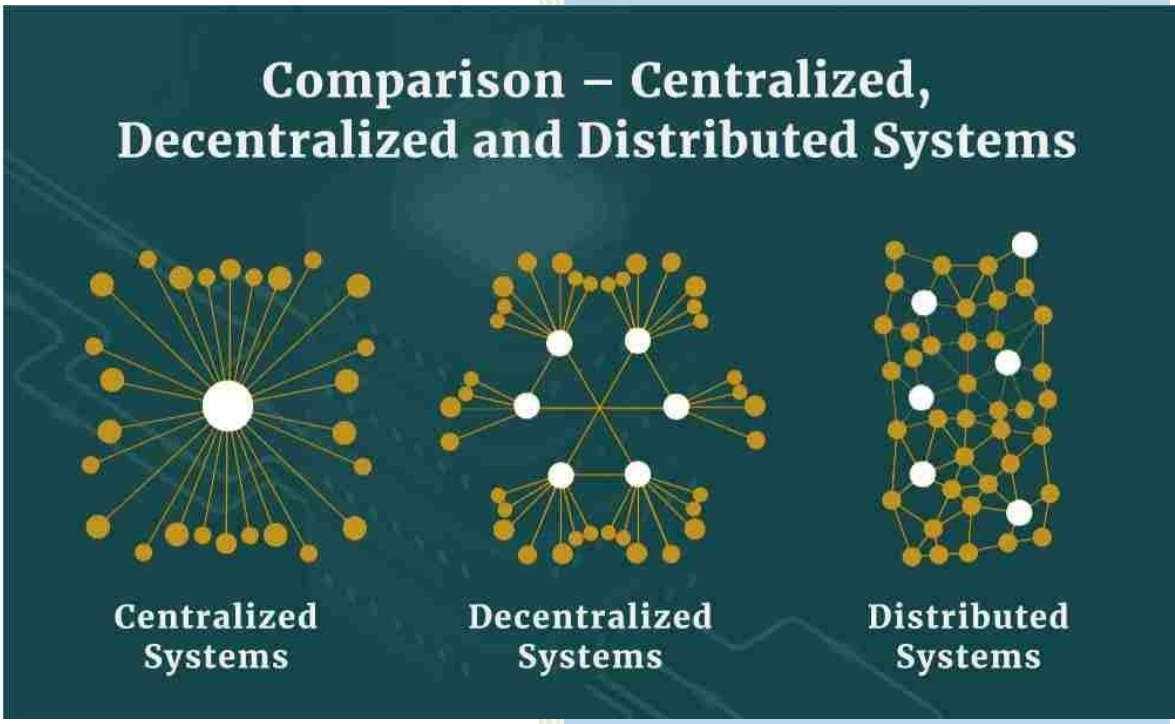
2022

Database 371



Ofentse Mapheto
Belgium campus iTversity
5/28/2022

Table of Contents

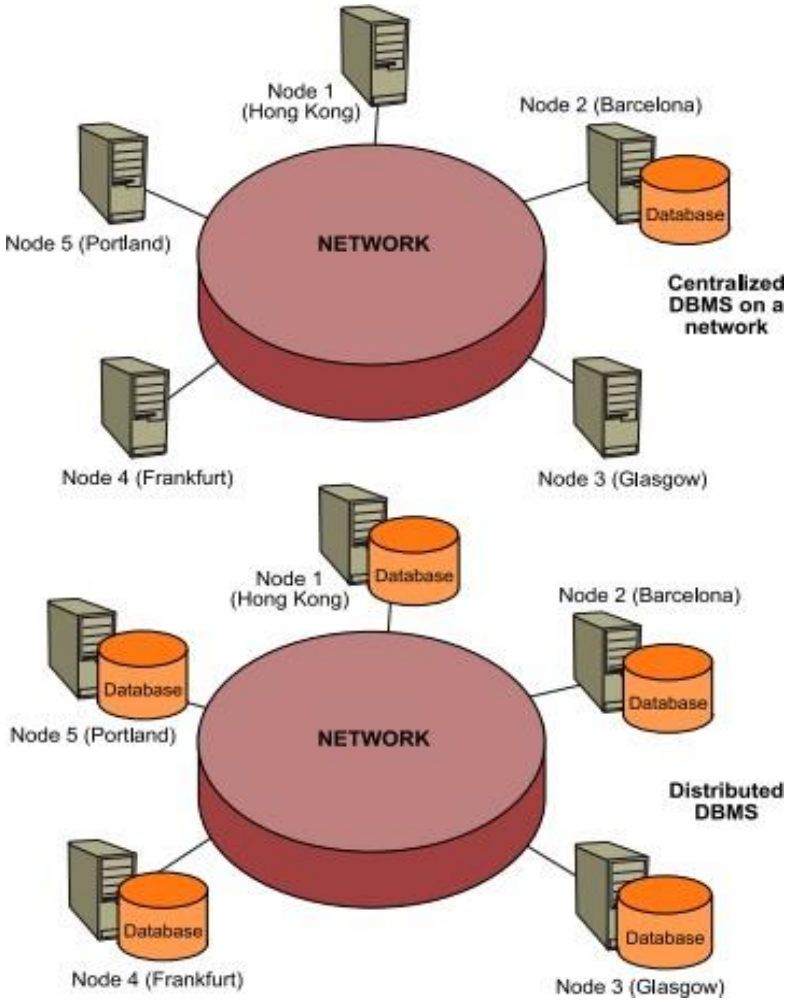
Centralized database vs Distributed database:

