

COMPUTER ENGINEERING DEPARTMENT

ASSIGNMENT NO-02

SUB: Database Management System

COURSE: T.E.

Year: 2020-2021

Semester: V

DEPT: Computer Engineering

SUBJECT CODE: CSC502

SUBMISSION DATE: 24/11/2020

=====

Name: Amey Thakur

Roll No.: 50

Batch: B3

Class: TE COMPS B

Q. No	Questions	CO Mapping
1.	Difference between DML and DCL commands.	CO4
2.	What is Normalization? Explain 1NF, 2NF, 3NF& BCNF.	CO5
3.	Write a short note on ARIES & Shadow paging.	CO6

Q.1 Difference between DML and DCL Commands.

Ans:

Basic Terms	DML	DCL
Full Form	Data Manipulation Language	Data Control Language
Used in	SQL (RDBMS)	SQL (RDBMS)
Commands	SELECT, INSERT, UPDATE, DELETE, MERGE, CALL	REVOKE, GRANT
Definition	DML statements are used for managing data within schema objects.	A Data Control Language is a syntax similar to a computer programming language used to control access to data stored in a database. (Authorization)
Classification	Procedural and Non-Procedural DMLs.	None.

Q.2. What is Normalization? Explain 1NF, 2NF, 3NF, BCNF.

Ans:

Normalization

Normalization is a process of designing a consistent database by minimizing redundancy and ensuring data integrity through decomposition which is lossless.

Goals of Database Normalization.

- (I) Ensure data integrity.
- (II) Prevents redundancy in data
- (III) To avoid data anomaly.

First Normal Form (1NF)

Definition.

- 1NF states that all attributes in relation must have atomic (indivisible) values and all attributes in a tuple must have a single value from the domain of that attribute.
- A relation is in 1NF, if every row contains exactly one value for each attribute.
- In short rules for data in 1NF is, A column in a table should contain only indivisible data.

Example :

- Consider an employee table with columns,
- The relational schema not in 1NF is represented as,

Employee Table

Employee-Id	Ename	Salary	Ecity
-------------	-------	--------	-------

- The state of Employee relational schema is as given below and it contains the Ecity which is non atomic (divisible) domain.

(Non-Normalized) Employee Table

Employee - Id	Ename	Salary	Ecity
10	Mahesh	50000	Mumbai, Pune
12	Suresh	25000	Mumbai
15	Ganesh	26000	Pune
18	Kasturi	50000	Mumbai, Delhi

- To convert relational schema in 1 NF, the ecity attribute is divided in atomic domains it may introduce some data redundancy.

1 NF Employee Table

Employee_id	Ename	Salary	Ecity
10	Mahesh	50000	Mumbai
10	Mahesh	50000	Pune
12	Suresh	25000	Mumbai
15	Ganesh	26000	Pune
18	Kasturi	50000	Mumbai
18	Kasturi	50000	Delhi

Information present in a relation is not redundant.

Second Normal Form (2NF)

Definition:

- A relation is in 2NF, if it is in 1NF and all non-key attributes in relation are fully functionally dependent on the primary key of the relation.

Example:

Employee (Employee-ID, Ename, Salary)

Employee-ID \rightarrow Ename, Salary

Employee-ID	Ename	Salary
10	Mahegh	50000
12	Suresh	25000
15	Ganesh	26000
18	Mahegh	50000

Third Normal Form (3NF)

Definitions

- A relation R is in 3NF if all non-prime attributes are,
 - ① Full, functionally dependent on primary key.
 - ② Non-transitive dependent on every key.
- A relational schema R is in 3NF, if non-trivial functional dependency $X \rightarrow A$ holds true where X is a superkey and A is a prime attribute.

