**Amritsar College of Engineering and Technology**

**Department of Computer Science and Engineering**

**Relational Database Management System II**

**BTCS 602**

**Short Answer Questions**

1.What is PL/SQL ?

Ans :- PL/SQL is a procedural language which has interactive SQL, as well as procedural programming language constructs like conditional branching and iteration.

2. Differentiate between % ROWTYPE and TYPE RECORD.

Ans :- % ROWTYPE is used when a query returns an entire row of a table or view.

TYPE RECORD, on the other hand, is used when a query returns column of different tables or views.

Eg.  TYPE r\_emp is RECORD (sno smp.smpno%type,sname smp sname %type)

e\_rec smp ROWTYPE

Cursor c1 is select smpno,dept from smp;

e\_rec c1 %ROWTYPE

3. Explain uses of cursor.

Ans :- Cursor is a named private area in SQL from which information can be accessed. They are required to process each row individually for queries which return multiple rows.

4. Show code of a cursor for loop.

Ans :- Cursor declares %ROWTYPE as loop index implicitly. It then opens a cursor, gets rows of values from the active set in fields of the record and shuts when all records are processed.

Eg.  FOR smp\_rec IN C1 LOOP

totalsal=totalsal+smp\_recsal;

ENDLOOP;

5. Explain the uses of database trigger.

Ans :- A PL/SQL program unit associated with a particular database table is called a database trigger. It is used for :

1)Audit data modifications.

2)Log events transparently.

3)Enforce complex business rules.

4)Maintain replica tables

5)Derive column values

6)Implement Complex security authorizations

6. What are the two types of exceptions.

Ans :- Error handling part of PL/SQL block is called Exception. They have two types : user\_defined and predefined.

7. . Explain Raise\_application\_error.

Ans :- It is a procedure of package DBMS\_STANDARD that allows issuing of user\_defined error messages from database trigger or stored sub-program

8. . Explain two virtual tables available at the time of database trigger execution.

Ans :- Table columns are referred as THEN.column\_name and NOW.column\_name.

For INSERT related triggers, NOW.column\_name values are available only.

For DELETE related triggers, THEN.column\_name values are available only.

For UPDATE related triggers, both Table columns are available.

9. How is a process of PL/SQL compiled?

Ans :- Compilation process includes syntax check, bind and p-code generation processes.

Syntax checking checks the PL/SQL codes for compilation errors. When all errors are corrected, a storage address is assigned to the variables that hold data. It is called Binding. P-code is a list of instructions for the PL/SQL engine. P-code is stored in the database for named blocks and is used the next time it is executed.

10. Differentiate between Syntax and runtime errors.

Ans :- A syntax error can be easily detected by a PL/SQL compiler. For eg, incorrect spelling.

A runtime error is handled with the help of exception-handling section in an PL/SQL block. For eg, SELECT INTO statement, which does not return any rows

11. . Explain Commit, Rollback and Savepoint.

Ans :- For a COMMIT statement, the following is true:

* Other users can see the data changes made by the transaction.
* The locks acquired by the transaction are released.
* The work done by the transaction becomes permanent.

A ROLLBACK statement gets issued when the transaction ends, and the following is true.

* The work done in a transition is undone as if it was never issued.
* All locks acquired by transaction are released.

It undoes all the work done by the user in a transaction. With SAVEPOINT, only part of transaction can be undone.

12. What is the importance of SQLCODE and SQLERRM?

Ans :- SQLCODE returns the value of the number of error for the last encountered error whereas SQLERRM returns the message for the last error.

13. If a cursor is open, how can we find in a PL/SQL Block?

Ans :- the %ISOPEN cursor status variable can be used.

* The work done by the transaction becomes permanent.

A ROLLBACK statement gets issued when the transaction ends, and the following is true.

* The work done in a transition is undone as if it was never issued.
* All locks acquired by transaction are released.

It undoes all the work done by the user in a transaction. With SAVEPOINT, only part of transaction can be undone.  
  