A centralized database is a type of database that contains a single database located at one location in the network. A distributed database is a type of database that contains two or more database files located at different locations in the network.

Centralized database is a database stored on a single location. It can be accessed, maintained and modified only from that specific location i.e. all the transactions are conducted on a central location. This location is thus mainly any database system or a centralized computer system. The centralized location is accessed via an internet connection (LAN, WAN, etc).

A distributed database is the term used to describe a set of databases stored on multiple computers (Farnham, 2021). A distributed database contains two or more database files located at different locations in the network. Distributed database does not mean that the user is presented with multiple databases, they are still presented with a single database as the centralized database would have. The data that is stored on 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.

The diagrams below depict both the centralized and distributed databases respectively:



(image source: http://openaccess.uoc.edu/webapps/o2/bitstream/10609/50441/4/Architecture%20Database_Unit3_Distributed%20Databases.pdf)

Price to pay (disadvantages) as per database choice:

Centralized database:

Centralized database has a highly congested traffic. Any level of failure disrupts the whole system. An attack on that central location puts the whole database or organization at a high level of risk. Centralized database requires constant maintenance. Centralized database has a higher usage, this will minimize the productivity. When the server goes down all system will be destroyed. The client is not able to access the database.

Distributed database:

Distributed database is very costly and it is difficult to maintain because of its complexity. it is difficult to provide a uniform view to user since it is spread across different physical locations. It is use because it is very large and complex. When one node becomes fail it is very difficult to detect because there are many nodes and responsible for their own behavior. The server accepts request from nodes in the system but it is actually difficult to find out which node indeed served the request (t4ttutorialsv, 2021).

Candidate applications:

Distributed database management system:

- Cluster computing is a technology that connects many computers to, which can achieve a common purpose. The computer group behaves like a computer.
- **Grid Computing:** In this type of computing, all resources are pooled together for sharing, thus transforming the system into a powerful supercomputer.

Centralized database:

- For personal use.
- Data analysis-When all the data is in one place and ready for analysis, data analysis is very simple.
- Application development: setting up a central server and sending client inquiries is very simple. Today's technology has pre-installed test servers, which can be started with about instructions. For example, the Express server and the Django server are two examples.

DATABASE MANAGEMENT SYSTEMS:

A database is a collection of data that allows users to easily create, read, update or delete records. The relational database is homogeneous and uses a traditional table format. There are various database concepts, namely MySQL, SQL Server, Oracle and others.

Distributed:

The collection database is logically interconnected. They are usually used to represent a single logical database. The data is actually stored in various locations. The DBMS can manage data in a site independently of other sites. The network connects the processor at each location (tutrialpoint, 2021).

They do not have any type of multi-processor configuration. Distributed databases are different from the loosely linked file system. Although the transaction processing is included in the distributed database, it is not the same as the transaction processing system (assignmenthelp4me, 2021).

Centralized:

A centralized database is a database stored in one place like a central computer. It can only be maintained and edited from there, and can usually be accessed via a LAN or WAN connection. Organizations such as universities, companies, and banks use a unified database.

