Security Issues in Distributed Database System

[1] Sangeetha Radhakrishnan, [2] Dr.A.Akila
Department of computer science
Vels University Chennai, India

Abstract— Distributed database is a collection of databases that can be stored at different computer networks and can be access from different locations. For a distributed environment, development of a secured database is a critical issue. Nowadays, distributed database became more and more popular and hence the importance of ensuring security of database in distributed environment became more significant. In this paper, the features of the distributed database system and its security issues are mentioned. Security issue may compromise the access control and the integrity of the system. Some solution for security aspects like access control, reliability, integrity and confidentiality are also mentioned.

Keywords: - Distributed database management system, Security of distributed database, Architecture of distributed database, Retrieval problems of distributed database.

I. INTRODUCTION

Organized collection of data can be referred as a database. Distributed database [1] is a collection of databases distributed and stored on multiple computers in same physical location or over a network of interconnected computers. All computers in the distributed environment has full control over managing the data. Each database [2] in the distributed environment must have a unique database name which identifies a database server in a distributed system. User can access the data at their location without interfering the work of others. A distributed database management system is defined as the software system that permits the management of the distributed database system (DDBS) and makes the distribution transparent to the users. A distributed database system is not a collection of files that can individually stored. To form a DDBS, files should be logically not related but should have a structure and access should be via a common interface.

II. DISTRIBUTED DATABASE

A. Need for distributed database

1) Organizational and economic reasons: Many organizations are not centralized, and a distributed database approach fits more naturally the structure of the organization. The organizational and economic motivations [3] are probably the most important reason for developing distributed databases.

2) Interconnection of existing databases: When several databases exist in an organization and there is a necessity of performing global applications, distributed databases are the natural solution. The distributed database is created from the already existing local databases. This process may require restructuring and the effort which is required is much less than creating a new centralized database.

3) Incremental growth: An organization grows by adding new organizational units such as new branches, new warehouses etc., then distributed database approach supports an incremental growth with a minimum degree of impact on already existing units.

4) Reduced communication overhead: Many applications are local that reduces the communication overhead with respect to a centralized database.so the maximization of

5) Performance considerations: The existence of several autonomous processors result in the increase of performance. The advantage of distributed databases is that the data decomposition reflects application dependent criteria which maximize application locality; and thus, the mutual interference between different processors is minimized. The load is shared between the different processors, and critical bottlenecks, suchas communication network or common services are avoided. This effect is a consequence of the requirement of autonomous processing capability for local applications stated in the definition of distributed database.

6) Reliability and availability: The distributed database approach especially with redundant data can be used in order to obtain higher reliability and availability.

B. Types of Distributed Database System

1)   Homogeneous Distributed database system [4]: data is distributed but all the servers run on same database management system.

2)   Heterogeneous distributed database system: different site run under the control of different DBMS.

C. Advantages of Distributed database

- Data is distributed, so network traffic reduced.
- Local database still works even if the company network is temporarily broken.
- If there is problem in one branch, then it does not affect working of another branch.
- Queries are local, so high performance.

D. Disadvantages of Distributed database

- In Distributed system, it is more complex to make sure that data and indexes are not corrupted.
- Distributed System is not efficient if there is heavy interaction between sites.
- It is more complicated to maintain Distributed database as compare to centralized database.
- Security problem as data are distributed.

E. Architecture of Distributed Database System

In distributed database system, the database is stored across different computers which are connected through a communication network in different locations. Computers may be of different operating systems and each fragment of the database may be managed by a different DBMS.
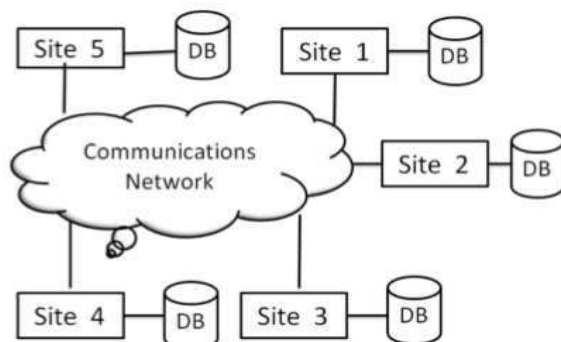

Fig 1: Distributed Database System

The basic design of a distributed database is similar to centralized database. There are some additional factors have to be considered in the distributed database as follows.

1) Data Fragmentation: database is broken up into fragments [5] which will be stored at different sites.
a)   Horizontal Fragmentation: A horizontal fragment of a table is a subset of rows in it. It divides a table horizontally by selecting the relevant rows and these fragments can be assigned to different sides in the b) Vertical fragmentation: In a vertical fragmentation, a table vertically divided by columns. There is necessary to include the primary key of the table in each vertical fragment to rejoin reconstructed.
c)   Mixed fragmentation: The original table can be reconstructed be applying union and natural join operations in the appropriate order.

2)   Data Replication: In replication, a copy of fragment is maintained at several sites. It is the design process of deciding which fragments will be replicated.
3)   Data Allocation: Each fragment has to be allocated to one or more than one sites, where it will be stored. There are three strategies regarding the allocation of data.
a)   Fragmented / Partitioned: The database is partitioned into fragments which are disjoint, with each fragment assigned to one site. This is also called 'non-redundant allocation'.
b)   Complete replication: A complete copy of the database is maintained at each site. Here, storage costs and communication costs for updates are most expensive.
c)   Selective replication: This scheme is a combination of fragmentation and replication.

