J I O M
Journal of Information and Operations Management
ISSN: 0976–7754 & E-ISSN: 0976–7762, Volume 2, Issue 1, 2011, pp-21-24
Available online at http://www.bioinfo.in/contents.php?id=55

DISTRIBUTED DATABASE-AN RELEVANCE TO BUSINESS ORGANIZATION

PATIL H.S.[1], MUKHTAR YAKUB SAHEB[2], PAWAR S.U.[2], PATIL D.W.[3]
[1]School of Mgt. Sci.,S.R.T.M.U.,Subcentre,Latur,Email:hanumanthpatil@ymail.com
[2]SSBES's, ITM, Nanded
[3]Swami Ramanand Teerth Marathwada University, SubCentre, Latur
*Corresponding Author: Email- hanumanthpatil@ymail.com

Received: December 01, 2011; Accepted: December 12, 2011

Abstract: Today, almost all organizations depend on their database systems for the crucial information they need to run their business. In every industry across the board, from retail chain stores to financial institutions, from manufacturing enterprises to government departments, and from airline companies to utility businesses, database systems have become the norm for information storage and retrieval. Database systems form the centerpiece of the growing and maturing electronic commerce. Database and Web technologies have merged. The traditional centralized database in today's and in coming days scenario is commonly implemented in the form of distributed database system that allows applications to access data from local and remote sites. . The primary role of databases is to support managerial decision making at all levels of the organizationviz.,Top Level → Strategic Decisions, Middle Management → Tactical Decisions & Operational Management → Daily, Operational Decisions.
Keywords: collaborating server, Query Decomposition, 2PL, OLAP

1. Introduction

As opposed to centralized database in which all the data is maintained at a single site, and assumed that the processing of individual transactions is essentially sequential. Distributed databases use a **client/server** architecture to process information requests.
This distributed database system is broadly classified in following two main categories:
1) **Homogenous distributed database system**, the system comprised of same database application i.e on all computers Oracle is used as DBMS system.
2) **Heterogeneous distributed database system**, at least one of the databases is not an Oracle Database.
It is quite often that distributed database is misunderstood to be alike replicated database. Upto some extent it may seems to be same but the distinction is based on it's purity. In a **pure** (that is, not replicated) distributed database, the system manages a single copy of all data and supporting database objects. Typically, distributed database applications use distributed transactions to access both local and remote data and modify the global database in real-time.

2. Motivations behind use of DDBS

The increasing trends in developing and using distributed database is due to following motivational factors.

1. Performance – as multiple resources can be used performance can be increased.
2. Increased availability – if one site is not in order then other site can be available.
3. Distributed Access to Data – data can be accessed from remote site if the required data is not available on local site.
4. Analysis of Distributed Data – for organization it is possible to access and analyze the data from multiple sites.

Features

1. Distributed Data Independence – users should be able to ask queries without specifying where the referenced relations, copies or fragments of the relations are located.
2. Distributed Transaction Atomicity – Users should be able to write transactions that access and update data at several sites as they would write transactions like local data.

3. Architecture of Distributed DBMS

The architecture of distributed DBMS involves

1.Client-Server,
2. Collaborating Server and
3. Middleware.

3.1 Client-Server Architecture

In this architecture one or more server processes running on mainframe catering services to one or many client processes running on personal computers. The main three reasons behind using this system are i) easy separation of client and server. ii) neither the expensive server or client is underutilized and iii) users can run GUI to which they are familiar

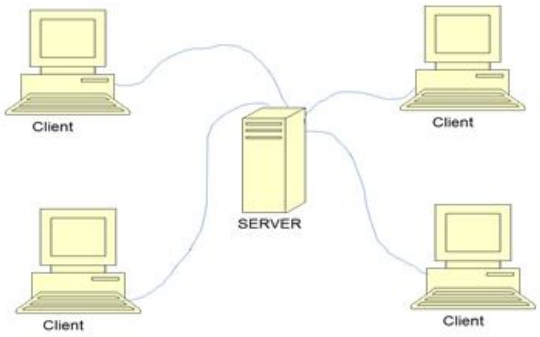

Fig 3.1-Architecture of Client-Server

3.2 Collaborating Server

Many times when data to be accessed from multiple servers the query decomposition is required to form subqueries which needs extra capabilities in client-server system. This limitation can be easily eliminated using collaborating server systems.

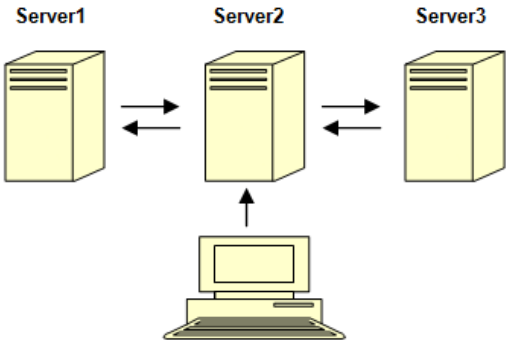

Fig 3.2- collaborating server

In this array of database servers able to run transactions spanning on multiple servers is used. But in this cost of optimization,, network communication and local processing to be considered.

