Available online at www.sciencedirect.com

ScienceDirect

Procedia Social and Behavioral Sciences 15 (2011) 2673–2677

Procedia Social and Behavioral Sciences

WCES-2011

The use of distributed databases in e-learning systems

Iacob (Ciobanu) Nicoleta – Magdalena[a] *

[a]*Phd student, Doctoral School, University of Pitesti, Department of Computer Science, Str. Targu din Vale, nr.1, Arges and Postcode 110040, Romania*

Abstract

E-learning: The modern technologies in education and research aims to build a society based on knowledge. In this paper we will describe the advantages of e-learning education and the importance of data allocation in the process of designing a distributed database system. In addition, we are going to analyze the available strategies for optimizing the query execution in a distributed environment, in order to obtain the smallest system response time. The purpose of this paper is to present the advantages offered by the use of distributed databases in the education field by universities with geographically distributed locations.

Keywords: e-learning; architecture; design; allocation; query optimization; strategies.

1. The general issues about e-learning

The information and communications technology and development of multiple connections, anywhere and anytime, allowed a new approach about education, by more geographic efficient distribution of information, knowledge and practice. Thus a new concept appeared, called **e-learning**, which means using new multimedia technologies and Internet to improve the quality of learning and evaluation process. **E-learning courses** include both content (that is, information) and instructional methods (that is, techniques) that help people learn the content (Clark & Mayer, 2008).

Beyond the psycho-pedagogical theories, the computer assisted instruction changes student attitudes, but also the teacher's to learning, the student becomes the learning center approach, his role is active because he builds knowledge, and the teacher mediates the learning filling his personal experience with what it offers the e-learning platform and educational software included. In the computer-assisted instruction, the access to educational objects must be addressed in terms of developing applications for reading information from different files (Isăilă & Smeureanu, 2010).

* Iacob (Ciobanu) Nicoleta – Magdalena. Tel.: +40-757-011-895 .

E-mail address: nicoleta.iacob_2007@yahoo.com .

This article was produced under the project: "Development of the doctoral schools by providing fellowships for young PhD students with frequency"- ID 52826 .

1877–0428 © 2011 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.sbspro.2011.04.168

E-learning is the interaction of the teaching/learning process and information technologies - ICT (Information and Communication Technology) -, covering a wide spectrum of activities, from the computer-aided education until the education conducted entirely in online manner.

The concept of e-learning provides several **advantages** to educational organizations which use this technology, including short and effective training, flexibility and modularization. **The disadvantages** of e-learning education are: high drop-out rate of students, working in small groups, high costs for design and maintenance. With all these disadvantages or limitations, the experience of e-learning platforms already utilized has shown that participants in education process using the new e-learning technologies soon became familiar with the virtual environment and showed us that this way of learning is a modern and efficient type of education.

The e-learning system has brought great changes in education process resulting in the reorganization and transformation of education in order to provide learning and training opportunity for all those concerned with their career evolution, regardless of age, gender, income or geographical area.

2. The distributed database definition and objectives

The need for distributed systems has been determined by studies, research and concerns in educational field by the universities with geographically distributed locations where the specific organizational structure promotes a decentralized research model. For institutions that are expanding globally, the exchange of data between multiple databases and applications has become very important.

A **distributed database (DDB)** is a collection of logically interrelated databases, but physically distributed over several computers (a network of computers).

To ensure compliance with **the DDB's specific objectives** such as: increasing of system reliability and of data availability, decentralization of system resources, better use of their and increase system adaptability to changes in organizational structure, in the DDB design aims to ensure the following *principles*:

- *Maximizing local processing of data* can be done by placing data closer to the applications they require. It is noted that in a properly designed DDB approximately 90% of the total data should be locally accessed and only 10% from remote locations. The advantage of the fully local execution of the application is given not only by the reduction of remote accesses but also by simplification of the control in application execution.
- *Enabling a high level of data security and availability* can be done by data replication to numerous sites. In this way, the system can use an alternative copy when the one which had to be accessed under regular circumstances is not available.
- *Parallel data processing* can be done by the efficient using of CPU capacity of each station.

The DDB systems present a number of a*dvantages* mainly related to the increase of system performance, by reducing the relative costs, processing and transmission of data.

