

IMPACT OF DISTRIBUTED DATABASE MANAGEMENT SYSTEM TO INDIVIDUALS, INSTITUTIONS AND ORGANIZATIONS ALL OVER THE WORLD

¹Seji Happy Alfred, ²Okungbowa Aihanuwa Augusta and ³Lambo Kenibanemi Egi
¹Department of Computer & Robotic Education, Federal College of Education
(Technical) Omoku, Rivers State, Nigeria, ²Street 7, Flat 3 Federal Housing Estate,
Ikpoba Hill, Benin City, Edo State, Nigeria, ³Bayelsa State Polytechnic, Eleibiri,
Bayelsa State.

Email: sejihappyalfred@gmail.com, aihanuwa75@gmail.com, lamboegi@gmail.com

ABSTRACT

Distributed Database Management System is a special area of concentration that enables the input of data, processing of data and management of data in a coherent form, therefore, it is indeed an area of concern that is beneficial to both individuals, institutions and organizations all over the world. This is due to its unique position to coordinate records, fields and files in a very modern form which means that the coordination of databases of individuals and unique organizations and that of institutions are well structured and organized in the Database Management System. In view of this, it is pertinent to note that this concept distributed Database Management System is a unique area that conforms to the concept of arranging data and data types in a coherent and organized manner to give a realistic result whenever the need arises. To this end, Distributed Database Management System is so compatible with information generation that needs to drive the organizational structures and institutional based structures to an enviable height that leads to global achievement and growth since it tend to incorporate all the units of the system in a well-organized manner. Distributed Database management system is indeed a reckoned concept all over the world in recent times that is the reason why organizations, institutions and even individuals hinge on it for developmental and organizational coordination in achieving its aims and objectives. In view of the above, data management, data output and data storage of information is, thereby, repositioned to improve organizations and institutions alike.

KEYNOTES: *Concept, Coordination, Distributed, Information, Generation, Development, Driving, Storage, Functionality Database, Management, System, Unique, Organization, Institution, Input, Output, Individuals, Position, Coherent, Conform, Fields, Records.*

INTRODUCTION

The impact of Distributed Database Management System is known all over the world because of its position in the field of computer science and the role it plays in organization of input processed data and the generated information which is organized, managed by Database management system. It is on this note that this work is done to realistically reveal the position of Distributed Database Management System in recent times because individuals, organizations and institutions hinges on it for its daily records and file management whether it is a single or a large entity of Database, it tend to coordinate in the actual manner that produces result. In view of the above, it is known that without Distributed Database Management System, institutions and organizations would have lost the concept of managing their data integrity. On this note, the position of Distributed Database Management System become so important because of the role it plays in harnessing the units of organizations and institutions to an enviable height in generating output and good result.

Query Processing

This is a technical issue in Distributed Database Management System which has to do in providing, designing algorithms that analyze queries and convert them into a series of data manipulation and operations. Beside the methodological issues important aspect of query processing is query optimization which is specifying on how to decide on the strategy for executing each query over the network in the most cost effective way, however, cost is defined. The factors to be considered are the distribution of data, communication cost and lack of sufficient locally available information, the objective is to optimize where the inherent parallelism of the distributed system is used to improve the performance of executing the query, subject to the above mentioned constraint which is the approaches are usually heuristic in nature. M. Tamer (2003).

User Access to shared Databases

Sharing of databases are formulated as transactions which are units of execution that satisfies four properties namely: atomicity, consistency, isolation, durability, jointly known as the ACID properties. Atomicity means that the transaction is an atomic unit and either the effect of all of its actions are reflected in the database, or none of them are consistency generally referred to the correctness of the individual transaction i.e., that the transaction does not violate any of the integrity constraints that have been denied over the database. Isolation addresses the concurrent execution of transaction that specifies that actions of concurrent transactions do not impact each other. Finally, durability concerns the persistent of data changes in the face of failures. The ACID properties are enforced by means of concurrency control algorithms and reliability protocols. Concurrency control involves the synchronization of accesses to the distributed database such that the utility of the database is maintained. M. Tamer (2003).

The concurrency control problem in a distributed context is somewhat different that in a centralized framework. One not only has to worry about the integrity of a single database but also about the consistency of the multiple copies of the database. The condition that requires all the values of multiple copies of every data item to converge to the same value is all mutual consistency.

