Relational Database Management Systems – April 2012

I. Section-A: 5 X 4 =20 Marks

1. Write about evolution of DBMS

The sophistication of modern database technology is the result of a decades-long evolution in data processing and information management. Tugged on one side by the needs and demands of management and restrained on the other by the limits of technology, data access technology has developed from the primitive methods of fifties of the powerful, integrated systems of today.

Management's expectations have grown in parallel to the evolution of technology. The early data processing systems performed clericals tasks that reduced paper handling. More recent systems have expanded to production and management of information, which has come to be viewed as a vital company resource. Today the most important function of database systems is to provide the basis for corporate management information systems.

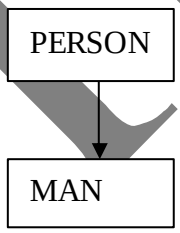
Implementation of technological change has been guided by genuine business needs. Management will only authorize a new computer system when it sees a clear benefit to offset the system's cost. And despite pitfalls and risks, benefits have been realized in many cases. Moreover, the end is not yet in sight and won't be for some time to come. Moreover, the end is not yet in sight and won't be for some time to come. New technology, such as object-oriented databases and client/server platforms, addresses new problems and will result in more powerful systems for the future.

The close relationship between database technology and business needs may be easier to understand if a closer look at the experience of International Product Distribution is taken into consideration.

2. Differentiate between Generalization and Specialization.

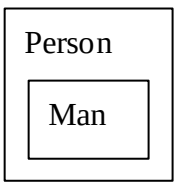
Generalization: An object set that is a superset of another object set.
Specialization: An object set that is a subset of another object set.

Ex:



Man is contained with person. This means that every man also a person, i.e, Man is continued within person. Then man is the specialization (subset) of Person.
On other hand, person is a generalization (superset) of Man.

An alternate representation of specialization is



Prasanth Kumar K
(Head-Dept of Computers, IIMC)

3. Write about Group by clause and Having clause.

Group by: Group By clause indicates that rows should be grouped on a common value of specified column or columns.

```
Select ename, deptno, avg(sal)
From emp
Group by deptno;
```

Having: Having clause places conditions on groups.

```
Select ename, deptno, avg(sal)
From emp
Group by deptno
Having sal>5000;
```

4. What is database integrity?

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

5. How do you classify the DBMS feature requirement?

The necessary security and control functions are satisfied by the following features:

1. Access control provided for programs and users.
2. Password-protected security tables.
3. Ability to abort transactions in process.
4. Automatic logging of failed attempts to access data.
5. Utilities to create and maintain password tables.

By defining required features that can be verified, the analyst avoids the ambiguities of attempting to evaluate whether security capability is good, bad or somewhere between.

Mandatory Feature: A DBMS feature that must be provided.
Important Feature: A DBMS feature that is not mandatory but makes the DBMS more attractive.
Optional Feature: A DBMS feature that is of secondary importance.
Unnecessary Feature: A DBMS feature that contributes nothing to the value of the DBMS.
Undesirable Feature: A DBMS feature that detracts from its value to the organization.

6. What is database partitioning?

Database is distributed such that there is no overlapping, or replication of data maintained at the various locations. Since there is no duplication of data, costs associated with storing and maintaining redundant data are avoided. Data availability may be limited, however if the same segment of data is used at more than one location. Reliability may also be affected, since a failure of one computer system means that the data which are stored at the location are not available to users anywhere in the system.

Prasanth Kumar K
(Head-Dept of Computers, IIMC)

Because of distributed environment of the DDBMS allows the database to be physically partitioned, data security may also be improved, particularly when the partitioned segments have different security needs.

The most straightforward way of implementing a partitioned database is to treat it as a series of independently operated database system with remote-access capability. An alternative, which is somewhat more complex is one in which the database systems are integrated such that a single query by the user may require access to more than one of the database systems. While the underlying complexities should be transparent to the user, the actual operations may be quite involved. For example, a relational JOIN operation whose result requires tables which are maintained in two distinct database locations.

7. What are the functions of DBA?

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

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

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

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

1. Analysis and routing of trouble reports: - A formal trouble-reporting system was established in order to report all errors to the DBA. Trouble reports are analyzed to determine the likely cause of each reported problem. The reports are then routed to the appropriate manager or user group for disposition. Each trouble report contains a complete log and descriptive information. Each report requires a formal response to the report's initiator specifying how the problem has been resolved.
2. Monitoring of hardware and software: The status of all hardware and software is regularly monitored, and reports of failures and consequent action are made to appropriate mangers and user groups. Periodic analysis of hardware and software requirements is made, forming the basis for decisions on replacement and upgrading, including needs for additional database storage media.