Example:

- Consider an Employee Table with FDs.
- Employee (Employee-ID, Ename, Salary, Department-ID)
 $\text{Employee-ID} \rightarrow \text{Ename, Salary, Department-ID}$

3NF Employee Table

Employee-ID	Ename	Salary	Department-ID
10	Mahesh	50000	C1
12	Suresh	25000	E2
15	Ganesh	26000	C1
18	Mahesh	50000	E2

Department (Department-ID, Dname)

$\text{Employee-ID} \rightarrow \text{Dname}$

3NF Department Table

Department-ID	Dname
C1	IT
E2	HR

Boyce - Codd Normal Form (BCNF)

Definition

- A relation R is said to be in BCNF, if and only if every determinant is a candidate key.
- A relational schema is in BCNF, if a non-trivial functional dependency $X \rightarrow A$ is true then X is a superkey of relation R.
- In 3NF, A should be prime attribute, which is not the case in BCNF.

Example:

- Consider an Employee Table in which Employee can work in more than one department,
- The relational schema not in 2NF is represented as,
- Consider an Employee Table with following FDs,
 $\text{Employee-Id} \rightarrow \text{Ename, Salary, Department-Id}$
 $\text{Department-Id} \rightarrow \text{Dname}$.
- The state of Employee relational is,

Employee Table

Employee-Id	Ename	Department-Id	Dname	Dtype
10	Mahesh	C1	IT	Technical
12	Ganesh	E2	HR	Skill
12	Ganesh	C1	IT	Technical
10	Mahesh	E2	HR	Skill
13	Satish	E1	Ts.	Technical

To normalize above schema to BCNF, decompose table as,

Employee (Employee-Id, Ename)

Employee-Id \rightarrow Ename

The determinant Employee-Id is candidate key.

Q.3. Write a short note on ARIES & Shadow Paging

Ans:

ARIES

- ARIES is a recovery algorithm that is designed for no-force type of backup approach.
- Recovery manager is generally called when there is a crash.
- Restart can be proceeded in three different phases.

Principles of ARIES algorithm

(i) Write-ahead logging

- Any change to a database object is first recorded in some log file.
- The record in the log must be written to any of stable storage before the change in database is written to disk.

(ii) Repeating history during Redo.

- ARIES finds all operations done by DBMS before the crash and restores system back to the same state that it was in at the first time of the crash.
- Then, it aborts all actions of transactions those are still there in active state at the time of the crash.

(iii) Logging changes during Undo.

- We make changes to the database during restoring database all transactions are logged in same order.
- So, it ensures same action is not repeated in the event of repeated restarts.

Phases of ARIES algorithm

(I) Analysis Phase

- It first finds dirty pages (data changes those are not committed to database) in the available buffer pool.
- It also identifies all active transactions at the time of the system crash.
- Identify Redo LSN from which redo should start.

(II) Redo

- In order to restore database system will repeat all actions performed on database from start of log or from any selected point in log or from Redo LSN.
- Then it restores the database state to state at which it was at the time of the system crash.
- RecLSN and Redo LSN avoid Redo action already reflected on page.