Reliability Protocol

This deals with the termination of transactions in particular, their behavior in the face of failures in addition to the typical failures (that is transaction failures and system failures) Distributed Database Management System have to account to communication network failures as well. The implication of communication failures is that, when a failure occurs on various sites becomes either inoperable or inaccessible, the databases are the operational sites remains consistent and up to date. This complicates the picture, as in actions of these sites have to be eventually reconcile with those of failed ones. Therefore, recording protocol coordinate determination of transactions so that they terminate uniformly, that is, they either abort or they commit at all the sites where they execute furthermore where the computer system or network covers from the failure, the distributed database management system should be able to recover and bring the databases at the failed site up to date. This may be especially difficult in the case of network partitioning where the sites are divided into two or more groups with no communication among them. Distributed database management system are typically replicated, that is, a number of the data items resides at more than one site, replication improves performance since data access can be localized in availability since the failure of a site does not make a data item inaccessible. However, management of replicated data requires that the values of multiple copies of a data item are the same. This is called the one copy equivalent property. Therefore, distributed database management system that allows replicated data implements, replication protocol to enforce one copy equivalent. M. Tamer (2003).

How does Distributed Database Management System Works

A distributed Database management System manages the distributed database in a manner so that it appears as one single database to users. A distributed database is a collection of multiple interconnected databases which are spread physically across various locations that communicate via a computer network channel.

Based on the above, distributed database management system is easy to expand as it already spreads across multiple systems and it is not too complicated to harm the system. The distributed database management system can have the data arranged according to levels of transparency, that is, data with different transparency levels can be stored at different locations.

What are the benefits of Distributed Database Management System?

The following are the advantages of distributed database over the centralized database system. Modular development, more reliable, dual communication cost, improved performance, improved share ability and local autonomy.

Modular Development

If the system needs to be expanded to new locations or new units, in centralized database systems, the action requires substantial efforts and disruption in the existing functioning.

However, in distributed databases, the work simply requires adding new computers and local data to the new site and finally connecting them to the distributed system, with no interruption in current functions.

More Reliable

In case of database failures, the total system of centralized databases comes to a halt. However, in distributed systems, when a component fails, the functioning of the system continues may be at a reduced performance. Hence, distributed database management system is more reliable.

Better Response

If data is distributed in an efficient manner, then user requests can be met from local data itself, thus providing faster response. On the other hand, in centralized systems, all queries have to pass through the central computer for processing, which increases the response time.

Lower Communication Cost

In distributed database system, if data is located locally where it is mostly used, then the communication costs for data manipulation can be minimized. This is not feasible in centralized systems.

Improved Performance

As the data is located near the site of greatest demand and given the inherent parallelism of distributed database management systems, speed of database access may be better than that achievable from a remote centralized database.

Furthermore, since each site handles only a part of the entire database, there may not be the same contention for CPU and I/O services as characterized by a centralized database management system.

Improved share ability and Local Autonomy

The geographical distribution of an organization can be reflected in the distribution of the data users at one site can access data stored at other sites.

Data can be placed at the site close to the users who normally use that data. In this way, users have local control of the data, and they can consequently establish and enforce local policies regarding the use of this data.

A global database administrator (DBA) is responsible for the entire system. Generally, part of this responsibility is assigned to the local level, so that the local DBA can manage the local database management. Bench Partner (2021).

Despite the benefits of distributed database management system, the ills of distributed database management system is as follows:

Complexity, cost, security, integrity control more difficult, lack of standards, lack of experience, database design more complex.

Complexity

A distributed database management system that hides the distributed nature from the user and provides an acceptable level of performance, reliability, availability is inherently more complex than a centralized database management system.

The fact that data can be replicated also adds an extra level of complexity to the distributed database management system. If the software does not handle data replication adequately, there will be degradation in availability, reliability and performance compared with the centralized system, and the advantages we cites above will become disadvantages.

Cost

Increased complexity means that we can expect the procurement and maintenance costs for a distributed database management system to be higher than those for a centralized database management system. Furthermore, a distributed database management system requires additional hardware to establish a network between sites.

There are ongoing communication costs incurred with the use of this network. There are also additional labor costs to manage and maintain the local database management systems and the underlying network.