Prasanth Kumar K
(Head-Dept of Computers, IIMC)

3. Testing: Performance acceptance testing is conducted to evaluate all new procedures, software, and hardware. Structural and consistency checks of the database are conducted on a regular basis.
4. Security: security classifications are implemented that identify which user groups are authorized to access specific data elements sin the database and what actions may be performed thereon. Computer operations area frequently monitored to assure that these access controls are functioning in the intended way.
5. Backup and recovery: Backup and recovery procedures are tested regularly to assure their effectiveness in restoring the database after any disruption of service/ a disaster plan has been drawn up and is tested periodically to make sure it works.
6. Performance evaluation: Priorities have assigned to activities that compete for database resources, such as processing transactions, generating reports, and processing queries, system performance is monitored by collecting statistics on transaction volume, response time, error rates, and hardware utilization. Input is elicited from system users to monitor their satisfaction with the system's performance. Database size and growth is also tracked. File expansion programs are run and database reorganizations are performed as necessary. Activity logs and abnormal termination logs are reviewed and summaries prepared for management evaluation.

8. Differentiate between Authorization and Authentication.

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

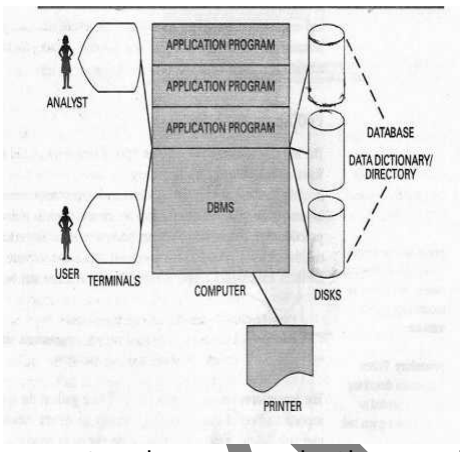
Authorization : - Authorization is the process of verifying the privileges of a valid user. Authorization in SQL is implemented with views. A view is a means of providing a user with a personalized model of the database. It is also a useful way of limiting a user's access to various positions of the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

II. Section-B: 5 X 10 =50 Marks

9 (a). What is DBMS? Explain the components of database System.

DBMS: Database Management System is a collection of interrelated data and set of programs to access those data. The DBMS is a general purpose software system that facilitates the process of defining constructing and manipulating databases for various applications.

Components of Database System:



A complete database system in an organization consists of four components.

(i). Hardware: The hardware is the set of physical devices on which a database resides. It consists of one or more computers, disk drives, CRT terminals, printers, tape drivers, connecting cables, etc.

The computers used for processing the data in the database may be mainframe, mini computers or personal computers. Mainframe and mini computers have traditionally been used on a stand-alone basis to support multiple users accessing a common database. Personal computers are often used with stand-alone databases controlled and accessed by a single user.

Disk drivers are the main storage mechanism for databases. Desktop computers, CRT terminals and printers are used for entering and retrieving information from the database.

The success of the database system has been heavily dependent on advances in hardware technology. A very large amount of main memory and disk storage is required to maintain and control the huge quantity of data stored in a database.

(ii). Software: A database system includes two types of software:

- General Purpose database management software usually called the database management system (DBMS).
- Application software that uses DBMS facilities to manipulate the database to achieve a specific business functions.

Application software is generally written by programmers to solve a specific company problem. It may be written in languages like COBOL or C or it may be written in a language supplied by DBMS like SQL. Application software uses the facilities of the DBMS to access and manipulate data in the database providing reports or documents needed for the information and processing needs of the company.

The DBMS is system software similar to an operating system. It provides a number of services to end users and programmers.

DBMS typically provides most of the following services.

- A central data definition and data control facility known as a data dictionary/directory or catalog.

- Data security and integrity mechanisms.
- Concurrent data access for multiple users.
- User-oriented data query, manipulation and reporting capabilities.
- Programmer-oriented application system development capabilities.

(iii). Data: No database system can exist without data. Data can be collected and entered into the database according to the defined structure.

(iv). People: Two different types of people concerned with the database.

