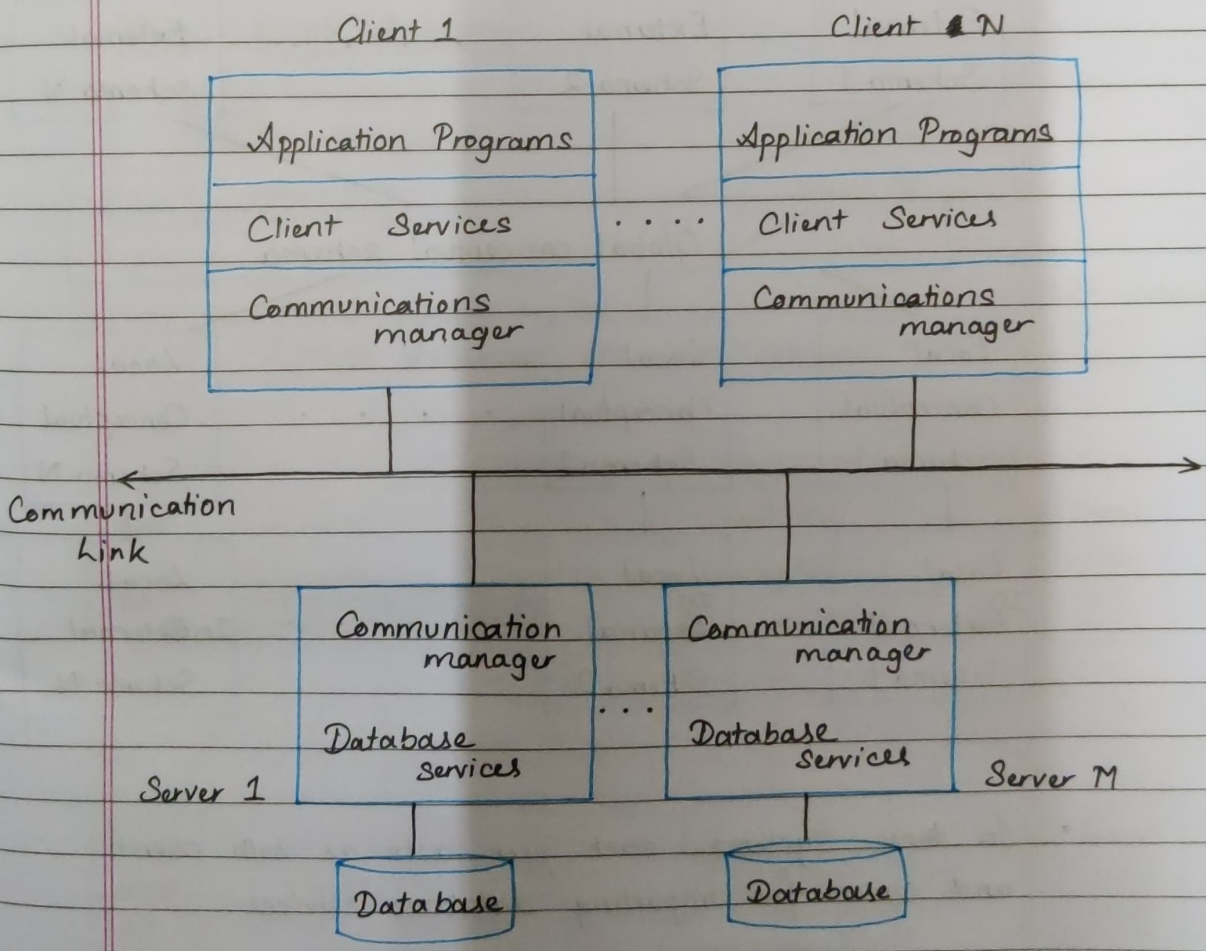
Unit - 1

Q1. Architecture Models

Some of the common architectural models are →

- Client - Server architecture for DDBMS
- Peer - to - Peer architecture for DDBMS
- Multi DBMS architecture

