1. Who is a DBA? What are the responsibilities of a DBA? April/May-2011

A database administrator (short form DBA) is a person responsible for the design, implementation, maintenance and repair of an organization's database. They are also known by the titles Database Coordinator or Database Programmer, and is closely related to the Database Analyst, Database Modeller, Programmer Analyst, and Systems Manager.

The role includes the development and design of database strategies, monitoring and improving database performance and capacity, and planning for future expansion requirements. They may also plan, co-ordinate and implement security measures to safeguard the database

2. What is a data model? List the types of data model used. April/May-2011

A **database model** is the theoretical foundation of a database and fundamentally determines in which manner data can be stored, organized, and manipulated in a database system. It thereby defines the infrastructure offered by a particular database system. The most popular example of a database model is the relational model.

types of data model used
- Hierarchical model
- Network model
- Relational model
- Entity-relationship
- Object-relational model
- Object model

3.Define database management system.

Database management system (DBMS) is a collection of interrelated data and a set of programs to access those data.

4.What is data base management system?

- A database management system (DBMS) is a software package with computer programs that control the creation, maintenance, and the use of a database.
- It allows organizations to conveniently develop databases for various applications by database administrators (DBAs) and other specialists.
- A database is an integrated collection of data records, files, and other database objects.
- A DBMS allows different user application programs to concurrently access the same database. DBMSs may use a variety of database models, such as the relational model or object model, to conveniently describe and support applications.
- It typically supports query languages, which are in fact high-level programming languages, dedicated database languages that considerably simplify writing database application programs.
- Database languages also simplify the database organization as well as retrieving and presenting information from it.
- A DBMS provides facilities for controlling data access, enforcing data integrity, managing concurrency control, recovering the database after failures and restoring it from backup files, as well as maintaining database security.

5.List any eight applications of DBMS.

a) Banking
b) Airlines
c) Universities

d) Credit card transactions
e) Tele communication
f) Finance
g) Sales
h) Manufacturing
i) Human resources

6.What are the disadvantages of file processing system?

The disadvantages of file processing systems are
a) Data redundancy and inconsistency
b) Difficulty in accessing data
c) Data isolation
d) Integrity problems
e) Atomicity problems
f) Concurrent access anomalies

7.What are the advantages of using a DBMS?

The advantages of using a DBMS are
a) Controlling redundancy
b) Restricting unauthorized access
c) Providing multiple user interfaces
d) Enforcing integrity constraints.
e) Providing back up and recovery

8.Give the levels of data abstraction.

a) Physical level
b) Logical level
c) View level

9.Define instance and schema.

Instance: Collection of data stored in the data base at a particular moment is called an Instance of the database.

Schema: The overall design of the data base is called the data base schema.

10.Define the terms of Data base schemas.

1) Physical schema
2) logical schema.

Physical schema: The physical schema describes the database design at the physical level, which is the lowest level of abstraction describing how the data are actually stored.

Logical schema: The logical schema describes the database design at the logical level, which describes what data are stored in the database and what relationship exists among the data.

11.What is conceptual schema?

The schemas at the view level are called subschema's that describe different views of the database.

12.Define data model.

A data model is a collection of conceptual tools for describing data, data relationships, data semantics and consistency constraints.

13.What is storage manager?

A storage manager is a program module that provides the interface between the low

level data stored in a database and the application programs and queries submitted to the system.

14.What are the components of storage manager?
The storage manager components include
a) Authorization and integrity manager
b) Transaction manager
c) File manager
d) Buffer manager

15.What is the purpose of storage manager?
The storage manager is responsible for the following
a) Interaction with the file manager
b) Translation of DML commands in to low level file system commands
c) Storing, retrieving and updating data in the database

16.List the data structures implemented by the storage manager. .
The storage manager implements the following data structure
a) Data files
b) Data dictionary
c) Indices

17.What is a data dictionary?
A data dictionary is a data structure which stores meta data about the structure of the database ie. The schema of the database.

18.What is an entity relationship model?
The entity relationship model is a collection of basic objects called entities and relationship among those objects. An entity is a thing or object in the real world that is distinguishable from other objects.

19.What are attributes? Give examples.
An entity is represented by a set of attributes. Attributes are descriptive properties possessed by each member of an entity set.
Example: possible attributes of customer entity are customer name, customer id, Customer Street, customer city.

20.What is relationship? Give examples
A relationship is an association among several entities.
Example: A depositor relationship associates a customer with each account that he/she has.
21.Define the terms i) Entity set ii) Relationship set
Entity set: The set of all entities of the same type is termed as an entity set.
Relationship set : The set of all relationships of the same type is termed as a relationship set.

22.Define single valued and multi valued attributes.
Single valued attributes: attributes with a single value for a particular entity are called single valued attributes.
Multi valued attributes : Attributes with a set of value for a particular entity are called multivalued attributes.

23.What are stored and derived attributes?
Stored attributes: The attributes stored in a data base are called stored attributes.
Derived attributes: The attributes that are derived from the stored attributes are called

derived attributes.

