.
3. Low-level or physical data models.

High level data models provide the concepts that are close to the way many users perceive data. Representational data models are provide concepts that provide the concepts that may be understood by end users but that are not too far removed from organization of data in the database. Physical data models describe the details of how data is stored in the computers.

What is data ware housing & OLAP?

Data warehousing and OLAP (online analytical processing) systems are the techniques used in many companies to extract and analyze useful information from very large databases for decision making.

What is real time database technology?

These are all the techniques used in controlling industrial and manufacturing processes.

What is program-data independence?

Unlike in the traditional file sys. the structure of the data files is stored in the DBMS catalog separately from the access programs . This property is called program-data independence i.e., We needn't to change the code of the DBMS if the structure of the data is changed .Which is not supported by traditional file sys.

What is program-operation independence?

An operation is specified in two parts.

1. Interface (operation name and data types of its arguments).
2. Implementation (the code part)

The implementation part can be changed without affecting the interface. This is called program-operation independence.

What is a view?

A view may be a subset of the database or it may contain virtual data that is derived from the database files but is not explicitly stored.

What is OLTP?

Online transaction processing is an application that involves multiple database accesses from different parts of the world. OLTP needs a multi-user DBMS s/w to ensure that concurrent transactions operate correctly.

What is the job of DBA?

A database administrator is a person or a group responsible for authorizing access to the database, for coordinating and monitoring its use, and for acquiring s/w and h/w resources as needed.

Who are DB designer?

Data base designers are responsible for identifying the data to be stored in the database and for choosing appropriate structure to represent and store this data.

What are different types of end users?

1. Casual end-users
2. Naive or parametric end users
3. Sophisticated end users
4. Stand alone users.

What is schema?

The description of a data base is called the database schema, which is specified during database design and is not expected to change frequently. A displayed schema is called schema diagram .We call each object in the schema as schema construct.

What are types of schema?

1. Internal schema.
2. Conceptual schema.
3. External schemas or user views.

What is Data independency?

Data independency is defined as the capacity to change the conceptual schema without having to change the schema at the next higher level. We can define two types of data independence:

1. Logical data independence.
2. Physical data independence.

LDI is the capacity to change the conceptual schema without having to change external schemas or application programs. PDI is the capacity to change the internal schema without having to change conceptual (or external) schemas.

What is an entity?

An entity is a thing in the real world with an independent existence.

What are attributes?

These are the particular properties that describe an entity.

What are diff. types of attributes?

1. Composite Vs simple attributes.
2. Single valued Vs multi-valued attributes.

3. Stored Vs derived attribute.
4. Null valued attributes.
5. Complex attributes.

What is domain value or value set of an attribute?

It is the set of values that may be assigned to that attribute for each individual entities.

What is degree of a relationship?

The no of entities participating in that relation.

What is recursive relationship?

It is the relationship where both the participating entities belong to same entity type .

What are relationship constraints?

1. Cardinality ratio.
2. Participation constraints.

What is Cardinality ratio?

The cardinality ratio for a binary relationship specifies the number of relationship instances that an entity can participate in.

What is a Participation constraint?

The participation constraint specifies whether the existence of an entity depends on its being related to another entity via the relationship type. This is of two types:

1. Total participation.
2. Partial participation.

What is a weak entity types?

The entity types that do not have key attributes of their own are called weak entity types. Rests are called strong entity types .The entity that gives identity to a weak entity is called owner entity. And the relationship is called identifying relationship. A weak entity type always has a total participation constraint with respect to its identifying relationship.

What is an ER Diagram?

This data model is based on real world that consists of basic objects called entities and of relationship among these objects. Entities are described in a database by a set of attributes.

What is specialization?

It is the process of defining a set of subclasses of an entity type where each subclass contain all the attributes and relationships of the parent entity and may have additional attributes and relationships which are specific to itself.

What is generalization?

It is the process of finding common attributes and relations of a number of entities and defining a common super class for them.

What are constraints on generalization and specialization?

1. Disjointness constraints.
2. Completeness constraints.

Disjointness constraint specifies that the subclasses of the specialization must be disjoint .i.e. an entity can be a member of at most one of the subclasses of the specialization. The reverse of it is overlapping.

Completeness constraint is a participation constraint which may be

1. Total
2. Partial

Total specialization constraint tells that each entity in the super class must be a member of some subclass in the specialization. And partial specialization constraint allows an entity not to belong to any of the subclasses .Thus we do have the following 4 types of constraints on specialization:

1. Disjoint, total
2. Disjoint, partial
3. Overlapping, total
4. Overlapping, partial

What is a ternary relationship?

A relationship with a degree 3 is called a ternary relationship.

What is aggregation and association?

Aggregation is an abstraction concept for building composite objects from their component objects. The abstraction of association is used to associate objects from several independent classes.

What is RAID Technology?

Redundant array of inexpensive (or independent) disks. The main goal of raid technology is to even out the widely different rates of performance improvement of disks against those in memory and microprocessor. Raid technology employs the technique of data striping to achieve higher transfer rates.

What is Hashing technique?

This is a primary file organization technique that provides very fast access to records on certain search conditions. The search condition must be an equality condition on a single field, called hash field of the file.