14. Define Implicit and Explicit Cursors.  
Ans :- A cursor is implicit by default. The user cannot control or process the information in this cursor.  
If a query returns multiple rows of data, the program defines an explicit cursor. This allows the application to process each row sequentially as the cursor returns it.  
  
15. Explain mutating table error.  
Ans :- It occurs when a trigger tries to update a row that it is currently using. It is fixed by using views or temporary tables, so database selects one and updates the other.  
  
16. When is a declare statement required?  
Ans :- DECLARE statement is used by PL/SQL anonymous blocks such as with stand alone, non-stored procedures. If it is used, it must come first in a stand alone file.  
  
17. How many triggers can be applied to a table?  
Ans :- A maximum of 12 triggers can be applied to one table.  
  
18. What is the importance of SQLCODE and SQLERRM?  
Ans :- SQLCODE returns the value of the number of error for the last encountered error whereas SQLERRM returns the message for the last error.  
  
19. Show the two PL/SQL cursor exceptions.  
Ans :- Cursor\_Already\_Open  
Invaid\_cursor

20. What are two disadvantages of dbms?

Ans :- Two disadvantages associated with database systems are listed below.

a. Setup of the database system requires more knowledge, money, skills, and time.

b. The complexity of the database may result in poor performance.

21.Explain the difference between physical and logical data independence.

Ans :- • Physical data independence is the ability to modify the physical scheme without making it necessary to rewrite application programs. Such modifications include changing fromunblocked to blocked record storage, or from sequential to random access files.

• Logical data independence is the ability to modify the conceptual scheme without making it necessary to rewrite application programs. Such a modification might be adding a field to a record; an application program’s view hides this change from the program.

22. List five responsibilities of a database management system. For each responsibility, explain the problems that would arise if the responsibility were not discharged.

Ans :- A general purpose database manager (DBM) has five responsibilities:

* interaction with the file manager.
* integrity enforcement.
* security enforcement.
* .backup and recovery.

23. What are five main functions of a database administrator?

Ans :- Five main functions of a database administrator are:

• To create the scheme definition

• To define the storage structure and access methods

• To modify the scheme and/or physical organization when necessary

• To grant authorization for data access

• To specify integrity constraints.

24.List six major steps that you would take in setting up a database for a particular enterprise.

Ans :- Six major steps in setting up a database for a particular enterprise are:

• Define the high level requirements of the enterprise (this step generates a document known as the system requirements specification.)

• Define a model containing all appropriate types of data and data relationships.

• Define the integrity constraints on the data.

• Define the physical level.

• For each known problem to be solved on a regular basis (e.g., tasks to be carried out by clerks or Web users) define a user interface to carry out the task, and write the necessary application programs to implement the user interface.

• Create/initialize the database

25. Explain the distinctions among the terms primary key, candidate key, and superkey.

Ans :- A superkey is a set of one or more attributes that, taken collectively, allows us to identify uniquely an entity in the entity set. A superkey may contain extraneous attributes. If K is a superkey, then so is any superset of K. A superkey for which no proper subset is also a superkey is called a candidate key. It is possible that several distinct sets of attributes could serve as candidate keys. The primary key is one of the candidate keys that is chosen by the database designer as the principal means of identifying entities within an entity set.

26. Explain the difference between a weak and a strong entity set.

Ans :- A strong entity set has a primary key. All tuples in the set are distinguishable by that key. A weak entity set has no primary key unless attributes of the strong entity set on which it depends are included. Tuples in a weak entity set are partitioned according to their relationship with tuples in a strong entityset. Tuples within each partition are distinguishable by a discriminator, which is a set of attributes

27. A weak entity set can always be made into a strong entity set by adding to its attributes the primary key attributes of its identifying entity set. Outline what sort of redundancy will result if we do so.

Ans :- The primary key of a weak entity set can be inferred from its relationship with the strong entity set. If we add primary key attributes to the weak entity set, they will be present in both the entity set and the relationship set and they have to be the same. Hence there will be redundancy.

28. Explain the distinction between condition-defined and user-defined constraints. Which of these constraints can the system check automatically? Explain your answer.

