

# DISTRIBUTED DATABASE

HITHIKSHA.P  
SAMRITHA.V  
MADHUMITHA.RG

## Abstract

A distributed database is one in which hardware and software at network communication. It coordinates their movement only by passing the messages. In short distributed database is the collection of database that can be stored at different location or different network site. A distributed database is one in which memory location are not all attached to a common processing unit for instance the CPU. This paper presents a complete view on the distributed database. This paper contains various types like replication, fragmentation and various problems that can be in distributed database system

## 1. Introduction

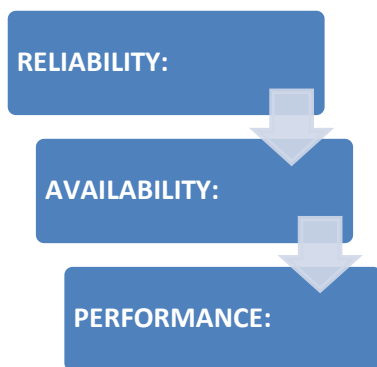
A distributed database management system has defined the software system that manages a distributed database in such a manner as if it were all stored in a single location. A distributed database is a collection of multiple consistent databases. It is spread physically across various location that communication via a computer network. In distributed database there are a number of database that may be geographically distributed all over the world. A distributed database management system manages the data in such a manner

that the data appears as a single database to the user.

## 2. Features of Distributed Databases

- ❖ It is used to create, retrieve, update and delete distributed database.
- ❖ It ensures that the data modified at any site is universally update
- ❖ it maintains confidentiality and data integrity of the database
- ❖ Data is physically stored across multiple sites.
- ❖ Data in each site can be managed by DBMS independent of the other sites.
- ❖ It is used in application areas where large volume of data are processed and accessed by numerous users simultaneously.
- ❖ A distributed database is not a loosely connected file system.
- ❖ A distributed database incorporates transaction processing.
- ❖ It is not synonymous with a transaction processing system.
- ❖ It is designed for heterogeneous database platform.

## 3. Goals of Distributed Database



### 3.1 RELIABILITY

In the distributed database if one system stops working the other system can complete the work or the task.

### 3.2 AVAILABILITY:

It means that in the database the given resources can be accessed by the users. In other words availability means the user of the database is available to access.

### 3.3 PERFORMANCE:

Performance can be achieved by distributing database over different locations, it makes the data available in every location which is easier to maintain.

## 4. Advantages of Distribution Data

Compared to a single, centralized system distributed system provides advantages. They are

- ❖ The database is already spread across multiple systems and it is not too complicated to add a system.
- ❖ In the distributed database data can be stored or arranged in different levels of transparency.
- ❖ These data's can be stored in different locations.

- ❖ The data can be stored to the according departmental information in an organization.
- ❖ It makes the organization easier to hierarchical access
- ❖ If the data nodes goes offline , the rest of the data can continue its function
- ❖ It the cheapest way to create a network of system containing a part of the database.
- ❖ This data can also be easily increased or decreased.

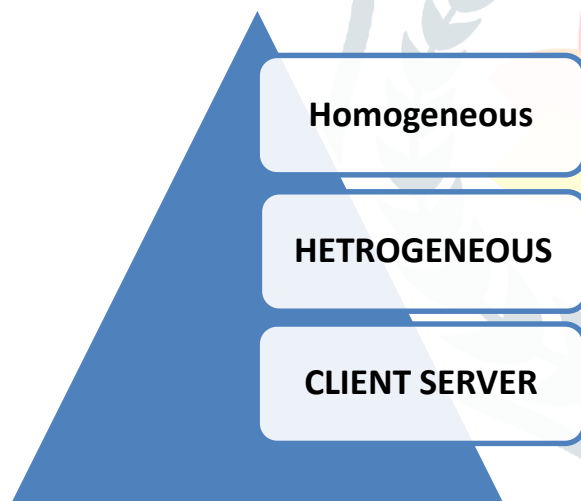
## 5. Disadvantages of Distributed Database