1. Internal hashing
2. External hashing
3. Extendible hashing
4. Linear hashing
5. Partitioned hashing

What are different types of relational constraints?

1. Domain constraints
2. Key constraints
3. Entity integrity constraints
4. Referential integrity constraints

Domain constraints specify that the value of each attribute must be an atomic value from the domain of the attributes. Key constraints tell that no two tuples can have the same combination of values for all their attributes. Entity integrity constraint states that no primary key value can be null. Referential integrity constraints states that a tuple in one relation that refers to another relation must refer to an existing tuple in that relation it is specified between two relations and is used to maintain the consistency among tuples of the two relations.

What is difference between a super key, a key, a candidate key and a primary key?

A super key specifies a uniqueness constraint that no two distinct tuples in a state can have the same value for the super key. Every relation has at least one default super key. A key is a minimal super key or the subset of the super key which is obtained after removing redundancy. A relation schema may have more than one key .In this case each key is called a candidate key. One of the candidate key with minimum number of attributes is chosen as primary key.

What is a Catalog?

A catalog is a table that contains the information such as structure of each file, the type and storage format of each data item and various constraints on the data. The information stored in the catalog is called Metadata. Whenever a request is made to access a particular data, the DBMS s/w refers to the catalog to determine the structure of the file.

What is a transaction?

A transaction is a logical unit of database processing that includes one or more database access operations.

What are the properties of transaction?

1. Atomicity
2. Consistency preservation

3. Isolation
4. Durability (permanence)

What are the basic data base operations?

1. Write_item(x)
2. Read_item(x)

What is Table ?

A table is the basic unit of data storage in an ORACLE database. The tables of a database hold all of the user accessible data. Table data is stored in rows and columns.

What is a "join"?

'join' used to connect two or more tables logically with or without common field.

What is Self Join?

This is a particular case when one table joins to itself, with one or two aliases to avoid confusion. A self join can be of any type, as long as the joined tables are the same. A self join is rather unique in that it involves a relationship with only one table. The common example is when company have a hierarchal reporting structure whereby one member of staff reports to another.

What is Cross Join?

A cross join that does not have a WHERE clause produces the Cartesian product of the tables involved in the join. The size of a Cartesian product result set is the number of rows in the first table multiplied by the number of rows in the second table. The common example is when company wants to combine each product with a pricing table to analyze each product at each price.

What is a Scheduled Jobs or What is a Scheduled Tasks?

Scheduled tasks let user automate processes that run on regular or predictable cycles. User can schedule administrative tasks, such as cube processing, to run during times of slow business activity. User can also determine the order in which tasks run by creating job steps within a SQL Server Agent job. E.g. back up database, Update Stats of Tables. Job steps give user control over flow of execution. If one job fails, user can configure SQL Server Agent to continue to run the remaining tasks or to stop execution.

What is normalization?

Database normalization is a data design and organization process applied to data structures based on rules that help build relational databases. In relational database design, the process of organizing data to minimize redundancy. Normalization usually involves dividing a database into two or more tables and defining relationships between the tables. The objective is to isolate data so that additions, deletions, and modifications of a field can be made in just one table and then propagated through the rest of the database via the defined relationships.

What are different normalization forms?

- **1NF: Eliminate Repeating Groups**

Make a separate table for each set of related attributes, and give each table a primary key. Each field contains at most one value from its attribute domain.

- **2NF: Eliminate Redundant Data**

If an attribute depends on only part of a multi-valued key, remove it to a separate table.

- **3NF: Eliminate Columns Not Dependent On Key**

If attributes do not contribute to a description of the key, remove them to a separate table. All attributes must be directly dependent on the primary key

- **BCNF: Boyce-Codd Normal Form** If there are non-trivial dependencies between candidate key attributes, separate them out into distinct tables.

- **4NF: Isolate Independent Multiple Relationships**

No table may contain two or more 1:n or n:m relationships that are not directly related.

- **5NF: Isolate Semantically Related Multiple Relationships**

There may be practical constraints on information that justify separating logically related many-to-many relationships.

- **ONF: Optimal Normal Form**

A model limited to only simple (elemental) facts, as expressed in Object Role Model notation.

- **DKNF: Domain-Key Normal Form**

A model free from all modification anomalies. Remember, these normalization guidelines are cumulative. For a database to be in 3NF, it must first fulfill all the criteria of a 2NF and 1NF database.

What is "normalization"? "Denormalization"? Why do you sometimes want to denormalize?

Normalizing data means eliminating redundant information from a table and organizing the data so that future changes to the table are easier. Denormalization means allowing redundancy in a table. The main benefit of denormalization is improved performance with simplified data retrieval and manipulation. This is done by reduction in the number of joins needed for data processing.

What is a "constraint"?

A constraint allows you to apply simple referential integrity checks to a table. There are four primary types of constraints that are currently supported by SQL Server: PRIMARY/UNIQUE - enforces uniqueness of a particular table column.

DEFAULT - specifies a default value for a column in case an insert operation does not provide one. FOREIGN KEY - validates that every value in a column exists in a column of another table.

CHECK - checks that every value stored in a column is in some specified list. Each type of constraint performs a specific type of action. Default is not a constraint.