Ans :- In a generalization–specialization hierarchy, it must be possible to decide which entities are members of which lower level entity sets. In a condition defined design constraint, membership in the lower level entity-sets is evaluated on the basis of whether or not an entity satisfies an explicit condition or predicate.User-defined lower-level entity sets are not constrained by a membership condition; rather, entities are assigned to a given entity set by the database user.

Condition-defined constraints alone can be automatically handled by the system. Whenever any tuple is inserted into the database, its membership in the various lower level entity-sets can be automatically decided by evaluating the respective membership predicates. Similarly when a tuple is updated, its membership in the various entity sets can be re-evaluated automatically.

29. Explain the distinction between total and partial constraints.

Ans :- In a total design constraint, each higher-level entity must belong to a lower-level entity set. The same need not be true in a partial design constraint. For instance, some employees may belong to no work-team

30. Design a relational database for a university registrar’s office. The office maintains data about each class, including the instructor, the number of students enrolled, and the time and place of the class meetings. For each student–class pair, a grade is recorded.

Ans :- Underlined attributes indicate the primary key.

student (student-id, name, program)

course (courseno, title, syllabus, credits)

course-offering (courseno, secno, year, semester, time, room)

instructor (instructor-id, name, dept, title)

enrols (student-id, courseno, secno, semester, year, grade)

teaches (courseno, secno, semester, year, instructor-id)

requires (maincourse, prerequisite)

31. Describe the differences in meaning between the terms relation and relation schema using example.

Ans :- A relation schema is a type definition, and a relation is an instance of that schema. For example, student (ss#, name) is a relation schema and ss# name

123-45-6789 Tom Jones

456-78-9123 Joe Brown

is a relation based on that schema.

32. List two reasons why null values might be introduced into the database.

Ans :- Nulls may be introduced into the database because the actual value is either unknown or does not exist. For example, an employee whose address has changed and whose new address is not yet known should be retained with a null address. If employee tuples have a composite attribute dependents, and a particular employee has no dependents, then that tuple’s dependents attribute should be given a null value.

33. Write an SQL trigger to carry out the following action: On delete of an account, for each owner of the account, check if the owner has any remaining accounts, and if she does not, delete her from the depositor relation.

Ans :- create trigger check-delete-trigger after delete on account referencing old row as row for each row

delete from depositor where depositor.customer-name not in ( select customer-name from depositor

where account-number <> orow.account-number )

end

34. Explain Timestamping:

Ans :- Basic timestamping is a concurrency control mechanism that eliminates deadlock. This method doesn’t use locks to control concurrency, so it is impossible for deadlock to occur. According to this method a unique timestamp is assigned to each transaction, usually showing when it was started. This effectively allows an age to be assigned to transactions and an order to be assigned. Data items have both a read-timestamp and a write-timestamp. These timestamps are updated each time the data item is read or updated respectively.

35. [Explain the ETL process in Data warehousing.](http://careerride.com/DB-ETL-process.aspx)

Ans :- Extraction, Transformation and loading are different stages in data warehousing.

36..[What is Data mining?](http://careerride.com/DB-data-mining.aspx)

Ans :- Data mining is a process of analyzing current data and summarizing the information in more useful manner..

37.[What are indexes?](http://careerride.com/DB-indexes.aspx)

Ans :- Index can be thought as index of the book that is used for fast retrieval of information. Index uses one or more column index keys and pointers to the record to locate record

38..[Explain the types of indexes?](http://careerride.com/DB-types-of-indexes.aspx)

Ans :- Clustered index  
Non-clustered 

39.[Define referential integrity.](http://careerride.com/DB-referential-integrity.aspx)

Ans :- It is the rules that are applied when the relationships are created. It ensures integrity of data and prevents inconsitent data into the tables

40.[Define Primary key .](http://careerride.com/DB-primary-key-and-foreign-key.aspx)

Ans :- A column or combination of columns that identify a row of data in a table is Primary Key.

41.[Define alternate key.](http://careerride.com/DB-alternate-key.aspx)

Ans :- There can be a key apart from primary key in a table that can also be a key. This key may or may not be a unique key

42.Delete vs. Truncate table.

