Efficient Data Management to support Educational Activities

C. Bouras[1,2], P. Destounis[1,2], J. Garofalakis[1,2], A. Gkamas[1,2], G. Sakalis[2], B. Sakopoulos[2], J. Tsaknakis[1,2], Th. Tsiatsos[1,2]

[1]Computer Technology Institute, Greece, Computer Engineering and Informatics Dept., Univ. of Patras, Riga Ferraiou 61, GR 262 21 Patras, Greece - P.O. Box 1122, GR 261 10, Patras, Greece

[2]Computer Engineering and Informatics Dept., Univ. of Patras, Riga Ferraiou 61, GR 262 21 Patras, Greece - P.O. Box 1122, GR 261 10, Patras, Greece

Abstract

In this paper we present data management issues faced during the design and development of an Open Distance Learning system for the University of Patras, Greece. In order to handle data efficiently, as required in a Web Tele-training Application, for each type of information maintained different strategies must be deployed according to their behaviourand structure.

The diversity and complexity of data, the network aspect of the application and web deficiencies impose an architecture design incorporating a plethora of technologies and tools that must be integrated in such a fashion that they efficiently organise these data preserving their relationships. This presents a software engineering challenge requiring coherence of solutions at all levels: structures, consistency, security, models, and protocols.

The paper presents the data components of a open and distance learning system that access the information stored in a database and the file system, their underlying technology, their interaction with the network services, and features regarding the ways they address issues faced in a open vendor-independent distance learning environment and outlines the system's overall architecture.

Key words: Web Query Languages, Data model for web data, Web site management, Teletraining

Introduction

An Open Distance Learning Information System (ODLIS) can be used in order to provide virtual lectures, virtual conferences, collaborative work on projects which are shared among institutions, exchange of useful material, experiences among teachers, research results and conclusions as well as versatile and more attractive presentations of the subjects taught. In addition, such a system should manage the educational material, its users and information that is useful for the educational procedure. The above requirements create two problems for the system: The first problem is the management of data in web-based applications, which have different characteristics and complexity. The management of information has many limitations of the file-oriented hypertext protocol. The second problem is the integration of different technologies and tools that support educational activities.

In our approach, all application's sensitive information is stored in a relational database management system that provides the mechanisms for the efficient indexing and ensures their consistency. More static data are stored directly on the file system while the database maintains annotation about them and pointers to their location.

In this paper we present the architecture and the components of an ODLIS and particularly the data access components and the business logic regarding the manipulation of information. In addition we catalogue the functional characteristics based on the user requirements, describe the network that supports the

transport of information. Furthermore we explain the types of data in our system, relationships between data, the issues of consistency, manageability, scalability, security, partitioning, transparency and we describe the characteristics of the mechanisms that used for the reliable storing. Finally we present the services provided by the system and our case study.

General Architecture of the Open Distance Learning Information System

The architecture of the ODLIS is based on the 3-tier architecture model. The ODLIS consists of the Database (1st layer - Data Access component) which is responsible for providing the first-born information, the Server (2nd layer - Business component) which is responsible for processing this information and the User Interface (3rd layer - Presentation component) which is responsible for the presentation of the results and the interaction with the users.