(III) Undo,

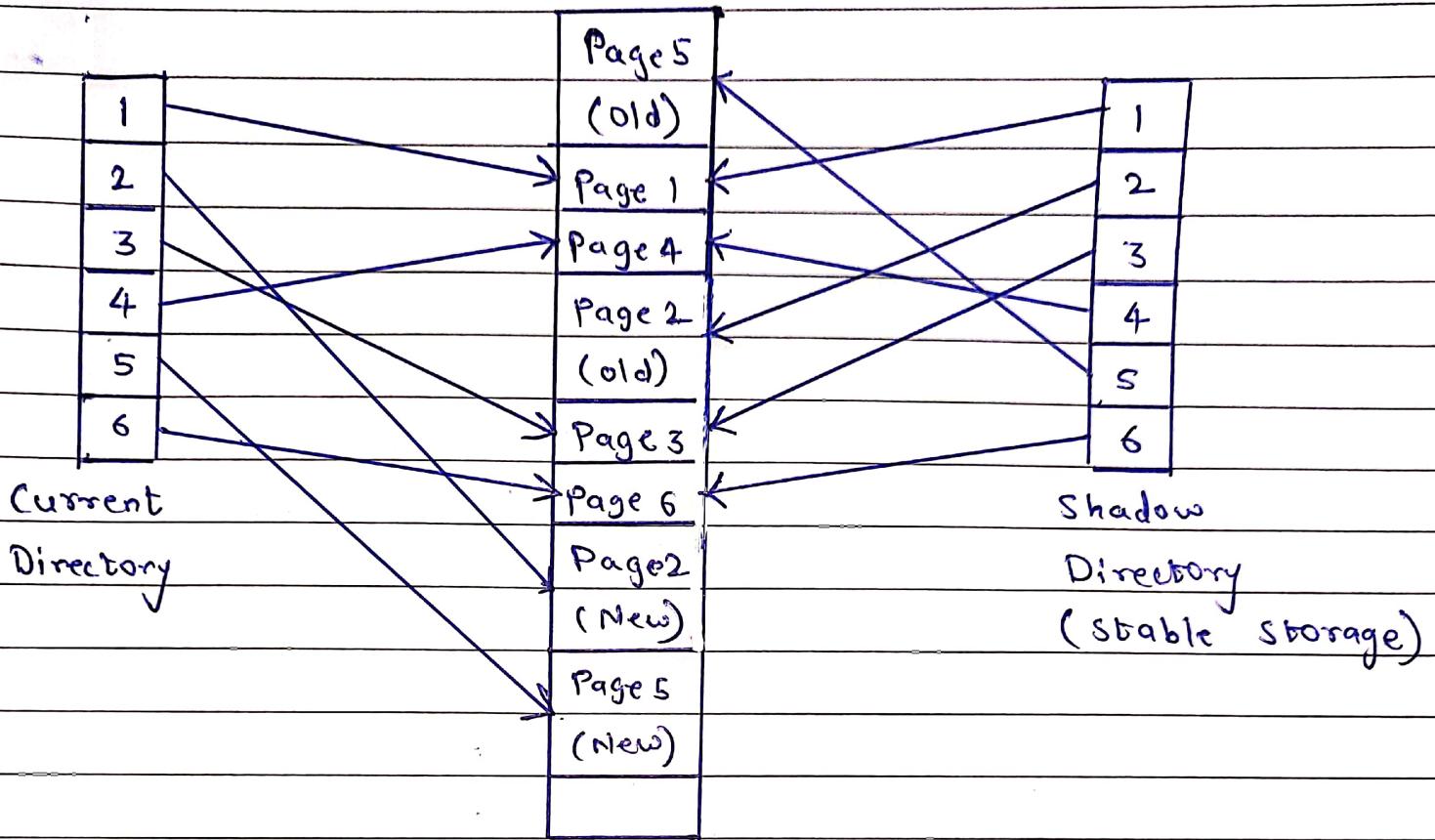
- It reverses or undoes all operations of transactions which are not committed.
- So after above action now database only reflects actions which are committed transactions.

Shadow Paging

- It is not always convenient to maintain logs off all transactions for the purpose of recovery.
- An alternative is to use a system of shadow paging.
- This is where the database is divided into pages that may be stored in any order on the disk.
- In order to identify the location of any given page, we use something called a page table.

Method

- During the life of a transaction two page tables are maintained as below,
 - I Shadow page table
 - II Current page table
- When a transaction begins both of these page tables point to the same locations (are identical).
- During the lifetime of a transaction the shadow page table doesn't change at all.
- However during the lifetime of a transaction update values etc. may be changed.
- For pages updated by the transaction, two versions are kept. The old version is referenced by the shadow directory and the new version by the current directory.
- So whenever we update a page in the database we always write the updated page to a new location.
- This means that when we update our current page table it reflects the changes that have been made by the transaction.



Pagetable - shadow paging

- The shadow page table shows the state of the database just prior to a transaction, and the current page table shows the state of the database during or after a transaction has been completed.