- Client - Server architecture →

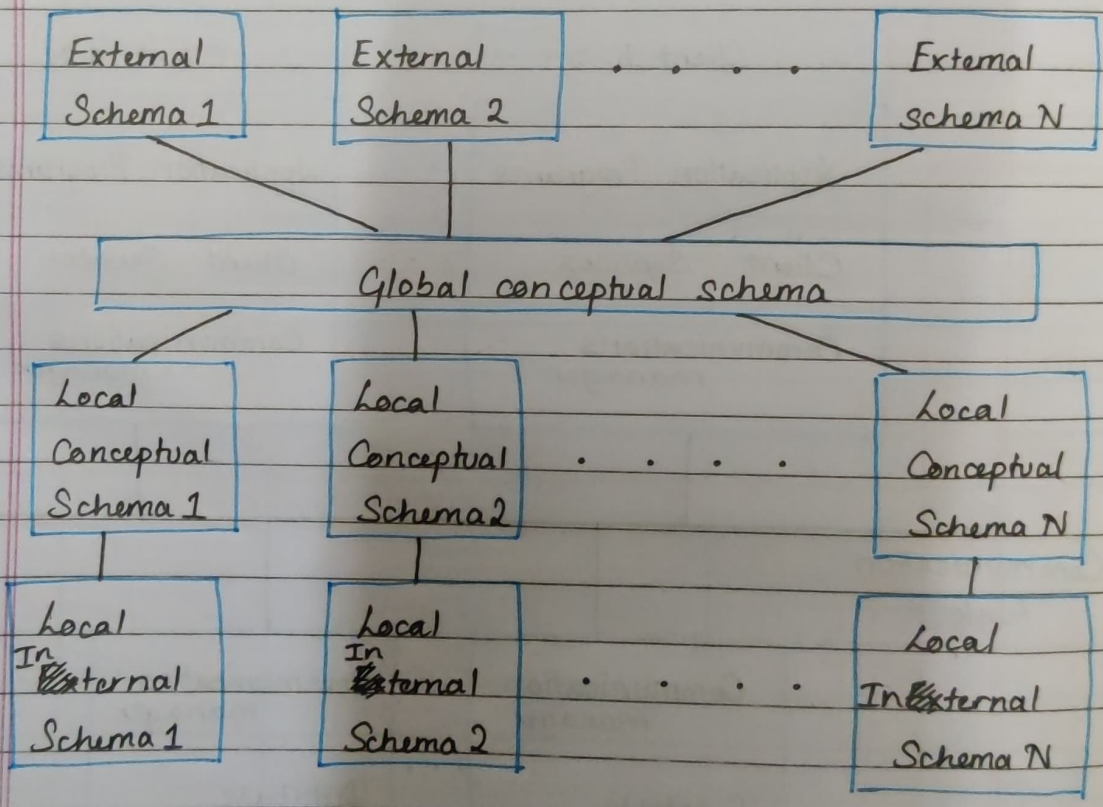


- This is a two level architecture where functionality is divided into servers and clients
- Client functions mainly include the user interface

- Server functions mainly include data management, transaction management, query processing etc.

- Two types of client-server architecture
  - → single server multiple client
  - → multiple server multiple client

- Peer-to-Peer architecture →



- In these systems, each peer acts as both client and server for imparting database services