24. What are composite attributes?
    Composite attributes can be divided in to sub parts.

25. Define null values.
    In some cases a particular entity may not have an applicable value for an attribute or if we do not know the value of an attribute for a particular entity. In these cases null value is used.

26. Define the terms i) Entity type ii) Entity set
    Entity type: An entity type defines a collection of entities that have the same attributes.
    Entity set: The set of all entities of the same type is termed as an entity set.

27. What is meant by the degree of relationship set?
    The degree of relationship type is the number of participating entity types.

28. Define the terms.
    i) Key attribute
    ii) Value set
    Key attribute : An entity type usually has an attribute whose values are distinct from each individual entity in the collection. Such an attribute is called a key attribute.
    Value set: Each simple attribute of an entity type is associated with a value set that specifies the set of values that may be assigned to that attribute for each individual entity.

29.What does the cardinality ratio specify?
    Mapping cardinalities or cardinality ratios express the number of entities to which another entity can be associated. Mapping cardinalities must be one of the following:
- One to one
- One to many
- Many to one
- Many to many

30. Define weak and strong entity sets.
    Weak entity set: entity set that do not have key attribute of their own are called weak entity sets. Strong entity set: Entity set that has a primary key is termed a strong entity set.

31.What are the two types of participation constraint.
- **Total:** The participation of an entity set E in a relationship set R is said to be **total** if every entity in E participates in at least one relationship in R.
- **Partial:** if only some entities in E participate in relationships in R, the participation of entity set E in relationship R is said to be **partial.**

32. Define the terms i) DDL ii) DML
    DDL: Data base schema is specified by a set of definitions expressed by a special language called a data definition language.
    DML: A data manipulation language is a language that enables users to access or manipulate data as organized by the appropriate data model

UNIT II SQL & QUERY OPTIMIZATION

1.What is embedded SQL? What are its advantages? April/May-2011

    Embedded SQL is a method of combining the computing power of a programming language and the database manipulation capabilities of SQL. Embedded SQL statements are SQL

statements written in line with the program source code of the host language. The embedded SQL statements are parsed by an embedded SQL preprocessor and replaced by host-language calls to a code library. The output from the preprocessor is then compiled by the host compiler. This allows programmers to embed SQL statements in programs written in any number of languages such as: C/C++, COBOL and Fortran.

2.What is the difference between tuple relational calculus and domain relational calculus? April/May-2011

The tuple-oriented calculus uses a tuple variables i.e., variable whose only permitted values are tuples of that relation. E.g. QUEL. The domain-oriented calculus has domain variables i.e., variables that range over the underlying domains instead of over relation. E.g. ILL, DEDUCE....The tuple-oriented calculus uses a tuple variables i.e., variable whose only permitted values are tuples of that relation. E

3.Write short notes on relational model.

The relational model uses a collection of tables to represent both data and the relationships among those data. The relational model is an example of a record based model.

4.Define tuple and attribute.

- **Attributes:** column headers

- **Tuple** : Row

5.Define the term relation.

Relation is a subset of a Cartesian product of list domains.

6.Define tuple variable.

Tuple variable is a variable whose domain is the set of all tuples.

7.Define the term Domain.

For each attribute there is a set of permitted values called the ***domain*** of that attribute.

8.What is a candidate key?

Minimal super keys are called ***candidate keys.***

9.What is a primary key?

Primary key is chosen by the database designer as the principal means of identifying an entity in the entity set.

10.What is a super key?

A ***super key*** is a set of one or more attributes that collectively allows us to identify uniquely an entity in the entity set.

11.Define- relational algebra.

The relational algebra is a procedural query language. It consists of a set of operations that take one or two relation as input and produce a new relation as output.

12.What is a SELECT operation?

The ***select*** operation selects tuples that satisfy a given predicate. We use the lowercase letter ss to denote selection. ss

13**.What is a PROJECT operation?**

The project operation is a unary operation that returns its argument relation with certain attributes left out. Projection is denoted by pie (pp pp).

14. Write short notes on tuple relational calculus.

The tuple relational calculation is anon procedural query language. It describes the desired information with out giving a specific procedure for obtaining that Information.

A query or expression can be expressed in tuple relational calculus as

{t | P (t)}

which means the set of all tuples 't' such that predicate P is true for 't'.

Notations used:
- t[A] the value of tuple 't' on attribute, A
- t r tuple 't' is in relation 'r'
- there exists

Definition for 'there exists' ( ):

t r(Q(t))

which means there exists a tuple 't' in relation 'r' such that predicate Q(t) is true.

- for all

Definition for 'for all' ( ):

t r(Q(t))

which means Q(t) is true for all tuples 't' in relation 'r'.

- Implication

Definition for Implication ( ):

P Q means if P is true then Q must be true.

15. Write short notes on domain relational calculus.

The domain relational calculus uses domain variables that take on values from an attribute domain rather than values for entire tuple.

16. Define query language.

A query is a statement requesting the retrieval of information. The portion of DML that involves information retrieval is called a query language.

