DHANALAKSHMI SRINIVASAN INSTITUTE OF RESEARCH AND TECHNOLOGY

SIRUVACHUR, PERAMBALUR – 611 113

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CS6302-DATABASE MANAGEMENT SYSTEMS

PART-B (16 MARK QUESTIONS)

UNIT: 1

1. EXPLAIN ABOUT DATABASE SYSTEM STRUCTURE.

- Storage manager
- Authorization and integrity manager
- Transaction manager
- File manager
- Buffer manager

Storage manager implements several data structure as a part of physical system implementation

- Data function
- Data dictionary
- Indices

The query processor

- DDL interpreter
- DML
- Query evaluation engine

2. DESCRIBE RELATIONAL MODEL.

Structure of relational data base

- Basic structure
- Database schema
- Keys
- Schema diagram
- Query languages

3. BRIEFLY EXPLAIN RELATIONAL ALGEBRA.

Fundamental operations

- Unary operations
- Binary operations

Select operations

σbranchname='perryridge' (loan)

The project operation

Лloannumber,amount (loan)

Composition of relational operations

□ □ Relational algebra expressions

Union operations

- r Ụ s
- r and s must be a same arity.
- They must have the same no of attributes.

The set difference operations

r-s produce a relation containing those tuples in r but not in s.

The Cartesian product operations

The rename operations

4. WHAT IS DATA MODELS? EXPLAIN IT DETAIL.

Entity relationship model

- Rectangles
- Ellipse
- Diamonds
- Lines

Relational model

Relational model use a collection of tables to represent both data and the relationships among those data. Each table has a multiple columns and each columns has unique name

Other data models

- Object oriented data model
- Object relational data model
- Network data model
- Hierarchical data model

5. DRAW AN ER DIAGRAM FOR A BANKING ENTERPRISE

- • Data requirements
- • Entity sets
- • Relationship sets
- • ER diagram

6. BRIEFLY DESCRIBE RELATIONAL CALCULUS WITH SOME EXAMPLES.

The tuple relational calculus

A query in a tuple relational calculus is expressed as {t| P(t)}

- Example Queries
- Formal definition
- Safety of expressions
- Expressive power of languages

The domain relational calculus

- Example Queries
- Formal definition
- Safety of expressions
- Expressive power of languages

UNIT-2

1. DESCRIBE INTEGRITY AND SECURITY.

Domain constraint

Referential integrity

A value that appears in one relation for a given set of attributes also appear for a certain set of attributes in another relation. This condition is called referential integrity.

- Referential integrity and E-R models
- Database modification
- Referential integrity in SQL

2. WHAT IS AGGREGATE FUNCTION? BRIEFLY DESCRIBE IT.

Aggregate functions are functions that take a collection of values as input and return a single value. SQL offers 5 built-in aggregate functions:

- Average: **avg**
- Minimum:**min**

- Maximum: **max**
- Total:**sum**
- Count:**count**

Average: avg

```
Select avg (balance)
From account
        Where branch-name='perryridge'
```

Count :**count**

```
select branch-name,count(distinct customer-name)
from depositior,account
            where depositor.account-number=account.account-number
             groupby branch-name
```

3. WHAT IS DATA DEFINITION LANGUAGE? EXPLAIN IT IN DETAIL.

The SQL DDL allows specification of not only a set of relations, but also information after each relation, including

- The schema for each relation
- The domain of values associated with each attribute
- The integrity constraints
- The set of indices to be maintained for each relation
- The security and authorization information for each relation
- The physical storage structure of each relation on disk

Domain Types in SQL

Char(n), varchar(n), int, small int, numeric(p,d), real, double, precision, float(n), date, time, timestamp.

Schema Definition in SQL

- Primary key
- Check

4. EXPLAIN MECHANISM OF NESTED QUERIES.

SQL provides a mechanism for nesting subqueries. A subquery is a select from where expression that is nested within another query. A common use of sub queries is to perform tests for set membership, make set comparisons, and determine set cardinality.

Set membership

```
(select customer-name from depositor)
```

Set comparison

```
select distinct T.branch-name
from branch as T,branch as S
where T.assets > S.assets and S.branch-city='Brooklyn'
```

Test for Empty Relations

```
Select customer-name
from borrower
where exits (select *
from depositor
where depositor.customer-
name=borrower.customer-name)
```

Test for the Absence of Duplicate Tuples

5. WRITE SHORT NOTES ON MODIFICTION OF THE DATABASE.

Definition

delete from r where P

- **Insertion**

```
insert into account
values ('A-9732','perryridge',1200)
```

- **Updates**

```
update account
set balance=balance*1.05
```

- **Update of a view**
- **Transaction**

UNIT-3

1. DESCRIBE LOG BASED RECOVERY.

The most usably structure for recording data base modification is the LOG.