NOT NULL is one more constraint which does not allow values in the specific column to be null. And also it the only constraint which is not a table level constraint.

What types of index data structures can you have?

An index helps to faster search values in tables. The three most commonly used index-types are:

- B-Tree: builds a tree of possible values with a list of row IDs that have the leaf value. Needs a lot of space and is the default index type for most databases.
- Bitmap: string of bits for each possible value of the column. Each bit string has one bit for each row. Needs only few space and is very fast.(however, domain of value cannot be large, e.g. SEX(m,f); degree(BS,MS,PHD)
- Hash: A hashing algorithm is used to assign a set of characters to represent a text string such as a composite of keys or partial keys, and compresses the underlying data. Takes longer to build and is supported by relatively few databases.

What is a "functional dependency"? How does it relate to database table design?

Functional dependency relates to how one object depends upon the other in the database. for example, procedure/function sp2 may be called by procedure sp1. Then we say that sp1 has functional dependency on sp2.

What is a "trigger"?

A trigger is a SQL procedure that initiates an action when an event (INSERT, DELETE or UPDATE) occurs. Triggers are stored in and managed by the DBMS.Triggers are used to maintain the referential integrity of data by changing the data in a systematic fashion. A trigger cannot be called or executed; the DBMS automatically fires the trigger as a result of a data modification to the associated table. Triggers can be viewed as similar to stored procedures in that both consist of procedural logic that is stored at the database level. Stored procedures, however, are not event-drive and are not attached to a specific table as triggers are. Stored procedures are explicitly executed by invoking a CALL to the procedure while triggers are implicitly executed. In addition, triggers can also execute stored Procedures.

Nested Trigger: A trigger can also contain INSERT, UPDATE and DELETE logic within itself, so when the trigger is fired because of data modification it can also cause another data modification, thereby firing another trigger. A trigger that contains data modification logic within itself is called a nested trigger.

What is View?

A simple view can be thought of as a subset of a table. It can be used for retrieving data, as well as updating or deleting rows. Rows updated or deleted in the view are updated or deleted in the table the view was created with. It should also be noted that as data in the original table changes, so does data in the view, as views are

the way to look at part of the original table. The results of using a view are not permanently stored in the database. The data accessed through a view is actually constructed using standard T-SQL select command and can come from one to many different base tables or even other views.

What is Index?

An index is a physical structure containing pointers to the data. Indices are created in an existing table to locate rows more quickly and efficiently. It is possible to create an index on one or more columns of a table, and each index is given a name. The users cannot see the indexes, they are just used to speed up queries. Effective indexes are one of the best ways to improve performance in a database application. A table scan happens when there is no index available to help a query. In a table scan SQL Server examines every row in the table to satisfy the query results. Table scans are sometimes unavoidable, but on large tables, scans have a terrific impact on performance. Clustered indexes define the physical sorting of a database table's rows in the storage media. For this reason, each database table may have only one clustered index. Non-clustered indexes are created outside of the database table and contain a sorted list of references to the table itself.

What is Stored Procedure?

A stored procedure is a named group of SQL statements that have been previously created and stored in the server database. Stored procedures accept input parameters so that a single procedure can be used over the network by several clients using different input data. And when the procedure is modified, all clients automatically get the new version. Stored procedures reduce network traffic and improve performance. Stored procedures can be used to help ensure the integrity of the database.

What are the advantages of using Stored Procedures?

- Stored procedure can reduced network traffic and latency, boosting application performance.
- Stored procedure execution plans can be reused, staying cached in SQL Server's memory, reducing server overhead.
- Stored procedures help promote code reuse.
- Stored procedures can encapsulate logic. You can change stored procedure code without affecting clients.
- Stored procedures provide better security to your data.

What is the difference between Function and Stored Procedure?

UDF can be used in the SQL statements anywhere in the WHERE/HAVING/SELECT section where as Stored procedures cannot be. UDFs that return tables can be treated as another rowset. This can be used in JOINS with other tables. Inline UDF's can be thought of as views that take parameters and can be used in JOINS and other Rowset operations.

What types of Joins are possible with SQL Server?

Joins are used in queries to explain how different tables are related. Joins also let you select data from a table depending upon data from another table.

Types of joins: INNER JOINS, OUTER JOINS, CROSS JOINS. OUTER JOINS are further classified as LEFT OUTER JOINS, RIGHT OUTER JOINS and FULL OUTER JOINS.

What is De-normalization?

De-normalization is the process of attempting to optimize the performance of a database by adding redundant data. It is sometimes necessary because current DBMSs implement the relational model poorly. A true relational DBMS would allow for a fully normalized database at the logical level, while providing physical storage of data that is tuned for high performance. De-normalization is a technique to move from higher to lower normal forms of database modeling in order to speed up database access..

What is the difference between VARCHAR() and VARCHAR2()?

Varchar means fixed length character data (size) i.e., min size-1 and max-2000

Varchar2 means variable (Dynamic) length character data i.e., min-1 to max-4000

Which command displays the SQL command in the SQL buffer, and then executes it?

LIST or L command is used to get the recent one from SQL Buffer.

What is difference between DELETE and TRUNCATE?