Security

In a centralized system, access to the data can be easily controlled. However, in a distributed management system not only does access to replicated data have to be controlled in multiple locations but also the network itself has to be made secure. In the past, networks were regarded as an insecure communication medium. Although this is still partially true, significant developments have been made to make networks more secure.

Integrity control more difficult

Database integrity refers to the validity and consistency of stored data. Integrity is usually expressed in terms of constraints, which are consistency rules that the database is not permitted to violate.

Enforcing integrity constraints generally requires access to a large amount of data that defines the constraints. In a distributed database management system, the communication and processing costs that are required to enforce integrity constraints are high as compared to centralized system.

Corporation. Ingres and IBM BS12 followed, other examples of an RDBMS include DB2, SAP Sybase ASE.

SQL Server

SQL Server developed by Microsoft, SQL Server is a relational database management system built on top of structured query language (SQL).

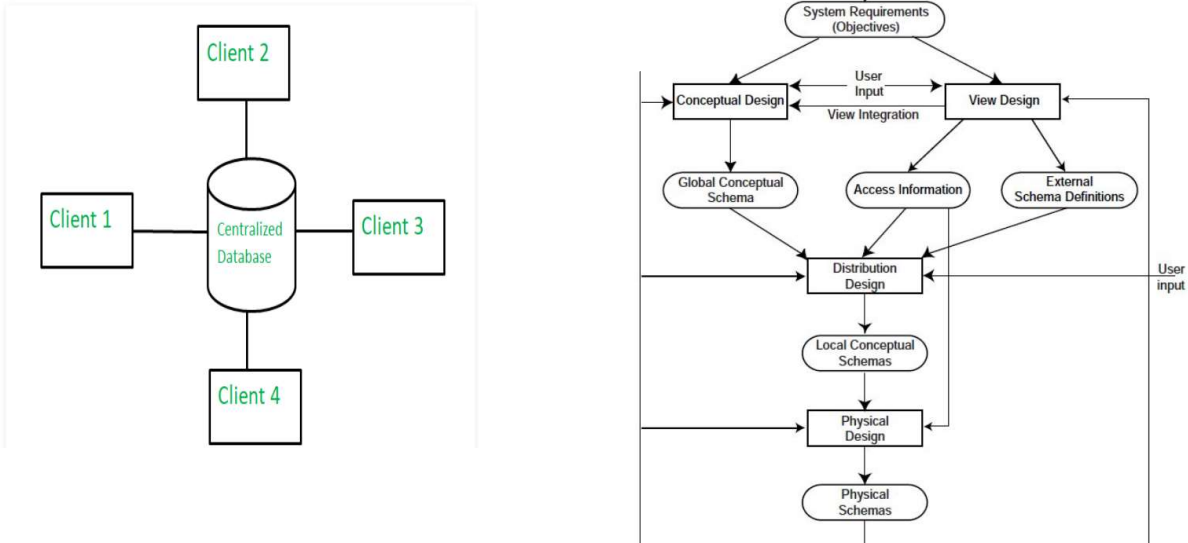


Fig. 3. & Fig. 4. Sample diagram of a Homogenous Distributed Database Management System Oracle Cooperation (1999)

Distributed DBMS – Concepts

For proper functioning of any organization, there’s a need for a well-maintained databases used to be centralized in nature. However, with the increase in globalization, organizations tend to be diversified across the globe. They may choose to distribute data over local servers instead of a central database. Thus, arrived the concept of Distributed Database.

This chapter gives an overview of databases and Database Management Systems (DBMS). A database is an ordered collection of related data. A DBMS is a software package to work upon a database. A detailed study of DBMS is available in our tutorial named “Learn DBMS”. In this chapter, we revise the main concepts so that the study of DDBMS can be done with ease. The three topics covered are database schemas, types of databases and operations on databases.