- ❖ The distributed data base is quite complex and complex too.
- ❖ In this the user must make sure that gets a uniform view of the database because it is spread across the multiple locations
- ❖ This database is very complex and very difficult to maintain.
- ❖ It is difficult to provide security to the database as it is located at different locations it must be secured.
- ❖ Moreover in the distributed database the infrastructure connecting all the nodes must also be secured.
- ❖ In the distributed database it is difficult to maintain data integrity because of its nature.
- ❖ In this database there may be redundancy in the database it is stored at multiple locations
- ❖ It is difficult to find people with the necessary experience who can manage and maintain it.

## 6. Objectives of Distributed Database:

- ❖ Should understand all the advantages and disadvantages.
- ❖ To know the issues involved in design distributed database
- ❖ Control and recovery the techniques of distributed concurrency.
- ❖ Identify and explain the interaction between different components is of client or server architecture as applied to database system.
- ❖ Query optimization as applied to distributed database.

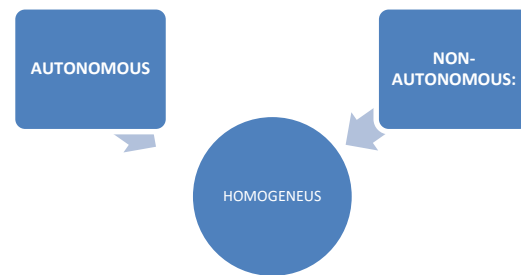
## 7. Types of Distributed Database System



Distributed database can be broadly classified into homogeneous and heterogeneous

- ❖ **Homogeneous**
- ❖ All the sites are identical in homogeneous distributed database.
- ❖ Each site is aware of all the other sites to process user request.
- ❖ If it is a single database then the database is accessed through single interface
- ❖ These sites use similar software.

The homogeneous databases are further divided into two types:

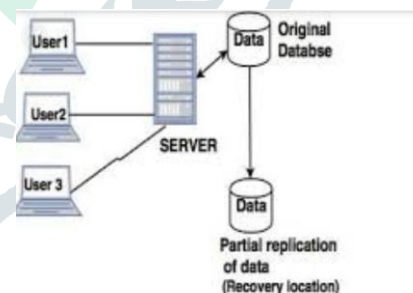


### ❖ AUTONOMOUS:

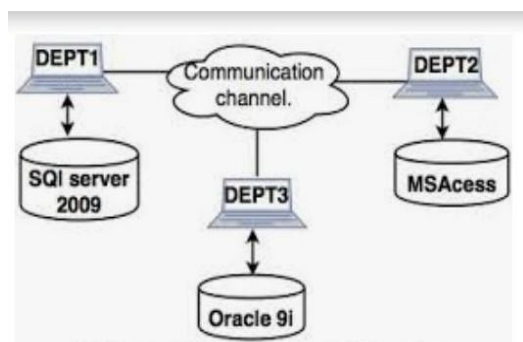
Each database is independent and functions on its own. The data updates are shared through message passing. The controlling applications are integrated.

### ❖ NON-AUTONOMOUS:

The central or master of the DDBMS coordinates data updates across the sites. Data is distributed across homogeneous nodes.

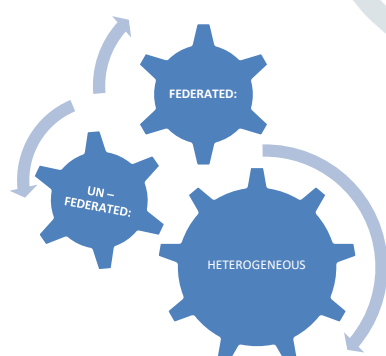


## 7.2 HETEROGENEOUS:



- ❖ Different sites have different operating system, DBMS products and data models.
- ❖ There may be limited co-operation in processing user request because the sites may not be aware of other sites
- ❖ Different site uses dissimilar schemas and software.
- ❖ Due to dissimilar software transaction processing is complex
- ❖ Relational, network, hierarchical or object oriented are the variety of DBMS which system may compose.
- ❖ Query processing is complex due to dissimilar software.