Both the commands are used to delete or remove the data from table, where as TRUNCATE is a DDL command and it can't be rolled back. DELETE is a DML command and it can be roll back.
TRUNCATE deletes the data permanently while using DELETE we can again get back the data.(ROLL back).

What is advantage of using trigger?

Trigger used when we need to validate a DML statement that modifies a table. Also to update automatically when DML command is executed. Triggers can be used to enforce constraints.

What is use of WITH GRANT OPTION and GRANT command?

These commands give privilege to grant his/her privileges to other users.

Write a query to find second maximum value from a table?

```
select max (sal) from tname where sal= (select max (sal) from tname where sal< (select max (sal) from tname));
```

Which is the subset of SQL commands used to manipulate Oracle Database structures, including tables?

Data Definition Language (DDL)

What operator performs pattern matching?

LIKE operator

What operator tests column for the absence of data?

IS NULL operator

Which command executes the contents of a specified file?

START <filename> or @<filename>

What is the parameter substitution symbol used with INSERT INTO command?

&

Which command displays the SQL command in the SQL buffer, and then executes it?

RUN

What are the wildcards used for pattern matching?

_ for single character substitution and % for multi-character substitution

State true or false. EXISTS, SOME, ANY are operators in SQL.

True

State true or false. !=, <>, ^= all denote the same operation.

True

What are the privileges that can be granted on a table by a user to others?

Insert, update, delete, select, references, index, execute, alter, all

What command is used to get back the privileges offered by the GRANT command?

REVOKE

Which system tables contain information on privileges granted and privileges obtained?

USER_TAB_PRIVS_MADE, USER_TAB_PRIVS_RECD

Which system table contains information on constraints on all the tables created?

USER_CONSTRAINTS

TRUNCATE TABLE EMP;

DELETE FROM EMP;

Will the outputs of the above two commands differ?

Both will result in deleting all the rows in the table EMP.

What is the difference between TRUNCATE and DELETE commands?

TRUNCATE is a DDL command whereas DELETE is a DML command. Hence DELETE operation can be rolled back, but TRUNCATE operation cannot be rolled back. WHERE clause can be used with DELETE and not with TRUNCATE.

What command is used to create a table by copying the structure of another table?

CREATE TABLE .. AS SELECT command

To copy only the structure, the WHERE clause of the SELECT command should contain a FALSE statement as in the following.

```
CREATE TABLE NEWTABLE AS SELECT * FROM EXISTINGTABLE WHERE 1=2;
```

If the WHERE condition is true, then all the rows or rows satisfying the condition will be copied to the new table.

Which TCP/IP port does SQL Server run on? How can it be changed?

SQL Server runs on port 1433. It can be changed from the Network Utility TCP/IP properties.

What are the difference between clustered and a non-clustered index?

A clustered index is a special type of index that reorders the way records in the table are physically stored.

Therefore table can have only one clustered index. The leaf nodes of a clustered index contain the data pages.

A non clustered index is a special type of index in which the logical order of the index does not match the physical stored order of the rows on disk. The leaf node of a non clustered index does not consist of the data pages. Instead, the leaf nodes contain index rows.

What are the different index configurations a table can have?

A table can have one of the following index configurations:

- No indexes
- A clustered index
- A clustered index and many nonclustered indexes
- A nonclustered index
- Many nonclustered indexes

What are different types of Collation Sensitivity?

Case sensitivity - A and a, B and b, etc.

Accent sensitivity: Kana Sensitivity - When Japanese kana characters Hiragana and Katakana are treated differently, it is called Kana sensitive.

Width sensitivity - A single-byte character (half-width) and the same character represented as a double-byte character (full-width) are treated differently than it is width sensitive.

What is OLTP (Online Transaction Processing)?

In OLTP - online transaction processing systems relational database design use the discipline of data modeling and generally follow the Codd rules of data normalization in order to ensure absolute data integrity. Using these rules complex information is broken down into its most simple structures (a table) where all of the individual atomic level elements relate to each other and satisfy the normalization rules.

What's the difference between a primary key and a unique key?

Both primary key and unique key enforces uniqueness of the column on which they are defined. But by default primary key creates a clustered index on the column, where unique creates a nonclustered index by default. Another major difference is that, primary key doesn't allow NULLs, but unique key allows one NULL only.

What is difference between DELETE and TRUNCATE commands?

Delete command removes the rows from a table based on the condition that we provide with a WHERE clause. Truncate will actually remove all the rows from a table and there will be no data in the table after we run the truncate command.

TRUNCATE:

-TRUNCATE is faster and uses fewer system and transaction log resources than DELETE.

- TRUNCATE removes the data by deallocating the data pages used to store the table's data, and only the page deallocations are recorded in the transaction log.
- TRUNCATE removes all rows from a table, but the table structure, its columns, constraints, indexes and so on, remains. The counter used by an identity for new rows is reset to the seed for the column.
- You cannot use TRUNCATE TABLE on a table referenced by a FOREIGN KEY constraint. Because TRUNCATE TABLE is not logged, it cannot activate a trigger.
- TRUNCATE cannot be rolled back.
- TRUNCATE is DDL Command.
- TRUNCATE Resets identity of the table

DELETE:

- DELETE removes rows one at a time and records an entry in the transaction log for each deleted row.
- If you want to retain the identity counter, use DELETE instead. If you want to remove table definition and its data, use the DROP TABLE statement.
- DELETE Can be used with or without a WHERE clause
- DELETE Activates Triggers.
- DELETE can be rolled back.
- DELETE is DML Command.
- DELETE does not reset identity of the table.

