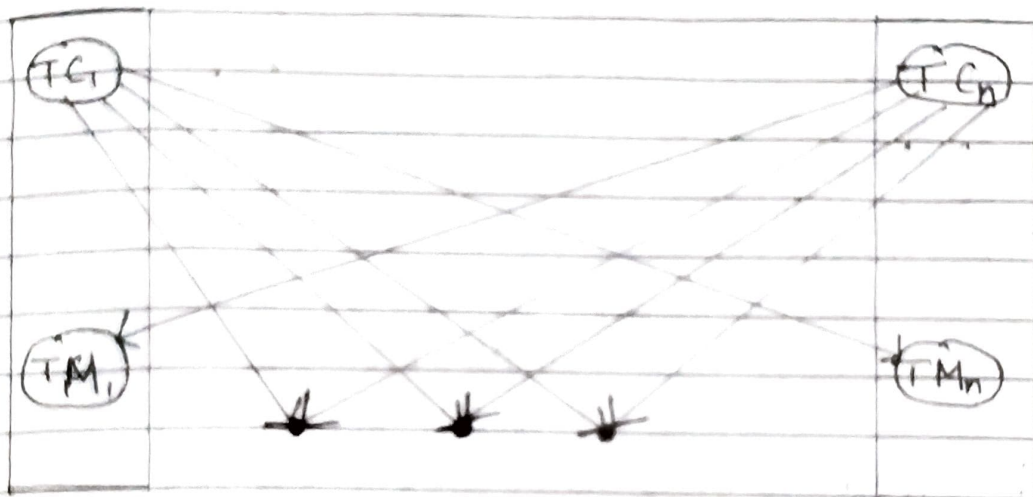


DBMS

ASSIGNMENT - II

Q1.

1. DISTRIBUTED DATABASE



A database which is not constraint to a single system, which is spread over multiple sites is known as Distributed system.

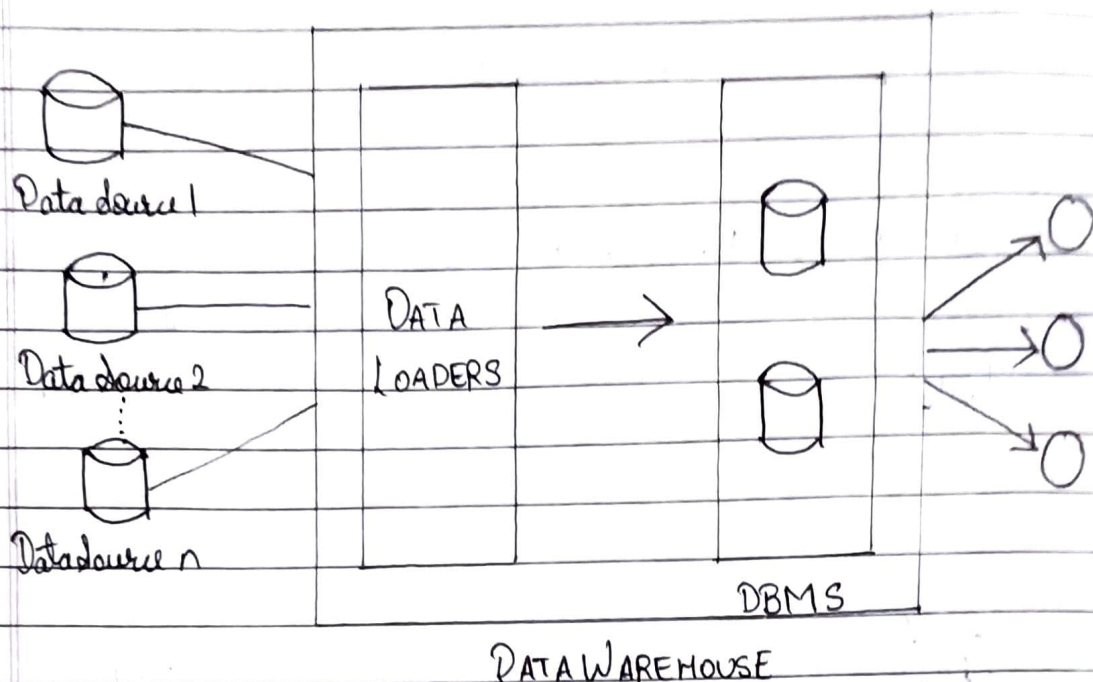
There are 2 Types of Distributed Systems:

- **HOMOGENOUS DATABASE**: All sites store identical sites database. OS, DBMS and data structures are same at all sites, Hence are easy to manage.
- **HETEROGENOUS DATABASE**: Different sites or computers may use different OS, different application database. They may even use different data models for database. Hence, translation will be required for communication between sites.

ADVANTAGES

- Modularity
- Fault Tolerance
- High Performance
- Data Sharing
- Low Cost Components

2 DATA WAREHOUSING



Data Warehouse stores huge amount of data which collected from multiple Heterogeneous Sources. It aims to produce statistical results ~~the~~ to improve the decision-making.

Eg: An employee ~~the~~ want to check the amount of ~~money~~ supplies to a given company of various competition company.

3 DATA MINING

Data Mining is the extraction of data from a data warehouse. Its main purpose is to gather information which can be used to predict future trends etc.

It has to be done in 3 steps namely:

- DATA PRE-PROCESSING: Data cleaning, integrating ~~the~~ selection & transformation takes place.
- DATA EXTRACTION: Extraction of a particular data.
- DATA EVALUATION AND PRESENTATION

Q2 - IT MAKES DATA UNREADABLE

It can only be read or updated again by first decrypting it with a corresponding decryption key. As a result, the data becomes strongly secured and unauthorized access can be easily prevented.

- HACKERS CAN'T MANIPULATE DATA SINCE DATA INTEGRITY

Even if hackers access the encrypted data for manipulation or misuse. They cannot do it as the encrypted data doesn't make sense.

Q3. SQL INJECTION is one of the most common web hacking techniques which might destroy the database. It is basically placement of malicious code in SQL statement via a webpage input. The web servers communicate with database servers anytime they need to retrieve or store user data. Attackers design the SQL statement such that they can be executed when the fetching of content is in process by web-servers.
Eg: For an entity named SAP-ID if the attacker input data 0=0 which is true then the user data is compromised.

Q4. Authentication is the process of verifying a user's credentials before he can log into the network. Normally performed using a username & password. This provides a secure means of limiting the access of unwanted intruders of the network. It's the process of determining whether someone or something is actually what it had told to be.

Q5.- AUTHORIZATION TO READ DATA

Used for confidential data which should only be seen by selected personnel.

- AUTHORIZATION TO INSERT DATA

Used for confidential data which should only be added by selected personnel.

- AUTHORIZATION TO UPDATE DATA

Used for confidential data which should only be updated by selected personnel.

- AUTHORIZATION TO DELETE DATA

Used for confidential data which should only be deleted by selected personnel.

Q6. DATA SWARM

It is a framework used for writing data processing pipelines. It has an extensible library of operations. Eg: moving data, executing query etc. developers define dependency graph of the task to be executed & data swarm takes care of that.

MAC

~~Media~~ Access Control Address also known as physical & hardware address. It is uniquely formatted into hexadecimal format & given later to each computer or network device on a computer network. Eg: D3-BA-D8-7C-4B-9B.