Access to data:

Centralized Database:

The centralized database is stored in a single location, such as a central computer. It can only be maintained and modified from this location, and can usually be accessed via an Internet connection (such as a LAN or WAN). The centralized database is used by universities, companies, banks and other organizations. All information in a centralized database can be easily accessed from the same location and at the same time (tutorialpoint, 2021).

All database connectivity is handled by Intelligence Server, which provides centralized control of database access. All connections to the system database are established through Intelligence Server. This means that only Intelligence Server computers need to establish a connection with the database. It also eliminates the need to rely on identically configured connections on the client and server computers. This makes it easy to set up, deploy and manage large systems (doc-archive, 2021).

Connecting to or disconnecting from the database will incur a small amount of overhead, which can cause a small but significant drop in performance on a highly concurrent system. Through connection cache, Intelligence Server can reuse database connections. This minimizes the overhead associated with repeated connection and disconnection of the database.

Intelligence Server can map Micro-Strategy users and user groups to data store login IDs. This allows multiple users to use a single database login to access the database. You can set the access priority by user, project, estimated labor cost, or any combination of these. The ability to run in multithreaded mode means that if one process fails, such as a lost or suspended database access thread, other processes will not be affected (Pedamkar, 2021).

Distributed database system:

Distributed database systems (such as centralized database systems) seek to provide data integration. However, distributed systems deal with this problem from a distributed perspective. In other words, data is no longer stored in a single node (Òscar Romero, 2021).

Distributed systems facilitate the exchange of data and the continuous autonomy of the different elements involved. Distributed processing is considered to be the best way to deal with large-scale data management problems, and is considered a direct application of the divide-and-conquer method.

There are two important differences between accessing data stored on a single node in a distributed system and accessing data in a centralized system. The advantage of distributed systems is the potential for parallel data processing and transmission. An effective data access strategy must take these differences into account (Alan R. Hevner, 1978).

The distributed database allows local users to manage and access the data in the local database, while providing some form of global data management, thus providing global users with a global view of the data. This global vision allows us to merge data from different sources that may not have been integrated before, thus providing the potential to discover new insights (McClean, 2003). The

composite local database can be homogeneous or it can be part of a design to distribute data storage and processing for greater efficiency.

Implementation cost:

The total implementation cost of any type of database is determined by many factors, including the cost of hardware, software, development, support, and maintenance. Large services that require multiple users to share access rights to the database and have the ability to view and modify data at the same time will require the database to be installed on a server where other computers can access through the network.

Transactions in DBMS can be divided into conflicting transactions and non-conflicting transactions, but also based on local transactions and non-local transactions or global transactions (depending on whether they are executed in a single site or local transactions) divide. Transactions or multiple websites (global transactions), the cost of transaction processing in distributed DBMS can be modeled as the cost of local and non-local (or global) transaction execution. These costs are weighted by terms reflecting the number of local and global transactions in the system (Badal, 1982).

The implementation cost of centralized distributed database management system is fairly lower than that of distributed simply because in centralized database, everything is hosted in-house and may require a single server and can be accessed wirelessly i.e. no need for servers.

Promise to the future:

Centralized:

The promise to the future for any type of database needs to be looked into based on the advantages brought forward by the type of the database system.

In the face of the 4IR, centralizing data may seem like the best option and most people are adapting to that. Having the data centralized poses a huge risk to the organization because of possible hackers.

Centralized data provides better data preservation than other types of databases due to often-included fault-tolerant setup. Data integrity is maximized and data redundancy is minimized, as the single storing place of all the data also implies that a given set of data only has one primary record. This aids in the maintaining of data as accurate and as consistent as possible and enhances data reliability. Easier for using by the end-user due to the simplicity of having a single database design. Generally bigger data security, as the single data storage location implies only a one possible place from which the database can be attacked and sets of data can be stolen or tampered with.

Now most people are managing corporate data and more data will be added. The enterprise cloud is still relatively immature - most data is still managed and stored internally, and most data managers have relatively little experience managing the cloud (D.J.ReidM.E.Orlowska, 1996).

Distributed DBMS:

Distributed systems are similar to decentralized systems in that nodes are independent of each other and are responsible for their respective behaviors, and there is no central node to receive and respond to requests. However, unlike decentralized systems, the results of distributed systems are based on the aggregate results obtained from each node in the system.