When is the use of UPDATE_STATISTICS command?

This command is basically used when a large processing of data has occurred. If a large amount of deletions any modification or Bulk Copy into the tables has occurred, it has to update the indexes to take these changes into account. UPDATE_STATISTICS updates the indexes on these tables accordingly.

What is the difference between a HAVING CLAUSE and a WHERE CLAUSE?

They specify a search condition for a group or an aggregate. But the difference is that HAVING can be used only with the SELECT statement. HAVING is typically used in a GROUP BY clause. When GROUP BY is not used, HAVING behaves like a WHERE clause. Having Clause is basically used only with the GROUP BY function in a query whereas WHERE Clause is applied to each row before they are part of the GROUP BY function in a query.

What are the properties and different Types of Sub-Queries?

Properties of Sub-Query

- A sub-query must be enclosed in the parenthesis.
- A sub-query must be put in the right hand of the comparison operator, and
- A sub-query cannot contain an ORDER-BY clause.
- A query can contain more than one sub-query.

Types of Sub-Query

- Single-row sub-query, where the sub-query returns only one row.
- Multiple-row sub-query, where the sub-query returns multiple rows, and
- Multiple column sub-query, where the sub-query returns multiple columns

What is SQL Profiler?

SQL Profiler is a graphical tool that allows system administrators to monitor events in an instance of Microsoft SQL Server. You can capture and save data about each event to a file or SQL Server table to analyze later. For example, you can monitor a production environment to see which stored procedures are hampering performances by executing too slowly.

Use SQL Profiler to monitor only the events in which you are interested. If traces are becoming too large, you can filter them based on the information you want, so that only a subset of the event data is collected. Monitoring too many events adds overhead to the server and the monitoring process and can cause the trace file or trace table to grow very large, especially when the monitoring process takes place over a long period of time.

What are the authentication modes in SQL Server? How can it be changed?

Windows mode and Mixed Mode - SQL and Windows. To change authentication mode in SQL Server click Start, Programs, Microsoft SQL Server and click SQL Enterprise Manager to run SQL Enterprise Manager from the Microsoft SQL Server program group. Select the server then from the Tools menu select SQL Server Configuration Properties, and choose the Security page.

Which command using Query Analyzer will give you the version of SQL server and OS?

SELECT SERVERPROPERTY ('productversion'), SERVERPROPERTY ('productlevel'), SERVERPROPERTY ('edition').

What is SQL Server Agent?

SQL Server agent plays an important role in the day-to-day tasks of a database administrator (DBA). It is often overlooked as one of the main tools for SQL Server management. Its purpose is to ease the implementation of tasks for the DBA, with its full- function scheduling engine, which allows you to schedule your own jobs and scripts.

Can a stored procedure call itself or recursive stored procedure? How much level SP nesting is possible?

Yes. Because Transact-SQL supports recursion, you can write stored procedures that call themselves. Recursion can be defined as a method of problem solving wherein the solution is arrived at by repetitively applying it to subsets of the problem. A common application of recursive logic is to perform numeric computations that lend themselves to repetitive evaluation by the same processing steps. Stored procedures are nested when one stored procedure calls another or executes managed code by referencing a CLR routine, type, or aggregate. You can nest stored procedures and managed code references up to 32 levels.

What is Log Shipping?

Log shipping is the process of automating the backup of database and transaction log files on a production SQL server, and then restoring them onto a standby server. Enterprise Editions only supports log shipping. In log shipping the transactional log file from one server is automatically updated into the backup database on the other server. If one server fails, the other server will have the same DB and can be used this as the Disaster Recovery plan. The key feature of log shipping is that it will automatically backup transaction logs throughout the day and automatically restore them on the standby server at defined interval.

Name 3 ways to get an accurate count of the number of records in a table?

```
SELECT * FROM table1
SELECT COUNT(*) FROM table1
SELECT rows FROM sysindexes WHERE id = OBJECT_ID(table1) AND indid < 2
```

What does it mean to have QUOTED_IDENTIFIER ON? What are the implications of having it OFF?

When SET QUOTED_IDENTIFIER is ON, identifiers can be delimited by double quotation marks, and literals must be delimited by single quotation marks. When SET QUOTED_IDENTIFIER is OFF, identifiers cannot be quoted and must follow all Transact-SQL rules for identifiers.

What is the difference between a Local and a Global temporary table?

A local temporary table exists only for the duration of a connection or, if defined inside a compound statement, for the duration of the compound statement. A global temporary table remains in the database permanently, but the rows exist only within a given connection. When connection is closed, the data in the global temporary table disappears. However, the table definition remains with the database for access when database is opened next time.

What is the STUFF function and how does it differ from the REPLACE function?