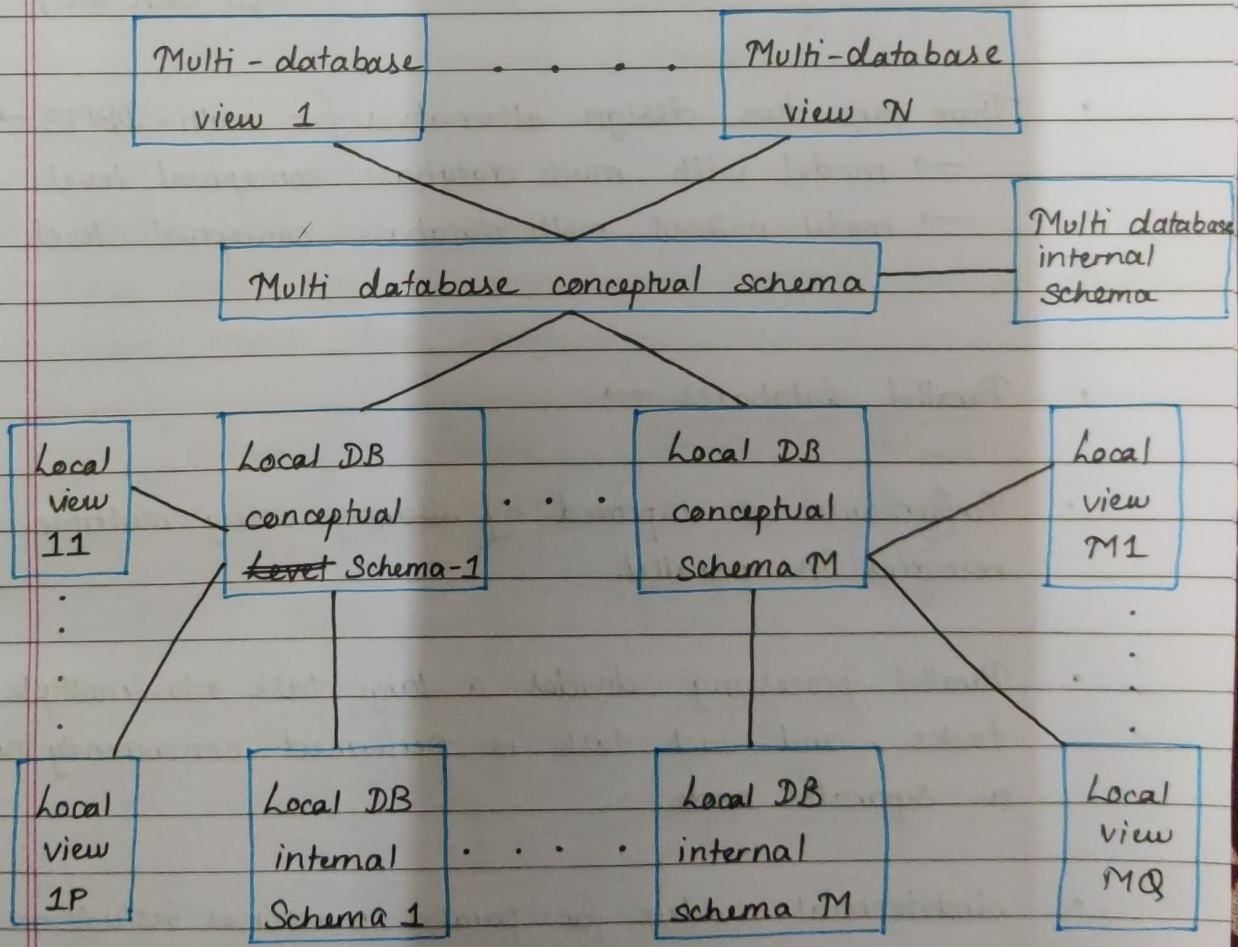
- The peers share their resources with other peers and coordinate their activities

- The architecture has 4 levels of schemas -

→ External schemas (Support user applications and user accers)
→ Global conceptual schema (enterprise view of data)
→ Local conceptual schema (local organization of data at each site)
→ Local internal schema (individual internal schema definition)

- Major components of a peer-to-peer system
  → User processor - user interface handler, global query optimizer
  → Data processor - Local query optimizer

- Multi - DBMS Architecture →



- This is an integrated database system formed by the collection of two or more autonomous database systems

- Multi - DBMS has 6 levels of schemas -

→ Multi - database view level (depicts multiple users views)
→ Multi - database conceptual schema (depicts integrated multi database)
→ Multi - database internal schema (depicts data distribution across different sites
→ Local database view level (depicts public view of local data)
→ Local database conceptual schema (local data organization at each site)
→ Local database internal schema (physical data organization at each site)

- There are two design alternatives for multi - DBMS →
  → model with multi database conceptual level
  → model without multi database conceptual level

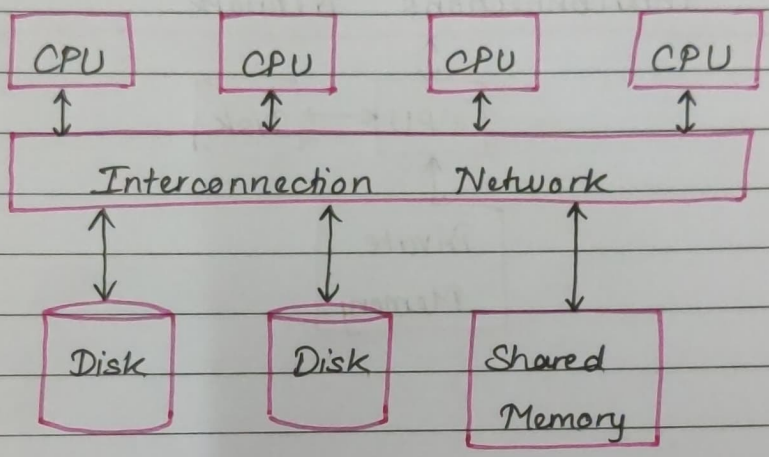
- Parallel databases →

- Performance is improved by ~~sharing~~ using multiple resources in parallel.

