**Distributed Database System and its Types**

A distributed database is like having pieces of a puzzle spread across different computers instead of just one. Each computer holds some of the data, and they can be on the same network or separate. This way, instead of cramming everything into a single database, the data is split up and stored in different places that don't rely on each other physically.

A distributed database is like a big team of computers that work together. Each computer, or node, stores some of the data, and together they make up the whole database.

With a distributed database, data is spread out across these computers, and they share the work of storing and managing it. Even if there's a problem with one computer or the network, the system keeps the data safe and available for users.

The main aim of a distributed database is to make sure big applications can handle lots of data smoothly, without crashing or slowing down.

With the use of distributed databases, we spread the data across multiple locations. This solves a bunch of problems. If one part of the system fails, the rest can keep working. Plus, we can make things run faster and handle more users because the workload is shared.

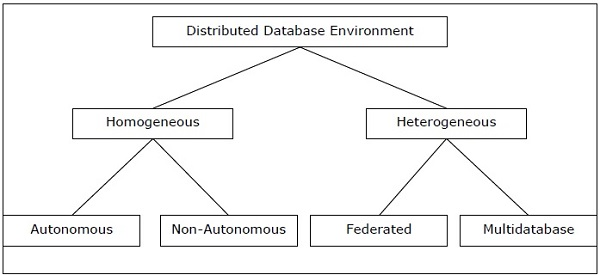
So, distributed databases help with things like keeping things running smoothly, even if something breaks, handling lots of users at once, and making sure everything works quickly.

**Distributed databases have several key features:**

1. **Location independence:** Data is stored at different sites and managed by separate Distributed Database Management Systems (DDBMS). Each site operates independently.
2. **Network connectivity:** All distributed databases within a network are linked, allowing them to communicate with each other. This enables seamless data exchange between different locations.
3. **Distributed query processing:** This involves handling queries, particularly read operations on large datasets, across a distributed environment. It includes converting high-level queries (like SQL) into execution plans and carrying out these plans effectively.
4. **Hardware neutrality:** The sites where data is stored are not dependent on specific hardware. They can operate independently of the physical equipment, often achieved through virtualization techniques.
5. **Distributed transaction management:** Distributed databases ensure consistency through various mechanisms such as commit protocols, recovery methods, and concurrency control techniques. These ensure that transactions remain reliable even in the face of failures across different parts of the system.

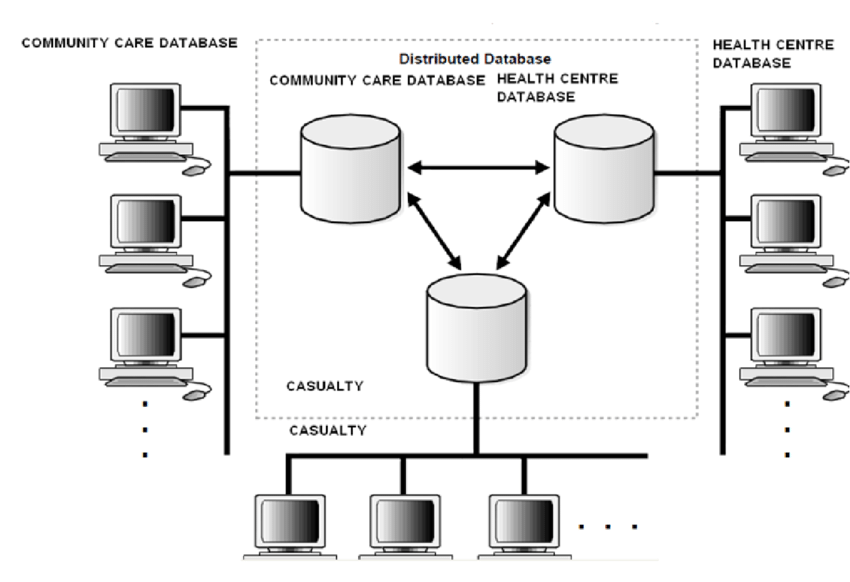
**There are two main types of distributed databases:**

1. **Homogeneous distributed database.**
2. **Heterogeneous distributed database.**

****

**Homogeneous Distributed Database:**

A homogeneous distributed database is a network of identical databases stored on multiple sites. This means all the databases store data in the same way, using the same operating system, Distributed Database Management System (DDBMS), and data structures. This uniformity across all sites makes them easier to manage.