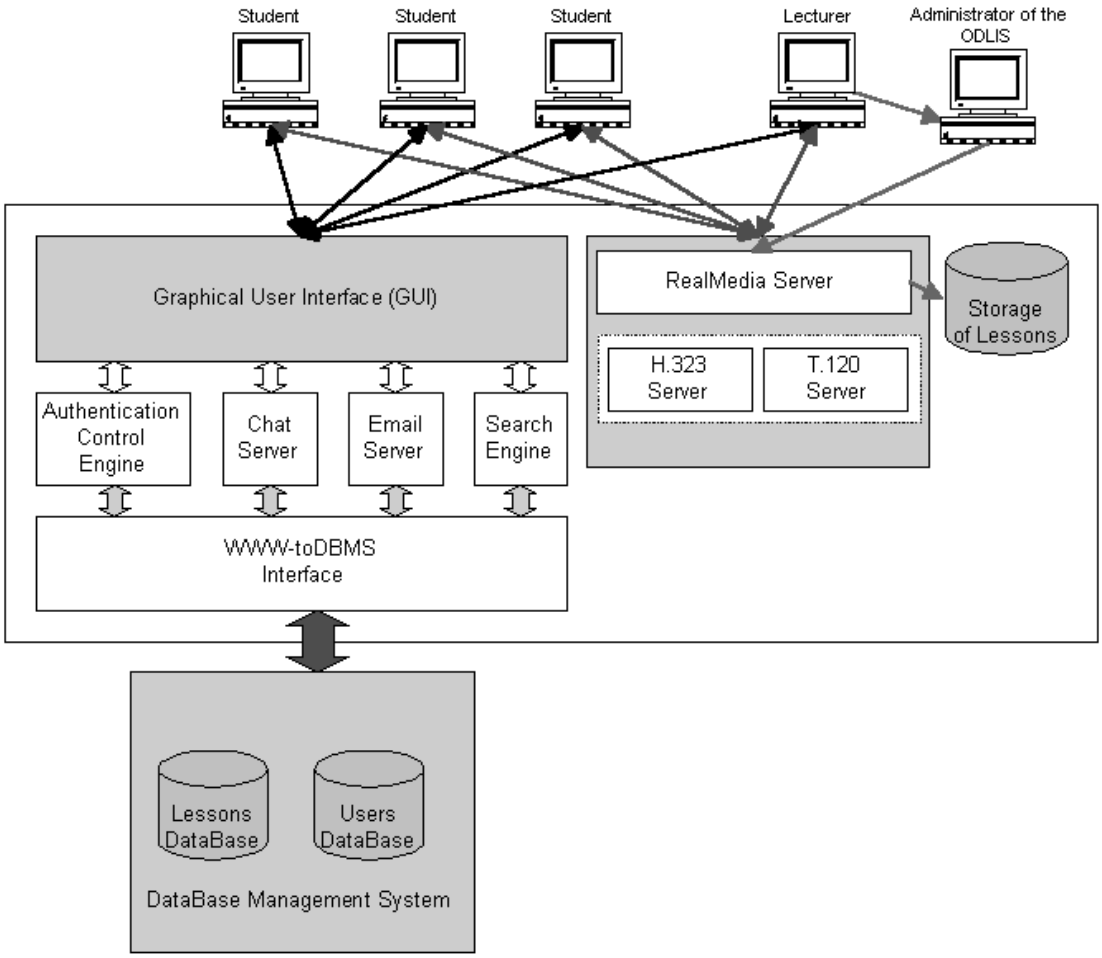

Figure 1 General Architecture of the ODLIS

The environment of the ODLIS consists of three modules that supply various functionalities. These modules are the followings:

- The module for the administration of the environment of the ODLIS offers two basic functions. The first is the administration of the users of the ODLIS and the second is the administration of the lessons of the ODLIS. This module provides capabilities to create, delete or modify a lesson, the educational material of a lesson and the characteristics of a user. In addition this module offers the capability to search the database of the ODLIS.
- The module for the administration of the educational procedure. This module is responsible for the interaction between the students and the teachers, which is based on a discussion list with thread capabilities, web based shared workspace and e-mail. This module is also responsible for the submission of exercises to the students by the teacher and the observation of a file with the grades of each student.
- The module which is responsible for providing the synchronous and asynchronous lectures over the network. This module is responsible for the interaction between the students and the teachers during a synchronous lesson, which is based on video communication, sharing Whiteboard, chat and “ask floor” capability and for the attendance of the asynchronous lessons by the students.

Functional Specifications and Requirements

The Open Distant Learning in Asynchronous mode problem

The services that are offered contained quite a few troubling issues in the database management and data transfer level.

- All asynchronous discussions are stored in a special database that keeps log for all participants and their statements. The database had to keep track of people as objects and be able to categorised their statements both in people's personal data and in discussion data during the whole process of the discussion
- Workspace sharing application had to be able to accept answers to the projects and furthermore keeping track of questions and remarks given by students. The project faced troubles concerning the user list synchronisation between the authentication level and workspace database users, especially when the power users (such as "instructor") created new workspace without the intervention of the administrator.
- Announcement database issues majored in the indexing and organisation of the data (announcements with advance searching tool). Announcements are a high level tool mainly used during the off line distant learning which meant that database structure should maintain a clear and distinct way of keeping and distributing information.

In the synchronous part of the project issues in totally different perspective had to be dealt with

- Logging for file transfer and chat-like communication had to be provided. The logs had to be explicit especially whatever concerned the members involved in the communication.
- Scheduling of the meeting had to be available. The meeting-time database had to update the announcements and the calendar of the instructor.