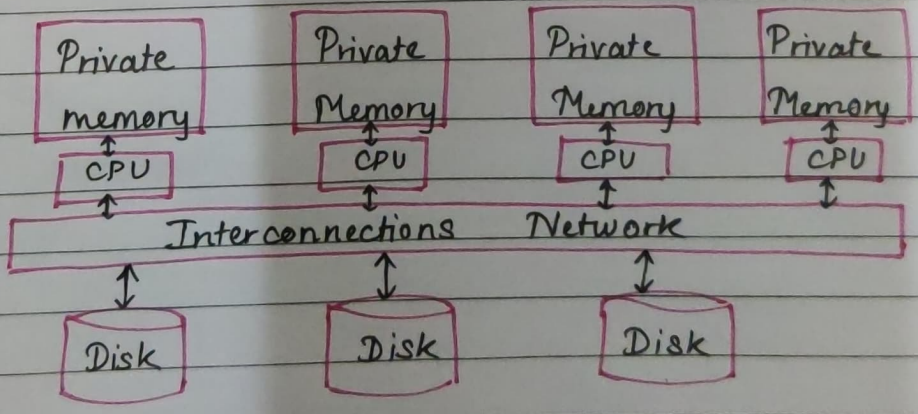
- Parallel processing divides a large task into multiple tasks, and each task is performed concurrently on a separate node.

- Architectural models for parallel machines include →

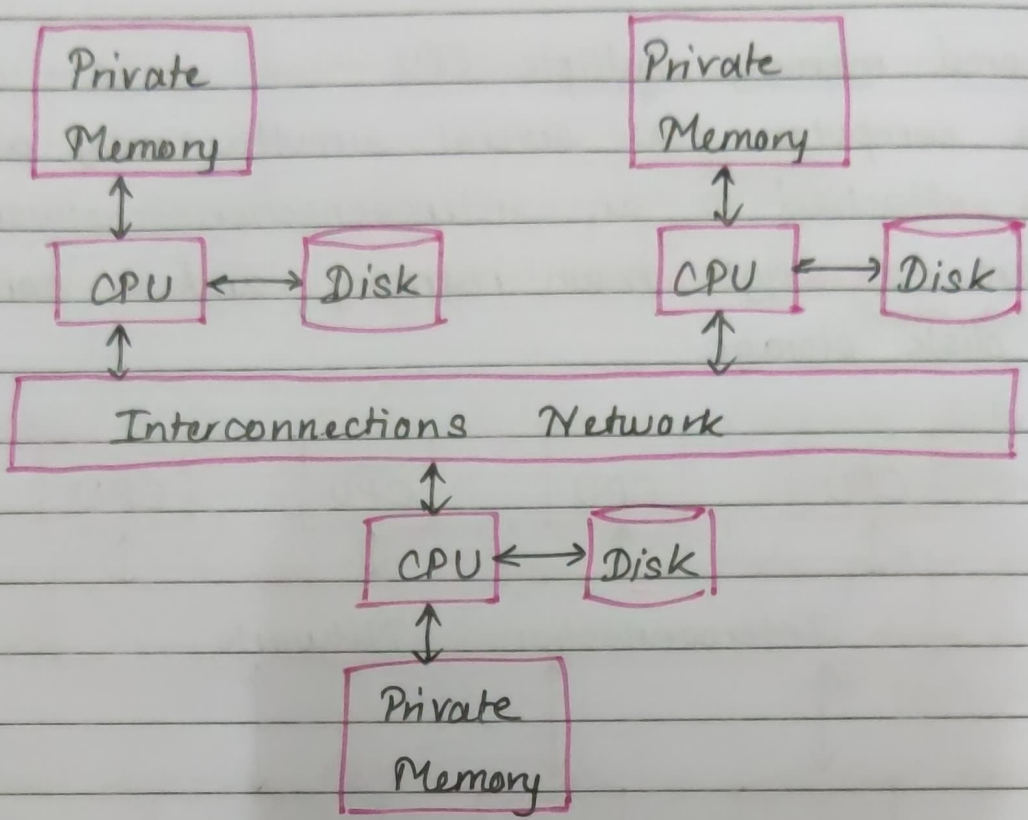
1. Shared memory multiple CPU –
The computer has several simultaneously active CPUs that are attached to an interconnections network and share a single main memory and a common array of disk storage