They are:

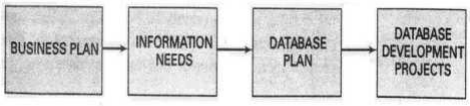
- Users: Executives, Managers, Staff, Clerical personnel.
- Practitioners: Database Administrators, Programmers.

9(b). Explain the strategic database planning.

Database Planning is a strategic corporate effort to determine information needs for an extended period. A successful database planning project will precede operational projects to design and implement new databases to satisfy the organization's information needs.

The Need for Database Planning: Database planning has significant advantages:

- It expresses management's current understanding of the information resource.
- It identifies and justifies resource requirements.
- It identifies opportunities for effective resource management including collaboration among departments or divisions within the organizations.
- It specifies action pans for achieving objectives.
- It can provide a powerful stimulus and sense of direction to employees at all levels, focusing their efforts, increasing their productivity and making them feel that they are a genuine part of the enterprise.



The Database Planning Project: Strategic Database Planning is initiated by senior management. They allocate resources and identify personnel to participate in the project. With their commission from management, team members have resources needed to carry out a successful project.

The project team should be extensive experience in information systems and other functional areas of the company. A group of four full-time members, two from information systems and two acquainted with most other areas of the company. All team members should be skilled and respected employees, since their work will have a major impact on the organization for many years. If they are not skilled in a methodology for carrying out the study, an outside consultant should be employed as an advisor to train the team in a suitable methodology. The project team leader should be a consultant but a permanent employee and possibly the head of the database administration.

During the project, the team interacts with senior managers from all the primary user areas. The senior end users identify the principal processes, activities, and entities used in manual or automated information processing. The project team synthesizes these data into a corporate information model included as part of the comprehensive database plan.

A report covering at least the next five should be delivered to senior management. This report will include analyses of the following:

- Information needs of the functional areas.

- Information needs of different management levels.
- Information needs of the geographical locations.
- A model of this information needs.
- Anticipated data volumes by geographical location projects for the period under study.
- A preliminary estimate of costs associated with system upgrades.
- Recommendations for detailed development of new or enhanced databases with schedules.

10(a). State all the Normal-Forms. Compare 3NF with Boyce-Codd Normal Form.

Normalization: Normalization is a process of evaluating and correcting table structures to minimize data redundancies, thereby reducing the likelihood of data anomalies. The normalization process involves assigning attributes to the tables based on the concept of Relational Data Model.

The objective of normalization is to ensure that each table conforms to the concept of well-formed relations. Normalized tables have the following characteristics.

- Each table represents a single subject. For example, a course table will contain only data that directly pertains to courses. Similarly, a student table will contain only student data.
- No data item will be unnecessarily stored in more than one table. The reason for this requirement is to ensure that the data are updated in only one place.
- All non-prime attribute in a table are dependent on the primary key. The reason for this requirement is to ensure that the data are uniquely identifiable by a primary key value.
- Each table is void of insertion, update or deletion anomalies. This is to ensure the integrity and consistency of the data.

1st Normal Form: A table is in first normal form if all the key attributes have been defined and it contains no repeating groups

2nd Normal Form: A table is in second normal form (2NF) if and only if it is in 1NF and every non key attribute is fully functionally dependent on the whole of the primary key (i.e. there are no partial dependencies).

3rd Normal Form:

A relation is said to be in 3NF if
(1). It is in 2NF
(2). There are no transitive dependencies

Step1: Identify each new determinant.

For every transitive dependency, write its determinant as a PK for a new table. A determinant is any attribute whose value determines other values within a row. If there are three dependencies, then there are three different determinants.

Write the determinant for this transitive dependency as:
JOB_CLASS

Step2: Identify the dependent attributes.
Identify the attributes that are dependent on each determinant identified in Step1 and identify the dependency.
JOB_CLASS -> CHG_HOUR
Step3: Remove the Dependent attributes from Transitive dependency.

Prasanth Kumar K
(Head-Dept of Computers, IIMC)

Eliminate all dependent attributes in the transitive dependencies from each table that have such transitive relationship.

Eliminate CHG_HOUR from the employee table to leave the EMPLOYEE table dependency definition as:

EMP_NUM - > EMP_NAME, JOB_CLASS

Note that the JOB_CLASS remains in the EMPLOYEE table to serve as the FK.

