

**THE BUCHAREST ACADEMY OF  
ECONOMIC STUDIES**

ISSUE

**5**

# Database Systems Journal

---

**ISSN: 2069 – 3230**

**Volume II (September 2011)**

---



**Journal edited by Economic  
Informatics Department**

## **Database Systems Journal BOARD**

### **Director**

Prof. Ion LUNGU, PhD - Academy of Economic Studies, Bucharest, Romania

### **Editors-in-Chief**

Prof. Adela Bara, PhD - Academy of Economic Studies, Bucharest, Romania

Prof. Marinela Mircea, PhD- Academy of Economic Studies, Bucharest, Romania

### **Secretaries**

Assist. Iuliana Botha - Academy of Economic Studies, Bucharest, Romania

Assist. Anda Velicanu Academy of Economic Studies, Bucharest, Romania

### **Editorial Board**

Prof Ioan Andone, A. I. Cuza University, Iasi, Romania

Prof Emil Burtescu, University of Pitesti, Pitesti, Romania

Joshua Cooper, PhD, Hildebrand Technology Ltd., UK

Prof Marian Dardala, Academy of Economic Studies, Bucharest, Romania

Prof. Dorel Dusmanescu, Petrol and Gas University, Ploiesti, Romania

Prof Marin Fotache, A. I. Cuza University Iasi, Romania

Dan Garlasu, PhD, Oracle Romania

Prof Marius Guran, Polytechnic University, Bucharest, Romania

Prof. Mihaela I. Muntean, West University, Timisoara, Romania

Prof. Stefan Nithchi, Babes-Bolyai University, Cluj-Napoca, Romania

Prof. Corina Paraschiv, University of Paris Descartes, Paris, France

Davian Popescu, PhD., Milan, Italy

Prof Gheorghe Sabau, Academy of Economic Studies, Bucharest, Romania

Prof Nazaraf Shah, Coventry University, Coventry, UK

Prof Ion Smeureanu, Academy of Economic Studies, Bucharest, Romania

Prof. Traian Surcel, Academy of Economic Studies, Bucharest, Romania

Prof Ilie Tamas, Academy of Economic Studies, Bucharest, Romania

Silviu Teodoru, PhD, Oracle Romania

Prof Dumitru Todoroi, Academy of Economic Studies, Chisinau, Republic of Moldova

Prof. Manole Velicanu, PhD - Academy of Economic Studies, Bucharest, Romania

Prof Robert Wrembel, University of Technology, Poznań, Poland

### **Contact**

Calea Dorobanților, no. 15-17, room 2017, Bucharest, Romania

Web: <http://dbjournal.ro>

E-mail: [editor@dbjournal.ro](mailto:editor@dbjournal.ro)

## Contents:

MySQL databases as part of the Online Business, using a platform based on Linux .....	3
Ion-Sorin STROE.....	3
Grid Computing Technology .....	13
Georgiana MARIN.....	13
Business Process Management Solutions Performance Tuning and Configuration.....	23
Alexandra FLOREA .....	23
Administration Interface in an e-Testing Software Platform.....	33
Diana BUTUCEA .....	33
Data Acquisition and Storage in Engineering Systems .....	43
Cezar Liviu CERVINSCHI.....	43
The development of an electronic business based on the MySQL technology .....	53
Andreea IONESCU, Andreea ISTOC.....	53

## MySQL databases as part of the Online Business, using a platform based on Linux

**Ion-Sorin STROE**

Romanian Academy of Economic Studies  
Romana Sq, no 6, 1<sup>st</sup> district, Bucharest, Romania  
Email: soso@soso.ro

*The Internet is a business development environment that has major advantages over traditional environment. From a financial standpoint, the initial investment is much reduced and, as yield, the chances of success are considerably higher. Developing an online business also depends on the manager's ability to use the best solutions, sustainable on a long term. The current trend is to decrease the costs for the technical platform by adopting open-source license products. Such platform is based on a Linux operating system and a database system based on MySQL product. This article aims to answer two basic questions: "A platform based on Linux and MySQL can handle the demands of an online business?" and "Adopting such a solution has the effect of increasing profitability?"*

**Keywords:** Online Business, Profitability, Database Systems, Technical Platform, Linux, Oracle, MySQL, Internet

### **1** Introduction

The goal of any physical or legal entity is to be efficient, meaning long-term profitability. Profitability is the most important defining element; there is no economical reason for a business to continue its activity as long as the costs are constantly higher than the income. This main feature is not exclusive, meaning that one must take into consideration finding a competitive model that follows all the rules and eventually anticipates them.

Growth and maintaining profitability is achieved through continuous optimization of the factors that make up the business. Thus, the online component is used to increase performance, whether by applying it at the same time with a classical business, whether applied as a standalone item.

Comparing two businesses, a classical one and an online one, the second may have a higher success rate [1], due to how it can expand: a large number of potential clients readily available at low cost, high financial return which leads to performances for clients and

automatically in favor of business.

Among the most profitable companies in the world top, compiled by Fortune Magazine in 2010 [2], *Google* (activity carried out exclusively online) is ranked 19. This shows that the online environment is ideal for growing business. Moreover, there is an international support to this: preoccupation in the field (there is a strong tendency to support online environment by all major international companies), the legislative stimulating base, the context for research, special funds (including European funds).

The element with a strong impact in business stability is the cost given by business needs: human needs, technical needs, current expenses, miscellaneous costs, etc. When we speak of optimization the purpose is reducing costs for sub-assemblies, meaning reducing the total cost.

The optimization can only occur if actual estate of the online business is known, meaning if an internal audit that addresses to system classification and finding the vulnerabilities is realized, such as those nonperforming elements that can be

improved. Knowledge will be dynamic because the business has dynamic elements that require repetitive analysis and cycles of knowledge in a certain period of time. One example refers to the hardware part; the processing requirements should be known (the traffic that is recorded on the site) in order to meet them immediately, without users noticing a decrease in the quality of the provided service.

This document wants to provide ways to reduce the cost of the technical elements in an online business, in this case the ones with a database system by adopting qualitative, efficient, secure and grounded solutions. Thus, it is proposed a technical solution based on a Linux operating system and MySQL database. This proposal aims to increase profitability by reducing the costs with the technical platform.

## 2. Technical Platform

The platform is basic for an online business, integrating all those elements that participate directly or indirectly in this activity. One of these factors is the technical part and includes hardware and software ensembles with which the business is visible on the Internet. Thus, the technical platform has the physical server, the operating system (together with adjacent elements) and the main software product – the site itself.

### 2.1. Hardware Platform

The visibility on the Internet of the software product, also called web site, is provided by specialized computer - servers, which are designed to offer bi-directional support between the client and the site.

In qualitative terms, meaning physical location of the hardware platform in a specialized environment, there are the following solutions:

- Shared hosting: a web server hosting a number of sites; the processing

resources are limited and shared. This solution is for online businesses that have minimal sites, the main advantage being the cost. The major disadvantage is that the processing resources are limited and used in common. So, there is the possibility that a particular site uses all available resources, the other ones becoming inaccessible.

One of the biggest companies that offer such a service is *HostGator* (<http://www.hostgator.com>) and according to what they produce, customers have unlimited storage space and bandwidth at a modest monthly price. At first analysis, the offer is attractive, but if studied, it you will be noticed that there are limitations such as the number of files (existing nodes) - even if unlimited data is stored, dividing them into documents should not exceed maximum number of nodes allowed (example: 10 files can be stored, each having 200GB, but not 1000 files that cumulatively do not exceed 1GB ); processing resources - an insignificant percentage of available memory can be used and the processor cannot be used over a short period of time. If those limitations are exceeded, the company reserves the right to suspend the hosted account.

- Hosting on a Virtual Personal Server (VirtualPrivateServer-VPS): the server resources are shared, each customer has processing capabilities provided with the possibility of using all available resources, if not used by other customers. The advantage of this type of hosting is that it ensures the minimum necessary for each entity, the solution being used by low complexity sites.

As with shared hosting solutions, some limitations appear, in terms of processing data that is related to Central Processing Unit (CPU) usage. If on the disk space and memory there is a well-defined

share, the processor can cause problems because a minimum insufficient resource is provided and if there are additional requirements, those will be satisfied only if the processor is not used by other customers - chances this to happen are low. This is a transitional solution to the Dedicated Server or Cloud Computing hosting.

- Hosting on a Dedicated Server: a dedicated server is intended only for a single client that can fully use all its capabilities. The disadvantage is the high price that the customer must pay. Also, what is related to the management server comes in the customer's duty: installing and configuring the operating system, the software products, startup, maintenance and updates. It should be noticed that this service is not a scalable hosting, hardware sizing cannot be dynamically made and, if done, may endanger the integrity and availability of the stored information on the website.

- Cloud Computing Hosting: this version is the best choice in terms of price / resource allocation. In Cloud Computing the needed resources are established by the client.

The VexxHost company (<http://www.vexxhost.com>) has implemented a system for this matter (see Figure 1.), which enables customers to increase processing and storage capabilities, as needed.

Thus, if at some point the website is accessed by a large number of users, processing capacities are dynamically resized depending on the requirements. As an example, if the online business management knows the state of the internal processes, it will propose increasing the hardware capabilities in the holiday months, knowing that the offered services will be of interest to much more clients, dramatically increasing the web traffic.

The advantage of this solution is given by the immediate access to additional resources and the cost that modifies according to the use of the hardware components.

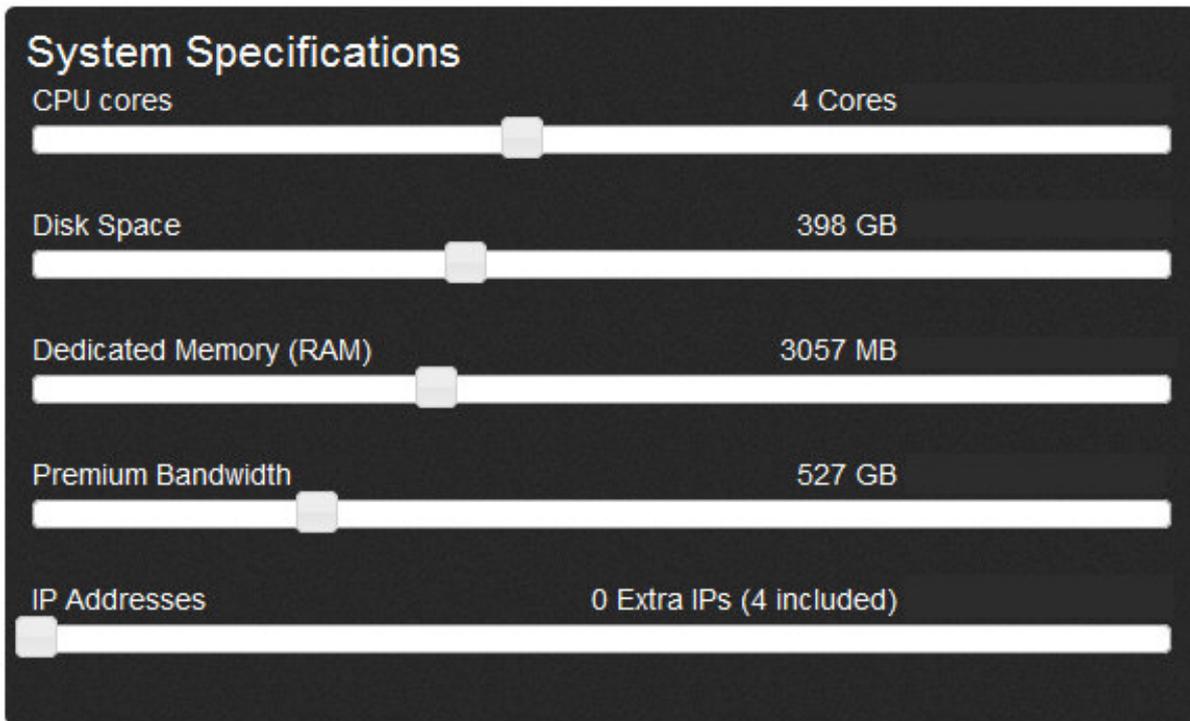


Figure 1. Establishing technical resources in a Cloud Computing system (vexxhost.com)

The Cloud Computing solution brings major advantages in terms of databases, such as: *“Improved availability: If there is a fault in one database system, will only affect one fragment of the information, not the entire database.”* [3] and *“Improved performance: Data is located near the site with the greatest demand and the database systems are parallelized, which allow the load to be balanced among the servers”* [3].

### 2.2 Software Platform

The basic component in the software platform is the operating system and its capacity to manage hardware and software sub-modules resources related to the server.

Most operating systems on the market can be installed on a Web server, but there are two specialized products:

- *Microsoft Server*: developed by Microsoft; is a product based on a license.
- Linux based operating systems, free, such as *Ubuntu Server*, *Fedora*, *CentOS* or license based

like *Red Hat Enterprise Linux*.

The proposed operating system within the platform is the *Community Enterprise Operating System* (CentOS) and it is using Red Hat Enterprise Linux product code, published by the Red Hat company, licensed under General Public License (GPL). The differentiation is made by the identification (logos), the legal references for clients support and the licensing costs that are eliminated.

CentOS is a highly stable solution, being sustained by a large number of volunteer programmers and testers that contribute on maintaining and constantly improving the product.

CentOS is the best solution for the online environment servers, which is based on Red Hat Enterprise Linux product - used by large companies, such as *London Exchange Stock*. What makes a server become a Web server is the Internet access and the suite of software installed. The main modules are:

- Web server such as Apache, IIS (Internet Information Services) or Apache Tomcat. Its role is to

mediate (see Figure 2) the connection between server and client. Thus, the web server takes the request from the user, redirects it to which software product it is addressed to, expects the outcome of the demand and then provides it to the client. Web

server is a node and deals with management input and output flows. This sub-component aggregates the subsystems and subservers, without the user noticing it, his feeling being that there is a unique system that communicates with him.

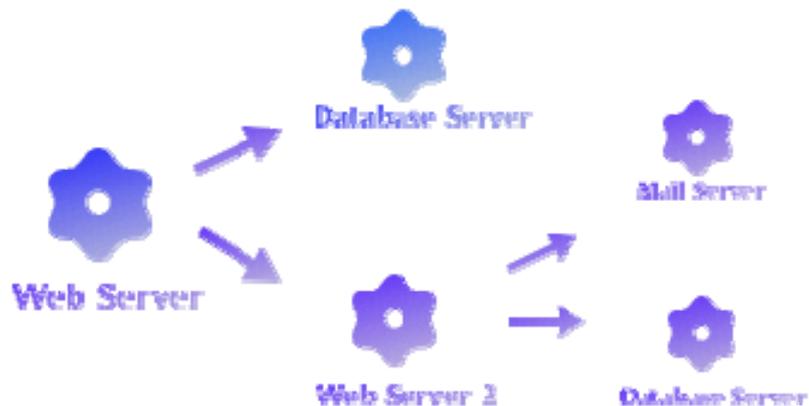


Figure 2. The Web Server as an interface

- Database server such as Oracle, MySQL Oracle, SQL Server, PostgreSQL or IBM DB2.  
This server handles the storage and data management, the performance being imperative. The database system must respond quickly to requests, manage and save data in an efficient manner, meaning that the interrogation should be done in an acceptable timeframe.
- Language Processor: PHP (Hypertext Preprocessor), Java, asp.NET, C #, etc. The supported programming language on the server has direct implication on other software products, by existing or not a committed working relationship. As an example, C #, as a Microsoft. Net component is optimized to work with Microsoft's Web server - Internet Information Services (IIS), as effectively as Hypertext Preprocessor (PHP) is working

with Apache web server.

### 3. Database System

There are many solutions for storing information using databases, but what should be taken into consideration is the optimal choice, to help the development of the owned online business.

The most widely used database systems are Oracle, Microsoft SQL Server, Oracle, MySQL and IBM IDB2.

The proposal of using a product that contains MySQL platform was motivated by the fact that this database system has increasingly become used in production because it offers advantages over the competition:

- There are no licensing costs.
- The product is open-source so, the community involvement makes the development progress to be steady.
- The functionalities are similar to the other database systems.
- Specialized companies (hosting, software selling companies, etc) provide default support for the

system database.

- It is not conditioned by a particular operating system. For example, SQL Server can only run on a Microsoft operating system.

MySQL can run on more than 20 operating systems [4], such as Windows (including Windows Server) or Linux

based systems. Performance, scalability and security system has made the greatest sites (see), in terms of popularity and traffic stored in the database, to adopt this solution. Among customers who use MySQL as a database system included in the individual platforms can be mentioned FaceBook.com, NetLog.com or Big Fish Games (see Figure 3).