2. Shared disk architecture –
Each node has it's own main memory, but all nodes share mass disk storage



3. Shared Nothing architecture –
Each node has it's own mass storage as well as main memory


Private Memory
Private Memory
CPU
Disk
CPU
Disk
Interconnections Network
CPU
Disk
Private Memory

UNIT-I

Data Fragmentation:-

* Fragmentation is a process of disintegrating relations or tables into several partitions in multiple sites.
* It divides a database into various subtables and sub relations so that data can be distributed and store efficiently
* Database fragmentation can be two types.
  → horizontal
  → Vertical

In horizontal fragementation each tuple of a relation r is assigned to one or more fragments.

In Vertical fragmentation, the scheme for a relation r is split numerous smaller schemas with a common candidate key and a special attribute

Methods of Data fragmentation of a Table

* Horizontal fragmenation
* Vertical fragmentation
* Hybrid fragmentation

* **Horizontal fragmentation:-**

Fragmenation should be done in a way so that original table can be reconstructed from the fragments. this is needed so that the original table can be reconstructed from the fragments.

Vertical Fragmentation

In Vertical Fragmentation, the fields or Columns of a table are grouped into fragments. In order to maintain reconstructiveness.

Ex:- STUDENT

Regd-No Name Course Address Semester Fee Marks

```
CREATE TABLE STD_FEES AS
SELECT Regd_NO, Fees
FROM STUDENT;
```

The Fee details are given.

Horizontal Fragmentation:-

It groups the tuples of a table in accordance to values of one or more Fields.

Ex

```
CREAT COMP_STD AS
SELECT * FROM STUDENT
WHERE COURSE = "Computer Science";
```

Hybrid Fragmentation

In hybrid Fragmentation a Combination of horizontal and Vertical Fragmentation techniques are used.

two ways

* At First, generate a set of horizontal fragments, then generate Vertical Fragments from one or more of the horizontal Fragments.
* At First, generate a set of Vertical fragments then generate horizontal Fragments from one or more of the Vertical Fragments.

Data Replication:-

* Distributed database Replication is the process of Creating and maintaining multiple Copies of data in differerit Sites.
* The main benefit it brings to the table is that duplication of data ensures faster retrieval
* ~~How~~ e

Adv of Data Replications

Reliability - In Case any failure of any site, the database system Continues to work since a Copy is available at another site.

Reduce in Network Load - Since local Copies of data are available, query processing Can be done with reduced network usage.

Quicker Respone — Availability of local Copies of data ensure quick query processing and quick response.

Simple Transactions - Transactions require less number of join of table located at diffient sites and minimal Coordination across network.

Disadvantages:

Increased storage Requirements — Maintaing multiple Copies of data is associated with Increase Storage Cost.

Increased Cost – Each time a data item is updated, the update need to be reflex in all the Copies of the data at the different sites.

Some Commonly used Replication

* Snapshot replication
* Near-real-time replication
* Pull replication.

Distributed Database (DDB)

→ Distributed Database (DDB) is a collection of multiple, logically interrelated databases distributed over a computer network.

→ Distributed Database is a database that stores data in multiple locations instead of one location.