17. Write short notes on Schema diagram.

A database schema along with primary key and foreign key dependencies can be depicted pictorially by schema diagram. Each relation appears as a box with attributes listed inside it and the relation name above it.

18. What is foreign key?

A relation schema r1 derived from an ER schema may include among its attributes the primary key of another relation schema r2.this attribute is called a *foreign key* from r1 referencing r2.

19. What are the parts of SQL language?

The SQL language h as several parts:
data - definition language
Data manipulation language
View definition
Transaction control
Embedded SQL
Integrity

Authorization

20. What are the categories of SQL command?

SQL commands are divided in to the following categories:
1. data - definition language
2. data manipulation language
3. Data Query language
4. data control language
5. data administration statements
6. transaction control statements

21. What are the three classes of SQL expression?

SQL expression consists of three clauses:
Select
From
where

22. Give the general form of SQL query.

Select A , A ………….., An
1 2
Fro m R , R ……………, R
1 2   m
Where P

23. What is the use of rename operation?

Rename operation is used to rename both r elations and a attributes. It uses the as clause, taking the form: Old-name **as** new-name

24. Define tuple variable.

Tuple variables are used for comparing two tuples in the same relation. The tuple variables are defined in the from clause by way of the **as** clause.

25. List the string operations supported by SQL.

- Pattern matching Operation
- Concatenation
- Extracting character strings
- Converting between uppercase and lower case letters.

26. List the set operations of SQL.

1) Union
2) Intersect operation
3) The except operation

27. What is the use of Union and intersection operation?

Union : The result of this operation includes all tuples that are either in r1 or in r2 or in both r1 and r2.Duplicate tuples are automatically eliminated.

Intersection: The result of this relation includes all tuples that are in both r1 and r2.

28.What are aggregate functions? And list the aggregate functions supported by SQL?

Aggregate functions are functions that take a collection of values as input and return a single value. Aggregate functions supported by SQL are

- Average: avg
- Minimum: min
- Maximum: max
- Total: sum
- Count: count

UNIT III TRANSACTION PROCESSING AND CONCURRENCY CONTROL

1.What is meant by lossless-join decomposition? APRIL/MAY-2011

We claim the above decomposition is lossless. How can we decide whether decomposition is lossless?

1. Let R be a relation schema.
2. Let F be a set of functional dependencies on R.
3. Let R_1 and R_2 form a decomposition of R.
4. The decomposition is a lossless-join decomposition of R if at least one of the following functional dependencies are in F^+:

   a. $R_1 \cap R_2 \rightarrow R_1$
   b. $R_1 \cap R_2 \rightarrow R_2$

2.List the disadvantages of relational database system.
- Repetition of data
- Inability to represent certain information.

3.What is first normal form?

The domain of attribute must include only atomic (simple, indivisible) values.

4.What is meant by functional dependencies?

Consider a relation schema R and a C R and ß C R. The functional dependency a ß holds on relational schema R if in any legal relation r(R), for all pairs of tuples t1 and t2 in r such that

t1 [a] =t1 [a], and also t1 [ß] =t2 [ß].

5.What are the uses of functional dependencies?

To test relations to see whether they are legal under a given set of functional dependencies. To specify constraints on the set of legal relations.

6.What meant by trivial dependency?

Functional dependency of the form a   ß is trivial if ß C a. Trivial functional dependencies are satisfied by all the relations.

7.What are axioms?

Axioms or rules of inference provide a simpler technique for reasoning about functional dependencies.

8.What is meant by computing the closure of a set of functional dependency?

+ The closure of F denoted b y F   is the set of functional dependencies logically implied by F.

9.What is meant by normalization of data?

It is a process of analyzing the given relation schemas based on their Functional Dependencies (FDs) and primary key to achieve the properties

Minimizing redundancy

Minimizing insertion, deletion and updating anomalies   .

10.Define Boyce codd normal form .

A relation schema R is in BCNF with respect to a set F of functional + dependencies if, for all functional dependencies in F of the form. a->ß, where a

11.List out the desirable properties of decomposition.

- Lossless-join decomposition
- Dependency preservation
- Repetition of information

12.What is 2NF?

A relation schema R is in 2NF if it is in 1NF and every non-prime attribute A in R is fully functionally dependent on primary key.

13.What is the use of group by clause?

Group by clause is used to apply aggregate functions to a set of tuples. The attributes given in the **group by** clause are used to form groups. Tuples with the same value on all attributes in the **group by** clause are placed in one group.

14.What is the use of sub queries?

A sub query is a select-from-where expression that is nested with in another query. A common use of sub queries is to perform tests for set membership, make set comparisons, and determine set cardinality.

15.What is view in SQL? How is it defined?

Any relation that is not part of the logical model, but is made visible to a user as a virtual relation is called a view.

We define view in SQL by using the **create view** command. The form of the **create view** command is **Create view** v as <query expression>

16.What is the use of with clause in SQL?

The **with** clause provides a way of defining a temporary view whose definition is available only to the query in which the **with** clause occurs.

17.List the table modification commands in SQ.