Figure 3. Different customers using MySQL

3.1 MySQL based solution

Qualitatively, the difference between MySQL and the most important database systems is minimal, meaning functionality and performance are similar, with few exceptions.

The main advantage is the cost; MySQL is distributed with a GNU General Public License (GPL), meaning no running costs.

Thus, choosing a platform based on Linux operating system and having a MySQL database, the profitability rate will increase.

Profit can be calculated with the following formula:

$$\text{Profit} = \text{Income} - \text{Cost}$$

In the *initial situation*, in which it is used a database with non-free license, cost is composed of several sub costs (see Figure 4).

$$\text{InitialCost} = \text{SoftwarePlatformCost} + \text{HardwarePlatformCost} + \dots + \text{OtherCost}$$

$$\text{SoftwarePlatformCost} = \text{OperatingSystemLicenseCost} + \text{LicenseCostDataBaseSystem} + \text{MiscellaneousCost}$$

Figure 4. Initial cost structure

After adopting a platform based on Linux operating system and a MySQL database,

freely available solutions, the cost structure is illustrated in Figure 5.

$$\text{FinalCost} = \text{InitialCost} - \text{OperatingSystemLicenseCost} - \text{LicenseCostDataBaseSystem}$$

Figure 5. The structure of the final cost

It can be observed that the total costs decrease by eliminating licensing costs with the operating system and database system. It should be mentioned that this decrease of the costs will bring long-term profit growth.

Another advantage of using MySQL is the online orientation; this is why it is one of the most used database systems, having native integration in software related systems, such as PHP processor.

### 3.2. MySQL Administration

Database management can be done with open source software. One such product is VirtualMin, for management applications (including database) installed on the Linux operating system (web server, mail server, operating system, system firewall, etc).

Installing and managing a database can be made:

- From the command line of the operating system [5], requiring technical skills.
- Using specialized software that has graphical user interface such

as Webmin, not requiring specific knowledge.

The installation is done by executing the VirtualMin script that handles itself the whole installation process. From the command line of the operating system, it should run the commands related to file downloading and running it. As an example, for a CentOS operating system, here are the steps:

- Downloading file: `wget - http://software.virtualmin.com/gpl/scripts/install.sh`
- Granting running rights : `chmod + x install.sh`
- Script installation : `./ install.sh`
- After executing the script, VirtualMin can be accessed on the server's address, port 10000, using a Web browser and having available a graphical interface.

From the admin panel, different modules can be installed, including MySQL. After installation, VirtualMin offers a submenu with options and information for managing the system database (see Figure 6).



Figure 6. Options for managing the database, offered by VirtualMin

Using VirtualMin can manage databases, tables, fields, settings related to running the MySQL server, different settings, without using a graphical interface that would involve knowing some specific

commands.

## 4. Putting into operation

The proposed platform consists of:

- Cloud Computing hosting solution

- CentOS operating system
- Apache Web Server
- MySQL database server
- VirtualMin and WeMin administration system
- Other adjacent products

Putting it into service requires installation, configuration and optimization of all used software.

*Steps:*

1. Installation Server. The company used as a hosting Cloud Computing provider is VexxHost. After registering a new account (by accessing the <http://cloud.vexxhost.com> website) a new server configuration application can be sent. The request will include details of chosen operating system (in this case CentOS) and related hardware data: sizes storage space, estimated monthly traffic, the number of processor cores, memory size,

number of additional IPs, etc. By sending the form, results computing the payment for the provided service, and after payment, building physical server to a node in the system.

2. Completing the server installation requires knowing the Internet Protocols (IPs) which identifies it on the Internet. Thus, accessing one of this IPs, makes possible managing the server remotely. Set up will be conducted through a Secure Shell (SSH) connection. The server management is made exclusively from the command line, graphical interface missing for reasons like lack of used resources.
3. The access to the server is done by using a utility that has implemented the SSH protocol (Putty user in this case - see Figure 7); the necessary data is: IP Address, admin username and password (data provided after installing the server).

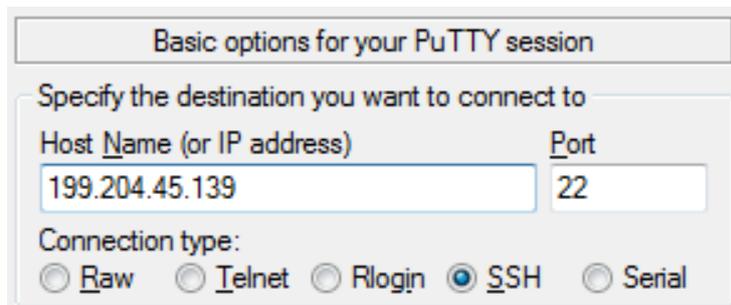


Figure 7. Putty New Connection

4. The next step is to install a software product that allows the server management to be realized using a graphical interface. This product is Webmin, containing VirtualMin module; the installation it described in 3.2 chapter. After installation, it can be accessed at the URL: <http://IPServer:10000> using the admin name and password.
5. Webmin and VirtualMin products simplify server management, the user using the Graphical user interface (GUI), so, there is no need for knowledge of specific Linux commands (except those related to the software installation).

## 5. Conclusions

Using open-source technologies will result in direct cost reduction, thus increasing profits. There is no general rule in this regard; such solutions should be analyzed, so as to see if in terms of quality are at the competition.

The platform based on Linux operating system and MySQL database is the optimum choice on helping develop a sustainable online business. This option can be used even if a shift is needed from the old solution, because long-term benefits are major. Besides the cost aspect, MySQL offers outstanding performance, so it can be said that there is no compromise.

The proposed platform is based on software products tested in specific conditions. The web server hosting the print service offered through the site must have certain performance features, this shown by the users experience on the site, meaning the degree of satisfaction. If a customer accesses the site and encounters technical issues, the chances that he quit the service increase.

The website quality (referring to the programming mode, the design and functionality) should be supported by the use of a suitable platform.

Using the presented platform brings a substantial cost reduction and increases business profitability through performance

brought as support service.

## 6. Acknowledgments

*This work was cofinanced from the European Social Fund through Sectoral Operational Programme Human Resources Development 2007-2013, project number POSDRU/107/1.5/S/77213 „Ph.D. for a career in interdisciplinary economic research at the European standards”*

## 7. References

- [1] M. Borges, N. Hoppen, and F. B. Luce, "Information Technology Impact on Market Orientation in E-business," *Journal of Business Research*, vol. 62, iss. 9, 2009.
- [2] F. Magazine. (2011, August 05) 2010 Top Companies: Most Profitable. [Online]. Available: <http://money.cnn.com/magazines/fortune/fortune500/2010/performers/companies/profits/>
- [3] T. Velte, A. Velte, and R. Elsenpeter, *Cloud Computing: a Practical Approach*, 1 ed., New York: McGraw-Hill Osborne Media, 2010
- [4] MySQL. (2011, August 05) MySQL :: Why MySQL? [Online]. Available: <http://mysql.com/why-mysql>
- [5] C. Bell, M. Kindahl, and L. Thalmann, *MySQL High Availability: [tools for Building Robust Data Centers]*, 1 ed., New York: O'Reilly Media, 2010



**Ion-Sorin STROE** graduated the faculty and master organized by the Faculty of Cybernetics, Statistics and Economic Informatics of the Academy of Economic Studies, with graduation and dissertation topic-themed works in the electronic business field. He is currently a Phd candidate in the same area, Economic Informatics.

From the environmental point of view he presents Word Wide Web interest in new technologies and platforms, both conceptually, software and hardware too, developing over the time many complex and profitable web projects that were highly appreciated.

## Grid Computing Technology

**Georgiana MARIN**

Romanian – American University Bucharest, ROMANIA

[georgiana.maryn@gmail.com](mailto:georgiana.maryn@gmail.com)

*This paper presents the grid computing technology, the recent developments in this field. The idea of grid computing has its origins in the early development of computer networks and technologies involved, given that CPU cycle use "free" or "unused" was seen as an optimal and cost-effective way to use all capabilities of the hardware resources which were very expensive machinery.*

*Keywords: grid computing, technology, virtualization, data base, file system*

### **1 Introduction**

The term grid is referring to a distributed computing infrastructure, able to provide resources based on the needs of each client. Grid technology can largely enhance productivity and efficiency of virtual organizations, which must face the challenges by optimizing processes and resources and by sharing their networking and collaboration. Grid computing technology is a set of techniques and methods applied for the coordinated use of multiple servers. These servers are specialized and works as a single, logic integrated system.

The grid has developed as a computing technology that connects machines and resources geographically dispersed in order to create a virtual supercomputer. A virtual system like this is perceived like it has all the computing resources, even if they are distributed and has a computing capacity to execute tasks that different machines cannot execute individually.

In the past few years grid computing is defined as a technology that allows strengthening, accessing and managing IT resources in a distributed computing environment. Being an advanced distributed technology, grid computing brings into a single system: servers, databases and applications, using a specialized software. In terms of partnership between organizations, grid

technology may include the same enterprise organizations as well as external organizations. Therefore, grids may involve both internal and external partners, as well as only internal ones. [1]

The complexity of the environment in which the grid will be developed and the requirements that have to be applied depend on the environmental impact of trade, defining the relationship of trust, security considerations, globalization and integration period, of the company involved, in the market. Based on the different levels of complexity for the enterprise, grids can be categorized as follows:

- **Infra-Grid** – this type of grid architecture allows optimizing the resource sharing within a division of the organization's departments. Infra-grid forms a tightly controlled environment with well defined business policies, integration and security.
- **Intra-Grid** – it's a more complex implementation than the previous because it's focused on integrating various resources of several departments and divisions of an enterprise. These types of grids require a complex security policies and sharing resources and data. However, because the resources are found in the same enterprise, the

focus is on the technical implementation of the policies.

- Extra-Grid – unlike intra-grid, this type of grid is referring to resource sharing to / from a foreign partner towards certain relationships are established. This grids extend over the administrative management of local resources of an enterprise and therefore mutual convention on managing the access to resources are necessary.
- Inter-Grid – this kind of grid computing technology enables sharing and storage resources and data using the Web and enabling the collaborations between various companies and organizations. The complexity of the grid comes from the special requirements of service levels, security and integration. This type of grid involves most of the mechanism found in the three previous types of grid.

Grid computing technology is not the only distributed technology; among other distributed technologies we can include: web technology, peer-to-peer, clustering and virtualization.

The similarities between the grid computing technology and the others distributed technologies are:

- A grid hides its complexity, interconnects different resources and provides a unified perception of the environment;
- Can share files and communicate directly through a central broker;
- Clusters and grids groups the resources in order to solve a problem and can have a tool for unified management of all components [2].

The differences between the grid computing technology and the others distributed technologies are:

- A grid allows machines to work together and collaborate in a way that involves more than simple communication.

- Grids can use the relations many to many to share a wide variety of resources, not just files.
- A cluster is made of computers of the same type located in a central location requiring a single system image – SSI. Unlike the cluster, where the emphasis is on performance of parallel processing, in a grid is on the resource sharing. SSI is not anymore a requirement, computers from grid are heterogeneous and geographically distributed.
- Unlike virtualization, which refers to a single system, a grid allows virtualization of heterogeneous resources everywhere, even global, to form a pool of IT services.

## 2. Grid computing requirements

The requirements of grid computing technology are: security, fault tolerance, global name spaces, scalability, adapting heterogeneity, persistence and non persistence, extensibility, complexity management and autonomy.

- Security - addresses a wide range of issues, such as: authentication, data integrity, access control and updates. Grid system and technologies, in order to earn a place on the corporations market and their use in important IT departments, is necessary the grid to be focused on the security issues. The security mechanism must allow the integration of applications with any degree of diversity and to support various security policies, adopted in each particular case of entities that integrate and implement a system like this. To ensure flexibility, security must be integrated in the environment of the grid from the beginning. Future changes of the system may lead to amendments to security needs so that the existing necessary security mechanism elements can avoid the need to

update and modify essential parts of the system and applications. The security mechanism must be sufficiently abstract to allow the definition and integration of any security policy. These conditions being given, the grid system must provide mechanisms that allow users and resources owners to use that policy that adapts to the requirements of security and performance and to local administrative requirements.

- Fault tolerance - Binding the developers to predict and deal with all the faults in a system leads to a significant increase of the application complexity. For this reason, grid computing would be provided with mechanisms for fault tolerant computing.
- Global name spaces - The absence of a global namespace that allows a convenient and consistent access to data and resources is a general issue in the fields that involves parallel and distributed calculations. Thus, one of the tendencies of grid computing is that each grid object must be able to access the grid in a transparent way without harming the data integrity and resources.
- Scalability - Grids adopt an architecture based on a review of non centralized resources. So, the architecture is based on the principle of distributed system: the service required by any component is independent from the number of system components.
- Adapting heterogeneity - A grid system must support interoperability between resources and heterogeneous hardware and software systems.
- Persistence and non persistence - One of the most important aspects of resource management is the aggregation in an advanced mechanism of detection algorithms and resource planning. This mechanism should ensure a high level of performance in order to respond to the high demands regarding service quality. In this regard, the resource management should firstly address to aspects concerning on the resources persistence and non persistence, because the environment proposed by the grid technology involves multiple administrative domains, a variety of heterogeneous resources whose behaviour in unpredictable even in the short term.
- Extensibility - means grid system flexibility so that they can meet the current and future requirements of users and also the ones that cannot be anticipated.
- Complexity management - Complexity is common in developing applications in a space with more dimensions, dimensions generated by: heterogeneity in using the resources and security policies, a wide variety of malfunctions, the requirements on the availability of resources, name spaces and the number of components. In this matter of abstraction that grid computing technology provides, must be sufficiently high to reduce the size of the units of measurements and allows them a efficient complexity administration.
- Autonomy - Grid systems may include resources from different organization involved in the distributed environment. Each resource owner wants to keep control over its own resources and therefore the grid system should give the opportunity to vary the level of access given to users, to specify when and how a resource can be used and also the possibility that each owner to organize their own resources according to needs, without relying on the grid system [3].

### 3. Advantages and disadvantages of Grid Computing Technology

Although grid computing has been developed in research laboratories, manufactures and companies have also started to adopt this technology on two of the most important benefits that grids brings: economy and performance. Grid computing aims to bring together operating systems and different hardware platforms into a single virtualized entity whose performances is higher on average then parts.

Grid computing saves financial resources both in capital and operating costs. This positive aspect is achieved by using all the computing resources of all components of the grid.

The second benefit of grid computing technology is performance in processing data integrity. By increasing the processing power, applications run faster the computing tasks and provide faster results. Even if the benefits of grids are real, this is still a specialized technology. Grid computing is suitable for organizations that already use in a way or another the high performance computing or are already oriented in some form towards distributed computing. Although there isn't a direct barrier to the passage of an organization to use a grid system, there are still issues to be taken into consideration before such a step. One of the reasons this happens is determined by the applications, which must provide the possibility of breakdown into smaller computing tasks in order to take advantage of parallel computing. If the operation currently executed depends on the previous tasks, then processing cannot be done in parallel on distributed resources so that the application cannot benefit from the grid advantages in terms of high performance computing.

Another issue that emerged over the time is the one that occurs in administration and accounting resources. This problem is more obvious as a solution involving a grid system involves multiple departments of

the same organizations and each department require proportionality between the grids contribution, through its own resources, and the gain obtained from it.

Besides resources accounting there's also a problem with software licensing. This is particularly important because licensing costs may eventually cancel the savings obtained from the use of grids.

Securing the servers and grid administration resource is an issue that should not be neglected. In many cases servers are public addressable because of data delivering and receiving from agents that are anywhere, geographically speaking. So, these servers are exposed to attacks and unauthorized access, even DoS – Deny of Service. Experts recommends that, in this case, the companies must eliminate all unnecessary services and to carefully monitor these machines. However, this brings an extra complexity problem.

One of the benefits of grid environments is that it allows the development of specific and on demand environments for various commercial environments. An important feature of on demand environments is the ability to respond to rapid changes in the market while reducing operational costs.

Grid computing helps distributing and sharing data, which enables a collaborative improvement at both inside and outside the enterprise. Thus, companies can reduce time spent placing products on the market, can quickly solve specific problems and can address immediately to the customers' requests. One of the troubles encountered when distributing the workload over several machines is the difficulty to trace the distribution process, which can lead to bandwidth problems. Is not only the bandwidth performance criterion but also the performances of the grid, of the way tasks are distributed and how they handle the distribution server. Bandwidth management is a serious problem that must be taken into account. A grid network bandwidth covers two aspects of any grid:

- CPU utilization. CPU bandwidth in a grid represents the maximum rate at which the grid can operate.
- Using the network. Available bandwidth of the grid and the one used internally by the grid.