3. The distributed architecture of an e-learning system

The distributed architecture of a system of e-learning means that there are four types of servers (Figure 1):

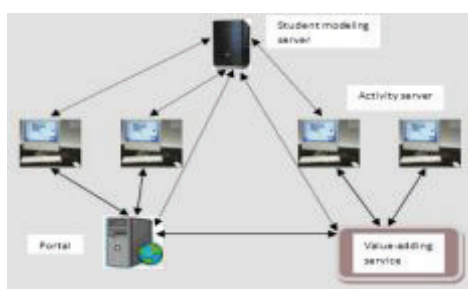

Figure 1. The architecture of an adaptive distributed e-learning system

- *The activities server* is a component focused on the needs of content and services.
- *The education portal* allows the instructor to structure the access to various learning objects in order to meet the requirements for students and thus providing a complete course.
- *The value-added service* is realized by providers of services independent of course, that can be reused, and combines the first two features, adding the adaptive sequencing, the annotation, the visualization and the content integration. It has the ability to query the activities server and access to them.
- *The server that stores the student model* is a component that represents students' needs and expectations of e-learning process. These servers allow custom content.

4. The design of distributed database

The design of distributed database represents a complex activity, which must take into consideration a lot of *factors* such as: the requests used to access the applications, the integrity restrictions defined at the database level, the configurations of the nodes and of the network. The first stage in distributed database design is to develop a data model in the same way as for a centralized system. The requirements of distribution may then mean that some changes need to be made to the data model. The second stage is to decide how to fragment and replicate the logical schema. A data fragment represents a horizontal and/or vertical fragment of tables in the database. The advantages of fragmentation are: the efficiency, the local optimization, the security, the ease the query. In a computer network is very important to not allow the unauthorized persons to gain access to the information sent between computers (Defta, 2010). The main aim of distributed database design is to reduce communications traffic. We wish to locate data close to its point of use. Hence a consideration of the topology of the communications network will influence the distributed design (Beynon-Davies, 2004). The distributed type of data involves apart from the activities which refer to the design centralized database new ones, fragmentation and data allocation geographically distributed. The allocation of fragments is closely related to the replication of data from DDB. Fragmentation and replication can be combined: A relation can be partitioned into several fragments and there may be several replicas of each fragment.

The replication involves the existence of a copy of the database on one or more nodes in the system. Any change in the site should be reflected in all copies of the site. Its main *advantages* are safety, availability and performance for query operations, but costs of storage are higher, so are the costs of update. Assuming replicas are mutually consistent, replication improves availability since a transaction can read any of the copies. Furthermore, replication provides more reliability, minimizes the chance of total data loss, and greatly improves disaster recovery (Rahimi & Haug, 2010).

5. Allocation problem definition

The problem of fragment allocation can be treated as an optimization problem of placement of each fragment. Efficient distribution of fragments requires a balance between costs (storage, processing and transmission of data), performance (especially response time) and data distribution restrictions.

The optimal solution for the distribution can be defined by two *objectives*:

a. *Minimize costs* involves calculating the total cost of a query execution.

b. *Maximizing performance*. The performances are measured, especially, by system response times when dealing with read/write data processing operations.

Finding an optimal or even a good solution to distributed data allocation is a complex optimization problem (Elmasri & Shamkant, 2006).

6. Optimization of distributed queries

Query optimization is the process of selecting the most efficient query-evaluation plan from among the many strategies usually possible for processing a given query, especially if the query is complex.

The factors that affect query evaluation in a distributed system are: the fragmentation details for the relations, the location of these fragments/tables in the system, and the speed of communication links connecting the sites in the system (Rahimi & Haug, 2010).

The problem of optimizing distributed query execution has the goal to obtain a system response time as small as possible but also to minimize total cost of execution.

The total cost of a query execution in a distributed system, (formula 1), consists of the CPU cost for processing (CCPU), the cost of accessing the discs (CI/O) and the cost of communication between nodes (consists of the cost of sending/receiving messages between nodes (CMSG)) and, respectively, the cost of transferring data between nodes (CTR).

Total_cost=CCPU*nr_instructions+CI/O*nr_I/O+CMSG*nr_messages+CTR*data_quantity (1)

The response time of distributed system, (formula 2), is calculated from the moment when the query begins until it receives response from the system.

Response_time=TCPU*nr_instructions+TI/O*nr_I/O+TMSG*nr_messages+TTR*data_quantity (2)

The minimization of the response time is achieved by increasing the parallelism degree of a distributed query execution. The higher the execution and communication time the lower the response time will be. This does not necessarily involve that the total cost will be minimized.

