

Distributed Databases

Manoj Kumar¹, Manju Kumari²

¹M.D.U. Rohtak

¹9466826078, ¹mkdhull89@gmail.com

²L.P.U. University Punjab

²9813300717, mk.siwach13@gmail.com

Abstract: *Distributed databases is a database in which the database are save on several computers in a network. The data on several computers can simultaneously access and edit using by a network. The users have access to the portion of the database at their location without interfering with the work of others users. Every database server in the distributed database is controlled by DBMS. The Distributed database system synchronizes whole data periodically and, in cases where multiple users must access the same data, ensures that updates and deletes performed on the data at one location will be automatically reflected in the data stored elsewhere. Distributed DBMS can use replicated components to eliminate single point failure. The users can still access part of the distributed database with “proper care” even though some of the data is unreachable. Distributed transactions facilitate maintenance of consistent database state even when failures occur. In a distributed environment, it is much easier to handle expansion. Latest sites can be add to the network without affecting the operations of other sites. This flexibility allows an organization to expand relatively easily. Adding processing and storage power to the network can usually handle the increase in database size.*

1. INTRODUCTION

In today's world of universal dependence on information systems, all sorts of people need access to companies' databases. In addition a company self employees, these include the company customers, potential customers, suppliers, and vendors etc. It is easy for a company to all of its databases concentrated at one mainframe computer site with worldwide access to this site provided by telecommunications networks, including Internet. The management of such a centralized system and its databases can be controlled in a well-contained manner and it poses some problems as well. If single site goes down, because this everyone is blocked from accessing the databases until the site comes backup again. Also the communications costs from the many PCs and terminals to the central site can be expensive. Solution for these problems is alternative design to the centralized database. The idea is that instead of having one, centralized database, we are going to spread the data out among the cities on the distributed network, each of which has its own computer and data storage facilities. All of the distributed data

is still considered to be a single logical database. When we process anywhere on the distributed network queries the database, the network or data being sought is located. Whenever user access the query, and the result is returned to the user. This feature is known as location transparency. This can become rather complex very quickly, and it must be managed by sophisticated software known as a distributed database management system. Collection of multiple, logically interrelated databases distributed over a network is known as distributed database. Distributed database management system is the software that manages the DDB, and provides an access mechanism that makes this distribution transparent to the data and applications.

2. ADVANTAGES

Reflects organizational structure

Many organizations are naturally distributed over several locations. For example, a bank has many offices in different areas. It is natural for databases use in such an application to be distributed over these areas. A bank may keep a database at each branch office containing details such things as the staff that work at that location, the account information of customers etc. The staff at a branch office will make local inquiries of the database. The company headquarters may wish to make global inquiries involving the access of data at all or a number of branches.

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 the data stored at other sites. Data can be placed at the site close to the users who normally use that data. So, the 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 whole system. Generally, part of this responsibility is assigned the local level, so that the local DBA can manage the local DBMS.

Improved availability

In a centralized DBMS, a computer failure terminates the applications of the DBMS. However, a failure at one site of a

DDBMS, or a failure of a communication link making some sites inaccessible, does not make the entire system inoperative. Distributed DBMSs are designed to continue to function despite such failures. If a single node fails, the system may be able to reroute the failed node's requests to another site.

Improved reliability

As the data may be replicated so that it exists at more than one site, the failure of a node or a communication link does not necessarily make the data inaccessible. So, it is reliable when compared to a Centralized database system.

Improved Performance

The data is located near the site of 'greatest demand', and given the inherent parallelism of distributed DBMSs, 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 DBMS.

Economics

It is now generally accepted that it costs much less to create a system of smaller computers with the equivalent power of a single large computer. This makes it more cost effective for corporate divisions and departments to obtain separate computers. It is also much more cost-effective to add workstations to a network than to update a mainframe system. The second potential cost saving occurs where database are geographically remote and the applications require access to distributed data. In such cases, the relative expense of data being transmitted across the network as opposed to the cost of local access, it may be much more economical to partition the application and perform the processing locally at every site.

Modular growth

In a distributed environment, it is much easier to handle expansion. New sites can be added to the network without affecting the operations of other sites. This flexibility allows an organization to expand relatively easily. Adding processing and storage power to the network can usually handle the increase in database size. In a centralized DBMS, growth may entail changes to both hardware (the procurement of a more powerful system) and software (the procurement of a more powerful or more configurable DBMS).

3. DISADVANTAGES OF DDBMS

Complexity

A distributed DBMS that hides the distributed nature from the user and provides an acceptable level of performance, reliability, availability is inherently more complex than a centralized DBMS. In this way the data can be replicated also adds an extra level of complexity to the distributed DBMS. If the software does not handle data replication adequately, there

will be degradation in availability, reliability and performance compared with the centralized system.

Cost

Increased complexity means that we can expect the procurement and maintenance costs for a DDBMS to be higher than those for a centralized DBMS. DBMS requires additional hardware to establish a network between the sites.

Security

In a centralized system, access to the data can be easily controlled. A distributed DBMS not only does access to replicated data have to be controlled in multiple locations but also the network itself has to be made secure. So, this is still partially true, significant developments have been made to networks more secure.

Integrity control more difficult

Database integrity indicates the validity and consistency of stored data. 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 DBMS, the communication and processing costs that are required to enforce integrity constraints are high as compared to centralized system.

Lack of Standards

A distributed DBMSs depend on effective communication, we are only now starting to see the appearance of standard communication and the data access protocols. There are also no tools to help users convert a centralized DBMS into a distributed DBMS.

Lack of experience

General-purpose distributed DBMSs have not been widely accepted. Many of the protocols and problems are well understood. Consequently, we do not yet have the same level of experience in industry with centralized DBMSs. For a prospective adopter of this technology, this may be a significant deterrent.

4. DATABASE DESIGN MORE COMPLEX

The normal difficulties of designing a centralized database, the design of a distributed database has to take account of fragmentation of the data, allocation of fragmentation to specific sites, and data replication.

5. CONCLUSION

Collection of multiple, logically interrelated databases distributed over a network is known as distributed database. The geographical distribution of an organization can be reflected in the distribution of the data, users at one site can access the data stored at other sites. In a centralized DBMS, a computer failure terminates the applications of the DBMS. In a distributed environment, it is much easier to handle

expansion. New sites can be add to the network without affecting the operations of other sites. Distributed DBMS depend on effective communication, we are only now starting to see the appearance of standard communication and the data access protocols.

REFERENCES

- [1] http://docs.oracle.com/cd/B10501_01/server.920/a96521/ds_concepts.htm
- [2] http://en.m.wikipedia.org/wiki/Distributed_database
- [3] <http://ecomputernotes.com/database-system/adv-database/ddbms>
- [4] http://www.teach-ict.com/as_a2_ict_new/ocr/A2_G063/334_applications_ict/distributed_database/miniweb/pg9.htm
- [5] <http://exploredatabase.blogspot.in/2014/08/advantages-and-disadvantages-of-dwistributed-databases.html?m=1>