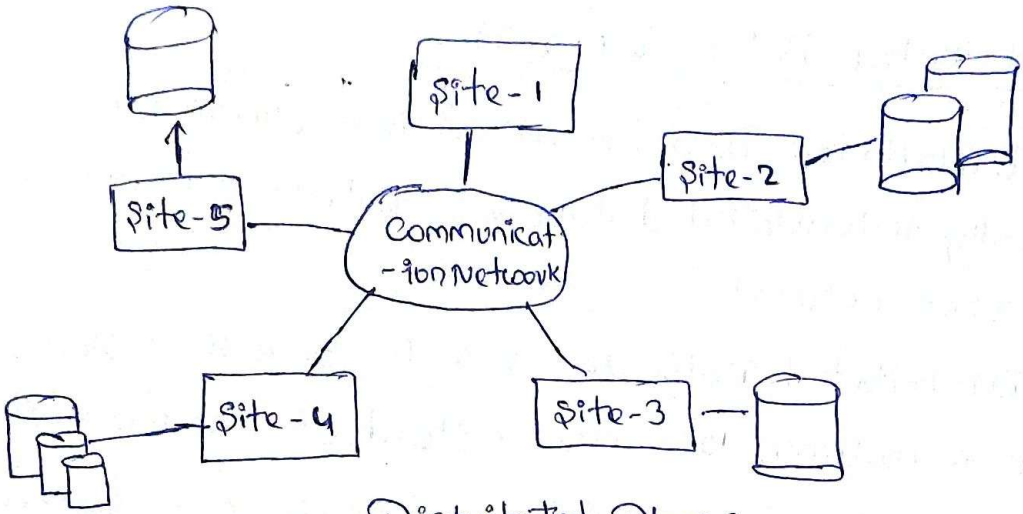
→ A Distributed Database management system (D-DBMS) is the software that manages the DDB and provides an access mechanism.

→ Distributed Database System (DDBS) =
DDB + D-DBMS.

→ Distributed databases
(Ex:- WWW, Cloud, Sensors, Mobiles, ...)

Features of Distributed ~~System~~ database

=> Location independent
=> Network independent
=> Distributed query processing.
=> Hardware independent.
=> DBMS independent.
=> Operating System independent.



Distvibuted Dbms environment

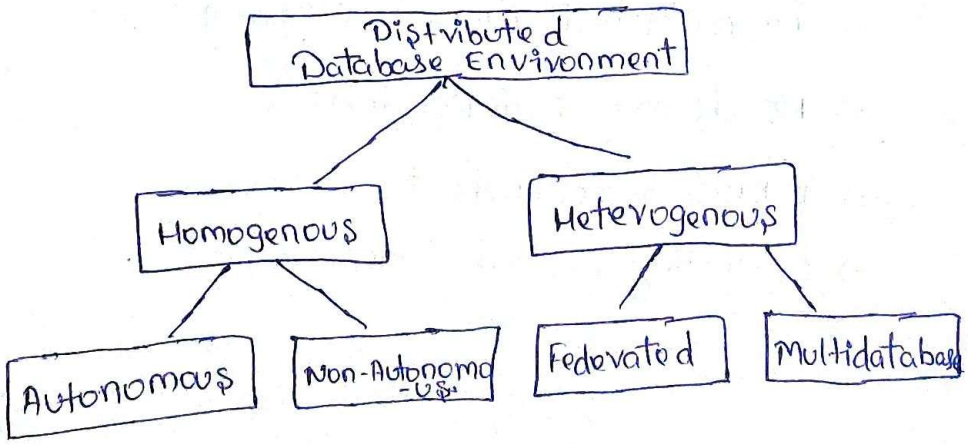
Advantages:-

→ Modular Development
→ More veliable
→ Better Response.
→ Lower Communication Cost.

Disadvantages:-

→ Comple x nature and expensive software.
→ Overall Costs
→ Securityissuses
→ Data integrity.

Types of Distvibuted Databases



Homogenous Distributed Database :-

In Homogenous Distributed Database, all the sites use idential DBMS and operating systems.

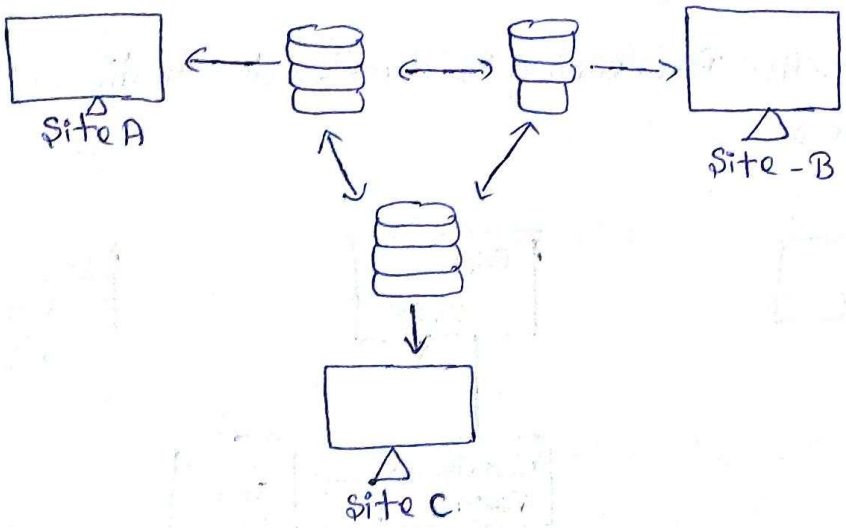
→ Homogenous Distributed database is a network of identical databases stored on the multiple sites.

