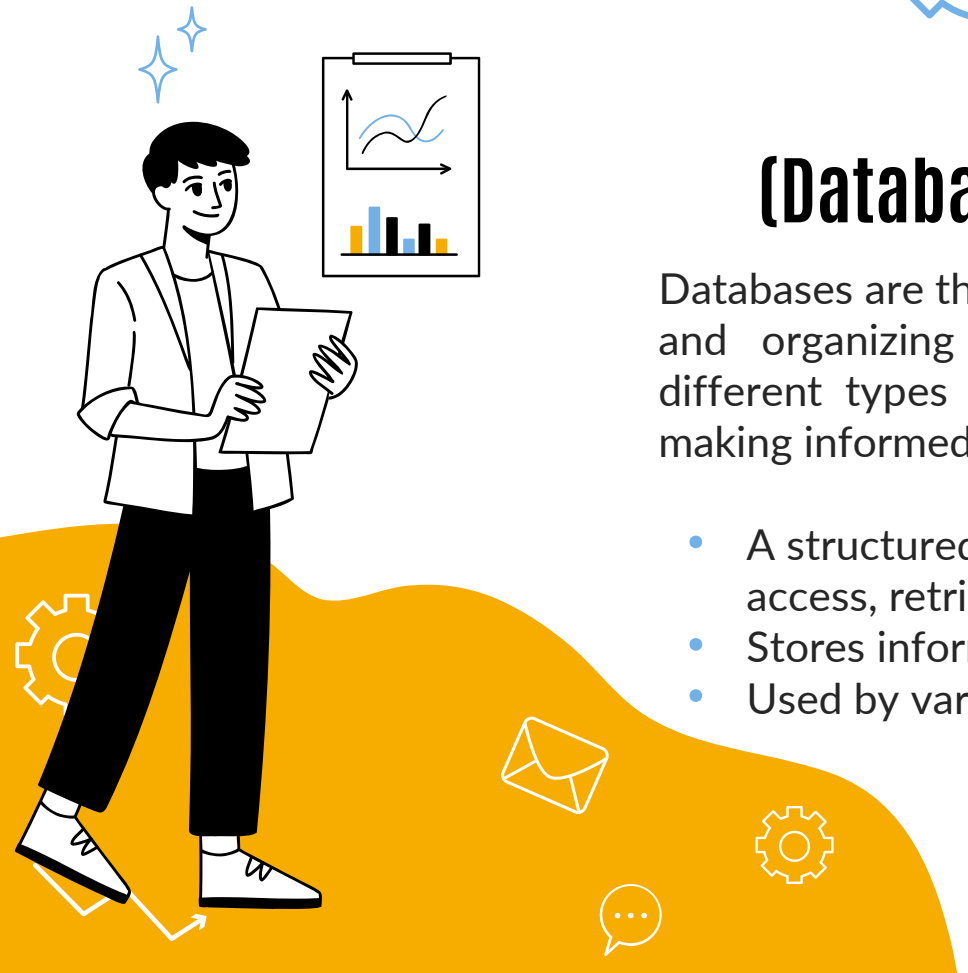
Centralized vs. Distributed Databases

Name : Surajit Jana
Roll : 34200121052
Stream : CSE
Year : 3rd
Semester : 6th
Subject : Distributed System
Subject Code : PEC-IT601B

INTRODUCTION

In the realm of data management, choices between centralized and distributed databases shape the core of IT strategies. Centralization emphasizes simplicity, while distribution thrives on flexibility.
In this presentation we explore the merits and trade-offs of these approaches, unveiling the dynamics that define our digital landscape. Let the comparison commence!



DBMS
(Database Management System)

Databases are the backbone of modern applications, storing and organizing critical information. Understanding the different types of database architectures is essential for making informed decisions about your data infrastructure.

- A structured collection of data organized for efficient access, retrieval, and management
- Stores information in tables with rows and columns
- Used by various applications to manage data.

Centralized Databases

A centralized database (CDB) is a database that is stored, located, and maintained in a single location. This location is usually a central computer or database system, such as a mainframe computer, desktop or server CPU. Centralized databases can be accessed, maintained, and modified from a single location.

- It stores all data in one location, typically on a single server.
- Access to the data is managed and controlled centrally, ensuring consistency.
- It minimizes data duplication and maintains data integrity.

Advantage

- Since all data is stored at a single location only thus it is easier to access and coordinate data.
- The centralized database has very minimal data redundancy since all data is stored in a single place.
- It is cheaper in comparison to all other databases available.