Ans :- Delete logs the deletion of each row whereas Truncate doesn't log deleted rows in the transaction log. This makes truncate command is bit faster than Delete command.

43.Define constraints.

Ans :- Constraints enforce integrity of the database. Constraints can be of following types

* Not Null
* Check
* Unique
* Primary key
* Foreign key

44.Define stored procedure.

Ans :- Stored procedure is a set of pre-compiled SQL statements, executed when it is called in the program.

45. Define Trigger.

Ans :- Triggers are similar to stored procedure except it is executed automatically when any operations are occurred on the table.

46.what are E-R diagrams?

Ans :- The logical structure of databases is illustrated in the ER diagrams. These diagrams use symbols for representing different types of information. Boxes are used for representing entities. Rhombuses are used for representing relationships and ovals are used for representing attributes.

For a good database creation it is essential to have ER diagram. These diagrams follow top-down approach. To focus on how the database really works with all data flow and interactions, the best pattern is ER Diagram

#### 47. What are the benefits of normalizing database?

Ans :- The benefits of normalization are:

- The process of searching, sorting and creating indexes is faster  
- More tables can be derived for clear and needed tables to be designed  
- Clustering indexes can be created which provides the flexibility in fine tuning queries.  
- Less redundant data and fewer null values will make the database more compact.  
- The indexes of tables make data modification commands execution much faster.  
- If redundant data is not maintained, the execution of triggers is quicker.  
- Normalization facilitates in reducing data modification anomalies.

#### 48. What is normalization?

Ans :- Normalization is the way of organizing data in a database by removing redundancy and inconsistent dependency. Database normalization has the rules to be followed while creating databases.  
Each rule is called as normal form.

First Normal form states - Remove duplicate column and identify each set of related data with a primary key.  
Second normal form - Create relationship between master and master detail tables using foreign key.  
Third normal form - Remove the fields that do not depend on the primary key.It is the process of organizing data into related table.

To normalize database, we divide database into tables and establish relationships between the tables.   
It reduces redundancy. It is done to improve performance of query.

#### 49. What is denormalization?

Ans :- Redundant data addition for optimizing the performance of a database, denormalization is the process. The inefficiencies those inherit in RDBMS software are addressed with denormalization. It is a process of combining data into a single table. Denormalization structures are used in databases for providing quick access for a specific user needs.

#### 50. What are the lock types?

Ans :- Main lock types:

* Shared: Applied to read only operations where the data is not modified. E.g.: Select statements.
* Update: Applied to resources which can be updated. It resolves dead locks in case of multiple sessions are reading, locking or updating resources later.
* Exclusive: Used for operations involving data modification. E.g.: Insert, Update, and Delete. This ensures that multiple updates are not made to the same data at the same time.
* Intent: Establishes a lock hierarchy. E.g.: Intent shared Intent exclusive and Shared with intentexclusive.
* Schema: Used when schema dependent operations are being executed. E.g.: Schema modification and Schema stability.
* Bulk update: Used while bulk copying of data and Tablock is specified.

#### 51. What is sub-query?

Ans :- A query within a query is defined as a sub-query. A sub query is executed by enclosing it within another query. Sub queries are used for returning single row as an atomic value. A sub query is nested in DML statements, and enclosed in parenthesis.

#### 52. What is Distributed Queries?

Ans :- Distributed queries can access data from different data sources. These sources can reside on the same server or a different server. This means that distributed queries can query multiple databases.

#### 53.What is a linked server?

Ans :- A linked server allows remote access. They have the ability to issue distributed queries, update, commands, and transactions across different data sources. A linked server has an OLE DB provider and data source.

#### 54. Define Distributed Query and Linked Server

Ans :- Linked servers allow SQL Server to access data from remote data sources. One can execute queries, perform data modifications, and also execute remote procedural calls using Linked servers. Queries executed through Linked servers are termed as distributed queries. These remote data sources are connected through an OLEDB provider. The types of distributed queries which can be processed is thus based on how the OLEDB provider was designed.

#### 55. What security features are available for stored procedures?

#### Ans :- Database users can have permission to execute a stored procedure without being granted permissions to access the database objects on which the stored procedure operates.