These two elements affects a number of different systems and it's important to balance these elements in order to avoid the risk of having an inefficient grid, a grid that is unable to manage the computing tasks, or customers to start overburden the resources that they have at their disposal.

Although there are several technologies that can be integrated into the grid or pattern which can form the basis of the mechanism of resource management and workload distribution and planning, however, they must be adapted to the needs of grids. We cannot avoid the artificial intelligence solutions to solve these type of problems, but must consider the complexity of the system and the additional resources that they are bringing. Categories of grid computing fits into several areas of research communities. These may include areas oriented in intensive computational calculus, peer to peer, utility, data and applications and collaborations. Each of these communities can use a different model on the adoption of grid technology.

Grid computing is a new concept for commercial industry and a large area of interest has developed around infrastructure virtualization through the manipulation of resources as utilities.

Adopting a grid solution in a commercial activity depends on the ability of technology to meet the needs of improving turnover. This involves the adoption of the grid models key factors, such as adaptation of existing resources, reducing operational costs, creating a flexible and scalable infrastructure, while accelerating development, reducing the period of development and marketing of products and increase customer satisfaction and business productivity.

An important issue in grid adoption models is the complexity of IT infrastructure needed to implement a grid system. The integration complexity of heterogeneous environments is a challenge and it must be taken into account factors such as activation of grid resources in homogeneous and heterogeneous environments, enabling resources in the form of services to external participants and porting applications to grid applications. These features allow a classification of grids into the following categories, with varying degrees of complexity at the level of integration:

- Grids developed to optimize computing infrastructure;
- Computing grids, with the virtualization of processing resources.
- Data grids with virtualization and data storage resources.
- Service grids with virtualized services for easy integration.
- Virtualized applications through the composition of resources, through service interfaces, applications from various partners.

Architecture and technology standards developed for grid computing presents a crucial role in adopting grids commercially. Because these standards are constantly evolving and are not mature enough to support subsequent stages, which can evolve into after the adoption of grids, the speed at which these stages will be achieved is reduced. Grid computing solutions have been adopted in several key areas (finance, education, telecommunication, research) and answer the implementation and evolution of environmental performance requirements that are integrated. From this we can conclude that the success of grid computing depends on the integration, service orientation and ability to break down skills and applications and then expose them through the services.

Although grid computing is an important step in information technology it does not

involve the disappearance of existing resources and advanced equipment such as supercomputers, mainframes, clusters etc. The grid has no purpose in replacing the current technology resources, performance management, but the efficiency in extracting optimum benefits from investments made in such resources. There are still many areas (particle physics, simulations, computational intensity processing) in which execution of tasks involves only supercomputer (or dedicated resources in general performance), only areas where the grid cannot cope, or rather does not provide the necessary desired performance in such cases.

Growth of technology and applications on this area, in the last period, may lead to a future in areas indispensability of enterprise grid computing. Such a critique must always be approached in the case of adoption of this technology to enable a more coherent adaptation to the needs for which a solution is desired grid.

#### **4. Technologies used in computational grids**

Performance criteria in a grid environment are dictated by several of the most important aspects are planning mechanisms and the mechanisms of load balanced and integrating resources and data necessary for their calculations on the same node. The choice in favour of another technology to support the role of a computational grid environment is a problem of decision that can be solved only after a careful analysis on a number of factors who are determining the particular environment in which to apply the solution and the specific requirements of grid.

Heterogeneity. Any facility planning software that is used in the grid should be able to support heterogeneous computing nodes. Even if the current implementation does not need them, yet this method is a useful strategy for the future expansion of the system. Thus the nodes of a service reservoir have to bear the same service or

set of services irrespective of the choice or implementation of their type.

Security. Mechanisms planning software must support the required levels of security specified in the particular needs of any environment that will be integrated grid system. However, many implementations of grid are located in the same administrative domain where security requirements are less stringent than in distributed environments that cover more organizations. Under these conditions many grids shows a low outside the additional security provided by the operating system of nodes, since all the nodes are behind the same firewall as well protected as any system in the network company. Thus a compromise must be made because the existence of an additional security level without a decrease in performance (speed of the executable is directly affected), but leads to a secure data protection.

Reliability. Since the planning provides some of the most important performance, the mechanisms must ensure the reliability levels needed to resolve defects. And in this case the solutions on this issue are not unique and their usefulness varies depending on the environment in which a solution will be integrated grid. Such technology should be appropriate and adapted to the user.

Thus, in some environments the emphasis is on completion of the total load of calculation, so that some form of recovery of computing tasks is sufficient. If a sub-task of calculation fails due to a defect node, network or software, then that burden is being passed back to another appropriate node. This mechanism provides some fault tolerance but without any warranty and / or improve the response time of the task, is known that it will grow with each retransmission of sub-tasks.

On the other hand, in other environment the response time is the most important criterion and the planner should be able to ensure the return of preserving the response but decreased during execution.

Such solutions usually involve a launch of multiple copies of sub-tasks to different nodes in grid computing. This technology is suitable for planning grid environments are those that present a high flexibility towards upgrades and additional features that ensure the fundamental planning.

Data Virtualization is an architectural approach that hides the location and data format for the user. Level virtualization architecture that addresses the data shows a direct proportion to the level of complexity with the complexity of the level of computing resources. Data handled in a grid environment can be structured or unstructured, ordinary files and other objects stored in file systems, also in case of a grid covering a development involving several organizations, external data sources can cover a wide large format and can be publicly accessible, available on a contract basis, can be shared between companies working together in a consortium or assumptions to be shared in an academic or business environment based on collaboration in a joint project, some data may support active interrogation while others do not. In these conditions of an environment that involves a high degree of dynamism so the question of criteria for a decision on the choice of technologies to fill the grid environment. The answer is not unique and the selection is based on a number of important factors and their relationship shows a coefficient related to the particular specifications of each environment that integrates a grid system. Thus Cluster file systems or distributed network can integrate successfully into the grid system. Also aggregation and data replication may be needed in certain grids. Cluster file systems is particularly suitable for constraints within a single administrative domain because they are usually sensitive to the degree of separation and differentiation between the cluster nodes, therefore cluster systems are less applicable in distributed environments that cover more administrative areas. In the same way file systems are network-

oriented environments such as logos or single-site campus.

Both solutions use the advantages of technological advances achieved in computer networks. Distributed file systems have been developed following several administrative areas covering a wide geographic dispersion. However only single file systems do not provide sufficient data management mechanisms, especially where large volumes of data are stored in databases or other structured storage resources. Thus we have an important role of federal databases and aggregates.

Aggregated databases are trying to achieve consolidation of all data formats from disparate systems into a single parallel database. Although this model provides scalability it is also oriented in the implementation, covering a secure administrative area.

Federative data bases are trying to keep track of data distribution in different areas but to provide unified interfaces that hide implementation and data source to user. A federative database provides a greater level of flexibility as opposed to aggregated databases and a higher potential performance problems when characteristics of remote data to the user. So, another mechanism may be added: data replication. Another aspect to be taken into account in choosing the virtualization of data to support the heterogeneity of data sources both in terms of platform (providing access and data) and in terms of technologies or techniques used for data storage.

Linking data and calculations is a fundamental problem and refers to their alignment on the same node in the grid to allow for optimal performance. Myriad of available technologies and products that aim to solve this problem shows a touchstone: the lack of collaborative approach based on open standards (this methodology to be approached from a considerable number of organizations in the business environment in recent years only). Thus, each technology presents a

remarkable intelligence, but within the limits of their own views on the environment and the grid components that involve, or it can integrate them in time. In this way each technology manifests its decisions regarding the settlement calculations with data alignment in a manner isolated from the others. In conditions of low complexity and where planning decisions are obvious mechanism for calculating the load (the entity charged with moving the appropriate entities for calculating the node) and the mechanism for moving data (the entity charged with moving the data entities) may relate so that both data and calculation takes place on the same node. In terms of complexity of a grid and especially the expansion in complexity over time, especially regarding the integration of geographically distributed resources with grades of spreading increasingly larger planning mechanisms linking computing tasks with mechanisms data handling is becoming more difficult because the decision conflicts are becoming more frequent. Alignment of data and calculations on the same node is a strategy that further studies should be conducted because it allows obtaining optimal performance in grid environments and systems that exceed the current average.

### **5. Principles of development of a Grid System Grid, applications and requirements**

Grid systems must ensure the transparency of the following ways: access, location, heterogeneity, failure, replication, scalability, concurrency and behaviour. Users and developers should not know exactly where an entity is located to interact with it and should not be forced to acknowledge the failure of components / resources of the grid system. This grid system must be equipped with self-recovery mechanisms in order to accomplish tasks. This represents the mask (transparency) on the vision system.

Grid architecture must be adaptable to as many solutions on the grid user needs and goals. A rigid system in which the known policies is limited, decision are pre-selected transaction cannot meet these demands. This grid system should allow greater flexibility for both semantic applications without imposing predetermined solutions to unique methods of implementation. Grid systems must also allow users to choose the desired type and level of functionality and appeal to their own compromise between functionality, cost and performance. This philosophy should be integrated into the grid by specifying the architecture and functionality but not implementation of core aspects of the system. The core should therefore be made of extensible components, which can be updated and replaced whenever possible / necessary while implementing the system should provide an implicit development of each functional component useful to normal users.

In general one can speak in a grid environment for the existence of four categories of users: application users, application developers, system administrators and managers. Grid system architecture and implementation must be realized that it allows users to applications and their developers is focused on their task and not involve their direct work on issues related to installation and implementation of grid infrastructure, but can also be provided and ensuring access to it the grid infrastructure and related details if required.

Another aspect of flexibility to be ensured to grids is to keep the machine's operating system will be integrated into the system. Such grid systems must be developed so as to work with and above the machine's operating system and involve such changes, on configuration, but fewer in number and complexity.

In developing an application grid security issues require consideration of a

comprehensive approach. Below are treated briefly these considerations:

- Single sign-on. ID Mapping over the system. GSI (Grid Infrastructure Security) provides authentication, authorization and secure communication. Therefore we must fully understand the implications and security management.
- Multiplatform. Although GSI is based on open standard software running on multiple platforms though different platforms security mechanism will not always consistent. For example, a security mechanism for reading, writing and execution on a UNIX system is different from the traditional system of Microsoft Windows. This should be taken into account the different platforms will run the application.
- Use of GSI. For any specific function of the application that may require authentication or a special permit application indicated that the default mechanism to use GSI as it simplifies the development process.
- Encryption. Although GSI together with data management mechanisms ensure secure communications and data encryption over the network still need to consider what happens once the data arrived at the destination. An example in this regard is the special case where data are transmitted to a resource to be processed and are stored on a local disk in non-encrypted format, and then there is the risk that other track users to have access to that data.

Influence of resource management issues on application development:

- Choice of appropriate resources. Choosing appropriate resources an application is made by the resource management mechanism that works in conjunction with a broker. This implies that the application to specify accurately the working environment required (OS, processor speed,

memory, etc.). The application shows how dependent fewer specific platform the greater the chances that an appropriate and available resource to be found and resolved quickly workload.

- Multiple computing sub-tasks. For applications involving multiple computing tasks to analyze the interdependencies between them should be made to avoid additional logic: the communication intercrosses, data sharing and tasks competing administration.
- Managing computing tasks. If the application has to provide a reactive response to the user or to release resources when the application should be designed so as to use the mechanisms of grid resource management to ensure consistency and integrity of an environment.

The mechanisms for data management in a grid to maximize the best use of limited storage space, the network bandwidth and computing resources. The following aspects of data management must be considered in the development of grid applications:

- Dataset size. If the application works with data sets of very large ineffective, if not impossible, for data to be moved on the system that will run and load calculation. In this case one possible solution is data replication (copying a subset of the data set) on the system to execute computing tasks.
- Geographical distribution of users, computing resources, data and storage. If the grid is geographically distributed environment and shows limited network speeds then the application must take into account aspects of design data access limited or low speed.
- WAN data transfer level. As grid environments and distributed networks involve extensive grid any application must take into account

issues of security, reliability and performance for handling data across the Internet or another WAN. Such applications require a logic performance for dealing with changing situations in which access to data may be slow or restricted.

- Planning data transfer. Planning involves data following two main issues: ensuring the transfer of data to the appropriate location and time at which this transfer is required, taking into account the number and size of competing data to / from any resource.

## 6. Conclusions

Spurred the development of both high-speed networks and the increasing computing power of microprocessors, processing grid has a remarkable impact not only at academic level, but increasingly more and enterprise in all fields. Despite technological advances, however, less than 5% of Windows servers processing power and desktops, respectively 15-20% for UNIX servers is used. Companies make profits and offers exceptional rates, with a return period of low and low total cost compared to other technological solutions.

## References

- [1] M. Velicanu, I. Lungu, I. Botha, A. Bara, A. Velicanu, E. Rednic – *Sisteme de baze de date evaluate*, Ed. Ase, Bucuresti 2009
- [2] A. Abbas - *Grid Computing: A Practical Guide to Technology and Applications*, Editura Charles River Media, 2005
- [3] Ian Foster, Carl Kesselman, Steven Tuecke - *The Anatomy of the Grid: Enabling Scalable Virtual Organizations*
- [4] W. Jia, W. Zhou - *Distributed Network Systems. From Concepts to Implementations*, Editura Springer, 2005
- [5] Virginia Center for Grid Research - [http://www.cs.virginia.edu/~jfk3w/grid\\_group/grid\\_overview.html](http://www.cs.virginia.edu/~jfk3w/grid_group/grid_overview.html)
- [6] Andrew Conry-Murray – *Grid Computing's Promises and Perils* <http://www.itutilitypipeline.com/showArticle.jsp?articleID=17603526>
- [7] [http://www1.ibm.com/grid/about\\_grid/ibm\\_grid/get\\_started.html](http://www1.ibm.com/grid/about_grid/ibm_grid/get_started.html) [http://www.cs.virginia.edu/~jfk3w/grid\\_group/grid\\_overview.html](http://www.cs.virginia.edu/~jfk3w/grid_group/grid_overview.html)



**Georgiana MARIN** graduated the Faculty of Computer Science for Business Management of the Romanian American University in 2003. In 2009 she graduated the Economic IT Masters Program of the Romanian American University. She is currently a PhD Candidate in Economic Informatics of the Academy of Economic Studies.

## Business Process Management Solutions Performance Tuning and Configuration

**Alexandra FLOREA**

Economic Informatics and Cybernetics Department  
Academy of Economic Studies  
Bucharest, ROMANIA  
alexandra.florea@ie.ase.ro

*Business Process Management (BPM) solutions provide full integration between business and IT requirements, ensuring that the business requirement adapt more easily to the environmental changes. Given that today's business environment is characterized by ever growing dynamism, the use of BPM solutions can provide additional consistency and opportunities for rapid alignment to change.*

*In this article we discuss some of the standards and technologies for BPM solutions, we present several proposals for the lifecycle of a business process and also the main stages of our own proposal for a methodology for developing a BPM solution. Also we analyze one of the most popular BPM solutions on the market (Web Sphere Business Process Management suite) and present some key tuning and deployment guidelines.*

**Keywords:** *Business process, Business process management, Process lifecycle, Tuning,*

### 1 Introduction

Studying current professional literature related to economics and management, a new principle emerges that "mankind is moving quickly toward a new kind of economy and thus a new kind of company and management"[1], this change being triggered specifically by developments in information technology and communications that led to profound transformations. These transformations mark the transition to a society based on information, communication and knowledge.

In the last decades, profound changes taking place globally reached all areas but mostly the business, so companies are forced to undertake a major reorganization of their management models and methods. The new competitive requirements of organizations require more attention given to customers, lower costs, quality and adaptability and technological developments continue to help organizations achieve these goals

Once organizations became aware of the need for change, they have begun to pay greater attention to key business processes with direct impact on customer profitability and growing importance. To be effective today, organizations must be able to define, analyze, implement, measure and control their processes, and this change is not possible without educating staff and implementing appropriate technologies. Business models continue to change as new technologies are introduced and each functional area of an organization is influenced by how well the information system transforms data into information or facts that help achieve business goals.

Mergers and acquisitions, changing business models, new business requirements and changes in customer expectations, all put multiple problems with processes in the organization. These problems can be solved by integrating Business Process Management (BPM) technologies, which enable organizations to manage changes in processes, changes that are required simultaneously in several areas of business.

Solutions in the Business Process Management category allow integration of very diverse systems that exist in an organization, providing additional advanced features for modeling and automation of business flows in the company. Business Process Management Solutions (BPMS) have visual tools that substantially simplify how complex business processes, that involve access to various data sources and interoperability of many applications, are defined and integrated. Through BPM solutions business processes can be modeled directly by financial and economic analysts without having to request support from the IT departments.