The user and lesson database issues had also an important role and needed special care especially as far as the educational part is concerned. The database system had to take into consideration all the educational aspects that were imposed by the academic view of the project. User database had to keep wait state status when registering a student for the cross certification with the secretariat of the department involved.

Moreover the database had to be ready for annual rebuilding - reutilization prepared for the next academic year. At that time all out-dated data had to be moved to the back up database keeping in mind all issues that may be needed in the following year for better indexing and quicker access and retrieval.

Network

The ODLIS uses a TCP/IP network in order to provide its services. There is a server in which can be connected all the users. All the users, who are connected to the Internet and have a login and a password, are able to connect to ODLIS. The topology of the network is the following:

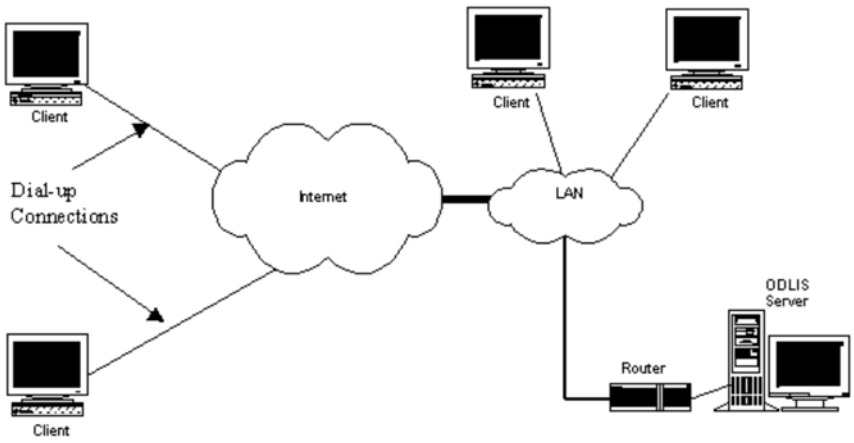

Figure 2: The topology of the network

The data are sent as packets with a TCP (Transmission Control Protocol) or UDP (User Datagram Protocol) connection in proportion of the packets. Both voice and video data use a UDP connection, because they require high speed and not high quality. On the other hand text and images require high quality and it is essential to use the TCP protocol.

The services of the ODLIS use the RTP/RTCP (Real Time Protocol/Real Time Control Protocol) protocols in order to provide end-to-end network operations that are essential for real time applications. In addition,

the RTSP (Real Time Streaming Protocol) is used. The RTSP is an application level protocol for the control, the real time transmission and on demand delivery of data such as voice and video.

Data manipulation issues

The data access components providing the interface to the information stored reside mainly on the web server and the database server. They access information that is stored on the file system or the relational database management system.

Interface to Web

The web server's data access components take the form of server-side scripts while those of the database server comprise the stored procedures used for the management of the information stored in the database. The technology used for the server-side scripts is PHP. PHP is a server-side cross-platform HTML embedded scripting language. The majority of the HTML document responses of the application's user interface are produced 'on the fly' according to user's access level, information accessed and user's request so the need for such a mechanism like PHP was essential [11].

The PHP offers mechanisms for the manipulation of files and folders as well access to most of the available RDBMSs. It provides mechanisms for maintain session information – information is stored in session variables that contain information during the user's visit. The session variables are used mainly for caching data that are stored in the database or the file system and are accessed frequently during a session.

Currently, the three most popular web servers Apache, Microsoft's IIS, and Netscape Enterprise offer server-side scripting capabilities, executing scripts and components on the server, without the performance limitations and development difficulties of CGI. HTML and HTTP do not by themselves provide a way to access databases or carry information about users from page to page. Server-side scripting accesses programs on the server that provide this necessary functionality behind the scenes to deliver Web applications and customized HTML for each user. Web server scripting, also, separates the content from the presentation for easier design and data management allowing the use of templates for creating HTML documents on the fly.

The contents of a page can come from anywhere – databases, plain text files, searches, calculations– and be dynamically inserted before it is sent to the user. Information can be managed in the most appropriate manner, and does not have to be stored in HTML pages that must be changed by hand whenever the data changes.

Types of information

The system is based on a mixed model. All data comprising the lectures' content are stored on the file system while all other kind of data are stored in the database. The reasons for following that model have to do with issues of manageability, scalability and transparency.