Database users can be restricted from reading SQL command in the stored procedure by encrypting it.

#### 56. What are the steps to process a single SELECT statement?

#### Ans :- Steps a. The select statement is broken into logical units b. A sequence tree is built based on the keywords and expressions in the form of the logical units. c. Query optimizer checks for various permutations and combinations to figure out the fastest way using minimum resources to access the source tables. The best found way is called as an execution plan. d. Relational engine executes the plan and processes the data

#### 57.What are the steps to process a single SELECT statement?

Ans :- SQL Server uses the following steps to process a single SELECT statement:

1. The parser breaks SELECT statement into logical units, such as keywords, expressions, operators, and identifiers.  
2. A query tree is built.   
3. The query optimizer then determines the best steps that can return fastest result while consuming the fewest resources.   
4. The relational engine begins executing the execution plan, requesting the storage engine to pass up data.  
5. The relational engine processes the data returned from the storage engine into the result set format and returns the result set to the client.

#### 58. Explain the storage models of OLAP.

Ans :- **MOLAP Multidimensional Online Analytical processing**   
In MOLAP data is stored in form of multidimensional cubes and not in relational databases.

Advantage  
Excellent query performance as the cubes have all calculations pre-generated during creation of the cube.

Disadvantages  
It can handle only a limited amount of data. Since all calculations have been pre-generated, the cube cannot be created from a large amount of data.

It requires huge investment as cube technology is proprietary and the knowledge base may not exist in the organization.

#### 60. What is the significance of NULL value and why should we avoid permitting null values?

Ans :-Null means no entry has been made. It implies that the value is either unknown or undefined. We should avoid permitting null values because

Column with NULL values can't have PRIMARY KEY constraints.   
Certain calculations can be inaccurate if NULL columns are involved.

61. List the ACID properties. Explain the usefulness of each.

Ans :- The ACID properties, and the need for each of them are:-

• Consistency:

Execution of a transaction in isolation (that is, with no other transaction executing concurrently) preserves the consistency of the database. This is typically the responsibility of the application programmer who codes the transactions.

Atomicity:

Either all operations of the transaction are reflected properly in the database, or none are. Clearly lack of atomicity will lead to inconsistency in the database.

• Isolation:

When multiple transactions execute concurrently, it should be the case that, for every pair of transactions Ti and Tj , it appears to Ti that either Tj finished execution before Ti started, or Tj started execution after Ti finished.

Thus, each transaction is unaware of other transactions executing concurrently with it. After a transaction completes successfully, the changes it has made to the database persist, even if there are system failures.

62. Suppose that there is a database system that never fails. Is a recovery manager required for this system?

Ans :- Even in this case the recoverymanager is needed to performroll-back of aborted transactions.

63. Consider a file system such as the one on your favorite operating system. What are the steps involved in creation and deletion of files, and in writing data to a file?

Ans :- There are several steps in the creation of a file. A storage area is assigned to the file in the file system, a unique i-number is given to the file and an i-node entry is inserted into the i-list. Deletion of file involves exactly opposite steps.

64. Explain how the issues of atomicity and durability are relevant to the creation and deletion of files, and to writing data to files.

Ans :- For the file system user in UNIX, durability is important for obvious reasons, but atomicity is not relevant generally as the file system doesn’t support transactions. To the file system implementor though, many of the internal file system actions need to have transaction semantics. All the steps involved in creation/deletion of the file must be atomic, otherwise there will be unreferenceable files or unusable areas in the file system.

65. Explain the distinction between the terms serial schedule and serializable schedule.

Ans :- A schedule inwhich all the instructions belonging to one single transaction appear together is called a serial schedule. A serializable schedule has a weaker restriction that it should be equivalent to some serial schedule. There are two definitions of schedule equivalence – conflict equivalence and view equivalence.

66. What is a cascadeless schedule?Why is cascadelessness of schedules desirable? Are there any circumstances under which it would be desirable to allow noncascadeless schedules? Explain your answer.