BPM applications provide full integration between business and IT, business requirements being adapted to changes in current business environment in a simpler manner. Given that today's business environment is characterized by growing dynamism, the use of BPM solutions can provide an additional type of consistency and rapid alignment possibilities to the changes that occur.

## **2 Standards and technologies for Business Process Management Solutions**

Integration of applications through BPM solutions is based on fundamental standards and technologies such as web services and XML. Web services represent a standardized way of distributing software, which uses the Internet and basic technology behind this network. Also, Web services enable interconnection of a wide array of applications available on different platforms and in various locations around the globe. In this way, we can say that new technology has opened the door to a new era of computing applications dominated by a high degree of intelligence, capable of making decisions and looking for information on the Internet as a support for most reasonable decisions.

EXtended Markup Language or XML is the basis of all elements which represent the foundation of Web services technology. Considering platform independency, XML is the engine that enables data transfer via the Internet, also constituting the foundation of Web services.

An important element to be mentioned in connection with BPM solutions and Web services is Service Oriented Architecture (SOA). SOA provides access to other applications. BPM uses SOA to include information on the best applications in the process. We can make an analogy saying that if SOA provides avenues for the flow of information then BPM is the machine that uses the infrastructure to obtain something useful.

Sometimes BPM can be a catalyst for developing a SOA strategy. In a world where executives try to obtain a concrete value from the investments they made in the IT field, SOA itself can be quite hard to sell because of the difficulty of explaining the added value it brings using concrete terms and quantifying this value. A strategy that can be used to counteract this shortcoming is to promote SOA as an element that supports business process management, because it has a concrete component and it's much easier to quantify the added value achieved [2].

The most important developed BPM standards are:

- Business Process Modeling Notation (BPMN) which focuses on graphical modeling of business processes. BPMN it's a standard for modeling business processes, proposed and developed by the Business Process Management Initiative and provides a graphical notation, easy to use and understand by all users involved in business processes [3].
- Business Process Execution Language (BPEL) is a standard based on XML and Web services that allows modeling and automation of business flows. Using this language, both the business flows and the business rules

can be defined in an intuitive manner, a high level of transparency in the conduct of these operations being assured. BPEL technology simplifies the integration of various applications and business processes

- Business Process Modeling Language (BPML) is an XML-based meta-language that describes business processes. It was originally intended to support business processes that can be executed by a BPMS [4]. The language was designed to handle processes with a long lifecycle, which persistence is supported in a transparent manner. XML exchanges occur between different participants, each with an established role, and components belonging to their partners, like BPEL constructions. BPML also includes transactional support and exception-handling mechanisms.
- Business Process Query Language (BPQL) - focuses on administrative and monitoring aspects [5]. BPQL is a management interface to a business process management infrastructure that includes a process execution facility (process server) and a process deployment facility (process repository).

A notable aspect is that each of these standards, like many others not mentioned, focuses only on some aspects of the elements that are typically addressed by a BPM product. The challenge will arise when integrating these standards in a life cycle for continuous process improvement through BPM.

### 3 Business process lifecycle.

There are several approaches to the representation of a process' life cycle. Some illustrate the chronology of the life cycle phases of development while others

focus only on the logical dependencies between different phases, considering that each of them can be resumed at different times.

According to [6] there are five stages through which a process passes throughout its life cycle, namely: modeling, implementation, execution, analysis and optimization, illustrated in Figure 1.

In the modeling stage new processes are designed or existing ones are modified according to the new requirements. Activities in this stage are carried out either by analysts or by the processes responsible and result in advanced graphics representations of tasks to be executed.

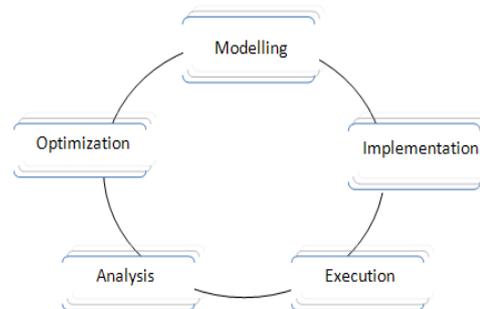


Figure 1. The life cycle of a business process. Adaptation from [6]

The purpose of the implementation phase is to obtain an executable model of the processes, based on existing abstract models. Basically, the theoretical model previously obtained must be adapted so that it can be used in the context of real business environment, taking into account the resources available and actors involved.

The next phase represents the actual implementation of the result of the previous stage (enabling the executable process model). This phase is ideally concluded by obtaining added value for the company and the desired results by the client.

Running processes are subject to careful monitoring and analysis in the phase which bears the same name, in order to identify opportunities for optimization. Once the analysis is completed, the data obtained are used in the optimization and redesign phase in order to obtain a more efficient process

Another approach is the one proposed by [7], which identifies four stages of the business process lifecycle, organized in a cyclical structure illustrating logical dependencies but not the temporal

sequence of their development. The four stages are analysis and design, configuration, execution and evaluation and they are represented in Figure 2.

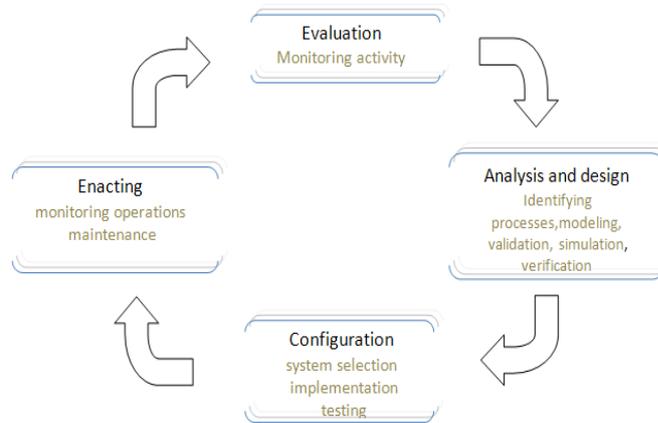


Figure 2. The life cycle of a business process. Adaptation from [7]

The life cycle of the process starts with an analysis and design phase. During this phase an assessment of the organization is made, in order to identify and understand the processes taking place within it and the organizational and technical environment in which they are carried out. Based on these evaluations, the processes are identified, reviewed, validated and represented using graphical tools resulting in business process models.

Once these models are built, the transition to the second stage takes place, namely to the configuration stage. Previously completed models must be implemented, and this can be achieved in two ways. Whether through the adoption of a set of policies and procedures that employees must comply with, in which case the process can be made without the assistance of a computer system, or by using a dedicated software.

In the latter case, the configuration phase involves the following activities: choosing a platform to implement, enriching the process model with technical information necessary for the enactment of the processes by the system and not least system configuration

according to the organizational features of the enterprise environment.

The enacting phase of business processes includes the actual execution of the processes. The process management system actively controls the execution of process instances according to the designed model. In this stage of the lifecycle, the process execution is monitored and it's ensured that activities are conducted according to implementing restrictions specified in the model.

In the evaluation phase the available information is used to assess and improve business process models and their implementation. Execution logs are evaluated using activities monitoring techniques, which aim to identify the quality level of process models and the adequacy of the execution environment.

After studying the presented business process lifecycle as well as other materials regarding this issue, we propose a personal view on the topic.

Under this proposal the life cycle of a business process consists of six steps, represented according to the execution cycle and the logical dependencies between them.

These six stages are:

- Investigation

- Modelling
- Simulation and implementation
- Execution
- Monitoring

- Evaluation and optimization.
- These proposal was developed after identifying all the steps in a BP lifecycle and it groups them according to logical

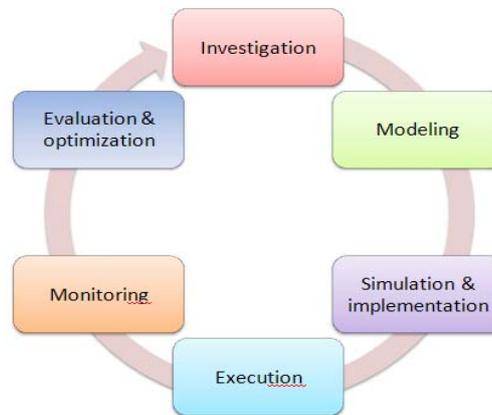


Figure 3. Proposed lifecycle of a business process

corelations. A graphical representation of the proposal is offered in Figure 3.

The business lifecycle begins with the investigation stage. This is the moment when the strategic objectives of the company are correlated with the shorter term business objectives in order to identify and fully understand the scope of the business processes that are enacted within the company. This stage is necessary considering the direct link between the business processes and the strategic objectives of the firm.

After the processes have been identified, we move on to the second stage of the lifecycle which is modeling. Graphical tools are used to represent the processes thus creating the business process model. Once a process has been designed it moves on in the 3rd stage: simulation and implementation. The business model obtained in the previous stage is further developed by adding technical information regarding the environment in which the processes will run. Once the business model has been completed the process enters its 4th lifecycle stage which is execution: enactment of the process in the production environment of the company.

Monitoring - running processes are monitored to measure key performance

indicators and other relevant metrics. The results of monitoring are evaluated either according to predetermined criteria or by using simulation tools to obtain information on how to optimize processes using real data, this being the final stage of evaluation and optimization.

#### 4. Methodology proposal for developing a BMP solution

Based on the described lifecycle, we propose an informal methodology that organizes the phases to be followed when developing a business process management solution.

This methodology is based on the life cycle stages of business processes summarized above, and uses a series of informal notations, its purpose being to provide a guide for those who are responsible for planning and implementing the project management business processes.

The methodology is organized in seven stages:

- Strategic Objectives
- Investigation
- Modelling
- Identification
- Simulation and implementation
- Execution and Monitoring
- Evaluation and optimization.

The first step (or stage) is to identify strategic and operational objectives of the company, which is necessary due to the direct link between the organization's strategy and business processes.

Business processes must work towards achieving the strategic objectives of the organization and as such they must be clearly identified and formulated. Processes contribute to realizing business goals which in turn contribute to meeting the strategies.

Once the organizational objectives have been identified it is necessary to study existing processes and to identify processes that should be implemented (to-be processes). The overall objective of this phase is collecting and organizing information on all aspects of business process management, to ensure that we have a complete overview. This phase was called Investigation.

The third phase is the phase of modeling, in which the theoretical processes model is designed. Information obtained in the previous phase are structured, analyzed and represented as the processes model.

Once we have obtained the processes model, we must realize a selection of the platform on which the model will be implemented. Selected platform can be technical or non-technical. In the first case a variety of platforms might be suitable for the implementation of business processes, including automated

platforms, such as enterprise application integration software or workflow management systems that support human interaction flows. In the second case, the business processes model is realized through various business policies and procedures.

This selection phase is called Identification and it supposes passing through a decision node after the phase was completed. If the results of the stage indicates the need for using a non-technical platform, then we jump on to the execution and monitoring stage, and if technological platform is needed then we first go through the simulation and implementation phase.

Simulation and implementation of business processes involves completing the business processes model with all the technical information necessary to adapt the model to selected technology and also the development of prototypes to verify the accuracy and efficiency of the model obtained, under production conditions.

Once this phase is completed we proceed to execution and monitoring, a phase which comprises the enacting of the modeled processes. Simultaneously with the process/processes execution a monitoring component of the system visualizes the state of process instances and collects information on them. Data is usually stored in log files and it represents the foundation for the next phase: evaluation and optimization.

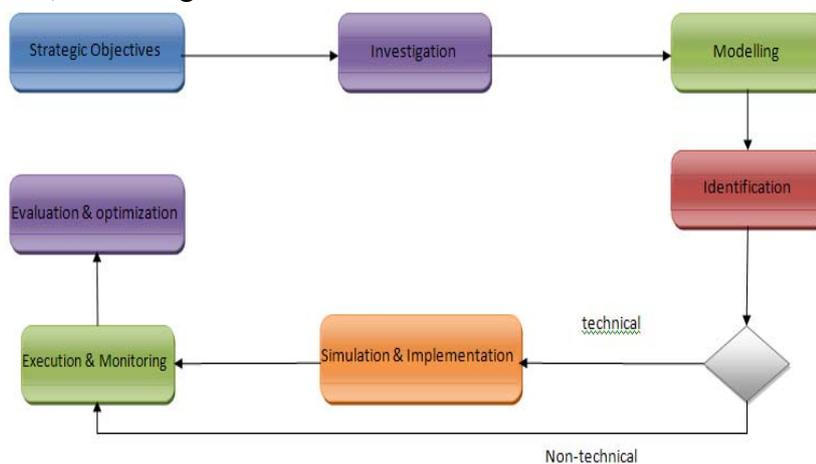


Figure 4. Proposed methodology for developing a BPM solution

It is approximated that currently there are over 160 BPM solutions on the market, and it is estimated that the market will grow by about 20% annually over the next two years. Opportunities on the BPM solutions market are very high. In 2000, most estimates showed that the BPM market was worth tens or hundreds of millions of dollars. Following analysis concluded that in 2011 the same market would reach a value between 4 and 6 billion dollars.

From the large palette of BPM solutions a few of them stand out and take a high percentage of the market. One of these is the Business Process Management solutions suite from IBM WebSphere® portfolio, which is dedicated to ensuring the growth and productivity of a company and represent key elements of IBM Business Process Management Platform. They enable advanced integration by modeling, monitoring and effective management of current or planned processes within the company or partner network. Furthermore, increases the visibility of these processes within the company and how to quantify and audit them.

New software applications can be exploited quicker and the market transactions can be performed more efficiently than the competition. Control of these processes moves from IT people to business people which means extensive possibilities to adapt companies to economic changes.

The Business Process Management suite comprises the following products: WebSphere Process Server, WebSphere Enterprise Service Bus (WebSphere ESB), WebSphere Integration Developer, WebSphere Business Monitor, WebSphere Business Modeler.

WebSphere Process Server is a comprehensive Service Oriented Architecture integration platform that allows the deployment of standards-based business integration. It is based on the robust J2EE 1.4 infrastructure and

platform services and includes, but is not limited to capabilities such as business process automation.

WebSphere ESB is a flexible connectivity infrastructure for integrating applications and services. It manages the flow of messages between service requesters and service providers and offers some mediation modules that handle mismatches between requesters and providers, including protocol or interaction-style, interface and quality of service mismatches

WebSphere Integration Developer (WID) is an integrated development environment for building applications based on service-oriented architecture and it's based on industry standards, most notably WSDL, XSD, BPEL, Java, and UML. WID uses a set of visual construction tools and higher level concepts, which lets one focus on the business problems instead of code issues.

WebSphere Business Monitor is a Web application that is deployed and run under WebSphere Process Server 6 and that provides the ability to monitor different aspects of business performance in real-time. Among the features it provides we mention: monitor and manage business performance indicators, visualize performance data such as KPIs and metrics, view business-critical information graphically and others.

WebSphere Business Modeler is IBM's premier business process modeling and analysis tool for business users. It offers process modeling, simulation, and analysis capabilities to help business users understand, document, and deploy business processes for continuous improvement.

## **6 WebSphere Business Process Management solutions performance tuning and configuration**

IBM engineers suggest the following key tuning and deployment guidelines as relevant in virtually all performance-sensitive customer engagements [8]:

- Use a high-performance disk subsystem. In virtually any realistic topology, a server-class disk

subsystem (for example, RAID adapter with multiple physical disks) will be required on the tiers that host the message and data stores to achieve acceptable performance.

- Set an appropriate Java heap size to deliver optimal throughput and response time. JVM™ verbosegc output will greatly help in determining the optimal settings.
- Where possible, utilize non-interruptible processes (microflows) instead of interruptible processes (macroflows). Macroflows are required for many processes (for example, if human tasks are employed or state must be persisted). However, there is significant performance overhead associated with macroflows.
- Use DB2 instead of the default Derby database management system (DBMS). DB2 is a high-performing, industrial strength database designed to handle high levels of throughput and concurrency. It scales well and delivers excellent response time.
- Tune the database for optimal performance. Proper tuning and deployment choices for databases can greatly increase overall system throughput.
- Disable tracing. Tracing is clearly important when debugging, but the overhead of tracing severely impacts performance.
- Configure thread and connection pools for sufficient concurrency. This is especially important for high-volume, highly concurrent workloads, since the thread pool settings directly influence how much work can be concurrently processed by the server.
- Use composite query tables for task and process list queries. Query tables are designed to produce

excellent response times for high-volume task and process list queries.

- Use work-manager-based navigation to improve throughput for long-running processes. This optimization reduces the number of objects allocated, the number of objects retrieved from the database, and the number of messages sent for Business Process Choreographer messaging.
- Avoid unnecessary usage of asynchronous invocations. Asynchronous invocation is often needed on the edges of modules, but not within a module. Utilize synchronous preferred interaction styles.
- Avoid too granular of transaction boundaries in Service Component Architecture (SCA) and Business Process Execution Language (BPEL). Every transaction commit results in expensive database or messaging operations. Design your transactions with care.