Lectures' content, HTML documents, images, audio, video, presentation slides are accessed by the Web server or the Real media server. They are static information that doesn't change frequently. They are stored directly to the file system. Storing them with another data manipulation mechanism like DBMS would make the system more complicated. Furthermore, storing and editing or authoring this information can be done easily, with the use of ordinary tools.

The rest of information is stored to the DBMS. To be more specific an R-DBMS is being used, as a relational model can easily represent the kind of information that is being stored more sufficiently than any other model like object-oriented or hierarchical model.

The reasons that a DBMS storing mechanism is used are many:

- A database is designed, built, and populated with data for a specific purpose. It has an intended group of users and some preconceived applications in which these users are interested.
- It has the ability to handle large amount of data easily.
- It is scalable. A developer can easily alter the database schema without changing all storing and reading mechanisms.
- A database can give access to the data to multiple users with different roles and different level of accessibility to each one.
- It provides indexing mechanisms that used for the efficient accessing of large amount of data.
- It also provides mechanisms to easily enforce integrity or referential constraints to data.

The information stored on the database consists of tables containing information about the lectures such

as keywords, pointers to the lectures' content, questions, authors or professors, and students. For each student a performance and attendance log is kept. Even though, the web server is responsible for user authorisation, each user rights are kept in the database.

Being more specific, there are curriculums, which contains groups of courses. Every course has a code name, start and end date, and description. Every course contains groups of lectures. These lectures are the static information like HTML documents, Images and Multimedia content. A lecture may have two parts. One asynchronous with static data, and one synchronous. Also a lecture has keywords so a student or a Teacher can easily find a lecture that is interesting for him. Moreover, there are questions at every lecture, and exercises at the end of every course.

The system also stores information about the users and the administrators. A user can be a teacher or a student or both. But if someone is teacher at one course he cannot be also a student to that course also. For every student, the system logs his attendance and his marks from questions and exercises. The Administrator has the authority to add and remove users, and to create new curriculums and add the material to the system.

A detailed description of the database is at Figure 3 :