Accuracy of result is high, because the average of results gathered from all the individual nodes are collectively described and output is produced. The overall throughput of the system is considerably low because there is widespread use of the nodes which leads to low latency. With the help of distributed systems, cluster computing can also be achieved. Moreover, it is possible as multiple devices collectively work together to attain a common objective, thus creating a cluster of computers (Quest, 2021).

Relationship among various issues:

Security:

A database management system provides security to the stored database and removes the redundant data. Thus it reduces the size of data and eventually increases its accuracy.

ERD tables:

A DBMS allows the administrator or developer to form tables based on the entities and relationships among them. A logical structure can be generated which allows a user to easily understand the architecture of a database.

Multiple views:

A database management system allows a database administrator to have multiple views of data stored, thus allowing users to concentrate on the desired data as per requirements.

ACID properties:

The database management system is controlled by ACID attributes, which are described as atomicity, consistency, isolation, and durability. These are mainly used in the case of transactions that cause changes to the database. Therefore, they maintain data quality based on attributes to maintain database efficiency.

Consistency:

A database management system (DBMS) creates a stronger sense of continuity, which greatly increases database efficiency. That's because it can identify and mitigate all inconsistencies in data transmissions.

Real-world entity:

Almost all the modern databases utilize real-world entities in order to offer a better reference for users. Thus, the purpose of an attribute can easily be identified without any major issue.

Multiple user access:

A database management system (DBMS) has the capability to allow multiple users to access data at the same time and make adjustments in parallel. However, if more than one user attempts to exploit the same data, access is automatically denied.

Query language:

The query language is used to perform data manipulations on a database management system, which makes it much more effective since all operations can be done with great precision. It also allows for many different filtering options when retrieving data.

References

Alan R. Hevner, S. B. (1978). Optimization of Data Access in Distributed Systems. *Purdue e-pubs*, 4 - 5.

assignmenthelp4me. (2021, May 20). *https://assignmenthelp4me.com/*. Retrieved from https://assignmenthelp4me.com/article-what-is-the-difference-between-centralized-and-distributed-dbms-337.html

Badal, D. Z. (1982). DISTRIBUTED VS. CENTRALIZED DATABASE SYSTEMS - TRANSACTION EXECUTION COST AND PERFORMANCE ANALYSIS. *Monterey* , 6 -7.

D.J.ReidM.E.Orlowska. (1996, February 19). *https://www.sciencedirect.com/*. Retrieved from https://www.sciencedirect.com/science/article/pii/0895717795002308

doc-archive. (2021, May 20). *https://doc-archives.microstrategy.com/*. Retrieved from https://doc-archives.microstrategy.com/producthelp/10.11/SystemAdmin/WebHelp/Lang_1033/Content/Benefiting_from_centralized_database_access_contro.htm

Farnham, K. (2021, January 11). *https://insights.diligent.com/*. Retrieved from https://insights.diligent.com/data-management/centralized-vs-distributed-databases

McClean, S. I. (2003, July 17). *https://www.sciencedirect.com/*. Retrieved from https://www.sciencedirect.com/topics/computer-science/distributed-databases

Òscar Romero, M. O. (2021). Distributed Database. *openaccess*, 6 - 7. Retrieved from http://openaccess.uoc.edu/webapps/o2/bitstream/10609/50441/4/Architecture%20Database_Unit3_Distributed%20Databases.pdf

Pedamkar, P. (2021, May 20). *https://www.educba.com/centralized-database/*. Retrieved from https://www.educba.com/centralized-database/

Quest. (2021, May 20). *https://www.quest.com/*. Retrieved from https://www.quest.com/whitepaper/database-professionals-look-to-the-future-2020-trends-in-database-admi8141411/

t4ttutorialsv. (2021, May 20). *https://t4tutorials.com/*. Retrieved from https://t4tutorials.com/comparison-of-centralized-and-distributed-database-with-advantages-disadvantages/

tutorialpoint. (2021, May 20). *https://www.tutorialspoint.com/*. Retrieved from https://www.tutorialspoint.com/Centralized-Database-Management-System#:~:text=A%20centralized%20database%20is%20stored,colleges%2C%20companies%2C%20banks%20etc.

tutrialpoint. (2021, May 20). *https://www.tutorialspoint.com/*. Retrieved from https://www.tutorialspoint.com/distributed_dbms/distributed_dbms_databases.htm