3.3 Middleware

In this, instead of having expensive multiple servers handling queries spanning multiple servers, only one such server – **middle server** is used which can decompose original query in multiple subqeury. All the other servers they just handles their own transaction as instructed by middle server.

4. Data Storage Issue

In distributed databases data can be stored either according to fragmentation or replication scheme.

In fragmentation either horizontal or vertical fragmentation is used.

Horizontal fragmentation – rows or relations divided to be stored in multiple sites.

Vertical fragmentation – column are divided to be stored on multiple sites.

In replication multiple copies of fragmented rows are stored on multiple sites so that maximum data availability and speedy evaluation of query can be achieved. The replication may be synchronous or asynchronous.

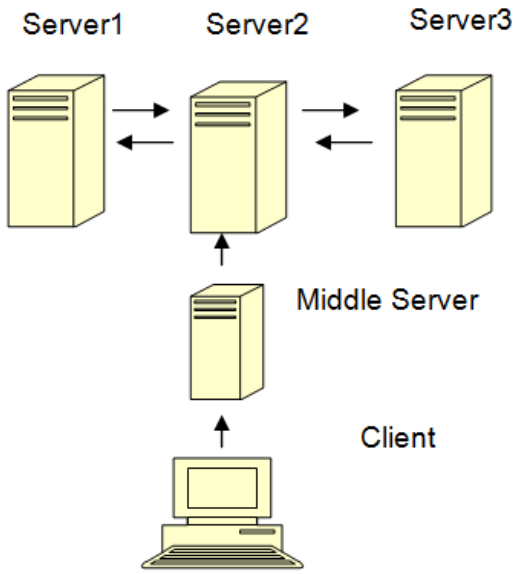

Fig 3.3- Middleware

5. Distributed DB working overview

Whenever user fires a query i.e. user request will be given to USER PROCESSOR which accepts it by User Interface handler. The Semantic Data Controller keeps database valid relative to specified set of constraints. The Global Query Optimizer performs the main job of query decomposition so that it can be handled by multiple appropriate servers with the help of Global Execution Monitor.

Each site or server having its own Data Processor which receives user request and processes it to generate the result set with the help of Local Recovery Manager and Runtime Support Processor. This Runtime Support Processor will interact with Global Execution Monitor to return the result in turn the Global Execution Monitor responds to User Interface Handler which gives response to User.

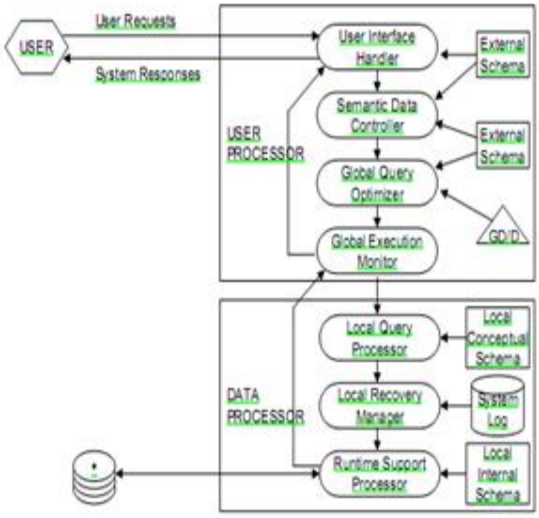

Fig 5.1-structure of working of Distributed DB

6. Distributed Database Transactions

In distributed environment any transaction for its completion uses data on local site as well as on other sites. During this concurrency control becomes a major concern which is dealt with locks and recovery measures.

i) Distributed Concurrency Control – it provides solution for handling locks for those objects which are stored on multiple sites and their unlocking.

ii) Distributed Recovery – this deals with the atomicity property of transaction that enforces either all actions of transaction at all sites to be saved or all actions to be discarded.

The above both measures i.e. concurrency control and recovery are implemented with the help of **Strict 2PL-Two-Phase Locking.**

It is the protocol widely used in locking which follows two rules.

(1) If a transaction *T* wants to *read & modify* an object, it first requests *exclusive* lock on the object.

The DBMS keeps track of the locks it has granted and ensures that if a transaction holds an exclusive lock on an object, no other transaction holds a shared or exclusive lock on the same object.

(2) All locks held by a transaction are released when the transaction is completed.

If transaction is complete, the DBMS releases the lock acquired by transaction on object. This is also done when "COMMIT" or "ROLLBACK" statement is fired.

7. Applications of Distributed Database

With emerging technologies distributed database finds applications in various fields such as:

- Data warehousing,
- OLAP (On-Line Analytical Processing),
- Mobile system and

Data warehousing is used by OLAP tools in decision support system.

A mobile system uses Geo-Positioning system to providing required data from any site. Like data of all Blood Banks is distributed and a doctor or patient with the help of internet connected system may need to find out availability of blood from any blood bank.