- Deletion
- Insertion
- Updates
- Update of a view

18.List the SQL domain Types.

SQL supports the following domain types.

1) Char(n)   2) varchar(n)   3) int   4) numeric(p,d)
5) float(n)   6) date.

19. What is the use of integrity constraints?

Integrity constraints ensure that changes made to the database by authorized users do not result in a loss of data consistency. Thus integrity constraints guard against accidental damage to the database.

20. Mention the 2 forms of integrity constraints in ER model.

- Key declarations
- Form of a relationship

21. What is trigger?

Triggers are statements that are executed automatically by the system as the side effect of a modification to the database.

22. What are domain constraints?

A domain is a set of values that may be assigned to an attribute .all values that appear in a column of a relation must be taken from the same domain.

23. What are referential integrity constraints?

A value that appears in one relation for a given set of attributes also appears for a certain set of attributes in another relation.

24. What is assertion? Mention the forms available.

An assertion is a predicate ex pressing a condition that we wish the database always to satisfy.
Domain integrity constraints.
Referential integrity constraints

25. Give the syntax of assertion.

Create assertion <assertion name> **check** <predicate>

26. What is the need for triggers?

Triggers are useful mechanisms for alerting humans or for starting certain tasks automatically when certain conditions are met.

27. List the requirements needed to design a trigger.

The requirements are

- Specifying when a trigger is to be executed.
- Specify the actions to be taken when the trigger executes.

28. Give the forms of triggers.

The triggering event can be insert or delete.
For updated the trigger can specify columns.
The referencing old row as clause
The referencing new row as clause
The triggers can be initiated before the event or after the event.

29. What does database security refer to?

Database security refers to the protection from unauthorized access and malicious destruction or alteration.

30. List some security violations (or) name any forms of malicious access.

- Unauthorized reading of data
- Unauthorized modification of data
- Unauthorized destruction of data.

31. List the types of authorization.

- Read authorization
- Write authorization
- Update authorization
- Drop authorization

32. What is authorization graph?

Passing of authorization from one user to another can be represented by an authorization graph.

33. List out various user authorization to modify the database schema.

- Index authorization
- Resource authorization
- Alteration authorization
- Drop authorization

34. What are audit trails?

An audit trail is a log of all changes to the database along with information such as which user performed the change and when the change was performed.

35. Mention the various levels in security measures.

- Database system
- Operating system
- Network
- Physical
- Human

36. Name the various privileges in SQL.

- Delete

- Select
- Insert
- Update

37.Mention the various user privileges.
- All privileges directly granted to the user or role.
- All privileges granted to roles that have been granted to the user or role.

UNIT IV TRENDS IN DATABASE TECHNOLOGY

1.What are the ACID properties? APRIL/MAY-2011

(*atomicity, consistency, isolation, durability*) is a set of properties that guarantee database transactions are processed reliably. In the context of databases, a single logical operation on the data is called a transaction. For example, a transfer of funds from one bank account to another, even though that might involve multiple changes (such as debiting one account and crediting another), is a single transaction.

2. What are two pitfalls (problem) of lock-based protocols? APRIL/MAY-2011
- Deadlock
- Starvation

3. What is transaction?

Collections of operations that form a single logical unit of work are called transactions.

4. What are the two statements regarding transaction?

The two statements regarding transaction of the form:
Begin transaction
End transaction

5. What are the properties of transaction?

The properties o f transactions are:
Atomicity
Consistency
Isolation
Durability

6. What is recovery management component?

Ensuring durability is the responsibility of a software component of the base system called the recovery management component.

7. When is a transaction rolled back?

Any changes that the aborted transaction made to the database must be undone. Once the changes caused by an aborted transaction have been undone, then the transaction has been rolled back.

8. What are the states of transaction?

The states of transaction are
Active
Partially committed
Failed
Aborted
Committed

Terminated

9. List out the statements associated with a database transaction.
Commit work
Rollback work

10. What is a shadow copy scheme?
It is simple, but efficient, scheme called the shadow copy schemes. It is based on making copies of the database called shadow copies that one transaction is active at a time. The scheme also assumes that the database is simply a file on disk.

11. Give the reasons for allowing concurrency.
The reasons for allowing concurrency is if the transactions run serially, a short transaction may have to wait for a preceding long transaction to complete, which can lead to unpredictable delays in running a transaction. So concurrent execution reduces the unpredictable delays in running transactions.

12. What is average response time?
The average response time is that the average time for a transaction to be completed after it has been submitted.

13. What are the two types of serializability?
The two types of serializability is
Conflict serializability
View serializability

14. Define lock.
Lock is the most common used to implement the requirement is to allow a transaction to access a data item only if it is currently holding a lock on that item.

15. What are the different modes of lock?
The modes of lock are:
    Shared
    Exclusive

16. Define deadlock.
Neither of the transaction can ever proceed with its normal execution. This situation is called deadlock.

17. Define the phases of two phase locking protocol.
Growing phase: a transaction may obtain locks but not release any lock.
Shrinking phase: a transaction may release locks but may not obtain any new locks.