## 7 Performance tuning methodology and checklist

For system performance tuning to be performed efficiently and successfully, serious and vast training and experience are required. Therefore we do not try to describe the activity in detail but only highlight key aspects that are particularly important.

A system-wide approach to performance tuning of a WebSphere BPM environment is recommended.

When tuning a system it is important to remember that the activity includes and regards all the elements of the deployment topology and in the case of the WebSphere BPM suite this includes :

- Physical hardware topology choices
- Operating system parameters tuning
- WebSphere Process Server, WebSphere Application Server, and ME tuning.

The tuning methodology proposed consists of 4 iterative steps that will be executed in a loop.

1. Selection of the initial parameters and run the system.
2. Monitorization of the system to obtain metrics that indicate system health and to assess the need for further tuning.
3. Use of data obtained in the previous stage in order to guide further tuning changes. In general, this phase requires the analyst to examine the collected monitoring data, detect performance bottlenecks, and do further tuning.
4. Repeat until done.

In order to make sure all the important steps in tuning a WebSphere BPM solution are met, [8] suggests a checklist of activities to be used as a guide, by those performing the tuning. The list is organized in several categories, according to the tool used:

#### 1. Common tuning actions:

- Disable tracing and monitoring when possible.
- Move databases from the default Derby to a high-performance DBMS such as DB2.
- If security is required use application security instead of Java2 security.
- Use appropriate hardware configuration for performance measurement
- If hardware virtualization is used, ensure that adequate processor, memory, and I/O resources are allocated to each virtual machine. Avoid over-committing resources.
- Do not run the production server in development mode or with a development profile.
- Do not use the Unit Test Environment (UTE) for performance measurement.

- Tune external service providers and external interfaces to ensure that they are not the system bottleneck.
  - ConFigure MDB Activation Specs.
  - ConFigure for clustering (where applicable).
  - ConFigure thread pool sizes.
  - ConFigure data sources
  - Consider using non-XA data sources for CEI data when that data is non-critical.
- #### 2. Business Process Choreographer actions
- Use work-manager-based navigation for long-running processes.
  - Optimize Business Flow Manager resources.
  - If work-manager-based navigation is used, also optimize message pool size and intertransaction cache size.
  - Optimize the database configuration for the Business Process Choreographer database (BPEDB).
  - Optimize indexes for SQL statements that result from task and process list queries using database tools like the DB2 design advisor.
  - Turn off state observers that are not needed (for example, turn off audit logging).
- #### 3. Messaging and message bindings actions
- Optimize activation specification (JMS).
  - Optimize queue connection factory (JMS, MQJMS, MQ).
  - ConFigure connection pool size (JMS, MQJMS, MQ).
  - Optimize listener port configuration (MQJMS, MQ).
  - ConFigure SIBus data buffer sizes.
- #### 4. Database actions
- Place database tablespaces and logs on a fast disk subsystem.
  - Place logs on separate device from tablespace containers.
  - Maintain current indexes on tables.
  - Update database statistics.
  - Set log file sizes correctly.
  - Optimize buffer pool size (DB2) or buffer cache size (Oracle®).

## 5. Java

- Set the heap/nursery sizes to manage memory efficiently.
- Choose the appropriate garbage collection policy.

## 6. Monitor

- ConFigure CEI.
- Set message consumption batch size.

## 8. Conclusions

Business Process Management Solution provide increased performance, productivity and reduce processing time by automating and optimizing complex processes by managing the workflow throughout the organization. Based on flexible customized standards, and with a large applicability area, BPM solutions can be implemented quickly and easily according to the needs of beneficiaries, by creating an infrastructure that connects users and applications.

One activity of utmost important in ensuring that the implemented BPM solution provides optimum results is tuning. The art of system performance tuning, requires training and experience and therefore it is difficult to synthesize it in just a few indications, so the intention of the present paper was only to provide a few key points of special importance.



**Alexandra Maria Ioana FLOREA** (born 1984) has graduated from the Faculty of Cybernetics, Statistics and Economic Informatics of the Academy of Economic Studies in 2007 and also from the Faculty of Marketing in 2008. Since then she is a PhD candidate, studying to obtain her PhD in the field of economic informatics.

At present she is pre-assistant lecturer at the Academy of Economic Science from Bucharest, Economic Informatics Department and her fields of interest include integrated information systems, information system analysis and design methodologies and database management systems.

## References

- [1] O. Nicolescu (coordonator general), *Abordari moderne în managementul si economia organizatiei*, Ed. Economică, Bucharest, 2004.
- [2] M. Cooper, P. Patterson - *Business Process Management (BPM) topics covering definition, objectives, systems and solutions.*, CIO electronic magazine, 2007, [www.cio.com](http://www.cio.com).
- [3] Business Process Management Initiative, <http://www.bpmi.org/>
- [4] L. Alboaie, S. Buraga – *Servicii Web. Concepte de bază și implementări*, Ed. Polirom, Iași, 2006, ISBN 973-681-522-6nbsp;
- [5] M. Owen, J. Raj - *BPMN and Business Process Management. Introduction to the New Business Process Modeling Standard*, 2004
- [6] F. Perkonigg – *Application of Business Process Management in Web-Based Education*, Master's thesis, Graz, 2006
- [7] M. Weske - *BPM concepts, languages, architecture*, Ed. Springer, 2007, ISBN: 3540735216.
- [8] WebSphere Business Process Management (BPM) V7 Performance Red Paper, [ibm.com/Redbooks](http://ibm.com/Redbooks), 2010.

## Administration Interface in an e-Testing Software Platform

**Diana BUTUCEA**

Academy of Economic Studies, Bucharest, Romania  
dianabutucea@gmail.com

*Administration interfaces represent a subject that requires an in-depth analysis. Due to the complex tasks it needs to execute, the developer of such interfaces must find a perfect balance between aesthetics, functionality, ergonomics and modularity of it. While web users become more and more selective into choosing the websites they browse, the interface of any application must be impeccable. Since e-testing software platforms require a complex content management system, the paper will provide arguments over functionality, aesthetics, ergonomics and modularity of such platforms. An in-depth comparison over four CMSs will be made and a model of implementation will also be presented focusing on the administration interface and the functionalities it offers.*

**Keywords:** administration interface, user interface, database, software system, e-testing

### 1 Introduction

Over the last centuries innovative efforts have produced impressive technological achievements: sophisticated medical cures, agricultural methods, new modes of transport, communication media, information technologies etc. These achievements keep fostering the optimism for prosperity, higher standards of living or, in a broader sense, better conditions of life. The cradle of the optimism goes back to the Enlightenment, an intellectual movement in the seventeenth and eighteenth century that strongly influenced the portrayal of mankind. It is the era of great scientists, philosophers and writers, like Descartes, Newton, Leibnitz, Locke, Kant, Voltaire and Diderot. They claim that man is rational and good by nature. Also Darwin should be mentioned, whose theory of evolution reflected the conflict between science and religion, while it rejected the idea of creation of life according to the Bible book of Genesis. Rather than the creationist belief that every species was created individually by God and is not subject to change or progress, Darwin claimed that life has developed in a

progressive way from primitive forms to complex organisms.

The Enlightenment marked the liberation from the medieval doctrines of magic, superstition, prejudices and the fear of God by replacing it with human rationality. The fear of God made way for a scientific description and explanation of the world. Beliefs weren't anymore accepted on the authority of priests, sacred texts, or tradition, but only on the basis of reason. Reinforced by the idea of natural regularity and material cause the Scientific Revolution successfully proclaimed the ideology of upward development, progress and improvement of the world, encouraged by an ever-increasing knowledge, understanding and control of nature's processes. It asserts that the individual as well as humanity as a whole can progress to perfection. Being tightly linked with the starting points of modern society, innovation is a necessary condition for all economic functioning. Innovations further the creation of new products, services and production processes, which will give an economic actor an advantage over its competitors. The predominant motto is "innovate or pine away" and the concepts of growth, progress, innovation and change seem to have become

self-evident within the context of our societal system [9].

In recent years the pace at which new technologies become available has increased rapidly. Internet, tablets, e-readers and smart phones are fast-growing markets. So far direct technology push on the educational system remained largely without effect. Educational institutions, positioned as public utilities rather than competitive business, lacked the stimulus for displaying innovative power and could easily ignore new technological trends and resist radical changes. Today, the conditions have changed: the pressure on education is high, new technologies are flooding the markets, and learners, be it schoolchildren or adult professionals, grow up immersed in new digital communication technologies and will demand high quality, flexible, modern and tailored learning services. [10]

Within the technology framework illustrated in Section 2, the paper wishes to provide a model of implementation for an administration module used in an e-testing software system. The proposed architecture supports fully interactive operation of the administrator over the server through the user interface and is open towards the integration of more advanced administration functionalities. Section 3 will provide the functional description of the application's components and design features whereas Section 4 concludes the paper.

## 2. User interfaces features

### 2.1. Design and aesthetics

Wiktionary [11] defines **aesthetics** (*uncountable*) as the [study](#) or [philosophy](#) of [beauty](#). Beauty in World Wide Web stands for a creative, original, functional and usable layout. An interface that would generate admiring comments, lots of traffic, link exchange and link passing (a term proposed for the

transmission of the link between Internet users through different communication channels, e.g. Yahoo! Messenger or Facebook). Aesthetics inside a web application is mainly due to the web designer's experience and creativity. He is the person who has the vision to *draw* a layout following a list of specifications but also putting his fingerprint on the project (his vision over the website theme and purpose). A good layout would guarantee at least one thing to a web application: that people would not instantly click the browser's *Close* button when they start browsing it; instead they would be tempted to continue using it. [5]

Regarding the user interface matter, the distributed information approach to display design identified three forms of interface communication: information retrieval, comparison, and interaction. This theory is seeded in traditional cognitive theory stemming from connectionist networks and mapping, but the resources model acknowledged that task success is partially display dependent. According to distributed information resources research, display design impacted information-processing tasks. Various studies have examined each type of interface communication and the role of display design concluded that attention to the technology components is essential for effective positive results. [1]

After checking and optimizing the design part, it becomes the developer's job to create the functionality as appealing as possible. Therefore, different technologies (both client-side and server-side) are used in order to achieve spectacular visual effects like multimedia files embedded into web pages, special features modal windows, hover, floating or fading effects, small details that make an interface user-friendly, functional and browse-appealing.

### 2.2. Ergonomics

Wiktionary [11] defines **ergonomics** (*plural ergonomics*) as the science of the

design of equipment, especially so as to reduce operator fatigue, discomfort and injury. Starting from this definition and extrapolating, on the web, ergonomics stands for ease of use and a page's property to give the user access to any information, link, button or functional need, in order to create a non-obstructive browsing experience. Therefore, this aspect is very important and must seriously be taken into consideration when starting to develop a complex administration interface.

### 2.3. Modularity

**Modularity** (uncountable) – the property of being modular [11] states, in terms of user interfaces, a CMS's ability to easily create custom or new modules. Modern software platforms need to meet this standard and even unexperienced users have to be able to create such modules.

A comparison over all of these features and not only has been made further in the paper.

## 3. Administration interfaces

### 3.1. Overview

Talking about an e-Testing software platform requires a good, in-depth analysis. Since developing such a system can become quite challenging in both overall and step-by-step development process, a solid ground in matters of documentation and system requirements must exist even before the project begins. In this phase, a developer must find the answer to several questions which include: *What is the main functionality of the platform?*, *Who is going to operate the platform?*, *What are the expected results?* or *Which are the system's limitations?* These questions have to find full answers in the analysis phase of development in order to go on with the project. It is the software

engineer's job to define its platform guidelines, to choose its models of implementation, to coordinate the development workflow and to assure smooth and in-time project hand-over.

In the development process of an e-testing software platform, the answers to the first set of questions were extremely clear – the developer must build up a software system that would generate real-time tests and deliver instant results to the examined students, therefore, a software platform that would complete these tasks is compulsory. The decision was to split the work into three separate modules – the user interface, the administration interface and the database system (as will be shown below).

### 3.2. Open Source CMSs vs. Custom made CMSs

An administration interface for a software platform has to meet a few criteria: good aesthetics (as shown in Section 2), ease of use (ergonomics), fully customisable modules, clean and fully customisable core code, non-restrictive administration interface and page templates. Considering these rather few criteria, a decision must be taken by the software developer over which administration interface to use for its platform. Therefore, a study over four major CMSs (Content Management Systems) has been made in order to decide which administration interface suits the platform best. Since the research wanted to point out the advantages and disadvantages of using these open-source tools, the paper will present the author's conclusions after using each of these interfaces and, in parallel, developing a custom one.

#### a. The Joomla! CMS

The Joomla! project is one of the most rapidly growing open-source Content Management Systems in the market. It has a large support community and lots of developers base their projects on it. Even though Joomla! has a good set of modules

that can be used by unskilled developers, it becomes extremely difficult for anyone to create a custom module and place it in its project. The five criteria defined before are not satisfied, since the administration interface is non-intuitive, the modules are un-customisable and the code is almost impossible to change. Considering these aspects, as well as Table 1, the decision was not to use Joomla! as an administration interface.

#### *b. The Mambo CMS*

Since Mambo is the predecessor of Joomla!, it has almost the same advantages and disadvantages for the end-user. The administration interface is good enough to develop template-based projects, it includes lots of modules, but does not allow too much customisation, neither on the administration interface and on the user interface. The CMS has proven itself not suitable for the e-testing software platform.

#### *c. The Drupal CMS*

The Drupal CMS is mostly preferred by enterprise corporations due to its scalability and power to process and store large amounts of data and files. Drupal has an easy to use administration interface, easy user management, lots of plugins available free and built-in modules, but, again, customising a module has proven to be a difficult task, therefore, even though this CMS gained

the best score (as shown in Table 1), it still didn't meet the requirements for the development of the e-testing software platform.

#### *d. The WordPress CMS*

Considering the aspects that have been taken into considerations with all of the previous CMSs, WordPress, as the last open source platform to be tested proved itself extremely user-friendly in terms of interface customisation but really restrictive in terms of code editing and also in terms of security (the upgrade cycle is too fast, therefore updating the versions is compulsory).

After testing these four administration interfaces, the results were as follows:

**Obs.** The metrics defined for the research has been chosen as follows:

- if the aesthetics of the administration interface is good, the CMS received one point; else, it received zero points;
- if the CMS has been considered ergonomic and intuitive, it received one point; else – zero;
- if the CMS had customisable modules – one point; else zero;
- if the CMS had customisable core code – one point; else zero;
- if the CMS had a non - restrictive administration interface – one point; else zero.

Table 1. CMSs' scores

Criteria	Joomla!	Mambo	Drupal	WordPress
Aesthetics	1	0	1	0
Ergonomics	1	1	1	1
Customisable modules	0	0	0	0
Customisable core code	0	0	0	0
Non-restrictive administration	0	0	1	1

<b>interface</b>				
<b>Total</b>	<b>2</b>	<b>1</b>	<b>3</b>	<b>2</b>

After analyzing the scores and the main features these types of administration interfaces can offer, the decision was taken over implementing a custom-made administration interface rather than use an existing one, since none of them has proven itself reliable for the software platform that needed to be developed. Even though the amount of work is considerably larger, a skilled developer can easily create its own administration interface, build his own modules and be able to obtain good results on both client-side and administrator interface.

#### 4. Model of implementation

##### 4.1. Web technologies integrated in the administration module

Considering the wide variety of web technologies and because each technology has its strong and weak parts, it remains to the developer's choice and work experience which one to use in its software platform. Depending on the tasks the application needs to execute and on the developer's work experience, a software system can be build using different web-based technologies.

**Server-side technologies** refer to scripting where a client generates a request to a server, the server analyzes this request and generates a response back to the client. The process is executed on the server and the response returns as HTML code to the client (the web browser). These technologies are mostly preferred by developers due to ease of integration within mark-up language and because of their capability to generate dynamic HTMLs. [2]

Since its release, in 1994, PHP (the acronym for Personal Home Page) became very popular due to its ease of nesting into HTML code, but started to

fully expand since its fourth release, when Zeev Suraski and Andi Gutmans rewritten the PHP engine under the acronym Zend.[8]

PHP can be regarded as a mature technology in the sense that it has a large user base, is widely supported and has many advanced features.

Opposed to server-side technologies, **client-side technologies** are the ones that generate the events directly on the client application (usually, the web browser), in order to cut out the time needed for the server's interrogation and response delay and, most important, to create visual and functional effects similar to desktop applications, for an excellent browsing experience to the user. The resemblance to desktop applications is quality in ease and use while the difference comes from the mobility and accessibility [3]. In this context, RIAs (Rich Internet Applications) have emerged in the last few years, providing developers a large set of tools to use.