The log is a sequence of log records recording all the update activities in the data base.

There are several types of log records. An update log records describes a single data base write it has these fields

- **Transaction identifier**
- **Data item identifier**
- **Old value**
- **New value**

The various types log records as.

□ □ < $T_i$start >. Transaction T_i has started

□ □ < T_i, T_x v_1,v_2 > .Transaction T_i has performed a right on data item

□ □ < $T_i$commit > Transaction T_i has committed

□ □ < $T_i$about > Transaction T_i has aborted

- Deferred data base modification
- Immediate data base modification
- Check point
- Shadow paging

2. WHAT IS SERIALIZABILITY?EXPLAIN ITS TYPES.

The data base system must control concurrent execution of transactions, to ensure that the data base state remains consistent.There are different forms of schedule equivalence they lead to the notions of

- Conflict serializability
- View serializability

Conflict serializability

We say that i_i and I_j conflict if they are operations by different transaction on the same data item and at least one of these instruction is a write operations

View serializability

The concept of view equivalence leads to the concept of View serializability we say that a schedules S is view serializable if it is view equivalent to a serial scheduler

3.WRITE SHORT NOTES ON TRANSACTION STATE.

A transaction may not always complete its execution successfully such A transaction is termed aborted.

A transaction must be in one of the following states

- Active
- Partially committed
- Failed
- Aborted
- Committed

4. BRIEFLY DESCRIBE CONCURRENCYEXECUTION.

- Lock – based protocols
- Locks

There are various modes in which a data item may be locked in this section we restrict our attention to two modes

- Shared
- Exclusive

```
T1 : lock – x(B );
read(B);
B:=B-50;
write(B);
unlock(B);
Lock-x(A);
read(A);
A:=A+50;
write(A);
unlock(A).
```

5. EXPLAIN CONCURRENCY CONTROL.

Concurrency control

Oracles multiversion concurrency control differs from the concurrency mechanism used by some other data base vendors.

Read only queries are given a read –consistent snapshot which is view if the data base as it existed at the specific point in time, containing all update that we were committed by that point in time and not containing any updates that were not committed at any point in time thus read clock are not used in read only queries don’t interfere with other data base activity in term of locking.

Managed stand by data base

- To ensure high availability oracle provide a managed stand by data base
- future
- A stand by data base is a copy of the regular data base ie in solved on the separate system.
- If a catastrophic failure occur on the primary system, the stand by system is activate and take over there by minimizing effect on failure on a availability.
- Oracle keeps the stand by data base up to date by constantly applying archived redo logs that are shipped from the primary data base the

backup data base can be brought online in read- only mode and used for reporting and decision support queries.

UNIT-4

1. DESCRIBE FILE ORGANISATION.

A file is organized logically as a sequence of records. These records are mapped onto disk blocks.

Fixed-Length Records

```
type deposit=record
Accountnumber:char(10);
branch name:char(22);
balance: real;
end
```

Variable length records

storage of multiple record types in a file
Record types that allow variable lengths for one or more fields
Record types that allow repeating fields
Byte string Representation
Fixed length representation

- Reserved space
- List representation

2. DEFINE RAID. BRIEFLY EXPLAIN IT.

A variety of disk organization techniques, collectively called redundant arrays of independent disks (RAID)

- Improvement of reliability via redundancy.
- Improvement in performance via parallelism
  1. Bit level striping
  2. Block level striping

RAID levels

- RAID level 0
- RAID level 1
- RAID level 2(memory style error correcting code)
- RAID level 3 (Bit interleaved parity organization)
- RAID level 4 (Block interleaved parity organization)
- RAID level 5 (Block interleaved distributed parity)
- RAID level 6 (P+Q redundancy)

3.WRITE SHORT NOTES ON INDEX STRUCTURE OF FILES.

There are two basic kinds of indices

- Ordered indices
- Hash indices

Each technique must be evaluated on the basis of these factors:

- Access types
- Access time
- Insertion time
- Deletion time
- Space overhead

Ordered indices

- Primary index
- Dense and sparse indices
- Multilevel index
- Index update
- Secondary indices

B$^+$-Tree index files

B$^+$-Tree index structure is the most widely used of several index structures that maintain there efficiency despite insertion and deletion of data.