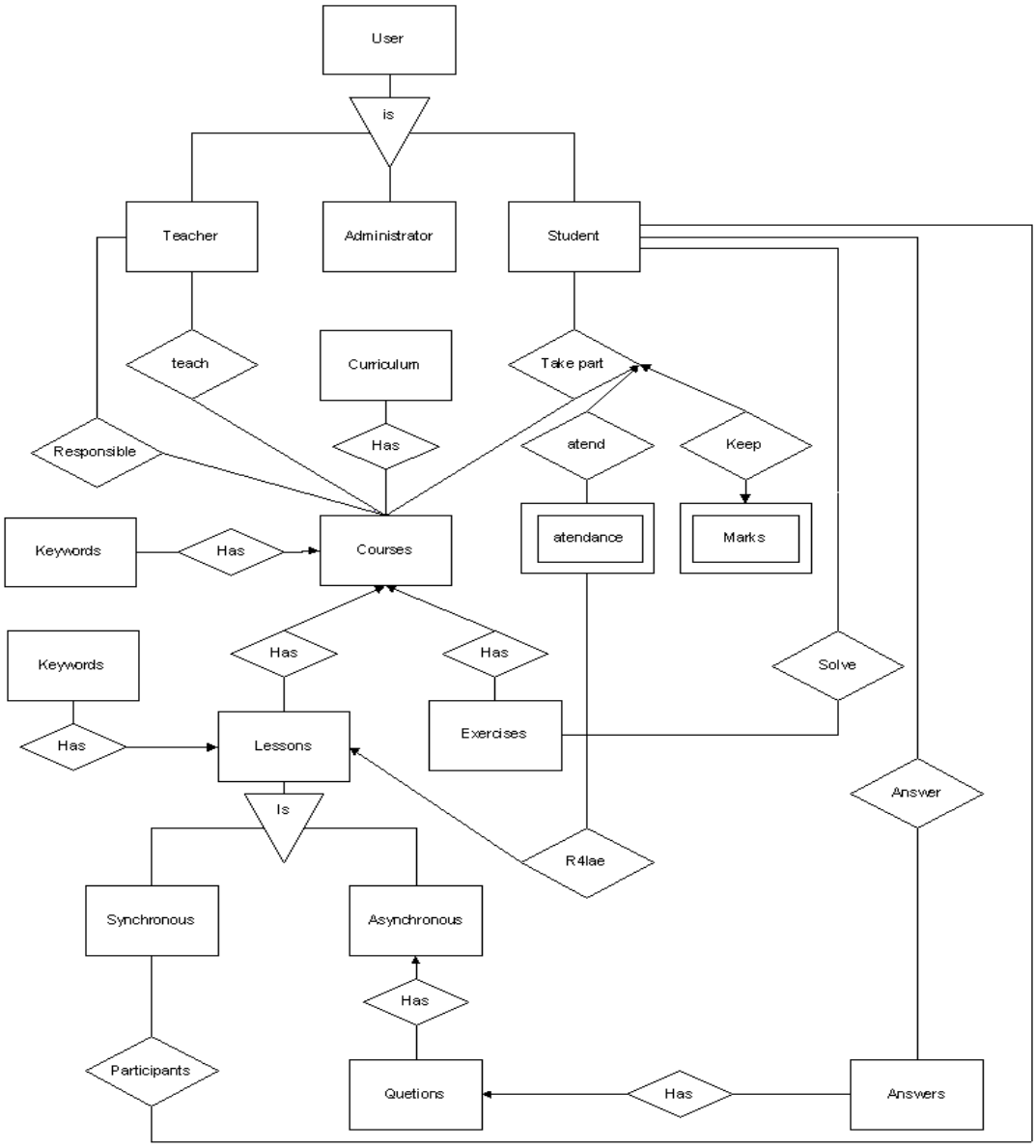

Figure 3. Relational Diagram

Services

Asynchronous Open and Distance Learning Service (AODL)

For the implementation of the of the AODL service we use the architecture of the Figure 4.

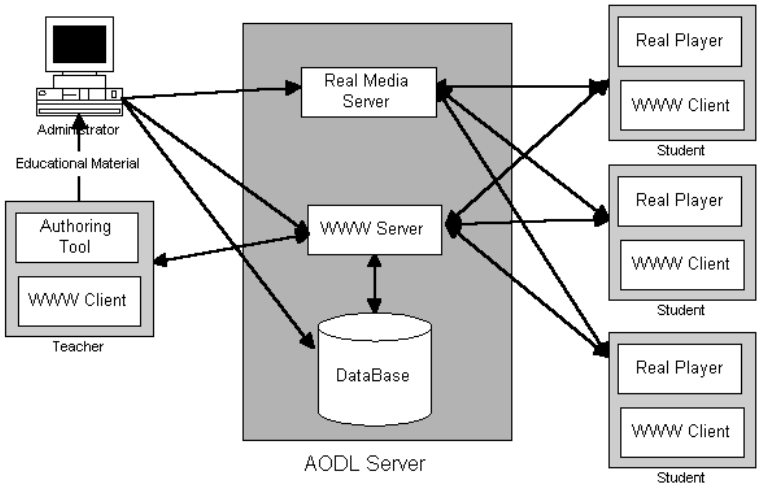

Figure 4 Implementation of the AODL service

The Server of the AODL consists of the Real Media Server, the Web Server of the ODLIS and the database of the ODLIS. The Real Media Server is responsible for the transmission of the video over the network, the Web Server is responsible for the transmission of the slides and the database is responsible for management of the data.

The above-described solution is using for the transmission of the video the streaming technology. The streaming technology with the achievement of the last years (for example the RTSP protocol) has many advantages for the transmission of multimedia data over a TCP/IP network like Quality of Service schemes (with the use of buffering).

The above-described architecture did not confine the way that an asynchronous lecture can be transmitted. For example a teacher can create an asynchronous lecture with tool of his choice (for example Macromedia) and suggest the student to install the appropriate software or plug-in to their computer in order to attend the asynchronous lecture.

Synchronous Open and Distance Learning Service (SODL)

The Figure 5 displays the implementation of the SODL with the use of H.323 reflector.