**Salient Features:**

* The sites use very similar software.
* The sites use identical DBMS or DBMS from the same vendor.
* Each site is aware of all other sites and cooperates with other sites to process user requests.
* The database is accessed through a single interface as if it is a single database.

**Types of Homogeneous Distributed Database**

**There are two types of homogeneous distributed databases −**

* Autonomous − Each database is independent and functions on its own. They are integrated by a controlling application and use message passing to share data updates.
* Non-autonomous − Data is distributed across the homogeneous nodes and a central or master DBMS coordinates data updates across the sites.

**Heterogenous Distributed Database**

A Heterogeneous distributed database is the opposite of a Homogeneous distributed database. Unlike the uniformity of Homogeneous databases, Heterogeneous databases use different schemas, operating systems, Distributed Database Management Systems (DDBMS), and data models, making them challenging to manage.

In a Heterogeneous distributed database, each site may operate independently and may not be aware of other sites. This lack of awareness limits cooperation in processing user requests. Therefore, translations are necessary to establish communication between sites, enabling them to exchange and understand data effectively.

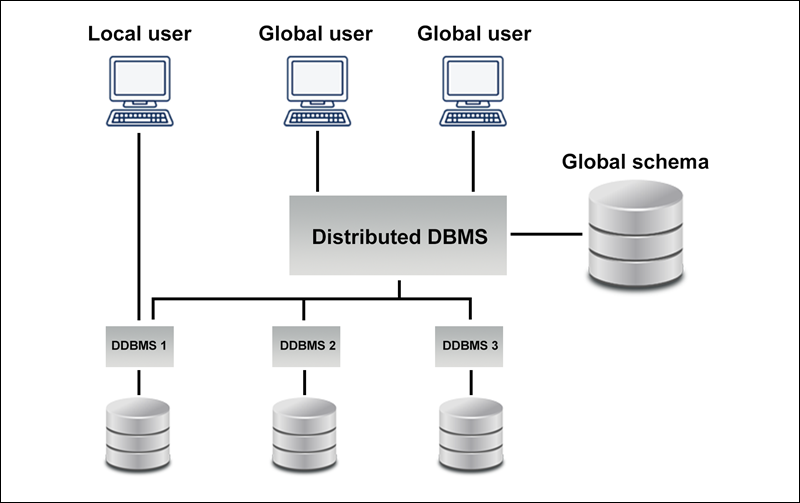
**Salient Features:**

* Different sites use dissimilar schemas and software.
* The system may be composed of a variety of DBMSs like relational, network, hierarchical or object-oriented.
* Query processing is complex due to dissimilar schemas.
* Transaction processing is complex due to dissimilar software.
* A site may not be aware of other sites and so there is limited co-operation in processing user requests.

**Types of Heterogeneous Distributed Databases**

**Federated −** The heterogeneous database systems are independent and integrated 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.



**Advantages of Distributed Database In Dbms**

1. **Better Reliability**: Distributed databases offer better reliability than centralized databases. When database failure occurs in a centralized database, the system comes to a complete stop. But in the case of distributed databases, *the system functions even when a failure occurs*, only performance-related issues occur which are negotiable.
2. **Modular Development**: This implies that the system can be expanded by *adding new computers and local data to the new site* and connecting them to the distributed system without interruption.
3. **Lower Communication Cost**: Locally storing data reduces communication costs for data manipulation in distributed databases. In centralized databases, local storage is not possible.
4. **Better Response Time**: As the data is distributed efficiently in distributed databases, this provides a better response time when user queries are met locally. In the case of centralized databases, *all the queries must pass through the central machine* which increases response time.

**Disadvantages of Distributed Database In DBMS**

1. **Costly Software**: Maintaining a distributed database is costly because we need to ensure *data transparency* and *coordination across multiple sites* which requires costly software.
2. **Large Overhead**: Many operations on multiple sites require complex and numerous calculations, *causing a lot of processing overhead*.
3. **Improper Data Distribution**: If data is not properly distributed across different sites, then responsiveness to user requests is affected. This in turn increases the response time.