18. Define upgrade and downgrade.
It provides a mechanism for conversion from shared lock to exclusive lock is known as upgrade.
It provides a mechanism for conversion from exclusive lock to shared lock is known as downgrade.

19. What is a database graph?
The partial ordering implies that the set D may now be viewed as a directed acyclic graph, called a database graph.

20. What are the two methods for dealing deadlock problem?
The two methods for dealing deadlock problem is deadlock detection and deadlock

recovery.

21. What is a recovery scheme?
An integral part of a database system is a recovery scheme that can restore the database to the consistent state that existed before the failure.

22. What are the two types of errors?
The two types of errors are:
- Logical error
- System error

23. What are the storage types?
The storage types are:
- Volatile storage
- Nonvolatile storage

24. Define blocks.
The database system resides permanently on nonvolatile storage, and is partitioned into fixed-length storage units called blocks.

25. What is meant by Physical blocks?
The input and output operations are done in block units. The blocks residing on the disk are referred to as physical blocks.

26. What is meant by buffer blocks?
The blocks residing temporarily in main memory are referred to as buffer blocks.

27. What is meant by disk buffer?
The area of memory where blocks reside temporarily is called the disk buffer.

28. What is meant by log-based recovery?
The most widely used structures for recording database modifications is the log. The log is a sequence of log records, recording all the update activities in the database. There are several types of log records.

29. What are uncommitted modifications?
The immediate-modification technique allows database modifications to be output to the database while the transaction is still in the active state. Data modifications written by active transactions are called uncommitted modifications.

30. Define shadow paging.
An alternative to log-based crash recovery technique is shadow paging. This technique needs fewer disk accesses than do the log-based methods.

31. Define page.
The database is partitioned into some number of fixed-length blocks, which are referred to as pages.

32. Explain current page table and shadow page table.
The key idea behind the shadow paging technique is to maintain two page tables during the life of the transaction: the current page table and the shadow page table. Both the page tables are identical when the transaction starts. The current page table may be changed when

a transaction performs a write operation.

33. What are the drawbacks of shadow-paging technique?
- Commit Overhead
- Data fragmentation
- Garbage collection

34. Define garbage collection.
Garbage may be created also as a side effect of crashes. Periodically, it is necessary to find all the garbage pages and to add them to the list of free pages. This process is called garbage collection.

35. Differentiate strict two phase locking protocol and rigorous two phase locking protocol.
In **strict two phase locking protocol** all exclusive mode locks taken by a transaction is held until that transaction commits.
Rigorous two phase locking protocol requires that all locks be held until the transaction commits.

36. How the time stamps are implemented?
- Use the value of the system clock as the time stamp. That is a transaction's time stamp is equal to the value of the clock when the transaction enters the system.
- Use a logical counter that is incremented after a new timestamp has been assigned; that is the time stamp is equal to the value of the counter.

37. What are the time stamps associated with each data item?
- W-timestamp (Q) denotes the largest time stamp if any transaction that executed WRITE (Q) successfully.
- R-timestamp (Q) denotes the largest time stamp if any transaction that executed READ (Q) successfully.

UNIT V ADVANCED TOPICS

1. What are the advantages and disadvantages of indexed sequential file? APRIL/MAY-2011
The advantage of ordering records in a sequential file according to a key is that you can then search the file more quickly. If you know the key value that you want, you can use one of the relatively fast searches. The disadvantage is that when you insert, you need to rewrite at least everything after the insertion point, which makes inserts very expensive unless they are done at the end of the file. An indexed file approach keeps a (hopefully) small part of each row, and some kind of "pointer" to the row's location within the data file. This allows a search to use the index, which is ordered by the index and (again hopefully) much smaller and therefore much faster than scanning the entire data file for the indexed data.

2.What is database tuning? APRIL/MAY-2011
Database tuning describes a group of activities used to optimize and homogenize the performance of a database. It usually overlaps with query tuning, but refers to design of the database files, selection of the database management system (DBMS), operating system and CPU the DBMS runs on.

3. Give the measures of quality of a disk.
- Capacity
- Access time
- Seek time
- Data transfer rate
- Reliability

- Rotational latency time.

4. Compare sequential access devices versus random access devices with an example.

sequential access devices	random access devices
Must be accessed from the beginning	It is possible to read data from any location
Eg:- tape storage	Eg:-disk storage
Access to data is much slower	Access to data is faster
Cheaper than disk	Expensive when compared with disk

5. What are the types of storage devices?
- Primary storage
- Secondary storage
- Tertiary storage

6.Draw the storage device hierarchy according to their speed and their cost.
- Cache
- Main memory
- Flash memory
- Magnetic disk
- Optical disk
- Magnetic tapes

7.What are called jukebox systems?

Jukebox systems contain a few d rives and numerous disks that can be loaded into one of the drives automatically.

8.What is called remapping of bad sectors?

If the controller detects that a sector is damaged when the disk is initially formatted, or when an attempt is made to write the sector, it can logically map the sector to a different physical location.