The heterogeneous databases are future divided into two types:



### ❖ FEDERATED:

The heterogeneous database systems are independent in nature and integrated in nature. They function as a single database system.

## ❖ UN –FEDERATED:

The database system employs a central coordinating module through which data can be accessed.

### 7.3 Client server

Client server systems are constructed that the database can be reside on central computer which is known as server. This can be shared among several users. The user can access the server through a client or a client application. In a large client/server system thousand of user may be connected to the sql server and can be installation at the same time. The sql server safeguards and prevent from the problems such as multiple user trying to update the same piece at the same time. This client server has the ability to use familiar tools on client machine.

## 8 Principles of distributed database

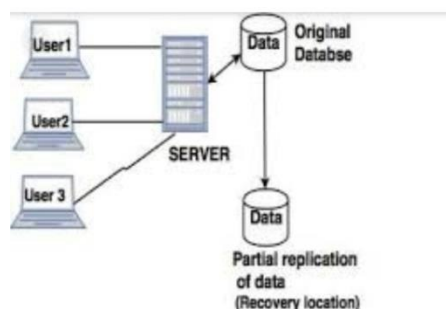
- ❖ Replication
- ❖ Fragmentation

### 8.1 REPLICATION:

Replication stores data in more than one site. It is very useful in improving the availability of the data. It is simply a copying of the data from a database from one server. So the user can use the same data without any inconsistency. For the result the user can access data relevant to their tasks without interfering in others work. However in data replication the data is stored in different locations but a particular relation has to

reside at only one location. The replication can be

- ❖ **Full replication:** Is the one in which the whole database is stored at every site.
- ❖ **Partial replication:** Only some frequently used fragments of the database are replicated and other are not replicated.



## 8.2 Fragmentation

It is a database server that allows you to control the data where the data is stored at the table level. It enables the user to define groups of rows or index keys within a table according to some algorithm or scheme. The user can store each group or fragment in a separate database space. And it must be associated with a special physical disk.

❖ **Horizontal Data Fragmentation:** horizontal data fragmentation divides a relation (table) horizontally into the group of rows to create subset of tables.

EXAMPLE: Account (acc no., balance, branch, name)

In this example if values are inserted in table branch name as Delhi, Chennai, Rajasthan.

The query will be written as: `select*from acc where branch name='Delhi';`

❖ **Vertical Data Fragmentation:** vertical data fragmentation divides a relation (table) vertically into the group of rows to create subset of tables.

EXAMPLE: Account (acc no., balance, branch, name)

In this example if values are inserted in table branch name as Delhi, Chennai, Rajasthan.

The query will be written as: `select*from acc _no`

## 9. Problems of DDBMS

The problematic area of database is

**9.1 The Ideal Situation:** The geographical or local in which every single application should be able to work transparently on data. That is

- The data must be spread across a variety of different DBM's.
- It should run on the variety of different machines.
- It must be supported by a variety of different operating system.
- Needs to be connected by a variety of different communication networks.

## 10. Conclusion

The distributed database is the current scenario of the fast changing world. There are many tools and techniques that can support the management of distributed database. Distributed database system also discussed the various techniques for concurrency control. Distributed database

has its own advantages and disadvantages and problems also. This paper presents a complete view on the distributed database system. The problem area describe in the paper are very useful while implementing distributed database system.

## 11. REFERENCE

<https://www.tutorialspoint.com>

<https://en.m.wikipedia.org>

[www.peter-lo.com](http://www.peter-lo.com)

[www.yugabyte.com](http://www.yugabyte.com)

<https://cs.uwaterloo.ca>