AJAX (Asynchronous JavaScript and XML) allows every element within a web interface to be individually and quickly updated without affecting the rest of the interface. This, of course, is not what most web users are accustomed to. Initiating an action within most web sites triggers the inevitable blank screen and page loading process. Though not very responsive, the full-page update makes it very clear to users that their action has resulted in a reaction and that a response will be available as soon as the page is refreshed. Because AJAX - based updates are very fast and incremental (often affecting only a small portion of the user interface), users may not notice them, especially when they are used to seeing full-page rewrites, but they make the web pages work more efficiently, increasing the browsing speed.

In matters of markup languages, the simplicity of HTML has captivated many

skillful programmers and creative web developers, enabling them to create all sorts of websites. Over the years, HTML has withstood many challenges and evolved with new features as the web technology advances at an astonishing pace. Today, HTML is still the predominant technology for website designs. These websites provide mostly browseable-only contents with form interaction using Common Gateway Interface (CGI).

Dynamic web content generation and interaction are achieved by using a database system as the backend data storage and retrieval system in conjunction with a client (web browser) and a web server. The majority of these types of Internet applications are most used by the e-commerce and entertainment sectors. However, most of these more sophisticated interactive applications still provide only simple interactions such as product navigation, shopping cart storage and form processing. More complex interaction is achieved by using Javascripts embedded in HTML or plug-ins. [6]

Two-tiers approaches to client / server architectures followed the file server approach. In this system, the client workstation is responsible for managing the user interface, including presentation logic, data processing logic and business rules logic and the database server is responsible for database storage, access and processing. Therefore, user authorisation, integrity checking, data dictionary maintenance, query and update processing are all performed at one location, on the database server. [7]

#### *4.2. Custom-made administration interface*

Since none of the four open source administration interfaces did not prove themselves efficient or customisable enough to meet the demanded requirements (as shown in

Section 3), an administration module for an e-testing software system has been developed. The solution has been implemented using PHP as server-side technology and AJAX as client-side scripting (technologies which are similar to the ones the four CMSs that have been tested use) and is composed of three major modules: the student module, the administration module and the database structure. What I needed was a tool that would dynamically generate tests, randomize questions and / or answers, generate instant results and have the capability to save and manipulate the tests' results. Therefore, the platform I have developed is fully customizable (the modules can be easily upgraded, if necessary), has an easy module implementation system and is able to manage large amounts of data and files. The implemented algorithms have proven themselves efficient and the database structure is stable and hasn't encountered errors of any kind.

The users that would interact with the application were divided in three separate categories: the master administrator, the administrator and the students.

- the master administrator is a special type of user and has special administrative roles, mostly into assigning roles and privileges to the application's users.
- after logging into the application, the administrator has many different management options, including disciplines, chapters, questions and answers, sections that define the content of the online tests; besides these sections, tests and exams management sections are different and manage the creation, distribution and manipulation of the online tests.
- students are only allowed to take the exams

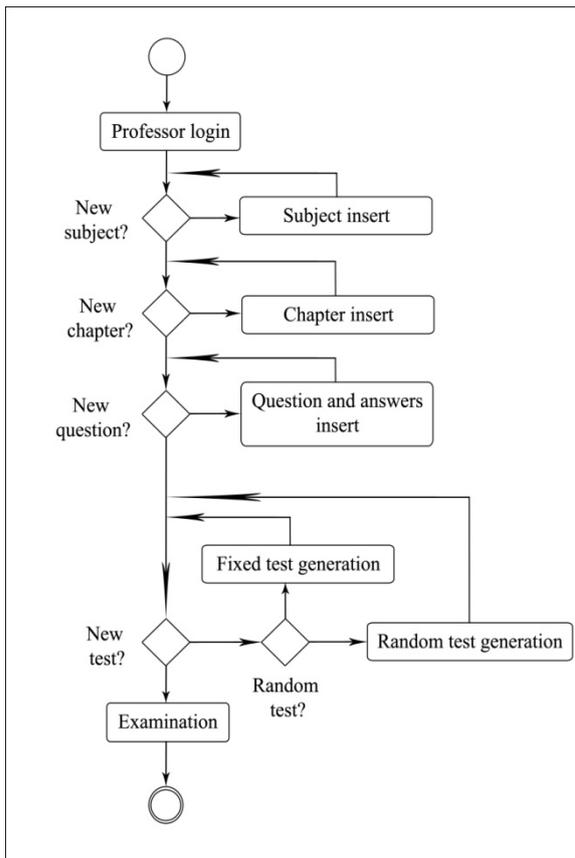


Figure 1. Activities diagram

Figure 1 shows the activities diagram and presents the workflow and the conditional steps an administrator does in order to deliver the tests to his students. As previously mentioned, after logging into the application, the administrator has to insert a new subject, link chapters to this new subject and then, in a cascade mode, insert his questions. After completing these steps, the administrator must create his test, add questions to it (if it is a fixed test) or just name it, if he wants to create a random-generated new test. All these options are available through the administration module presented in Figure 2.

After creating the tests, inside the Exams Management section, the administrator authorizes the students to take the exam, after they have previously registered through an online form. The proposed authorization algorithm counts the students that have registered, validates the list of students and, with the professor's permission, triggers the exam.



Figure 2. Administration module interface

If a student is not authorized by the professor, he would immediately be removed from the waiting list. A log file is automatically generated, allowing the teacher to have full control over his admission lists and his taken exams. Figure 2 presents the main interface of the administration module. Note that, if the test is a random-generated one, each student

will have different sets of questions, receiving their mark separately, according to their test.

Since this interface is a custom-built one and since every implemented module is known by me, I can clearly state that building this administration platform from scratch has been the best solution for my type of software system. Also, because classic CMSs are mostly targeted to build

commercial websites and to be used by unskilled users, the choice of building a custom administration interface has become compulsory in this situation.

#### 4.3. Future work

Being aware that the presented subject has not yet reached its full potential and that there still are a lot of aspects to cover, the research will continue by studying the fields of semantic web and ontologies. Since a standardization method has not yet been defined in this research area, I hope that my future studies and research would become useful materials for other researchers. Meanwhile, I wish to improve and develop the functionalities of the software model, and, in parallel with semantic web, to fulfil some objectives like implementation of an e-testing platform that would be used by blind people, implementation of a single sign on mechanism and gradual transition to web 3.0. The main interest would also imply improving the user interface, adding new question types, new test types and proposing new implementation algorithms to the present software system.

## 5. Conclusions

An administration interface represents the communication channel of a user with any application. In the development process of an interface, the developer must consider at least two aspects: the aesthetic of the interface and the functionality of it.

Recent studies have revealed a high correlation between how an application looks and the user's desire to continue browsing it [5], therefore the way an administration interface is

designed could easily define a user's will of using it.

In addition, the term administration interface strongly relates to the database it uses. Since all data is gathered from the database system, the information must be unaltered and ready to be delivered to the user when he needs it. A powerful content management system is the software platform that does this delivery whenever necessary. Even though, overall, all open source CMSs are working on the same technologies and use the same DBMS, major differences appear in terms of presentation and modularisation. Data flows are also different and modifying any of them proves to be a difficult task even for skilled developers.

Linking a database to an administration interface is usually done through configuration files, two examples of such files being listed below:

```
<?php
//php database configuration
file:
    $conn =
mysql_connect('host_name', 'user_name
', 'password') or die('db connection
failed');
    $db =
mysql_select_db('db_name');
?>

<!-- asp.net connection string -->
<connectionStrings>
    <add
        name="name_of_db_connection"
        connectionString="Data
Source=data_source_name;Initia
l Catalog=db_name;Persist
Security Info=True;User
ID=user_id;Password=user_passw
ord"
        providerName="System.Data.SqlClient"/>
</connectionStrings>
```

The paper presents the implementation model of an e-testing administration interface, analyzing some of its features, functional mechanisms and implementation algorithms. E-testing software systems represent a special kind of web applications. Because of their functional purpose, they need to be robust

and reliable. These two major characteristics are implemented through the web technologies used in the development process. In chapter 4.1, a set of arguments to support the decisions regarding the technologies used in this process have been presented. Nevertheless, although the market offers a wide variety of web technologies, in the end it comes to the developer's decision and work experience which one to use for its software platform [4].

While server-side technologies represent the basis of the software platform, client-side technologies manage to bring in front the sparks that make the difference between a static, dynamic, aesthetic or functional web application.

The paper also presents a comparative study made by the author over some of the most commonly used Content Management Systems. The study revealed a rather intuitive result (which has now been confirmed, through its research, by the author as well) – every CMS, no matter its name, producer or level of popularity, is a restrictive tool when it comes to building a custom-made web application. Even though four of these types of software platforms have been tested, in the end the author's custom-made administration interface has proven itself the best solution into building the desired e-testing software system.

In conclusion, a good balance between aesthetics, functionality, ergonomics and modularity of the administration interface, a solid structured database system, good security and reliable software development will create the premises for any administration module or interface to become widely used and appreciated.

## 5. Acknowledgments

This work was co financed from the European Social Fund through Sectorial Operational Programme Human Resources Development 2007-2013, project number POSDRU/107/1.5/S/77213 „Ph.D. for a career in interdisciplinary economic research at the European standards”.

## References

- [1] Austin, K., „Multimedia learning: Cognitive individual differences and display design techniques predict transfer learning with multimedia learning modules”, *Computers & Education* 53, 2009, pp. 1339–1354
- [2] Butucea, D., Cervinschi, C., “Development of e-learning software systems using web technologies”, *The Tenth International Conference on Informatics in Economy IE*, 2011
- [3] Castellanos-Nieves, D., Fernandez-Breis, J. T, Valencia-Garcia, R., Martinez-Bejar, R., Iniesta-Moreno, M., “Semantic Web Technologies for supporting learning assessment”, *Information Sciences*, No. 181, 2011, pp. 1517 – 1537
- [4] Cervinschi, C., Butucea, D., “Integration of web technologies in software applications. Is Web 2.0 a solution?”, *Database Systems Journal*, 1, 2010, pp. 39-45
- [5] Cervinschi, C., Butucea, D., “Impact of web technologies in functionality and aesthetics of web applications”, *18th International Economic Conference – IECS*, 2011
- [6] Chang, G., Hsieh, J., Calixto, P., „Components for Building Desktop-Application-Like Interface in Web Applications”, *APWeb, LNCS 3399*, 2005, pp. 960–971
- [7] Hoffer, A., Prescott, B., Topi, H., “Modern Database Management”, *Pearson Prentice Hall*, 2009, pp. 406-407

[8] Săulescu, O., “Creați ușor un Website dinamic cu PHP și MySQL”, Chip, Nr.2, 2003

[9] Westera, W., „Beyond functionality and technocracy: creating human involvement with educational

technology.”, Educational Technology and Society, 8, 2005, pp. 28–37

[10] Westera, W., „The eventful genesis of educational media”, Educational Technology and Society, Published online: 21 May 2011

[11] <http://en.wiktionary.org>



**Diana BUTUCEA** graduated in 2008 from the Economic International Relations Department of the Academy of Economic Studies in Bucharest, in 2009 from the Faculty of Automatics and Applied Informatics of the Transilvania University of Brasov and in 2010 from the Economic Informatics masters at the Academy of Economic Studies in Bucharest. Her parallel interests, in economy and software engineering, are now merging into her studies and researches since she is PhD student at the Academy of Economic

Studies, Bucharest, studying integrated software systems, web technologies and e-learning platforms.

## Data Acquisition and Storage in Engineering Systems

Cezar Liviu CERVINSCHI

Academy of Economic Studies, Bucharest, Romania  
cezar@symbolsoftware.ro

*Data storage represents a sensitive subject when it comes to a software system or a software application. Referring to engineering systems, data storage becomes even more challenging, since multiple functions must be accomplished and various tasks need real time response and high accuracy. While gathering information through data acquisition is relatively simple, with hardware and software equipment providing wide and complex documentation, data storage raises a set of issues, mainly due to communication channels, storage devices or software algorithms. In this context, the paper will focus on presenting some of the mostly used hardware devices and data transmission protocols in engineering applications, creating an overview over data manipulation and providing an implementation method, together with proposed solutions and a model of implementation regarding the topic.*

**Keywords:** data storage, engineering, data acquisition, database, system, microchip

### 1 Introduction

The term ‘*engineering*’, taken into consideration as it is, is void from the very beginning: the American Engineers’ Council for Professional Development (ECPD) defines the term as ‘*the creative application of scientific principles to design or develop structures, machines, apparatus, or manufacturing processes, or works utilizing them singly or in combination, or to construct or operate the same with full cognizance of their design; or to forecast their behaviour under specific operating conditions; all as respects an intended function, economics of operation and safety to life and property*’ [1, 4]. Considering this rather wide definition, it is obvious that the area of research must be clearly specified.

The paper will focus on the electrical engineering branch, studying data acquisition and storage methods for electronic devices, hardware components and will provide a functional example, together with other possible implementation scenarios regarding both acquisition and storage of information.

Since most engineering-related software concentrate on resolving only one specific task (the task the system is designed to do), each engineering system needs and uses its own software platform, requiring special resources for both the development and operating processes. However, with the massive growth of nano-technology and electronic devices, lots of issues have been removed, since high performance microchips have been created and powerful development boards incorporate all the needed functionality in terms of connectivity, communication interfaces, memory types, signal converters or graphic display [5]. In addition to these hardware break through-s, software development environments have adapted to these changes and provide specialized classes and objects that manipulate data transmitted through these devices (e.g. Microsoft Visual Studio has specialized classes for manipulating threads, message queuing or different types of networks).

Regarding the data storage support, it can be of various types, regarding the system put into discussion. Since most of the microcontrollers have embedded memory, it comes natural for the engineer to use the

built-in memory available onboard. It is a good strategy, but is mostly used into loading and running the software the microchip uses, leaving the gathered data in a third state logic, the 'unknown / undefined' state. This is where the developer, through his code, must decide what to do with the gathered data and one of the safest ways of avoiding both data loss and memory overloading is to redirect information to a database system.

Arguments over using database technology in engineering systems are the mechanisms they provide in managing the consistency and availability of data, standardization of input data or high storage capabilities. Reuse of data can be done at any time, a simple dump or query interrogation retrieving information that afterwards may be listed on external support (e.g. .xls files), used in statistic processing or in automatic creating of log files.

## 2. Engineering systems

As previously mentioned, an engineering system is void without clearly specifying its functional area. Process automation has always been a challenge for engineers and researchers, since new and more sophisticated electric and electronic devices appear, and automated processing lines tend to replace manual labour. Terms like microchip, multi-agent system, PLC (Programmable Logic Controller), development board, I/O (input - output) have emerged and human society must adapt to them.

### 2.1. Microcontrollers

Microcontrollers are integrated circuits based on different types of architectures and which, depending on the system's functionalities, incorporates different components. Since the majority of microprocessors require a lot of

discrete components and integrated circuits to execute relatively simple functions, large and complex hardware structure results [2]. Therefore, new controllers have been developed and they incorporate the hardware structure as a 'black box' with input and output signals and the microchip as the box (containing Analog/Digital converter, comparator, Pulse Width Modulation, serial communication, internal clock, RAM / EPROM / EEPROM / Flash memory, programmable I/O).

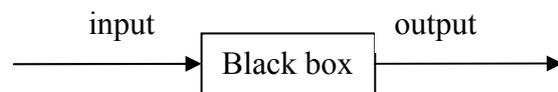


Figure 1. Black box structure

Loading the software on the microchip requires only a connection to a PC.

### 2.2. Microcontrollers' architectures

One of the most important elements that must be taken into consideration when analyzing a system is the Central Process Unit's (CPU) architecture [3].

#### a) Von Neumann architecture

Microcontrollers based on this architecture have a CPU with only one memory device used for both memorizing code and instructions and also the processed data. There is only one internal bus used for evaluating code functions (e.g. FETCH\_OP()) and data. The two operations are executed sequentially and derive into a delay on operations response.

#### b) Harvard architecture

This type of architecture has separate memory devices for program code and for data, therefore separate buses exist, making possible the parallel execution of both processes. Digital Signal Processing (DSPs) use this type of architecture and, because of its high costs, it is not widely used.

#### c) CISC (Complex Instruction Set Computer)

Almost all microchips' CPUs rely on the CISC concept, meaning that over 80 instruction functions are built-in the

microchip's CPU, operating with addresses, registers or only allowing certain operations. A big disadvantage would be the usage of small and memory consumption functions rather than a single macro-instruction.

d) *RISC (Reduced Instruction Set Computer)*

As its name states, this type of architecture uses a small set of instructions, making the CPU efficient and rapid. Its rapidness is increased by including a separate instruction pipeline, thus accelerating data transmission.