Boyce-Codd Normal Form:

A Relation is said to be in BCNF if

1. It is in 2NF
2. Every determinant is a key.

When a table consists of one attribute primary key, the 3NF and BCNF are equivalent. Most designers consider BCNF to be a special case of 3NF. In other words, a table is in 3NF when it is in 2NF and there are no transitive dependencies.

Another situation may occur because of BCNF. A non-key attribute determines the part of primary key.

1) A+B - > C, D
   C- >B

The table structure has no partial dependencies, nor it contain transitive dependencies. The condition C- >B indicates that a non-key attribute determines part of the primary key.

2) Another example:
   STUDENT(stuid,classcode,enrollment)
   STAFF(staffed,classcode)

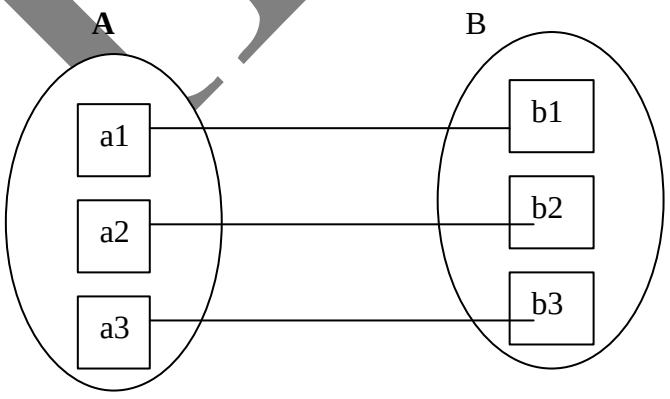
   The dependencies are:
   Stuid + staffid - > classcode,enrollmenr
   Classcode - >staffid

10(b) Define Cardinality. Explain the basic relationship cardinalities

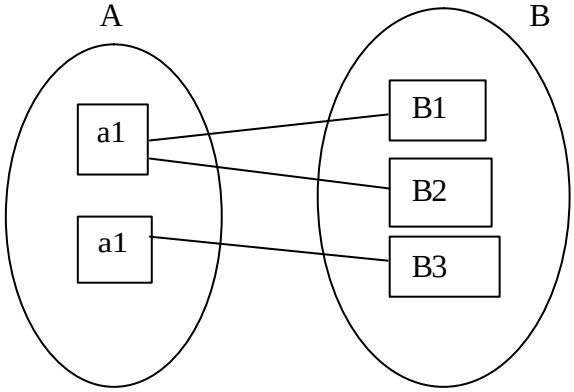
Mapping cardinalities express the number of entities to which another entity can be associated via a relationship set.

For a binary relationship set R between entity sets A and B, the mapping cardinalities must be one of the following.

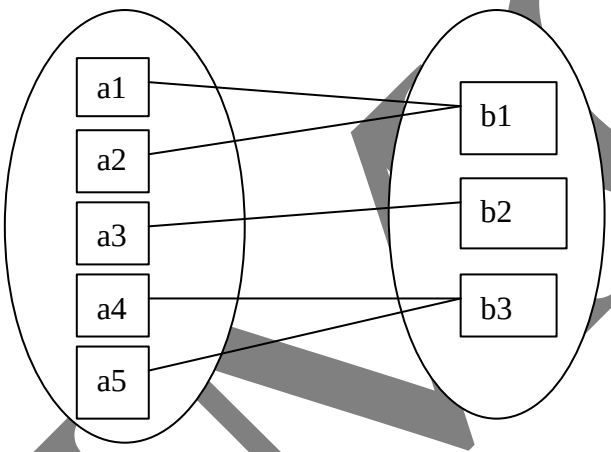
(i). <u>One to One:</u> An entity in A is associated with at most one entity in B and an entity in B is associated with at most one entity in A.



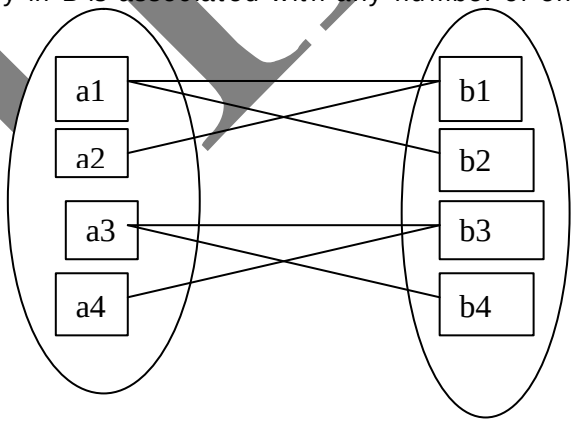
(ii). <u>One to many:</u> An entity in A is associated with any number of entities in B. An entity in B, however can be associated with at most one entity in A.