Ans :- A cascadeless schedule is one where, for each pair of transactions Ti and Tj such that Tj reads data items previously written by Ti, the commit operation of Ti appears before the read operation of Tj. Cascadeless schedules are desirable because the failure of a transaction does not lead to the aborting

of any other transaction

67. What benefit does strict two-phase locking provide? What disadvantages result?

Ans :- Because it produces only cascadeless schedules, recovery is very easy. But the set of schedules obtainable is a subset of those obtainable from plain two phase locking, thus concurrency is reduced.

68. What benefit does rigorous two-phase locking provide? How does it compare with other forms of two-phase locking?

Ans :- Rigorous two-phase locking has the advantages of strict 2PL. In addition it has the property that for two conflicting transactions, their commit order is their serializability order. In some systems users might expect this behavior.

69. In timestamp ordering,W-timestamp(Q) denotes the largest timestamp of any transaction that executed write(Q) successfully. Suppose that, instead, we defined it to be the timestamp of the most recent transaction to execute write(Q) successfully.Would this change in wording make any difference? Explain your answer.

Ans :- It would make no difference. The write protocol is such that the most recent transaction to write an item is also the one with the largest timestamp to have done so.

70. In multiple-granularity locking, what is the difference between implicit and explicit locking?

Ans :- When a transaction explicitly locks a node in shared or exclusive mode, it implicitly locks all the descendents of that node in the same mode. The transaction need not explicitly lock the descendent nodes. There is no difference in the functionalities of these locks, the only difference is in the way they are acquired, and their presence tested.

71. Use of multiple-granularity locking may require more or fewer locks than an equivalent system with a single lock granularity. Provide examples of both situations, and compare the relative amount of concurrency allowed.

Ans :- If a transaction needs to access a large a set of items, multiple granularity locking requires fewer locks, whereas if only one item needs to be accessed, the single lock granularity system allows this with just one lock. Because all the desired data items are locked and unlocked together in the multiple granularity scheme, the locking overhead is low, but concurrency is also reduced.

72. Define Validation

Ans :- If the probability that two concurrently executing transactions conflict is low, this protocol can be used advantageously to get better concurrency and good response times with low overheads. Not suitable under high contention, when a lot of wasted work will be done.

73. Define Multiversion timestamp ordering

Ans :- Use if timestamp ordering is appropriate but it is desirable for read requests to never wait. Shares the other disadvantages of the timestamp ordering protocol.

74. Under what conditions is it less expensive to avoid deadlock than to allow deadlocks to occur and then to detect them?

Ans :-: Deadlock avoidance is preferable if the consequences of abort are serious (as in interactive transactions), and if there is high contention and a resulting high probability of deadlock.

75. Explain the difference between the three storage types—volatile, nonvolatile, and stable—in terms of I/O cost.

Ans :- Volatile storage is storage which fails when there is a power failure. Cache, main memory, and registers are examples of volatile storage. Nonvolatile storage is storage which retains its content despite power failures. An example ismagnetic disk. Stable storage is storage which theoretically survives any kind of failure (short of a complete disaster!). This type of storage can only be approximated by replicating data. In terms of I/O cost, volatile memory is the fastest and non-volatile storage is typically several times slower. Stable storage is slower than non-volatile storage because of the cost of data replication.

76. Why and how Stable storage cannot be implemented.

Ans :- Stable storage cannot really be implemented because all storage devices are made of hardware, and all hardware is vulnerable to mechanical or electronic device failures.

77. Explain how database systems deal with this problem.

Ans :- Database systems approximate stable storage by writing data to multiple storage devices simultaneously. Even if one of the devices crashes, the data will still be available on a different device. Thus data loss becomes extremely unlikely.

78 .What is Checkpoint?

Ans :- Checkpointing is done with log-based recovery schemes to reduce the time required for recovery after a crash. If there is no checkpointing, then the entire logmust be searched after a crash, and all transactions undone/redone from the log. If checkpointing had been performed, then most of the log-recordsprior to the checkpoint can be ignored at the time of recovery.

79. Compare the shadow-paging recovery scheme with the log-based recovery schemes in terms of ease of implementation and overhead cost.