STUFF function is used to overwrite existing characters. Using this syntax, STUFF (string_expression, start, length, replacement_characters), string_expression is the string that will have characters substituted, start is the starting position, length is the number of characters in the string that are substituted, and replacement_characters are the new characters interjected into the string. REPLACE function to replace existing characters of all occurrences. Using the syntax REPLACE (string_expression, search_string, replacement_string), where every incidence of search_string found in the string_expression will be replaced with replacement_string.

How to get @@ERROR and @@ROWCOUNT at the same time?

If @@Rowcount is checked after Error checking statement then it will have 0 as the value of @@Recordcount as it would have been reset. And if @@Recordcount is checked before the error-checking statement then @@Error would get reset. To get @@error and @@rowcount at the same time do both in same statement and store them in local variable. `SELECT @RC = @@ROWCOUNT, @ER = @@ERROR`

What is a table called, if it has neither Cluster nor Non-cluster Index? What is it used for?

Unindexed table or Heap. Microsoft Press Books and Book on Line (BOL) refers it as Heap. A heap is a table that does not have a clustered index and, therefore, the pages are not linked by pointers. The IAM pages are the

only structures that link the pages in a table together. Unindexed tables are good for fast storing of data. Many times it is better to drop all indexes from table and then do bulk of inserts and to restore those indexes after that.

Can SQL Servers linked to other servers like Oracle?

SQL Server can be linked to any server provided it has OLE-DB provider from Microsoft to allow a link. E.g. Oracle has an OLE-DB provider for oracle that Microsoft provides to add it as linked server to SQL Server group.

What is BCP? When does it used?

BulkCopy is a tool used to copy huge amount of data from tables and views. BCP does not copy the structures same as source to destination. BULK INSERT command helps to import a data file into a database table or view in a user-specified format.

How to implement one-to-one, one-to-many and many-to-many relationships while designing tables?

One-to-One relationship can be implemented as a single table and rarely as two tables with primary and foreign key relationships. One-to-Many relationships are implemented by splitting the data into two tables with primary key and foreign key relationships. Many-to-Many relationships are implemented using a junction table with the keys from both the tables forming the composite primary key of the junction table.

What is an execution plan? When would you use it? How would you view the execution plan?

An execution plan is basically a road map that graphically or textually shows the data retrieval methods chosen by the SQL Server query optimizer for a stored procedure or ad-hoc query and is a very useful tool for a developer to understand the performance characteristics of a query or stored procedure since the plan is the one that SQL Server will place in its cache and use to execute the stored procedure or query. From within Query Analyzer is an option called "Show Execution Plan" (located on the Query drop-down menu). If this option is turned on it will display query execution plan in separate window when query is ran again.

What is sub-query? Explain properties of sub-query.

Sub-queries are often referred to as sub-selects, as they allow a SELECT statement to be executed arbitrarily within the body of another SQL statement. A sub-query is executed by enclosing it in a set of parentheses.

How to implement one-to-one, one-to-many and many-to-many relationships while designing tables?

One-to-One relationship can be implemented as a single table and rarely as two tables with primary and foreign key relationships. One-to-Many relationships are implemented by splitting the data into two tables with primary key and foreign key relationships. Many-to-Many relationships are implemented using a junction table with the keys from both the tables forming the composite primary key of the junction table.

What is Collation?

Collation refers to a set of rules that determine how data is sorted and compared. Character data is sorted using rules that define the correct character sequence, with options for specifying case-sensitivity, accent marks, kana character types and character width.

What is a Linked Server?

Linked Servers is a concept in SQL Server by which we can add other SQL Server to a Group and query both the SQL Server dbs using T-SQL Statements. With a linked server, you can create very clean, easy to follow, SQL statements that allow remote data to be retrieved, joined and combined with local data. Stored Procedure sp_addlinkedserver, sp_addlinkedsrvlogin will be used add new Linked Server.

What is cursors?

Cursor is a database object used by applications to manipulate data in a set on a row-by-row basis, instead of the typical SQL commands that operate on all the rows in the set at one time. In order to work with a cursor we need to perform some steps in the following order:

- Declare cursor
- Open cursor
- Fetch row from the cursor

Process fetched row
Close cursor
Deallocate cursor

What is the use of DBCC commands?

DBCC stands for database consistency checker. We use these commands to check the consistency of the databases, i.e., maintenance, validation task and status checks.

What is User Defined Functions?

User-Defined Functions allow to define its own T-SQL functions that can accept 0 or more parameters and return a single scalar data value or a table data type.

What kind of User-Defined Functions can be created?

There are three types of User-Defined functions in SQL Server 2000 and they are Scalar, Inline Table-Valued and Multi-statement Table-valued.

What are the authentication modes in SQL Server? How can it be changed?

Windows mode and mixed mode (SQL & Windows).

To change authentication mode in SQL Server click Start, Programs, Microsoft SQL Server and click SQL Enterprise Manager to run SQL Enterprise Manager from the Microsoft SQL Server program group. Select the server then from the Tools menu select SQL Server Configuration Properties, and choose the Security page.

Where are SQL server users names and passwords are stored in sql server?

They get stored in master DB in the sysxlogins table.

What are three SQL keywords used to change or set someone's permissions?

GRANT, DENY, and REVOKE.

What are the OS services that the SQL Server installation adds?

MS SQL SERVER SERVICE, SQL AGENT SERVICE, DTC (Distribution transac co-ordinator)

What is Identity?