(iii). <u>Many to One:</u> An entity in A is associated with at most one entity in B and an entity in B, however can be associated with any number of entities in A.



(iv).<u>Many to many:</u> An entity in A is associated with any number of entities in B and an entity in B is associated with any number of entities in A.



11(a) Explain in detail server data manipulation.

SQL Server and Oracle support standard SQL features as well as a number of useful enhancements in their interactive languages. In addition to these, both provide control structures allowing for batch-style programs in which sets or batches of statements can be executed without interruption. These control structures are part

of SQL Server's Control-of-flow and Oracle PL/SQL languages and include conditional statements and statements for building loops for iterative processing. Variables can also be defined for use in storing and calculating values.

Control-of-flow (or) PL/SQL:

1) BEGIN----------END statements, which define blocks of SQL. Statements to be treated as units for execution.
2) If ----------ELSE Statements for condition execution.
3) WHILE Statements for repeated orientation execution.
4) BREAK and CONTINUE Statements for early exit from WHILE loops.
5) DECLARE Statements allowing defining of local variables.
6) RETURN Statements
7) PRINT Statement provide for messages to be sent to the user.
8) Comments which allow internal documentation to be included with programs

1. BEGIN----------END:- Condition (IF) and interactive (WHILE) Statements control
the execution of a single statement (or) a block of statements.

```
BEGIN
SQL Statement
----------
SQL Statement
END
```

A Statement block is either a single SQL Statement without delimiters (or) set of two (or) more SQL Statements is limited by BEGIN----------END.

2. IF-----ELSE:

```
If<Conditional EXP>
<Statement block>
ELSE
<Statement block>
```

If can be used to determine conditions relative to the relation as a whole before applying updater to the relation as a whole.

3. WHILE Statement: A Statement which uses a condition to control the iterative execution of a Statement block

```
WHILE <Conditional EXP>
<Statement block>
```

This statement causes the statement block to be repeatedly executed, as long as the conditional evaluated to true.

```
WHILE( value> 0 )
BEGIN
STATEMENT1
STATEMENT2
END
```

4. BREAK AND CONTINUE: BREAK causes execution control to exit from an iterative while loop. CONTINUE causes execution control to return to the first statement in a WHILE LOOP.

5. DECLARE: In SQL Server, a local variable is declared by giving it a name beginning with @

```
declare @a int
```

6. RETURN: When the last statement of a procedure is executed, the procedure is finished and returns control to the calling procedure. The RETURN statement causes the procedure to terminate and immediately return the control to the calling procedure.

7. PRINT: A PRINT statement is used to print the output.

PRINT sum - > This statement is used to print the value of sum on the screen.

8. Comment: Comments are used for internal documentation of the program.
/* This is my first program */

11(b) Explain various DML Commands used in SQL with examples.

DML Commands:
1. Insert
2. Update
3. Delete
4. Select

(1) Insert: To get data into a database, we need to use the 'insert' statement. The general syntax is:
INSERT INTO <table-name> (<column1>,<column2>,<column3>,...)
VALUES (<column-value1>,<column-value2>,<column-value3>);

Eg. Insert into emp(empno,ename,sal) values(200,'abc',10000);

(2) Update: The update statement is used for changing records.
The general syntax is:
UPDATE <table-name>
SET <column> = <value>
[where condition]

E.g. Update emp
set sal=sal+1000;

The above statement increases the salary of all employees by Rs.1000.

e.g. Update emp
set sal=sal+1500
where job='clerk';

(3) Delete: The delete statement is used to remove rows from the table.
The general syntax:
Delete from <table-name>
[Where condition]

E.g. Delete from emp where sal>=50000;

(4)Select: The basic structure of SQL query consists of three clauses.

1) Select: The select clause list columns desired in result of the query.
2) Form: The form clause lists one or more tables to be referenced by the query.
3) Where: The where clause contains a condition for selecting rows from the table listed in from Clause.

A SQL query has the following form:

Select A1,A2…….An
From t1,t2,…….tn
Where Condition;
Here A1,A2……..An are attributes
t1,t2…….tn are tables.

E.g. Select eno from emp
Where sal>15000;

12(a) What is DBMS? Explain the components of database System.