9.Define access time.

Access time is the time from when a read or write request is issued to when data transfer begins.

10.Define seek time.

The time for repositioning the arm is called the seek time and it increases with the distance that the arm is called the seek time.

11. Define average seek time.

The average seek time is the average of the seek times, measured over a sequence of random requests.

12. Define rotational latency time.

The time spent waiting for the sector to be accessed to appear under the head is called the rotational latency time.

13. Define average latency time.

The average latency time of the disk is one-half the time for a full rotation of the disk.

14. What is meant by data-transfer rate?

The data-transfer rate is the rate at which data can be retrieved from or stored to the disk.

15. What is meant by mean time to failure?

The mean time to failure is the amount of time that the system could run continuously without failure.

16. What are a block and a block number?

A block is a contiguous sequence of sectors from a single track of one platter. Each request specifies the address on the disk to be referenced. That address is in the form of a block number.

17. What are called journaling file systems?

File systems that support log disks are called journaling file systems.

18. What is the use of RAID?

A variety of disk-organization techniques, collectively called redundant arrays of independent disks are used to improve the performance and reliability.

19.How the reliability can be improved through redundancy?

The simplest approach to introducing redundancy is to duplicate every disk. This technique is called mirroring or shadowing. A logical disk then consists of two physical disks, and write is carried out on both the disk. If one of the disks fails the data can be read from the other. Data will be lost if the second disk fails before the first fail ed disk is repaired.

20. What is called mirroring?

The simplest approach to introducing redundancy is to duplicate every disk. This technique is called mirroring or shadowing.

21. What is called mean time to repair?

The mean time to failure is the time it takes to replace a failed disk and to restore the data on it.

22. What is called bit-level striping?

Data striping consists of splitting the bits of each byte across multiple disks. This is called bit-level striping.

23. What is called block-level striping?

Block level striping stripes blocks across multiple disks. It treats the array of disks as a large disk, and gives blocks logical numbers.

24. What are the two main goals of parallelism?
- Load –balance multiple small accesses, so that the throughput of such accesses increases.
- Parallelize large accesses so that the response time of large accesses is reduced

25. What are the factors to be taken into account when choosing a RAID level?
- o Monetary cost of extra disk storage requirements.
- o Performance requirements in terms of number of I/O operations
- o Performance when a disk has failed.
- o Performances during rebuild.

26. What is meant by software and hardware RAID systems?

RAID can be implemented with no change at the hardware level, using only software modification. Such RAID implementations are called software RAID systems and the systems with special hardware support are called hardware RAID systems.

27. Define hot swapping.

Hot swapping permits the removal of faulty disks and replaces it by new ones without turning power off. Hot swapping reduces the mean time to repair.

28. Which level of RAID is best? Why?

RAID level 1 is the RAID level of choice for many applications with moderate storage requirements and high I/O requirements. RAID 1 follows mirroring and provides best write performance.

29. Distinguish between fixed length records and variable length records.

Fixed length records

Every record has the same fields and field lengths are fixed.

Variable length records

File records are of same type but one or more of the fields are of varying size.

30. What are the ways in which the variable-length records arise in database systems?

- Storage o f multiple record types in a file.
- Record types that allow variable lengths for one or more fields.
- Record types that allow repeating fields.

31. Explain the use of variable length records.

- They are used for Storing of multiple record types in a file.
- Used for storing records that has varying lengths for one or more fields.
- Used for storing records that allow repeating fields

32. What is the use of a slotted-page structure and what is the information present in the header?

- The slotted-page structure is used for organizing records within a single block.
- The header contains the following information.
- The number of record entries in the header.
- The end of free space .
- An array whose entries contain the location and size of each record.

33. What are the two types of blocks in the fixed –length representation? Define them.

- Anchor block: Contains the first record of a chain.
- Overflow block: Contains the records other than those that are the first Record o f a chain.

34. What is known as heap file organization?

In the heap file organization, any record can be placed anywhere in the file where there is space for the record. There is no ordering of records. There is a single file for relation.

35. What is known as sequential file organization?

In the sequential file organization, the records are stored in sequential order, according to the value of a "search key" of each record.

36. What is hashing file organization?

In the hashing file organization, a hash function is computed on some attribute of each record. The result of the hash function specifies in which block of the file the record should be placed.

37. What is known as clustering file organization?

In the clustering file organization, records of several different relations are stored in the same file.

38. What is an index?

An index is a structure that helps to locate desired records of a relation quickly, without examining all records.

39. What are the two types of ordered indices?
- Primary index
- Secondary index

40. What are the types of indices?
- Ordered indices
- Hash indices

41. What are the techniques to be evaluated for both ordered indexing and hashing?
- Access types
- Access time
- Insertion time
- Deletion time
- Space overhead

42. What is known as a search key?
An attribute or set of attributes used to look up records in a file is called a search key.

43. What is a primary index?
A primary index is an index whose search key also defines the sequential order of the file.

44. What are called index-sequential files?
The files that are ordered sequentially with a primary index on the search key are called index-sequential files.