- Structure of B$^+$-Tree
- Queries on B$^+$-Tree
- Update on B$^+$-Tree
- B$^+$-Tree file organization
- B-Tree index files

4. EXPLAIN HASH FILE ORGANIZATION.

Hash functions

- **The distribution is uniform**

Hash functions assign each bucket the same number of search – key values from the set of all possible search-key values

- **The distributed in random**

In the average case each bucket will have nearly same no of values assigned to it, regardless of the actual distribution of search-key values

- **Handling of bucket overflows**
    - Insufficient buckets
    - Skew

Open hashing

Under an alternative approach called open hashing

Close hashing

The form of hash structure that we have just described is something referred to as close hashing.

Hash indices

5. WHAT IS MAGNETIC DISKS?EXPLAIN IT.

- Magnetic disk provides the bulk of secondary storage of modern computer
- system.
- The disk capacity is growing at over 50% per year but the storage requirements of large applications has also been growing very fast and in some case every faster than the growth rate of disk capacities**.**
- A large data base may require 100 of disks.

Physical characteristics of disk

- Physical disks are relatively simple. Each disc platter has a flat circular shape
- We can call magnetic disk as

  Hard disk
  Floppy disk
- The read write head store information on a sector magnetically as reversals of the direction of magnetization of the magnetic material.
- There may be hundreds of concentric tracks on a disc surface, containing thousands of sectors.

UNIT 5

1. WRITE SHORT NOTES ON DATA WARE HOUSING.

Data ware housing applications requires the transformation of data from many sources into a cohesive consistent step set of data configured appropriately for use in dataware house operation.

Distributed Transformation services

Data ware housing is an approach to manage data in which heterogeneous data sources are migrated to a separate homogeneous database

Online Analytical processing services

OLAP services provide server and client capabilities to create and manage multidimensional OLAP data .

2. EXPLAIN NESTED RELATIONS.

Nested relations

- The assumption of INF is a natural one in the bank examples we have considered.
- However, not all applications are best modeled by INF relations.
- The nested relational model is an extension of the relational model in which domains may be either atomic or relation valued.
- We illustrate nested relations by an example from a library. Suppose we store for each book the following in formations
  - Book title
  - Set of authors
  - Publishers
  - Set of keywords

We can see that if we define a relation for the preceding information, several domains will be monatomic

Authors
Keywords
Publishers
Complex types
Collection and large object types

```
Create table books(
…
Keyword-set setoff(varchar(20))
…
)
```

Structure types
Creation of values of complex types

3. WHAT IS INHERITANCE? DESCRIBE IT IN DETAIL.

Inheritance

Inheritance can be at the levels of types, or at the level of tables We first consider inheritance of types, then inheritance at the level of labels.

Type inheritance

Suppose that we have the following type definition for people

```
create type person
(name varchar(20)
address varchar(20))
```

Table inheritance

```
Create table people of person
```

The consistency requirements for sub tables are

1. Each tuple of the sub table can correspond to at most one tuple in each of its
   immediate sub tables.
2. SQL:1999 has an additional constraint that all the tuples corresponding to each other must be derived from one tuple .

Overlapping sub tables

4. WHAT ARE THE TYPES OF REFERENCE? EXPLAIN IT WITH SUITABLE EXAMPLES.

Object oriented language provided the ability to refer the object attribute of the type can be referred to the specified type. We can define the type dept with a field name and a field head which is reference to the type person and a table dept of the type dept as follows

```
Create type dept
(Name varchar(20),
  Head ref(person)scope people)

  Create table dept of dept
```

The table definition must specify that the reverence is derived and must still specify a self referential attribute name. When interesting a tuple for deptwe can then use

```
Insert into dept
Values('CS','john')
```

5. DESCRIBE QUERIES WITH COMPLEX TYPES.

The present extension of the SQL query language deal with the complex Type.

Let us start with the simple example:

Find the title and the name of the publisher of each book this query carries out the task:

```
Select title, publisher.name
```

From books

path expression

The reference are dereference in 1999 by the –> simple

Select head-> name, head->address

From dept

- The transformation of the nested relation in to a form with fewer (or no) the relation –valued attribute value is called unnesting
- The reverse process of transformation a INF relation into a nested relation is called nesting**.**

- **Nesting and unnesting** An expression such as” head->name” is called the path expression.
- **Collection valued attributes**

6. EXPLAIN XML IN DETAIL

- Structure
- Document schema
- Xml schema

7. EXPLAIN DISTRIBUTED DATABASE IN DETAIL

- Data replication
- Fragmentation
- Transparency

8. EXPLAIN DATA MINING IN DETAIL

- Classification
- Decision tree classifiers
- Regression
- Association

9. EXPLAIN OODM IN DETAIL

- Object structure
- Object classes
- Inheritance
- Multiple inheritance identity