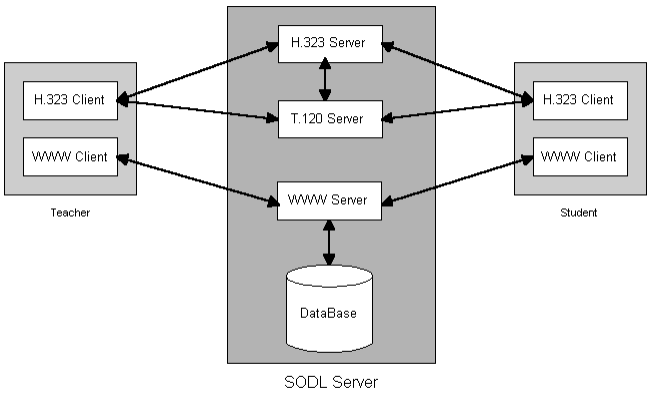

Figure 5 Implementation of the of the SODL service with the use of H.323 server

The SODL server consists of an H.323 server which is responsible for the transmission of the video, a T.120 server which is responsible for the application and the data sharing and the Web server of the ODLIS. The clients of the Teacher and the Students consist of a H.323 client and a web browser. In this implementation we use White Pine MeetingPoint Conference Server.

The above-described implementation has the advantage that follows international accepted standard (like ITU H.323 and ITU T.120). In addition this implementation give to the Teacher the flexibility to prepare the educational material with the software of his choice.

This implementation has the drawback that use tool designed for videoconference and not for distance learning. This may create problems to the educational procedure from malicious students.

Case study: The University of Patras in Greece

The Open and Distance Learning program of University of Patras in Greece will use the above described information system. In the first phase will operate three curriculums, two for undergraduate studies and one for postgraduate studies and. More particular will operate the postgraduate curriculum on Special Themes on Computer Science and the undergraduate curriculums on Computer Science and Neuro-Science. All the corriculums will use both asynchronous and synchronous lectures. In the future the ODL program of the University of Patras will offer more curriculums.

Conclusion

With the advancement of technologies, learning the future can become radically different from what it is today. Although no one can expect that educational networks will totally replace the traditional lecture, it is likely that such applications will obtain more and more positive ground in the future. On the other hand information systems that supports ODL can be useful tools especially when function in collaboration with traditional learning methods. In addition such systems gives the opportunity to users to access the educational material without time or place restrictions.

Our next step is to evaluate the usage of the ODLIS through the ODL program of University of Patras in Greece.

References

1. J. BOEGH, A. KREBS, L. PETERSEN & M. WAGNER, DELTA Danish Electronics, Light & Acoustics, Denmark, A Web based Virtual College, WebNet 98 — World Conference of the WWW, Internet & Intranet Orlando, Florida; November 7-12, 1998
2. E. S. BOS, A. KIKSTRA, C. M. MORGAN. Multiple Levels of use of the Web as a Learning Tool. Proceedings of ED-TELECOM 96, pages 31-36 - Boston, Mass USA, June 17-22, 1996.
3. Ch. Bouras, A. Gkamas, V. Kapoulas, P. Lampsas, Th. Tsiatsos, A platform for the implementation of the services of an educational network, 15th IFIP World Computer Congress TELETEACHING '98 in Vienna and Budapest 31 August - 4 September 1998
4. G. Eschelbeck. An Architecture for Multimedia Communication in a Distributed Education Environment. Proceedings of ED-MEDIA 95, pages 217-222 - Graz, Austria, June 17-21, 1995.
5. D. LEVIN, M. BEN-JACOB, Using Collaboration in Support of Distance Learning, WebNet 98 — World Conference of the WWW, Internet & Intranet Orlando, Florida; November 7-12, 1998.
6. D. PONTA, A. M. SCAPOLLA, M. TAINI. Telematics for Education: The Design of a Distributed Computer-Based Collaborative Learning System. Proceedings of ED-TELECOM 96, pages 252-257 - Boston, Mass USA, June 17-22, 1996.
7. A. F. SMEATON, F. CRIMMINS. Virtual Lectures for Undergraduate Teaching: Delivery Using RealAudio and the WWW. Proceedings of ED-MEDIA/ED-TELECOM 97, pages 990-995 Calgary, Canada, June 14-19, 1997.
8. K. STEFANOV, B. LOMEV, S. VARBANOV, R. NIKOLOV, "Distance learning course on business on the Internet: some implementation issues" 15th IFIP World Computer Congress TELETEACHING '98 in Vienna and Budapest 31 August - 4 September 1998
9. R. Wang, A. Karmouch. A Broadband Multimedia TeleLearning System. Proceedings of the Fifth IEEE International Symposium on High Performance Distributed Computing, pages:132-139.
10. Oracle Corp, http://www.oracle.com
11. PHP, http://www.php.net
12. RealNetworks, http://www.real.com