→ It is making them easy to manage.

Properties :-

→ Use very similar software

→ Sites use idential dbms or dbms from the same vendor.



There are two types of Homogeneous Distributed Database.

→ Autonomous - Each

Autonomous homogenous database is consist of nodes that operate independently and exchange information with eachother using message Parsing.

→ Non Autonomous

Non-Autonomous homogenous database are Co-ordinated by a central system across all the sites.

Heterogenous Distributed Database:-

→ It is opposite of Homogenous Distributed database.

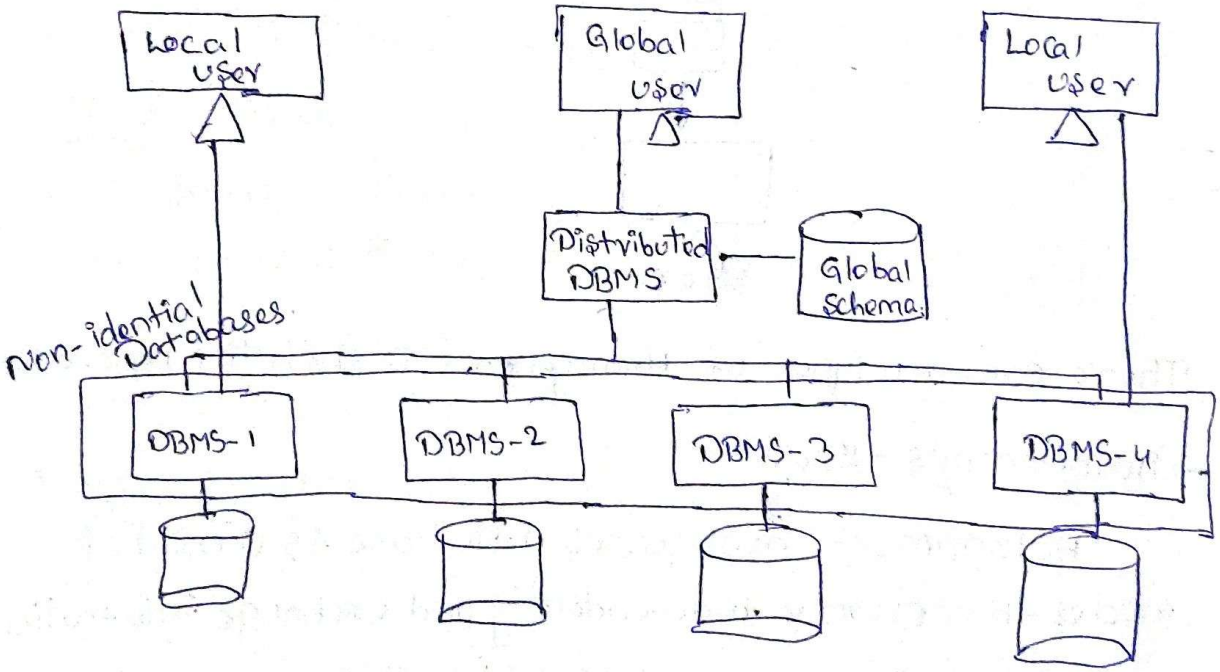
→ In Heterogenous Distributed Database, different sites have different operating systems, DBMS Products and data Model causing difficult to manage.

Properties

→ Different sites use dissimilar schemas and software

→ Query Processing is complex due to dissimilar schemas.

→ Transaction Processing is Complex due to dissimilar software's.



Types of Heterogenous Distributed Database.

Federated:-

The Heterogenous distributed database systems are independently in nature and integrated together so, that they function as a single database system.

un-federated:-

The database systems employ a central coordinating module through which the databases are accessed.