6.1. Strategies for optimizing the query execution

There are two strategies for optimizing the query execution when the necessary data is stored on multiple nodes: *queries - site* and *move - small,* in order to formulate and solve data allocation problems, and to minimize the quantity of data transferred between nodes, and consequently, the cost of communication, which is a significant part of the total cost.

a. **The query – site strategy** states that if we have a query that uses data that are found in several fragments located on different nodes, the best approach would be to transfer all those fragments at the node where the query was initiated, and to execute it there.
b. **The move – small strategy** states that if a binary operation involves two fragments that are on different nodes, then we must transfer the smallest fragment to the node where the larger fragment is located (Iacob, 2010).

7. Case study: "E-learning systems of academic management, using distributed databases

The e-learning is a technology that has revolutionized the traditional system of distance learning, and its opportunities have been seized and used not only by educational institutions, but also by different public or private organizations. In terms of structure, an *e-learning* system provides facilities for the transfer of knowledge through the development and publication of educational content in the form of courses or virtual libraries, as well as testing of knowledge using management simulations, scenarios or case studies for evaluation.

E-learning type courses are increasingly present in the offer of various universities or academies. The e-learning portals include a public section, generally available, for informational purpose, but also private section for the participants in the e-learning process. The distributed databases are suitable as a technical solution when there are several branches of a geographically dispersed university.

The portal will include a minimum of three sections, as follows:

a. *The section accessible to a student* that includes:
   - access to the information relevant to the courses they attend;
   - facilities for viewing and solving tasks.

b. *The private section of a teacher* must include facilities for:
   - *online* publishing of educational materials (courses, applications of seminar, laboratory, tests, etc..) and indicate useful resources their;
   - defining schedules of courses, seminars and laboratories;

- setting home tasks (requirements, scales, resources);
- management of exams and management of information about the online catalog.

c. *The section corresponding to a particular course* should be accessible only to students enrolled and teaching staff involved in that course. It usually includes the following:
- curriculum structure;
- information about the module of deployment of the course and other educational activities.

Thus, for each section of the portal there are some elements that must be fragmented or replicated between multiple sites. To better understand those concepts, we will give some specific examples.

To improve processing speed and data security it is necessary to partition database table which contains data about a university student, so that some records will be located in one site and others in another one. By this horizontal partitioning, the records of students who study in Paris will be stored in the database in Paris, since most requests of accessing table with students from Paris will be at the University of Paris. The disadvantage is that when an application needs to access all records of students in all towns within the university, it must collect data from each of these nodes. The second example is about the foreign languages courses, French for example, that must be found in the databases located in the countries where there is a branch of the university and where there are many requests for studying this foreign language, so there is a need for data replication. Distribution of these replicas has the objective to improve the speed of data processing operations and availability of transactions.

8. Conclusions

The use of distributed databases in e-learning systems has the goal to improve access to information and also rapid data collection. The aim is the creation of an educational network based on e-learning tools which allows greater flexibility in the training of persons in terms of efficiency in accordance with national standards.

The new information and communication technologies change the perspective of educational practice, complementing the educational system with modern learning methodologies involving information technology benefits. The e-learning does not want to replace the traditional educational systems, but to reinforce the learning process. This new way of learning using modern technologies is implemented in schools as an alternative to traditional education, and contributed to *the distance learning* form of study.

References

Beynon-Davies, P. (2004). *Database systems.* (3rd ed.). New York: Palgrave-Macmillan, (Chapter 37).

Clark, R., & Mayer, R. (2008). *E-learning and the science of instruction*. (2nd ed.). San Francisco: John Wiley and Sons, (Chapter 1).

Defta, L. (2010). Network security attacks. ARP Poisoning. Case study. *Proceedings of International Scientific Conference ECO-TREND*: Brâncuşi University Târgu Jiu, 617-622.

Elmasri, R., & Shamkant, B. (2006). *Fundamentals of Database Systems*. (5th ed.). USA: Addison Wesley, (Chapter 25).

Iacob, N. (2010). Distributed query optimization. *Proceedings of International Scientific Conference ECO-TREND*: Brâncuşi University Târgu Jiu, 635-640.

Isăilă, N., & Smeureanu, I. (2010). The accessibility of information in computer assisted learning process for persons with disabilities. *Proceedings of the 4th European Computing Conference (ECC'10)*: Politehnica University Bucharest, 139- 142.

Rahimi, S. K., & Haug , F. S. (2010). *Distributed database management systems*. A Practical Approach, IEEE, Commputer society, New Jersey: John Wiley & Sons, INC., (Chapter 4, 7).