Ans :- The shadow-paging scheme is easy to implement for single-transaction systems, but difficult for multiple-transaction systems. In particular it is very hard to allow multiple updates concurrently on the same page. Shadow paging could suffer from extra space overhead, but garbage collection can take care of that. The I/O overhead for shadow paging is typically higher than the log based schemes, since the log based schemes need to write one record per update to the log, whereas the shadow paging scheme needs to write one block per updated block.

80. Explain how the buffer manager may cause the database to become inconsistent if some log records pertaining to a block are not output to stable storage before the block is output to disk.

Ans :- If a data item x is modified on disk by a transaction before the corresponding log record is written to stable storage, then the only record of the old value of x is in main memory where it would be lost in a crash. If the transaction had not yet finished at the time of the crash, an unrecoverable inconsistency will result

81. Discuss the relative advantages of centralized and distributed databases.

Ans :- • A distributed database allows a user convenient and transparent access to data which is not stored at the site, while allowing each site control over its own local data. A distributed database can be made more reliable than a centralized system because if one site fails, the database can continue functioning, but if the centralized system fails, the database can no longer continue with its normal operation. Also, a distributed database allows parallel execution of queries and possibly splitting one query into many parts to increase throughput.

• A centralized system is easier to design and implement. A centralized system is cheaper to operate because messages do not have to be sent.

82 .Define Fragmentation

Ans :- With fragmentation transparency, the user of the system is unaware of any fragmentation the system has implemented. A user may formulate queries against global relations and the system will perform the necessary transformation to generate correct output.

83 Define Replication transparency

Ans :- With replication transparency, the user is unaware of any replicated data. The systemmust prevent inconsistent operations on the data. This requires more complex concurrency control algorithms

84. How might a distributed database designed for a local-area network differ

from one designed for a wide-area network?

Ans :- Data transfer on a local-area network (LAN) is much faster than on a wide-area network (WAN). Thus replication and fragmentation will not increase throughput and speed-up on a LAN, as much as in aWAN. But even in a LAN, replication has its uses in increasing reliability and availability.

85. Explain the difference between data replication in a distributed system and the maintenance of a remote backup site.

Ans :- In remote backup systems all transactions are performed at the primary site and the data is replicated at the remote backup site. The remote backup site is kept synchronized with the updates at the primary site by sending all log records. Whenever the primary site fails, the remote backup site takes over processing. The distributed systems offer greater availability by having multiple copies of the data at different sites whereas the remote backup systems offer lesser availability at lower cost and execution overhead.

86. What are the two types of time, and how are they different? Why does it make sense to have both types of time associated with a tuple?

Ans :- A temporal database models the changing states of some aspects of the real world. The time intervals related to the data stored in a temporal database may be of two types - valid time and transaction time. The valid time for a fact is the set of intervals during which the fact is true in the real world. The transaction time for a data object is the set of time intervals during which this object is part of the physical database. Only the transaction time is system dependent and is generated by the database system.

87. Suppose you have a relation containing the x, y coordinates and names of restaurants. Suppose also that the only queries that will be asked are of the following form: The query specifies a point, and asks if there is a restaurant exactly at that point. Which type of index would be preferable, R-tree or B-tree? Why?

Ans :- The given query is not a range query, since it requires only searching for a point. This query can be efficiently answered by a B-tree index on the pair of attributes (x, y).

88. List three factors that need to be considered in query optimization for mobile computing that are not considered in traditional query optimizers.

Ans :- Themost important factor influencing the cost of query processing in traditional database systems is that of disk I/O. However, in mobile computing, minimizing the amount of energy required to execute a query is an important task of a query optimizer. To reduce the consumption of energy (battery power), the query optimizer on amobile computer must minimize the size and number of queries to be transmitted to remote computers as well as the time for which the disk is spinning.

89. Explain how a TP monitor manages memory and processor resources more effectively than a typical operating system.

Ans :- In a typical OS, each client is represented by a process, which occupies a lot of memory. Also process multi-tasking over-head is high. A TP monitor is more of a service provider, rather than an environment for

executing client processes. The client processes run at their own sites, and they send requests to the TP monitor whenever they wish to avail of some service. The message is routed to the right server by the TP monitor, and the results of the service are sent back to the client.