III. SECURITY COMPONENTS IN DISTRIBUTED DATABASE

Security implies protection of information system from unauthorized access, modification and misuse of information. The purpose of distributed database security is to protect data from malicious attacks. Distributed system has four security components: security authentication, authorization, Encryption, and multi-level access control.

A.   Authentication: - While establishing a connection a user must provide the correct password to prevent unauthorized use of the database. Password are assigned

when user is created. User can change their password at any time. A database can store the password in data dictionary in an encrypted format

B. Authorization: - The purpose of authorization is to supply one secured access point enabling the users to access the authorized resources but can link up to the network once.

C. Encryption: - Technique of encoding data in such a way that only authorized users can understand it. Some standard encryption algorithms such as [6] RSA(Ron Rivest, Adi Shamir and Leonard Adleman), DES(Data Encryption Standard), PGP(Pretty Good Privacy) are useful for the encryption and decryption of data.

D. Multi-level access control: - User is limited from having complete access of data. It has access to relevant information only. Access policies for multi-level system are either open or closed. In an open system, all the data is considered unclassified unless access to a particular data element is expressly prohibited. In a closed system, access to all data is prohibited unless the user has specific access privileged. Distributed database system security [7] has focused on multi-level security.

IV. EMERGING SECURITY IN DISTRIBUTED DATABASE SYSTEM

Some emerging security tools used in distributed database system are data warehouses and data mining system, collaborative computing system, distributed object system and the web. Let us consider data warehousing systems. The major issues are enforcing appropriate access control technique when retrieving the data from warehouse, and ensuring that security is maintained in building a data warehouse from the backend database systems. For example, security policies of the different data sources that from the warehouse have to be integrated to form a policy for the warehouse. One has to maintain security rules during the transformation, the warehouse security policy has to be enforced and the warehouse has to be audited. The retrieval problem also becomes as issue here, For example the warehouse may store average salaries and a user can access that and then deduce the individual salaries in the data sources, which may be sensitive. So, the inference problem could become an issue for the warehouse. Till date, little work has been reported on security as well as the retrieval problem for the warehouse. This area needs much research intension.

Data mining causes serious security problems, For example, consider an user who has the ability to apply data mining tools. The user can assume various queries and infer a sensitive hypothesis which implements the retrieval problem through data mining. There are various ways to handle this problem. Given a database and a particular data mining tool, user can apply the tool and if the sensitive information can be deduced from unclassified information then there is a retrieve problem. One of the issue with this approach is user can apply only one tool but in reality, they may have several tools available. Furthermore, it is impossible to cover all of the ways that the retrieval issue could occur. Another solution to the retrieval issue is to build a retrieval controller that can detect the motive of users and prevent the issue. Such a controller lies between the data-mining tool and the data source or database, managed by a DBMS. Data mining system is being extended to function in a distributed environment. This system is called distributed data mining system and has received very little concentration. Other emerging technologies from distributed database are called collaborative computing system, the web and the distributed object management system. Much of the work on securing distributed object systems, there is a lot of work [8]by the object database can be applied to collaborative computing system's security. With respect to security of distributed management group's security special interest group. Much work has been done on securing the web as well. The main issue is ensuring the security of databases, the operating systems, web servers, the applications, the client and the network and also securely integrated.

V. CONCLUSION

Distributed database systems are getting more and more popular. Many organizations are now deploying distributed database system. This paper introduces the different aspects related to distributed database such as database system concept, architecture of distributed database and also some security issues in distributed database system. This paper also described the most common mechanism of security and the emerging security used in distributed system tools. A multimodal security with the traditional authentication system and biometric authentication will be deployed in the future to enhance the security in distributed database management system.

REFERENCES

[1] Bell, David and Jane Grisom, Distributed Database Systems.

Workinham, England: Addison Wesley, 1992.

[2] Charles P. Pfleeger and Shari Lawrence Pfleeger, Security in Computing, Prentice Hall Professional Technical Reference, Upper Saddle River, New Jersey, 2003.

[3] Stefano Ceri, Giuseppe Pelagatti: Distributed Databases: Principles and

Systems. McGraw-Hill Book Company 1984, ISBN 0-07-010829-3.

[4] Manoj Kumar Sah , Vinod Kumar and Ashish Tiwari,Security and Concurrency Control in Distributed Database System,IJSRM International Journal of scientific research and management,Vol. 2, Issue

12,pp:1839-1845, Dec 2014.

[5] M. Tamer Ozsu, Patrick Valduriez: Distributed database systems.

Second Edition, Prentice Hall of India, New Delhi 2001, ISBN 81-7808-

375-2.

[6] Simon Wiseman, DERA, Database Security: Retrospective and Way

Forward, 2001.

[7] Thuraisingham, Bhavani and William Ford, "Security Constraint Processing In A Multilevel Secure Distributed Database Management System," IEEE transactions on Knowledge and Data Engineering, v7 n2, pp. 274-293, April 1995.

[8] A.A.Akintola,G.A.Aderounmu and A.U.Osakwe," Performing Modeling of an Enhanced Optimistic Locking Architecture for Concurrency Control in a Distributed Database System",ACM vol.37, No.4, November 2005.