45. What are the two types of indices?
Dense index
Sparse index

46. What are called multilevel indices?
Indices with two or more levels are called multilevel indices.

47. What are called secondary indices?
Indices whose search key specifies an order different from sequential order of the file are called secondary indices. The pointers in secondary index do not point directly to the file. Instead each points to a bucket that contains pointers to the file.

48. What are the disadvantages of index sequential files?
The main disadvantage of the index sequential file organization is that performance degrades as the file grows. This degradation is remedied by reorganization of the file.

49. What is a B+-Tree index?
A B+-Tree index takes the form of a balanced tree in which every path from the root of the root of the root of the tree to a leaf of the tree is of the same length.

50. What is B-Tree?
A B-tree eliminates the redundant storage of search-key values .It allows search key values to appear only once.

51. What is hashing?
Hashing allows us to find the address of a data item directly by computing a hash function on the search key value of the desired record.

52. How do you create index in SQL?

We create index by he create index command.

Create index \<index name\> **on** \<relation name\> (\<attribute list\>)

53. Distinguish between static hashing and dynamic hashing?

Static hashing

Static hashing uses a h ash function in which the set of bucket adders is fixed. Such hash functions cannot easily accommodate databases that grow larger over time.

Dynamic hashing

Dynamic hashing allows us to modify the hash function dynamically. Dynamic hashing copes with changes in database size by splitting and coalescing buckets as the database grows and shrinks.

54. What is a hash index?

A hash index organizes the search keys, with their associated pointers, into a hash file structure.

55. What can be done to reduce the occurrences of bucket overflows in a hash file organization?

To reduce bucket overflow the number of bucket is chosen to be

(nr/fr)*(1+d).

We handle bucket overflow by using

- Overflow chaining (closed hashing)
- Open hashing

56. Differentiate open hashing and closed hashing (overflow chaining)

Closed hashing (overflow chaining)

If a record must be inserted in to a bucket b, and b is already full, the system provides an overflow bucket for b, and inserts the record in to the overflow bucket. If the overflow bucket is also full, the system provides another overflow bucket, and so on. All the overflow buckets of a given buckets are chained together in a linked list, overflow handling using linked list is known as closed hashing.

Open hashing

The set of buckets is fixed, and there are no overflow chains. Instead, if a bucket is full, the system inserts records in some other bucket in the initial set of buckets.

57. What is linear probing?

Linear probing is a type of open hashing. If a bucket is full the system inserts records in to the next bucket that has space. This is known as linear probing.

58. What is called query processing?

Query processing refers to the range of activities involved in extracting data from a database.

59. What are the steps involved in query processing?

The basic steps are:

Parsing and translation
Optimization
Evaluation

60. What is called an evaluation primitive?

A relational algebra operation annotated with instructions on how to evaluate is called an

evaluation primitive.

61. Define query optimization.
Query optimization refers to the process of finding g the lowest –cost method of evaluating a given query.

62. What is called a query –execution engine?
The query execution engine takes a query evaluation plan, executes that plan, and returns the answers to the query.

63. How do you measure the cost of query evaluation?
The cost of a query evaluation is measured in terms of a number of different resources including disk accesses, CPU time to execute a query, and in a distributed database system the cost of communication

64. List out the operations involved in query processing.
Selection operation
Join operations.
Sorting.
Projection
Set operations
Aggregation

65. What are called as index scans?
Search algorithms that use an index are referred to as index scans.

66. What is called as external sorting?
Sorting of relations that do not fit into memory is called as external sorting.

67. How to form the nested loop join?
Nested loop join consists of a pair of nested for loops.
Example: r||
T **s**
r is the outer relation and s is the inner relation.

68. What is meant by block nested loop join?
Block nested loop join is the variant of the nested loop join where every block of the inner relation is paired with ever y block of the outer relation. With in each pair of blocks ever y tuple in one block is paired with every tuple in the other blocks to generate all pairs of tuples.

69. What is meant by hash join?
In the h ash join algorithm a hash function h is used to implement partition tuples of both relations.

70. What is called as recursive partitioning?
The system repeats the splitting of the input until each partition of the build input fits in the memory. Such partitioning is called recursive partitioning.

71. What is called as an N-way merge?
The merge operation is a generalization of the two-way merge used by the standard in-memory sort-merge algorithm. It merges N runs, so it is called an N-way merge.

72. What is known as fudge factor?

The number of partitions is increased by a small value called the fudge factor, which is usually 20 percent of the number of hash partitions computed.