### 2.3. Communication

a) *RS232*

RS232 is a series of standards for serial binary single-ended data and control signals, connecting between a Data Terminal Equipment (DTE) and a Data Circuit-terminating Equipment (DCE). It is commonly used in computer serial ports, defining electrical characteristics, signals meanings and physical size and pinout of connectors.

A serial port is a serial communication physical interface through which information transfers from one device to another as a bit stream. A limitation of the standard is the maximum length of the data cable, which cannot be bigger than 10 meters.

The RS232 connector was originally developed to use 25 pins. In this **DB25** connector pinout, provisions were made for a secondary serial RS232 communication channel. On personal computers, the smaller **DB9** version is more commonly used (Figure 1) [4].

b) *RS485*

Similar to RS232, RS485 sends data streams through serial ports using a pair of wires to send differential signal over distances up to 1200 meters. The differential signals make RS485 very robust and more appropriate to industrial applications, while its noise immunity

and long-distance capability adapt perfectly to the industrial environment [6].

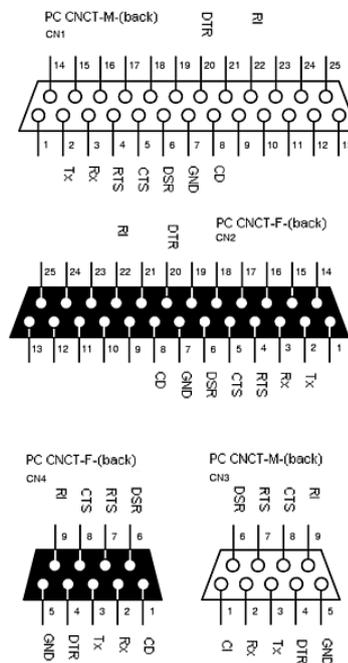


Figure 2. RS232 pinouts

## 3. Data acquisition and storage

### 3.1. Overview

The problem of data acquisition and storage is a complex one in engineering systems. While, as previously mentioned, microchips do own a small memory of their own (mostly used for storing the software application and functions the system performs), this memory, be it volatile or not, is not capable of storing all the data streams as they are registered and measured by the system.

Considering a high sampling rate for read data, the amount of information the microchip would have to memorize would increase within seconds, causing memory overflow and system crashes. To prevent such things, the software developer must decide what to do with the data.

First scenario states that the developer wants to discard the data, meaning that in this case he would just read the data, process it, if necessary, by holding

it into a memory buffer and, while new data flows through the system, the old data would disappear. The maximum output of this kind of approach would be receiving a graphical result of the data (e.g. measured parameters).

A second scenario would imply that the developer wants to have a hardcopy of the data flow somewhere on an external device. In this case, using a database management system represents a good approach, since no data would be lost, due to the DBMS's automatic backup tools. Therefore, as data flows through the system, the developer can set a sample rate (for example, the sample rate the system gathers data with) and insert into the database the desired information (data content, parameters, errors). This way, after stopping the device and analyzing data, engineers would know the parameters of the machine or the problems that occurred. This database-aided approach can be extremely useful in debugging an engineering system, since the user can track both system problems and productivity issues.

Depending on the user and on the type of the application he develops, the

Preamble (7 bytes)	Starting Delimiter (1 byte)	Destination Address (2 or 6 bytes)	Source Address (2 or 6 bytes)	Type or Length (2 bytes)	LLC Header and Information Field (0 – 1500 bytes)	Pad (0 – 46 bytes)	Frame Check Sequence (4 bytes)
-----------------------	-----------------------------------	---	--	--------------------------------------	--	--------------------------	---

**Table 1.** Ethernet frame format

The Preamble is used by the receiver to allow establishing the connection. The Starting delimiter is a simple flag that indicates the start of a frame. The type or Length field is the field that sets both the Information field and the Pad field (which is actually a compensation field for the Information field – if it has less than 46 bytes, the Pad field compensates it). Finally, the Frame Check uses a CRC-32 (Cyclic Redundancy Check) polynomial code to check the frame for

errors. In order to send a frame over a network through the Ethernet protocol, a station listens to the Ether. If it is busy, the station will enter a waiting queue and, after the thread gains access, the frame starts the transmission.

### 3.2. Transmission protocols

Having to choose which transmission protocol to use could prove to be a challenging task since the market provides several such protocols, each one of them meeting different demands, according to the engineering system they have to support.

#### a) The Ethernet protocol

Developed in the mid 70's as a part of the Ph. D. dissertation of Robert Metcalfe and standardized in the early 80's [7], when IEEE (Institute of Electrical and Electronics Engineering) started project 802 to standardize local area networks – LAN. Ethernet stations communicate through data frames, which have the following structure:

Similar to the Ethernet protocol, the GSM (Global System for Mobile Communications) protocol can be used in data transmission or acquisition systems.

#### b) The GSM protocol

Similar to the Ethernet protocol, the GSM (Global System for Mobile Communications) protocol can be used in data transmission or acquisition systems.

The standard has been developed by the European Telecommunications Standards Institute (ETSI) in order to describe second generation (or 2G) digital cellular networks' technologies. Clearly defined for mobile communications, GSM is a cellular network (data is sent through the network from one cell to another) that can easily be implemented on engineering systems that need data acquisition. It can be implemented using a GSM modem as a data dispatcher and connect to several other devices. Nowadays, with 3G networks expanding daily, this protocol should really be taken into consideration in any type of system that needs data transmission, although it has its limitations.

*c) The Wi-Fi protocol*

IEEE 802.11 is a set of standards for implementing wireless local area network (WLAN) computer communication in the 2.4, 3.6 and 5 GHz frequency bands [8]. Empirically, a WLAN network could be considered a wireless version of Ethernet. Over time, the protocol has improved, giving rise to different variations of it: 802.11a is a version that operates at 54 Mbps and is considered the favorite wireless LAN protocol for IP telephony, even though it has a major disadvantage – the network coverage only goes up to 30 m. A second version of the protocol, 802.11b, operates at 11 Mbps but has a major improvement in the network coverage – almost 100 m – making it suitable for public hotspots or small environments. An even newer version, 802.11g, works at 54 Mbps and is considered a high speed replacement for 802.11b.

*d) The ZigBee protocol*

ZigBee is a specification for a suite of high level communication protocols using small, low-power digital radios based on an IEEE 802 standard for personal area networks [8]. Since this protocol is extremely versatile, it has

been mostly implemented in engineering applications concerning fields like home automation, smart energy consumption systems, telecommunications, medical appliance or smart remote control. Thanks to its low-cost and low-power usage, the protocol allows longer life with smaller batteries, high reliability and extensive range of products. Standard ZigBee microchips have a built-in flash memory of 60 to 256 KB and they operate in all the major radio bands: industrial, scientific and medical (ISM) with data transmission rates varying from 20 to 250 kbps.

After analyzing this short and generic overview over some of the mostly used data transfer protocols, a conclusion can be made, and that is that electronic communication systems are evolving with an astonishing rhythm, from old-fashioned network cables and adapters, to cutting edge wireless technologies, high speed data transfer rates and glamorous GSM technologies for next generation gadgets.

In the same context, engineering systems become more and more sophisticated, integrates high-end technology and deliver real-time responses. Software support must keep up with the hardware devices, must create frameworks that would deliver adaptable software for these devices and integrate in the scientific process of technology development.

For an engineer, having to choose between a data acquisition protocol over another could become really difficult. Every decision must be fully compatible with the kind of task the system is projected to perform. Thus, choosing a DBMS to store the collected data would prove itself an even more challenging task, considering the compatibility issues that might occur.

#### **4. Case study**

For experimental purpose, I have developed a small robot whose task was to resolve a maze using proximity sensors and

the functions loaded inside the CPU. The maze had vertical walls and the robot would recognize black lines on the floor of it. Guided by these lines and by the sensors' received signals, the robot executes 180 degrees spins, left or right turns and stops, in order to get out of the maze.

#### 4.1. Solutions

The solution for resolving the maze proposes loading a Visual Studio C# program inside the microcontroller. When the robot starts running, the proximity sensors will search a wall on the right side, guiding the robot's moves. If the sensors detect no wall on the right side, the robot would turn right. If the sensors detect a wall on the right side and in front of the robot, he would turn left. The movement is set by the engine driver, which sends a signal to the two wheels. While moving, the communication is done through radio waves and real time information is sent to the external display device (a PC monitor). The emitter (placed on the robot) and receiver (connected to the RS232 PC interface) work on the same frequency with a similar baud rate, being fully compatible.

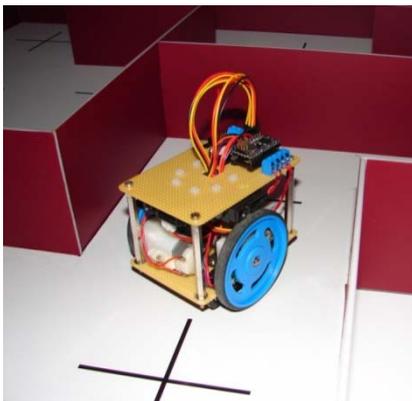


Figure 3. The robot inside the maze

#### 4.2. The sensorial system

In order to navigate through the maze, the robot uses three Sharp GP2D120 distance sensors for wall detection. They are highly precise and very commonly used in robotic systems. The detection distance keeps between 40 and 300 mm, allowing the robot to both keep its trajectory and to detect the maze's structure. The sensor is immune to the interference of ambient light and offers excellent indifference to the detected object's colour, therefore making the detection of a black wall in full sun light possible.

#### 4.3. The ATmega168 Microcontroller

The autonomous robot is built around a CPU Module who supports an AVR ATmega168 microcontroller (MC):

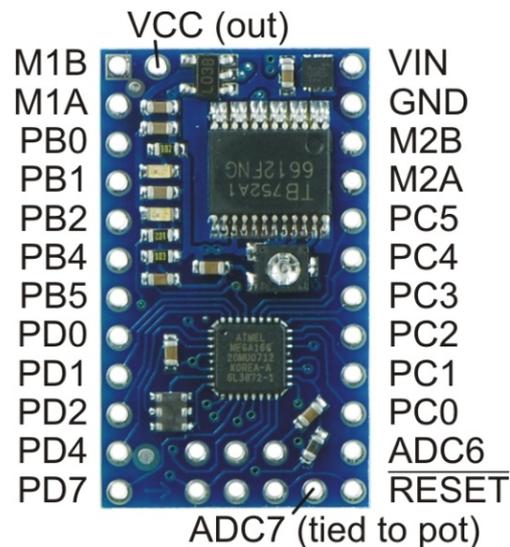


Figure 4. The CPU Module (pin's configuration)

The MC is based on the RISC architectural model. It has a 131 set of instructions and 32 x 8 general usage registers which are directly addressable by the Arithmetic and Logic Unit (ALU), allowing the access of two independent registers with one instruction, therefore the MC being up to 10 times faster than CISC-based conventional MCs.

The AVR architecture has two main memory spaces: Data Memory and Program Memory plus a Data Storage Memory. The



left, the robot must find a wall in front of him whereas, to turn 180 degrees, the system must find walls in both his left and right sides.

#### 4.7. Acquisition and storage

Regarding the data acquisition and storage issues, they remained at the developer's choice, meaning that he had the alternative to save or discard the received data. For the presented solution, since the purpose was creating a functional autonomous robotic system rather than a collection of data, data would not be saved on an external storage device.

While the robotic system is running through the maze, the alphanumeric display would indicate its state through the transmitted data. The robot moves inside the maze using the procedures described in section 4.6 and the SerialPort Terminal application displays the state messages on the computer.

A completely different scenario would have been if the data needed on-line or offline processing. Since the data acquisition software program has been developed using Microsoft's Visual Studio C#, a compatible and convenient DBMS for data storage would have been MS SQL Server.

Through an SQL Connection, when transmitting data, the program would automatically send the log messages and the parameters to a database table. Using a convenient baud rate, the software system would insert database recordings inside a specific table. Such a table structure is presented below:

Table 3. Proposed database table

Field Name	Data Type	Functionality
Acq_id	int	Recording identifier
Port_name	varchar	The opened port for the recording

Data_bit	varchar	Transmission data (bits)
Parity	int	Transmission parity
Baud_rate	varchar	Transmission baud rate
Stop_bit	bit	Transmission end bit (flag)

Using an administration module (for example, a web interface), any user could be able to manipulate and use the database records. Since every recording gives details over all the communication parameters, exporting the data set would give important information over the engineering system's functionality, over the status of the mechanisms, transmission parameters and system errors, if they occur. This data can easily be processed and the functionality of the engineering system significantly increased.

Extrapolating this idea, it is obvious that any engineering system can be improved by adding database storage support that can save all the parameters' statistics and functional log messages. Dealing with a large amount of information makes it almost mandatory that acquisition systems must have a backup support, mostly due to microchips' limited memory space. A database management system would most definitely assure disk space, backup utilities, data extraction tools and a completely functional work flow for the system: Data acquisition --> Database storage --> Data processing. This would also imply that system engineers would be aware at any time about the system's parameters and be able to take important decisions and might even change the system's structure, if the procedure calls it.

## 5. Conclusions

Building an engineering system is not necessarily a matter of quantity but a matter of quality. Such a system requires solid grounds in different areas of research,

but different tasks can be split between more specialists in order to achieve full potential from each engineering module. The paper has presented some of the most used microchips architectures and communication protocols. Automating a process proves, in this context, that understanding the system's requirements is essential into choosing a project development work flow.

After analyzing the functional needs and choosing the pieces that must be put together in order to develop the system, engineers must work together with software developers and system administrators at calibrating and tuning the whole system. Because the project to be delivered is not only a solution to a specific task, but might even be a worldwide innovative system, everything must go hand in hand and work perfectly. Examples of such automated systems can be found in medicine, home automation, smart energy consumption methods, remote control devices, building automation or retail services. All these research areas need data acquisition systems and storage space.

An important role in the development part comes to the system administrator, which has to decide which hardware platform to use, which operating system to install on the server and to decide the database management system to use for data storage. The DBMS is an important piece of the puzzle since a fraction of a second delay could compromise an entire period of data acquisition and research work or, worse, a company's productivity, therefore it must be carefully chosen in full compatibility with both hardware and software equipments.

In this context, the paper has presented possible ways of implementation for database management systems in data acquisition applications, relating the complexity of the engineering system to the DBMS to choose. In addition, a case study has been analyzed

from a functional point of view to data flow and data storage management. The research results state that a good engineering system needs a solid database backup support, together with high storage capacity and speed of response from the database server.

The presented case study represents an example of such a system, considering that autonomous robotic systems are studied intensely in matters of developing microcontrollers which, afterwards, can become the starting point of more complex systems such as robotic arms and other anatomical parts or intelligent devices for different areas of life. In the same time, the presented software application represents a good starting point for a more complex and highly developed data acquisition software system, together with the transmitter / receiver wireless connection (which, of course, could have been implemented using one of the listed protocols, not necessarily the radio wave one).

In conclusion, a good engineering system must work hand in hand with a database system since data must not be lost but processed. Important break-throughs have been made in the domains of nanotechnology and electronics resulting in high speed data flow devices, and this speed must now be transferred to databases in order to capture and store all data streams that flows inside these devices. Having implemented these features would make engineering, software development and system administration go hand in hand in all domains of life.

## References

- [1] "The Engineers Council for Professional Development", Science, Vol. 93, No. 2426, Jun. 1941, pp. 611-612.
- [2] F. M. A. da Silva, P. H. I. A. de Oliveira, "Development of a data acquisition system for light airplanes flight tests", 18th International Congress of Mechanical Engineering, 2005

- [3] G. Dudek and M. Jenkin, Computational Principles of Mobile Robotics, Cambridge University Press, Cambridge, 2010
- [4] Wikipedia, <http://en.wikipedia.org/>
- [5] MicroElectronika <http://www.mikroe.com/>
- [6] B&B Electronics <http://www.bb-elec.com/>
- [7] J. F. Shoch, Y. K. Dalal, D. D. Redell, R. C. Crane, Evolution of the Ethernet Local Computer Network, IEEE Computer, Vol. 2, No. 18, August 1982, pp. 14 – 26.
- [8] Wikipedia, <http://www.wikipedia.org/>



**Cezar Liviu CERVINSCHI** graduated from the Faculty of Automatics and Applied Informatics of the Transilvania University of Brasov in 2009. With a masters degree in Information Technology, his interest in web technologies, web development, integrated software systems and database administration brought the motivation for advanced research in this field of science, since his main target is building robust, reliable and complex software solutions. At present, he is a PhD candidate at the Academy of Economic Studies in Bucharest at the Economic Informatics department studying web technologies, databases, online security, e-business and e-marketing.