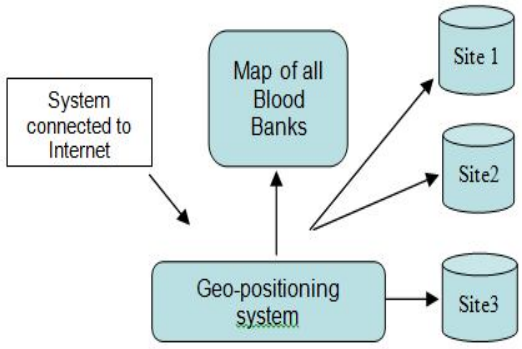

Fig 7.1-Use of Distributed DB

Some more application can also be sited as Data replication, Distributed situation monitoring applications.

8. Advantages & Disadvantages

- **Advantages**

By virtue of data being distributed on multiple sites, this database in distributed form gives following benefits:

i) Capacity and Incremental growth -If there is growth in organization which enforces the growth of database, then with little or no disturbance in existing database, the system can be upgraded.

ii) Reliability and availability - Even if one site (portion) is down or not working & if distributed database in replicated form is used, this make available the data at any instant and increased the reliability on system.

iii) Efficiency and flexibility - As the design of distributed database is such that piece of data is physically stored where it is mostly required hence it can be efficiently used as well as can be replicated easily.

iv) Distributed Database Sharing - Any user at any time can access data from it's own local as well as remote site.

v) Organizational Structure – the physical distribution of data is according to functional department organized in any organization.

vi) Local autonomy – any individual department can control their own site.

vii) Protection of valuable data – in case of disaster at one place, data at other sites is safe.

vii) Economics – it costs less to create a network of smaller computers with the power of single large computer.

viii) Modularity – System can be modified without affecting other modules/systems.

- **Disadvantages**

i) Complexity – continuous and extra attention of DBA is required to maintain the transparency of distributed database system.

ii) Economics – with the growth of organization and distributed sites complexity increases which also leads to labour costs.

iii) Security – each individual must be properly secured from unauthorized access & damage of data.

iv) Integrity Maintenance – in distributed database network each component must be feasible.

v) Lack of standards – yet no standards are laid either to convert existing database or create new distributed database.

vi) Concurrency control – very precisely by using locking and timestamping techniques concurrency must be controlled.

9. Cases

In international and national scenario distributed databases are being commonly implemented. The following enlisted international organizations are:

- **Sea Land Service Inc**

Sea-Land Service Inc. is the largest U.S.-based ocean carrier with a significant worldwide presence. Sea-Land uses its integrated network of ships, railroads, barge lines, and trucking operations to provide efficient containerized transportation to nearly any location in the world. To manage the large volumes of cargo data necessary to support this operation, Sea-Land recently

replaced much of its mainframe architecture with a client-server system based on 42 Microsoft SQL Server and Dell PowerEdge 6300 servers distributed at locations throughout the world.

Sea-Land's Shipment Management system continuously tracks the location of its 220,000 containers to provide real-time updates on the location of cargo and transportation assets. Sea-Land provides access to its customers through a web site that allows users to obtain shipment status and complete bookings online.

- **Northwest Airlines**

Northwest Airlines is the world's fourth largest airline, serving more than 400 cities. Its 415 airplanes fly over 1,700 flights daily. The airline uses a central DB2 database to track flight routes, schedules, and crew lists. To maximize the value of this information, Northwest Airlines had to make this data easily accessible to its various divisions from their local databases. This required a heterogeneous replication solution that would allow the operations data on DB2 to other DB2, Sybase, and Oracle databases. Initially the airline used a combination of customized programs, file transfer protocol (FTP), and middleware to implement a simple form of data replication. This project replaced Northwest's locally developed solution with one based on IBM's DataPropogator Relational and DB2 DataJoiner.

Some more international players of distributed databases are:

- **E-Plus Mobilfunk GmbH**
- **Surridge Dawson Ltd.**

In India, Damodar Valley Corporation – a Govt. of India Enterprise is in the business of power production and distribution, water management and soil conversion in estern India, having its various offices in New Delhi, West Bengal - Kolkata (HQ), Dhanbad, Bokaro (Jharkhand). To provide various services to it's customers as well inter-departmental employees, distributed database is care part in it's automation.

Conclusion

From above discussion it can be cleared that the organizational trend is inclined towards use of distributed databases as compare to traditional centralized databases. And with the growth of organization and it's customer base which demands more easy and fast services, surely distributed databases will be extensively used.

References

[1] Raghu Ramkrishnan / Johannes Gehrke: Database Management Systems by McGraw Hill Higher Education.

[2] Sheth and Larson J. (1990) *Computing Surveys*, 22(3):183{236.

[3] Tamer Özsu M. and Patrick Valduriez (1999) Principles of Distributed Database Systems. Prentice Hall.

[4] Quass D., Gupta A., Mumick I.S. and Widom J. (1996) Making Views Self-Maintainable for Data Warehousing, Proc. of Fourth Intl. Conf. on Parallel and Distributed Information Systems (PDIS '96).

[5] Gerald Bernbom (2000) Data Administration and Distributed Data Processing (Cause/Effect Vol. 14, Number 4, Winter 1992. http://www.lessaworld.com/dm/dm02.txt Internet6.

[6] Distributed Data Processing (2000) http://www.bus.okstate.edu/lhammer/AISweb/AIS9.HTM Internet.