Identity (or AutoNumber) is a column that automatically generates numeric values. A start and increment value can be set, but most DBA leave these at 1.

How do you load large data to the SQL server database?

BulkCopy is a tool used to copy huge amount of data from tables. BULK INSERT command helps to Imports a data file into a database table or view in a user-specified format.

Can we rewrite subqueries into simple select statements or with joins?

Subqueries can often be re-written to use a standard outer join, resulting in faster performance. As we may know, an outer join uses the plus sign (+) operator to tell the database to return all non-matching rows with NULL values. Hence we combine the outer join with a NULL test in the WHERE clause to reproduce the result set without using a sub-query.

Can SQL Servers linked to other servers like Oracle?

SQL Server can be lined to any server provided it has OLE-DB provider from Microsoft to allow a link.

E.g. Oracle has a OLE-DB provider for oracle that Microsoft provides to add it as linked server to SQL Server group.

How to know which index a table is using?

```
SELECT table_name,index_name FROM user_constraints
```

How to copy the tables, schema and views from one SQL server to another?

Microsoft SQL Server 2000 Data Transformation Services (DTS) is a set of graphical tools and programmable objects that lets user extract, transform, and consolidate data from disparate sources into single or multiple destinations.

Which virtual table does a trigger use?

Inserted and Deleted.

What is the difference between a "where" clause and a "having" clause?

"Where" is a kind of restriction statement. You use where clause to restrict all the data from DB. Where clause is using before result retrieving. But Having clause is using after retrieving the data. Having clause is a kind of filtering command.

What is the basic form of a SQL statement to read data out of a table?

The basic form to read data out of table is 'SELECT * FROM table_name;'. An answer: 'SELECT * FROM table_name WHERE xyz= 'whatever';' cannot be called basic form because of WHERE clause.

What structure can you implement for the database to speed up table reads?

Follow the rules of DB tuning we have to: 1] properly use indexes (different types of indexes) 2] properly locate different DB objects across different tablespaces, files and so on. 3] create a special space (tablespace) to locate some of the data with special datatype (for example CLOB, LOB and ...)

What are the tradeoffs with having indexes?

1. Faster selects, slower updates.
2. Extra storage space to store indexes. Updates are slower because in addition to updating the table you have to update the index.

Why can a "group by" or "order by" clause be expensive to process?

Processing of "group by" or "order by" clause often requires creation of Temporary tables to process the results of the query. Which depending of the result set can be very expensive.

What is "index covering" of a query?

Index covering means that "Data can be found only using indexes, without touching the tables". 16. What types of join algorithms can you have?

What is a SQL view?

An output of a query can be stored as a view. View acts like small table which meets our criterion. View is a precompiled SQL query which is used to select data from one or more tables. A view is like a table but it doesn't physically take any space. View is a good way to present data in a particular format if you use that query quite often. View can also be used to restrict users from accessing the tables directly.

Can a view be based on another view?

Yes.

What are the advantages of views?

- Provide an additional level of table security, by restricting access to a predetermined set of rows and columns of a table.
- Hide data complexity.
- Simplify commands for the user.
- Present the data in a different perspective from that of the base table.
- Store complex queries.

What is an Oracle sequence?

A sequence generates a serial list of unique numbers for numerical columns of a database's tables.

What is a synonym?

A synonym is an alias for a table, view, sequence or program unit.

What are the types of synonyms?

There are two types of synonyms private and public.

What is a private synonym?

Only its owner can access a private synonym.

What is a public synonym?

Any database user can access a public synonym.

What are synonyms used for?

- Mask the real name and owner of an object.
- Provide public access to an object
- Provide location transparency for tables, views or program units of a remote database.
- Simplify the SQL statements for database users.

What is an Oracle index?

An index is an optional structure associated with a table to have direct access to rows, which can be created to increase the performance of data retrieval. Index can be created on one or more columns of a table.

How are the index updates?

Indexes are automatically maintained and used by Oracle. Changes to table data are automatically incorporated into all relevant indexes.

What is a Tablespace?

A database is divided into Logical Storage Unit called tablespace. A tablespace is used to grouped related logical structures together

What is Rollback Segment ?

A Database contains one or more Rollback Segments to temporarily store "undo" information.

What are the Characteristics of Data Files ?

A data file can be associated with only one database. Once created a data file can't change size. One or more data files form a logical unit of database storage called a tablespace.

Define Transaction ?

A Transaction is a logical unit of work that comprises one or more SQL statements executed by a single user.

What is Read-Only Transaction ?

A Read-Only transaction ensures that the results of each query executed in the transaction are consistent with respect to the same point in time.

What is a deadlock ? Explain .

Two processes waiting to update the rows of a table which are locked by the other process then deadlock arises. In a database environment this will often happen because of not issuing proper row lock commands. Poor design of front-end application may cause this situation and the performance of server will reduce drastically. These locks will be released automatically when a commit/rollback operation performed or any one of this processes being killed externally.

What is a Schema ?

The set of objects owned by user account is called the schema.

What is a cluster Key ?

The related columns of the tables are called the cluster key. The cluster key is indexed using a cluster index and its value is stored only once for multiple tables in the cluster.