90. If the entire database fits in main memory, do we still need a database system to manage the data? Explain your answer.

Ans :- Even if the entire database fits in main memory, a DBMS is needed to perform tasks like concurrency control, recovery, logging etc, in order to preserve ACID properties of transactions.

91. Is a high-performance transaction system necessarily a real-time system? Whyor why not?

Ans :- A high-performance system is not necessarily a real-time system. In a high performance system, the main aim is to execute each transaction as quickly as possible, by having more resources and better utilization. Thus average speed and response time are themain things to be optimized. In a real-time system, speed is not the central issue. Here each transaction has a deadline, and taking care that it finishes within the deadline or takes as little extra time as possible, is the critical issue.  
  
92. Explain why it may be impractical to require serializability for long-duration transactions.

Ans :- In the presence of long-duration transactions, trying to ensure serializability has several drawbacks:-

a. With a waiting scheme for concurrency control, long-duration transactions will force long waiting times. This means that response time will be high, concurrencywill be low, so throughput will suffer. The probability of deadlocks is also increased.

b. With a time-stamp based scheme, a lot of work done by a long-running transaction will be wasted if it has to abort.

93. What is difference between Restoring and Recovery of database?  
Ans :- Restoring means copying the database object from the backup media to the destination where actually it is required where as recovery means to apply the database object copied earlier (roll forward) in order to bring the database into consistent state.

94. What is the difference between HOTBACKUP and RMAN backup?  
Ans :- For hotbackup we have to put database in begin backup mode, then take backup where as RMAN would not put database in begin backup mode. RMAN is faster can perform incremental (changes only) backup, and does not place tablespace in hotbackup mode

95. What happens actually in case of instance Recovery?

Ans :- Cache recovery: Changes being made to a database are recorded in the database buffer cache as well as redo log files simultaneously. When there are enough data in the database buffer cache, they are written to data files. If an Oracle instance fails before these data are written to data files, Oracle uses online redo log files to recover the lost data when the associated database is re-started. This process is called cache recovery.  
Transaction recovery: When a transaction modifies data in a database (the before image of the modified data is stored in an undo segment which is used to restore the original values in case the transaction is rolled back). At the time of an instance failure, the database may have uncommitted transactions. It is possible that changes made by these uncommitted transactions have gotten saved in data files. To maintain read consistency, Oracle rolls back all uncommitted transactions when the associated database is re-started. Oracle uses the undo data stored in undo segments to accomplish this. This process is called transaction recovery.

96. What is the benefit of running the DB in archivelog mode over no archivelog mode?  
Ans :- When a database is in no archivelog mode whenever log switch happens there will be a loss of some redoes log information in order to avoid this, redo logs must be archived. This can be achieved by configuring the database in archivelog mode.

97. If an oracle database is crashed? How would you recover that transaction which is not in backup?

Ans :- If the database is in archive log we can recover that transaction otherwise we cannot recover that transaction which is not in backup.

98. What is NoSQL?

Ans :- It stands for Not Only SQL and provides an alternative to relational databases. Instead of tabular data stores, they use graph stores, key-value stores, document databases, and wide-column stores. It is popular in the agile development world as developers don’t have to finalize the data model before storing information.

### 99. Why domain is of high importance?

Ans :- Domain describes possible values grouped together that can be given for an attribute. It is considered the same way as a constraint on the value of attribute. A domain can be attached to an attribute but only if the attribute is an element of specified set. For example: XYZ doesn’t fulfill the domain constraint but the integer value as 899 fulfills the criteria of domain constraint. Hence, domain is of high importance.

### 100. What are constraints in database?

Ans :- Constraints are kind of restrictions that are applied to the database or on the domain of an attribute. For example an integer attribute is restricted from 1-10 and not more than that. They provide the way to implement the business logic and the rules in database. In database it can be implemented in the form of check constraints that checks for the rules that haven’t been followed by the programmer. Constraint also used to restrict the data that can be stored in the relations. Domain constraint can be applied to check the domain functionality and keep it safe..