PART B:16 MARKS
UNIT I

1. a. With a neat diagram, explain the structure of a DBMS?(9)**-[April/May- 2011]**
   b. Draw an E-R diagram for a small marketing company database, assuming your own data requirements (7) – [April/May-2011]
2. a. Explain the architecture of DBMS (8)
   b. Compare File systems with database systems . (8)
3. What is meant by Relational calculus? Query examples for tuple and domain relational calculus?(16)
4. Explain all types of data models (16)
5. Explain E-R Model concept and extended E-R model. (16)
6. Define relational algebra. With suitable example. (16)
7. Roles of data base administrator and the details of aggregate function queries (16)
8. a. With relevant examples discuss the following in SQL
         i. DDL
         ii. DML
         iii. DCL
         iv. Views (8)
   b. Construct an ER diagram for a car insurance company that has a set of customers, each of whom owns one/more cars. Each car has associated with it zero to any number of recorded accidents. (8)
9. Explain the 3 schema architecture of DBS.Why do we need mappings between different schema levels
10. Consider the following tabls:
        Employee (Emp_no, Name, Emp_city) e
        Company (Emp_no, Company_name, Salary)
           i. Write a SQL query to display Employee name and company name.
           ii. Write a SQL query to display employee name, employee city ,company name and salary of all the employees whose salary >10000
           iii. Write a query to display all the employees working in 'XYZ' company.
11. Explain various DML commands with neat syntax.

UNIT II

1. a) What is data integrity? Explain the types of integrity constraints. **April/May-2011**

   b) What are the relational algebra operations supported in SQL? **April/May-2011**

2. a) Define class Hierarchies and Aggregation? (8)
3. a)Explain the strength and weakness of the Triggers mechanism (8)
              and update views in stored procedure.
    b)What is meant by Integrity and security? (8)
4. Consider the following relational schema
           Employee (empno,name,office,age)
           Books(isbn,title,authors,publisher)
           Loan(empno, isbn,date)
5. Write the following queries in relational algebra.
6. a. Find the names of employees who have borrowed a book

Published by McGraw-Hill. (4)
b. Find the names of employees who have borrowed all books
Published by McGraw-Hill. (4)
c. Find the names of employees who have borrowed more than
five different books published by McGraw-Hill. (4)
d. For each publisher, find the names of employees who have borrowed
More than five books of that publisher. (4)
7. Explain the details about embedded and dynamic SQL (16)
8. Draw a neat sketch to indicate the architecture of a distributed database system. With
an example explain the various form of data fragmentation used in DDB.
9. a) What is meant by Selection and projection used in Relational algebra? (8)
   b) How we made lossless and loss-join in relational model? (8)

UNIT III

1. a) Explain 1NF, 2Nf and BCNF with suitable example. **April/May-2011**
   b) Consider the universal relation R={ A,B,C,D,E,F,G,H,I} and the set of functional dependencies
F={(A,B)->{C],{A}>{D,E},{B}->{F},{F}->{G,H},{D}->[I,J}.what is the key for Decompose R into 2NF,the 3NF relations.**April/May-2011**
2. What are the pitfalls in relational database design? With a suitable example, explain the role of functional dependency in the process of normalization. **April/May-2011**
3. What is normalization? Explain all Normal forms.
4. Write about decomposition preservation algorithm for all FD's.
5. Explain functional dependency concepts.
6. Explain 2NF and 3NF in detail.
7. Define BCNF .How does it differ from 3NF.
8. Explain the codd's rules for relational database design.
9. Explain the ACID properties of a transaction.

UNIT IV

1. a) Explain about immediate update and deferred update recovery techniques. **April/May-2011**

   b) Explain the concepts of serializability. **April/May-2011**

2. a) Explain two-phase locking protocol. **April/May-2011**

   b) Describe about the deadlock prevention schemes. **April/May-2011**

3. a) How Transactions are possible in Distributed database? Explain briefly (8)
   b) What is Transaction state and its ACID properties? (8)
4. a) Discuss about two phase locking and commit protocol (8)
   b) Explain various recovery techniques during transaction in detail. (8)
5. a)How can we achieve concurrency control achieved in DBMS through Serializability? (8)
    b)Explain Recovery schemas in detail (8)
6. a) What is deadlock prevention and dead lock detection method (10)
   b) Explain the deadlock recovery technique (6)
7. Explain the following protocols for concurrency control
            i) Lock based protocols (8)
            ii) Time stamp based protocols DATABASE MANAGEMENT SYSTEM
8. What is concurrency control? how is it implemented in DBMS? (16)
9.a)Write short notes on shadow paging (8)
   b)Explain the deffered and immediate-modification version of the log-based recovery scheme.(8)

UNIT V

1. a) List the different levels in RAID technology and explain its features. **April/May-2011**
   b) Describe the different method of implementing variable length records. **April/May-2011**

2. a) Explain the various indexing schemes used in database environment. **April/May-2011**
3. a) What is Raid? List the different levels in Raid technology and explain its features (8)
   b) Discuss about primary file storage system (8)
4. Explain static and dynamic Hashing Techniques? (16)
5. Briefly describe about B+ tree index file structure. (16)
6. What are the steps involved in Query processing? How would you estimate the cost of query (16)
7. Discuss selection operation techniques (16)
8. Explain structure of file indices (16)
9. Explain different properties of indexes in detail (16)
10. With a neat diagram explain the steps involved in query process.
11. Explain in detail about B+ tree index files.
12. What is deadlock? Explain the various approaches used to recover from deadlock.