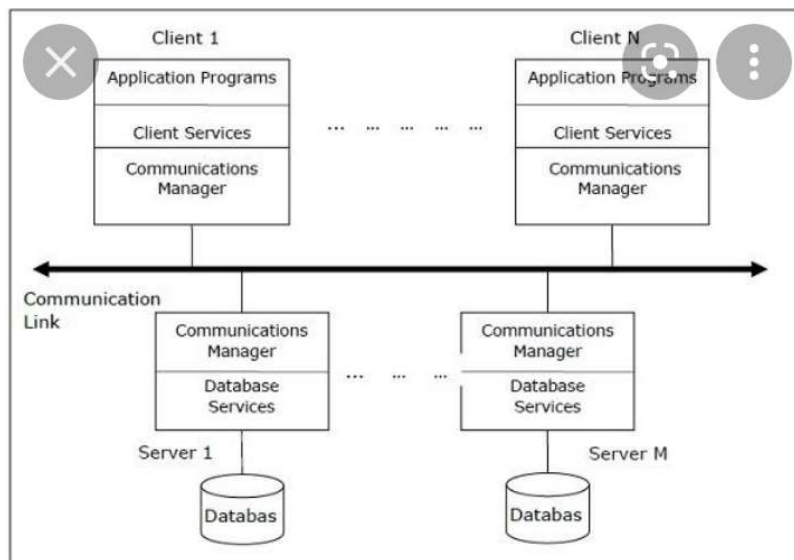


Fig. 2. Sample diagram of a Homogenous Distributed Database Management System Oracle Cooperation (1999)

Database and Database Management System

A database is an ordered collection of related data that is built for a specific purpose. A database may be organized as a collection of multiple tables, where a table represents a real world element or entity. Each table has several different fields that represent the characteristic features of the entity.

For example, a company database may include tables for projects, employees, departments, products and financial records. The fields in the Employee table may be Name, Company Id, Date of Joining, and so forth.

A database management system is a collection of programs that enables creation and maintenance of a database. DBMS is available as a software package that facilitates definition, construction, manipulation and sharing of data in a database. Definition of a database includes description of the structure of a database. Construction of a database involves actual storing of the data in any storage medium. Manipulation refers to the retrieving information from the database, updating the database and generating reports. Sharing of data facilitates data to be accessed by different users or programs.

Examples of DBMS Application Areas

Automatic Teller Machines
Train Reservation System
Employee Management System
Student Information System

Examples of DBMS Packages

MySQL
Oracle
SQL Server
dBASE
FoxPro
PostgreSQL, etc.

Database Schemas

A database schema is a description of the database which is specified during database design and subject to infrequent alterations. It defines the organization of the data, the relationships among them, and the constraints associated with them.

Databases are often represented through the three-schema architecture or ANSISPARC architecture. The goal of this architecture is to separate the user application from the physical database.

Internal Level having Internal Schema – It describes the physical structure, details of internal storage and access paths for the database.

Conceptual Level having Conceptual Schema – It describes the structure of the whole database while hiding the details of physical storage of data. This illustrates the entities, attributes with their data types and constraints, user operations and relationships.

External or View Level having External Schemas or Views – It describes the portion of a database relevant to a particular user or a group of users while hiding the rest of database.

Types of DBMS

There are four types of DBMS.

Hierarchical DBMS

In hierarchical DBMS, the relationships among data in the database are established so that one data element exists as a subordinate of another. The data elements have parent-child relationships and are modelled using the "tree" data structure. These are very fast and simple.

Hierarchical DBMS

Network DBMS

Network DBMS is one where the relationships among data in the database are of type many-to-many in the form of a network. The structure is generally complicated due to the existence of numerous many-to-many relationships. Network DBMS is modelled using "graph" data structure.

Network DBMS

Relational DBMS

In relational databases, the database is represented in the form of relations. Each relation models an entity and is represented as a table of values. In the relation or table, a row is called a tuple and denotes a single record. A column is called a field or an attribute and denotes a characteristic property of the entity. RDBMS is the most popular database management system.

For example – A Student Relation.

Relational DBMS

Object Oriented DBMS

Object-oriented DBMS is derived from the model of the object-oriented programming paradigm. They are helpful in representing both consistent data as stored in databases, as well as transient data, as found in executing programs. They use small, reusable elements called objects. Each object contains a data part and a set of operations which works upon the data. The object and its attributes are accessed through pointers instead of being stored in relational table models.

For example – A simplified Bank Account object-oriented database –

Object-oriented DBMS

Distributed DBMS

A distributed database is a set of interconnected databases that is distributed over the computer network or internet. A Distributed Database Management System (DDBMS) manages the distributed database and provides mechanisms so as to make the databases transparent to the users. In these systems, data is intentionally distributed among multiple nodes so that all computing resources of the organization can be optimally used.