Answer same as 9(a)

12(b) What is database recovery? Write about the sources of failure and recovery procedure.

Database Recovery:

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

Sources of Failure:

A useful classification of failure types includes the following:System errors: the system has entered an undesirable state, such as deadlock, which prevents the program from continuing with normal processing. This type of failure may or may not result in corruption of data files.

Hardware failures: Two of the most common types of hardware failure and loss of transmission capability over a transmission link. In the former case, the cause usually results from the disk read/write head coming in physical contact with the disk surface.

Logical errors: Bad data or missing data are common conditions that may preclude a program's continuing with normal execution.

Recovery Procedures: -

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

Aborted: A transaction may not always complete its process successfully. To be sure the incomplete transaction will not affect the consistent state of the database, such transactions must be aborted, and resorting the database to the state it was in

before the transaction in question began execution. Such restoration is achieved by rollback.

Committed: A transaction that successfully completes its processing is said to be committed. A committed transaction always leaves the database in a new consistent state.

The LOG is a history of all the changes made to the database as well as the status of each transaction. LOG information is stored on a mythical "stable storage" that survives all failures.
A recovery strategy can be pursued by one of two approaches
1) Logging with deferred updates: In this technique, all data is stored in database. after committed statement, then use redo ( ) operation. To perform commit statement again.

2) logging with Immediate database Modification: In this technique, all data is updated on database before commit statement is performed. When crash occurs before commit statement, then Undo operation is performed.

UNDO (Ti): which restores the value of all data items updated by transaction (Ti) to the old values.
REDO (Ti): which sets the values of all data items updated by transaction (Ti) to the new values.

These two operations are important in order to guarantee correct behaviour even if a failure occurs during the recovery process.

13(a) Explain distributed database system design.

The design of a distributed database system can be a complex task. Careful consideration must be given to the objectives and strategies to be served by the design and parallel decisions must be made as to how the data are to be distributed among the various network sites.

Strategies and Objectives: - Some of the strategies and objectives that are common to most implementation of distributed database systems are:

1) Location Transparency: - Location transparency enables a user to access data without knowing, or being concerned with, the site at which the data reside. The location of the data is hidden from the user.
2) Replication Transparency: - Replication transparency means that when more than one copy of the data exists, one copy must be chosen when retrieving data, and all copies must be updated when changes are made. Choosing one copy of the data for retrieval and always ensuring that all copies of the data are updated can be a burden on users. A DBMS should handle all such requirements, thereby freeing the user to concentrate on information needs.
3) Configuration Independence: - Configuration independence enables the organization to add or replace hardware without changing the existing software components of the DBMS. Configuration independence results in a system that is expandable when its current hardware is saturated.
4) No homogeneous DBMSs: - It is sometimes desirable to integrate database maintained by different DBMS on different computers. Often the DBMS are supplied by different vendors and may support different data models. One approach to integrating this database is to provide a single user interface that can be used to access the data maintained by the non-homogenous DBMS. The different data

models supported by the non-homogenous DDBMS are hidden from the user by this single, system wide interface.

5) Data Replication: - Data replication occurs if the system maintains several identical copies of a relation, R, with each copy being stored at a different site. Typically replication is introduced to increase the availability of the system: When a copy is unviable due to site failure(s), it should be possible to access another copy.

13(b) Explain the evolution model in DBMS.

The acquisition of DBMS reflects a major commitment by the firm. Methods of collecting and recording data may be affected. The acquisition of DBMS represents a commitment to using information to improve the way in which the firm does business. These commitments require that the process of choosing a DBMS include consideration of important DBMS features in a rational and consistent manner. The attainment of these objectives can be aided by the use of a formal methodology. The two formal methodologies are:

1. Scoring Model – It is easy to apply and has been widely used to aid in the DBMS selection process.
2. DEA – Data Envelopment Analysis is more powerful. It is also easy to apply.

Scoring Model: The Scoring Model has been widely used in practice. Many firms limit the recommended classification of requirements to just two – mandatory and desirable. As nearly as possible, the verification of mandatory requirements should not be subject to judgment or opinion. Desirable requirements may often do include features that are harder to measure.

Data Envelopment Analysis: DEA is a powerful methodology for evaluating decision alternatives such as DBMS selection, especially when there are several criteria bearing upon the selection. A basic understanding of the concepts and access to lineat programming software-widely available microcomputer use – is all that is required.