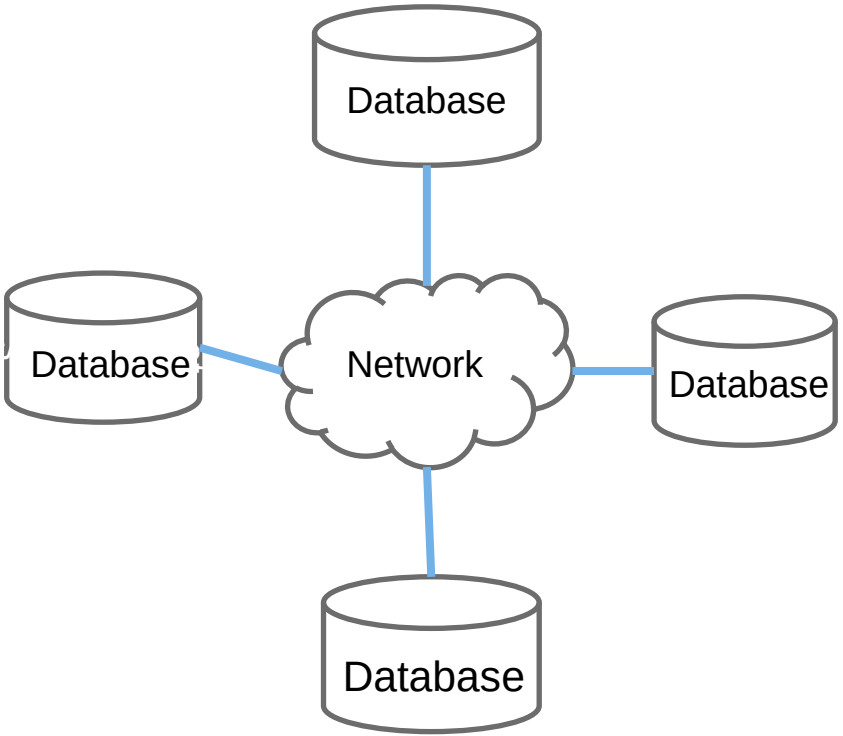
Disadvantage

- Data Bottleneck : High traffic may lead to slowdowns and bottlenecks in data access.
- Single point of failure : If any kind of system failure occurs in the centralized system then the entire data will be destroyed.
- Geographic limitations: latency for remote users

Distributed DBMS

A distributed database is basically a type of database which consists of multiple databases that are connected with each other and are spread across different physical locations. The data that is stored in various physical locations can thus be managed independently of other physical locations. The communication between databases at different physical locations is thus done by a computer network.

- Data is stored across multiple nodes, creating a distributed architecture.
- Multiple users and applications can access the distributed data concurrently.
- Distributed databases offer improved fault tolerance and resilience.



Advantage

- Scalability and Performance : Distribution allows for horizontal scalability and improved performance.
- Network Dependency : Reliability and performance depend on network stability and speed.
- Complex Management : Managing a distributed database architecture can be complex and resource-intensive.
- This database can be easily expanded as data is already spread across different physical locations.

Disadvantage

- This database is very costly and is difficult to maintain because of its complexity.
- it is difficult to provide a uniform view to users since it is spread across different physical locations.
- Complex management and administration
- Network Dependency : Reliability and performance depend on network stability and speed.
- Complex Management : Managing a distributed database architecture can be complex and resource-intensive.

Centralized Databases

Distributed Databases

Architecture overview

It is a database that is stored, located as well as maintained at a single location only.

It is a database that consists of multiple databases which are connected with each other and are spread across different physical locations.

The management, modification, and backup of this database are easier as the entire data is present at the same location.

The management, modification, and backup of this database are very difficult as it is spread across different physical locations.

Performance

The data access time in the case of multiple users is more in a centralized database.

The data access time in the case of multiple users is less in a distributed database.

	Centralized Databases	Distributed Databases
Scalability	Scaling is constrained by capacity of a single server , which can result in hardware limitations.	Can easily distribute the workload across multiple servers , allowing for more efficient utilization.
Data Consistency	This database has more data consistency in comparison	This database may have some data replications thus data consistency is less.
Cost	A centralized database is less costly.	This database is very expensive.
Fault Tolerance	The users cannot access the database in case of database failure occurs.	In a distributed database, if one database fails users have access to other databases.

Application of Centralized Databases

Small-to-Medium Businesses:
Managing customer data, inventory, sales, and financials within a single location or limited offices.

Enterprise Resource Planning (ERP)
Centralized data storage for integrated business processes and management.

Transaction Processing Systems
Efficient management of high-volume transaction data in banking and retail sectors.

Departmental or Team-Specific Applications:
Serving internal needs for HR, project management, or research data within a single organization.

Personal Databases:
Organizing tasks, contacts, finances, or collections for individual users.

Application of Distributed Databases

Cloud-Based Applications
Distributed data storage for scalable and resilient cloud-based services.

Content Delivery Networks (CDN)
Efficient distribution of multimedia content across global networks.

Real-time processing:
Powering autonomous vehicles, edge computing, & real-time collaboration.

Data resilience:
Protecting critical data like healthcare records & emergency response systems.

CONCLUSION

- Centralized and distributed databases each offer unique advantages and disadvantages.
- The best choice depends on your specific needs and priorities.
- Carefully evaluate your requirements and choose the architecture that best suits your current and future needs.

BIBLIOGRAPHY

- Principles of Distributed Database Systems, M.T. Ozsu and PValduriez, Prentice-Hall, 1991.
- Distributed Database Systems, D. Bell and J. Grimson, AddisonWesley, 1992.
- Database system concepts , Henry F. Korth

Thank YOU