Operations on DBMS

The four basic operations on a database are Create, Retrieve, Update and Delete.

CREATE database structure and populate it with data – Creation of a database relation involves specifying the data structures, data types and the constraints of the data to be stored.

Example – SQL command script

Example – SQL command to create a student table

```
CREATE TABLE STUDENT (ROLL INTEGER PRIMARY KEY); NAME VARCHAR 2(25), YEAR INTEGER, STREAM VARCHAR 2(10)
```

Once the data format is defined, the actual data is stored in accordance with the format in some storage medium.

Example of SQL command to insert a single tuple into the student table

```
INSERT INTO STUDENT (ROLL, NAME, YEAR, STREAM) VALUES (1, 'ANKIT JHA', 1, 'COMPUTER SCIENCE');
```

RETRIEVE information from the database – Retrieving information generally involves selecting a subset of a table or displaying data from the table after some computations have been done. It is done by querying upon the table.

Example – To retrieve the names of all students of the Computer Science stream, the following SQL query needs to be executed –

```
SELECT NAME FROM STUDENT WHERE STREAM = 'COMPUTER SCIENCE';
```

UPDATE information stored and modify database structure – Updating a table involves changing old values in the existing table's rows with new values.

Example – SQL command to change stream from Electronics to Electronics and Communications –

```
UPDATE STUDENT SET STREAM = 'ELECTRONICS AND COMMUNICATIONS' WHERE STREAM = 'ELECTRONICS';
```

Modifying database means to change the structure of the table. However, modification of the table is subject to a number of restrictions.

Example – To add a new field or column, say address to the Student table, we use the following SQL command.

```
ALTER TABLE STUDENT ADD (ADDRESS VARCHAR 2(50));
```

DELETE information stored or delete a table as a whole – Deletion of specific information involves removal of selected rows from the table that satisfies certain conditions.

Example – To delete all students who are in 4th year currently when they are passing out, we use the SQL command.

```
DELETE FROM STUDENT WHERE YEAR = 4;
```

Alternatively, the whole table may be removed from the database.

Example – To remove the student table completely, the SQL command used.

CONCLUSION

If these specifications are followed, it is obvious that the desires, aim of institutions and organizations in using distributed database management system will be fully and adequately achieved because it is definitely going to conform with standard because every result required to be processed by the distributed database management system would meet all specifications and achievements of institutions and organizations alike.

RECOMMENDATION

In view of the above detailed analysis on the impact of distributed database management system on institutions, organizations and other related areas of application I hereby wish to recommend as follows:

1. That distributed database management system should be regularly updated to conform to the needed standard of performance all over the world.
2. That data types and data inputs should be made in such a way that they align with system structure to produce good results at all time in organizations, institutions, etc.
3. That the monitoring and the regulation of database concept based on ISO standard must be maintained to enable distributed database management system function effectively.

REFERENCES

- Alexandru Boicea, Florin Radulescu, Ciprian-Octavian Trucia, Loredana Urse: Improving Query Performance in Distributed Database (2016).
- Daniel Closa, Alex Gardiner, Falk Giemsa, Joerg Machek: Database and Database Management System (2010).
- LING LIU, M. Tamer Ozsu: Encyclopedia of Database Systems (2009).
- M. Tamer Ozsu & Patrick Valduriez: Outline, Introduction, Background: Distributed DBMS Architecture (2001).
- M. Tamer Ozsu: Distributed Database Management systems (2003).
- Matthias Wiesmann, Fernando Pedone, Andre Schiper & Gustavo Alonso: Understanding Replication in Databases and distributed Systems (1999).
- Mopati Bernerdict Kekgathetse, Keletso Letsholo: A survey on database synchronization algorithms for mobile device (2016).
- Neil Matthew & Richard Stones: Beginning Databases with PostgreSQL: From Novice to Professional (2005).
- Paul Beynon-Davies: Distributed Database Systems (1996).
- R. Gudakesa, Made Sukarsa, I. Gusti Made Arya Sasmita: Two-ways database synchronization in homogeneous DBMS (2014).
- Yohannes Kelemu, Seema Patil: Hybridized Fragmentation of Distributed Database (2016).