What is Parallel Server ?

Multiple instances accessing the same database (Only In Multi-CPU environments)

What are roles? How can we implement roles ?

Roles are the easiest way to grant and manage common privileges needed by different groups of database users. Creating roles and assigning privileges to roles. Assign each role to group of users. This will simplify the job of assigning privileges to individual users.

What are Roles ?

Roles are named groups of related privileges that are granted to users or other roles.

What are the use of Roles ?

- REDUCED GRANTING OF PRIVILEGES - Rather than explicitly granting the same set of privileges to many users a database administrator can grant the privileges for a group of related users granted to a role and then grant only the role to each member of the group.
- DYNAMIC PRIVILEGE MANAGEMENT - When the privileges of a group must change, only the privileges of the role need to be modified. The security domains of all users granted the group's role automatically reflect the changes made to the role.
- SELECTIVE AVAILABILITY OF PRIVILEGES - The roles granted to a user can be selectively enable (available for use) or disabled (not available for use). This allows specific control of a user's privileges in any given situation.
- APPLICATION AWARENESS - A database application can be designed to automatically enable and disable selective roles when a user attempts to use the application.

What is Two-Phase Commit ?

Two-phase commit is mechanism that guarantees a distributed transaction either commits on all involved nodes or rolls back on all involved nodes to maintain data consistency across the global distributed database. It has two phase, a Prepare Phase and a Commit Phase.

What is Distributed database ?

A distributed database is a network of databases managed by multiple database servers that appears to a user as single logical database. The data of all databases in the distributed database can be simultaneously accessed and modified.

How can we reduce the network traffic?

- Replication of data in distributed environment.
- Using snapshots to replicate data.
- Using remote procedure calls.

What is SGA?

The System Global Area in an Oracle database is the area in memory to facilitate the transfer of information between users. It holds the most recently requested structural information between users. It holds the most recently requested structural information about the database. The structure is database buffers, dictionary cache, redo log buffer and shared pool area.

What is a shared pool?

The data dictionary cache is stored in an area in SGA called the shared pool. This will allow sharing of parsed SQL statements among concurrent users.

What is mean by Program Global Area (PGA)?

It is area in memory that is used by a single Oracle user process.

What is a data segment?

Data segment are the physical areas within a database block in which the data associated with tables and clusters are stored.

What are the disadvantages of not controlling concurrency?

1. Lost update problem
2. Temporary update (Dirty read) problem
3. Incorrect summary problem

What are serial, non serial?

A schedule S is serial if, for every transaction T participating in the schedule, all the operations of T is executed consecutively in the schedule, otherwise, the schedule is called non-serial schedule.

What are conflict serializable schedules?

A schedule S of n transactions is serializable if it is equivalent to some serial schedule of the same n transactions.

52. What is result equivalent?

Two schedules are called result equivalent if they produce the same final state of the data base.

What are conflict equivalent schedules?

Two schedules are said to be conflict equivalent if the order of any two conflicting operations is the same in both schedules.

What is a conflict serializable schedule?

A schedule is called conflict serializable if it is conflict equivalent to some serial schedule.

What is view equivalence?

Two schedules S and S' are said to be view equivalent if the following three conditions hold :

1. Both S and S' contain same set of transactions with same operations in them.
2. If any read operation read(x) reads a value written by a write operation or the original value of x the same conditions must hold in the other schedule for the same read(x) operation.
3. If an operation write1(y) is the last operation to write the value of y in schedule S then the same operation must be the last operation in schedule S'.

What is view serializable?

A schedule is said to be view serializable if it is view equivalent with some serial schedule.

What are the various methods of controlling concurrency?

1. Locking
2. Time stamp

Locking data item to prevent multiple transactions from accessing the item concurrently.

A time stamp is a unique identifier for each transaction, generated by the system.

What is a lock?

A lock is a variable associated with a data item that describes the status of the item with respect to the possible operations that can be applied to it.

What are various types of locking techniques?

1. a binary lock
2. Shared/Exclusive lock
3. Two phase locking

What is a binary lock?

A binary lock can have two states or values:

1. locked (1)
2. unlocked(0)

If locked it cannot be accessed by any other operations, else can be.

What is shared or exclusive lock?

It implements multiple-mode lock. Allowing multiple accesses for read operations but exclusive access for write operation.

Explain two phase locking?

All the locking operations must precede the first unlock operation in the transaction .It does have two phases:

1. expanding phase (Locks are issued)

2. Shrinking phase (Locks are released)

What are different types of two phase lockings (2pl)?

1. Basic
2. Conservative
3. Strict
4. Rigorous

this is the basic technique of 2pl described above.

Conservative 2pl requires a transaction to lock all the items it accesses before the transaction begins its execution, by pre-declaring its read-set and write-set. Strict 2pl guarantees that a transaction doesn't release any of its exclusive locks until after it commits or aborts. Rigorous guarantees that a transaction doesn't release any of its locks (including shared locks) until after it commits or aborts.

What is a deadlock?

Dead lock occurs when each transaction T in a set of two or more transactions is waiting for some item that is locked by some other transaction T' in the set. Hence each transaction is in a waiting queue, waiting for one of the other transactions to release the lock on them.