## The development of an electronic business based on the MySQL technology

**Andreea IONESCU, Andreea ISTOC**

Academy of Economic Studies, Bucharest, Romania

Academy of Economic Studies, Bucharest, Romania

University Hyperion, Bucharest, Romania

Bullguard, Bucharest, Romania

[andreea\\_ionescu\\_25\\_2011@yahoo.com](mailto:andreea_ionescu_25_2011@yahoo.com)

[andreea\\_istoc@yahoo.com](mailto:andreea_istoc@yahoo.com)

*This article aims to demonstrate the importance of using relational database management system in working with web applications. We chose MySQL technology like representative relational database management system because has: portability, scalability, speed, is easy to use, is open source, is widely used by web developers and provides good security.*

*We watched the analysis of an electronic business. We also presented the technologies that we have used: PHP, MySQL, HTML, CSS.*

**Keywords:** *relational database management system, electronic business, MySQL technology, business to consumer type, security*

### **1** Introduction

Development of the electronic business is in close contact with good maintenance of databases and with usage of relational databases management system able to do various transactions and to respond in due time needs customers. We observed that for big quantities of data MySQL server provides high performance of insert transactions [1], therefore we have chose to work with this technology.

Development of an electronic business involves ensuring of a good security.

In the future trust and security will lead to electronic business development [2].

To ensure development and improvement of business, e-business uses compliance and communication technologies in large organizations [3].

The current state of knowledge refers to the existence of many types of e-business on the market and the importance of the business models that lead to value creation and performance of the company [4].

In this article we analyze BCOMPUTERS an online business,

which it is based on a MySQL database.

### **2. The analysis of an electronic business**

We chose as case study analysis of the BCOMPUTERS online business, which is based on a MySQL database.

BCOMPUTERS site has the most diversified range of IT products on the Romanian market and it is found in the top of ecommerce sites. BCOMPUTERS is an electronic business to consumer which has: product catalog, a description of these products, products offered in promotions, shopping cart, online secure payment methods (Paypal).

The user is greeted with a friendly interface if he accesses the website. Since it can begin searching for a particular product, he will find products grouped by category. In the middle of the main page are presented products that are on promotions. On the left side of the web application are presented to customer, products that are grouped in categories and subcategories.



Figure 1. BCOMPUTERS electronic business

If the user chooses a category, there is a listing of all products in that category and the user has the option to filter the list of products even more, by choosing a subcategory. When a product attracts the attention of the buyer, he will access the details of that product:



Figure 2. Product's description

When the user accesses the search option from the menu, it appears a list of producers, if the client has an affinity for certain brands. He can choose from the listed brands.



Figure 3. List of Producers

When the user accesses information about a particular product, there are shown the product name, category from which it belongs, the manufacturer, the price and the warranty. In addition, under the product, its technical details are displayed that gives customers information that can

help in the decision to buy or not that product. When a user will be authenticated, he will introduce the username and password:



Figure 4. The Authentication based on username and password

The client uses the shopping cart to buy the products he wants:



Figure 5. Shopping cart

Payment of goods is done through three options: Cash, Card or PayPal, the customer can choose the payment's method that he wants:



Figure 6. Payment products

We also use the PHPMyAdmin management interface. The database that we worked is "calculatoarele1.sql" and consists 12 tables with the following structure: produse, producatori, comenzi, categorii, cos\_de\_cumparaturi, contact, detalii\_comenzi, disponibilitate, forum, promotii, categorii and users:

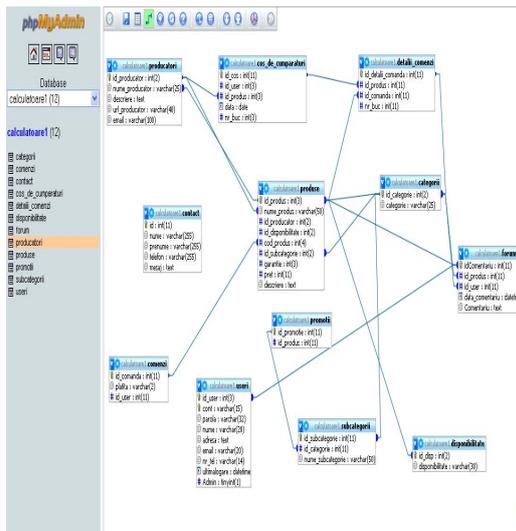


Figure 7. The MySQL database

The Produse table has the following structure:

```

calculatoare1.produse
# id_produs : int(3)
# nume_produs : varchar(50)
# id_producator : int(2)
# id_disponibilitate : int(2)
# cod_produs : int(4)
# id_subcategorie : int(2)
# garantie : int(3)
# pret : int(11)
# descriere : text
    
```

Figure 8. The produse table

The Useri table has the following structure:

```

calculatoare1.useri
# id_user : int(3)
# cont : varchar(15)
# parola : varchar(32)
# nume : varchar(20)
# adresa : text
# email : varchar(20)
# nr_tel : varchar(14)
# ultimalogare : date
# Admin : tinyint(1)
    
```

Figure 9. The useri table

Security of data refers to ensuring the protection and confidentiality of data [5]. The Data Security is ensured by using MD5 algorithm (Message Five Digest). MD5 encrypts the users' passwords in the database.

	id_user	cont	parola	nume
	1	student1	b9acd388cdd72928bc964d81d5858bb7	student

Figure 10. The encrypted password

The cos\_de\_cumparaturi table has the following structure:

```

calculatoare1.cos_de_cumparaturi
# id_cos : int(11)
# id_user : int(3)
# id_produs : int(3)
# data : date
# nr_buc : int(3)
    
```

Figure 11. The cos\_de\_cumparaturi table

The comenzi table has the the following structure:

```

calculatoare1.comenzi
# id_comanda : int(11)
# platita : varchar(2)
# id_user : int(11)
    
```

Figure 12. The comenzi table

The categorii table has the following structure:

```

calculatoare1.categorii
# id_categorie : int(2)
# categorie : varchar(25)
    
```

Figure 13. The categorii table

The Detalii\_comenzi has the following structure:

```

calculatoare1.detalii_comenzi
# id_detalii_comanda : int(11)
# id_produs : int(11)
# id_comanda : int(11)
# nr_buc : int(11)
    
```

Figure 14. The detalii\_comenzi table

The producatori table has the following structure:

```

calculatoare1.producatori
id_producator : int(2)
nume_producator : varchar(25)
descriere : text
url_producator : varchar(40)
email : varchar(100)

```

Figure 15. The producatori table

The disponibilitate table has the following structure:

```

calculatoare1.disponibilitate
id_disp : int(2)
disponibilitate : varchar(30)

```

Figure 16. The disponibilitate table

The forum table has the following structure:

```

calculatoare1.forum
idComentariu : int(11)
id_producator : int(11)
id_user : int(11)
data_comentariu : date
Comentariu : text

```

Figure 17. The forum table

The subcategorii table has the following structure:

```

calculatoare1.subcategorii
id_subcategorii : int(11)
id_categorie : int(11)
nume_subcategorii : varchar(50)

```

Figure 18. The subcategorii table

The contact table has the following structure:

```

calculatoare1.contact
id : int(11)
nume : varchar(255)
prenume : varchar(255)
telefon : varchar(255)
mesaj : text

```

Figure 19. The contact table

### 3. Technologies used:

To achieve this electronic business we used the XAMPP platform, that contains the Web Apache server, the MySQL database server, PHP and PERL.

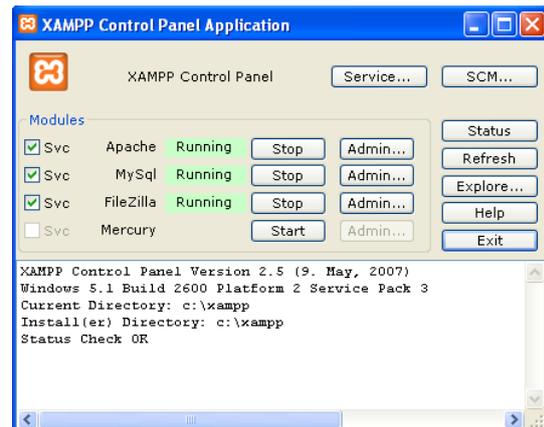


Figure 20. The XAMPP Control Panel

The XAMPP platform is easy to install and configure and it is implemented on a variety of operating systems such as Linux, Windows, Mac OS X, Solaris. For the development of BCOMPUTERS electronic business which is located at address: <http://bdcomputers.twostudio.ro> on the Internet, we used the following technologies: MySQL, PHP, HTML, CSS.

PHP (Hypertext Preprocessor) is one of the most popular programming languages who is used for development of web applications. He runs in server mode and is an open source product distributed. Currently reached version is 5.3.6.

The MySQL technology is an relational database management system.

We chose this technology because has many advantages: it is easy to use, can work with large databases. The current version is 5.0.8.

We used css styles to format the web pages:

```

<style type="text/css">
ul

```

```

{
align:center;
width:0;
padding:0;
margin:0;
style:Tahoma;
border-color:#000000;
}
a
{align:center;
style=Tahoma;
float:left;
width:140;
text-decoration:none;
color:black;
background-color:#3366CC;
border-color: 10 px solid black;
border-color: #000000;
}
a:hover {background-color:#3366CC;}
li {display:inline;
}
</style>

```

#### 4. The main commands in SQL

SQL has the first implementation in 1970 and it is based on the studies of E. F.Codd. SQL is an standardization language and it can be used to access databases: Oracle, SQL Server, DB2, MySQL.

SQL commands are grouped into five categories as follows:

- a) The query language
- b) The data manipulation language
- c) The data definition language
- d) The transactions control language
- e) The data control language

The query language allows retrieving of lines stored in database tables. The SELECT command is used. In SQL the query data stored in databases is done with the SELECT command. The SELECT command implements all operators of relational algebra. The SELECT command syntax is:

*SELECT attribute  
FROM object*

After SELECT keyword is specified list of attributes to be returned as a result of the query and after keyword FROM are specified objects (tables,

views, synonyms) from which we can select those attributes.

a) The data manipulation language allows you modifying the content of the tables. The used commands are:

-the INSERT command is used to add new lines in a table.

Also the INSERT command is used for adding new rows in a table or in tables previously created as basic views.

The INSERT command can be used in two ways:

b.1.) To enter data into a table, one record at a time. In this case the syntax is:

```
INSERT INTO table [(column1,
column2)] VALUES (value 1, value 2);
```

b.2.) To enter data into a table by copying many records in another table or group of tables, these recordings are the result of the SELECT command, in this case the syntax is:

```
INSERT INTO TABLE [(COLUMN 1,
COLUMN2, ...)] SELECT_COMMAND;
```

The UPDATE command- it is used to update data stored in a table.

The UPDATE command is used to modify existing values in a table or in a basic table and has the following syntax:

```
UPDATE TABLE [ALIAS]
SET award_columns, [award_columns,
...]
```

[WHERE CONDITION];

WHERE award column can have the one of forms:

-column= {(expression)|( subquery))  
or

( column [ ,column]...)=(subquery);

There are two ways for updating:

->providing explicitly for each value or expression to be modified;

-> obtaining values with a subquery;

The update command changes the values of records depending on the condition of where clause. In the given table in the absence of where clauses all of records will be updated.

DELETE command – it is used to delete lines in a table.

The DELETE command performs deleting records from a table or tables of a view depending of a certain condition and has the following general syntax:

```
DELETE FROM TABLE [WHERE
CONDITION]
```

The DELETE command delete some records depending on the condition of where clause. Without the where clause will be deleted all entries in the table given. In this clause may be included and subquery.

c) Data definition language allows the definition of a structure of tables that compose databases. The commands that we use to implement data definition language are:

CREATE command- for creating a database structure or a table.

In the first case for creating a database structure the syntax of CREATE command is:

```
CREATE          DATABASE
DATABASE_NAME;
```

For creating a table of a database the syntax of CREATE command is:

```
CREATE TABLE "table_name"
("column1"
```

```
"data_type_for_column_1",
```

```
"column2"
```

```
"data_type_for_column_2", ...)
```

The ALTER command for changing the structure of a database or a table.

The ALTER TABLE command:

-add a column;

-modify an existing column;

-defines a default value for a column;

The syntax of ALTER TABLE command is:

```
SQL>ALTER TABLE
      ADD      (column      datatype
[DEFAULT expression]
[, column datatype]);
```

```
SQL>ALTER TABLE table
MODIFY (column datatype [DEFAULT
expression]
      (, column datatype));
```

The ALTER TABLE is useful if you want to change a table structure. You can add new columns using the ADD clause:

The syntax is:

->table->table name;

->column->the name of new column;

->datatype-> the type date and length;

-Default expression->the specific default value for column;

You can modify an existing column using modify clause.

The DROP command- is used to delete the structure of a database.

The instruction SQL DROP TABLE is the SQL command that removes a table from a database.

The SQL DROP TABLE syntax is:

```
DROP <table_name> (<column_list>)
```

The RENAME command-is used to change the name of a table.

d) To implement the transactions control language we use the following commands:

-The changes made to database to become permanent we will use the COMMIT COMMAND;

-The ROLLBACK command-possible to dispense with the latest changes to the databases;

-The SAVEPOINT command –defines a rescue at which it can return and gives the changes made after this point to the database.

e)The Data control language allows the definition and modification rights on databases. The used commands are:

The GRANT command – for granting rights to other users of the database.

The REVOKE command – to cancel the rights of certain users.

## 5. The main functions

In tables from queries and reports we can define fields that contain results of calculations with data values in database

tables. These results come from the evaluation of arithmetic expressions that contain: numbers, names of table columns (attributes) and functions of the standard SQL language. Best known arithmetic operators for real numbers are: +, -, \*, / and for integer numbers are: +, -, \, mod.

The functions are classified in two types:

a) The functions on a several records-functions that calculate a value based on the values of a column of the table (depending on aggregate):

-The AVG() function - returns the average;

-The COUNT() function-returns the number of rows;

-The MAX() function -returns the maximum value;

-The MIN() function-returns the minimum value;

-The SUM() function -returns the sum of the values

-The VAR function-returns dispersion of the values from the respective column;

- The FIRST() function- returns the first value;

-The LAST() function- returns the last value;

b) The functions that works with the strings:

-The LCASE() function -converts a string to lowercase;

-The UCASE() function - converts a string to uppercase;

-The MID() function—extracts a substring from a string;

-The LEN() function –gives the length of a string;

-The FORMAT() function-formats a field;

-The CONCAT() function - is equivalent with the concatenation operator;

-The ROUND() function- rounds a numeric field to a specified number of

decimal places;

-The NOW() function-provides a string with current date and time;

-The format of the function is:

->format (column\_name, format);

-where format is a format string that gives the value of the column. For example the function:

->format ( now(), ‘YYYY-MM-DD’)

gives the current date as year-month-day;

The MID() function has prototype:

-mid(column\_name, index, [, length])

where:

-column\_name is the field which is extracted substring;

-index indicates position of the first character in the subsequence string (first character is string has index1);

-the optional length parameter gives the length of substring. If this parameter is not specified, the function gives the rest of the string.

The Round() function is:

->round(column\_name, number\_of\_decimal\_places)

->the function rounds to the nearest whole number;

The remaining functions have like argument the column name (attribute) of the table.

## 6. The E. F. Codd's rules

E. F. Codd (IBM researcher) has formulated 13 rules that express the maximum requirements for a database management system to be relational. The rules are useful for evaluating performance of relational database management system.

R0.The relationship-level data management languages used to operate with relationships (information unit).

R1. The logical representation of data: all information in the database must be stored and processed as tables.

R2. Guaranteeing access to data: the

data manipulation language must allow access to each atomic value in the database (table, column, key)

R3. The NULL value- declaration and processing must allow null values as missing or inapplicable data.

R4. The Metadata-information describing the database is stored in the dictionary and treated as tables, as well as actual data.

R5. The used languages- the databases management systems must allow the use of several languages, including at least one allowing define integrity constraints, data handling, access authorization, transaction processing.

R6. The updating the virtual tables- should be allowed as virtual tables and actually be updated (e.g. attribute "value" of a virtual table cannot be updated).

R7. Updates to the database- handling of tables must be done by retrieving and updating operations.

R8. The physical data independence, change the physical structure of data (mode of representation (organization) and how to access) does not affect programs.

R9. The logical data independence, changing the structure of data (logical) tables does not affect programs.

R10. The integrity constraints- they must be defined by data definition language and stored in the database dictionary.

R11. The geographical distribution of data: the data manipulation language must allow application programs to be the same for both distributed data and

for pooled data (the database management system will be in charge with data allocation and with location).

R12. Data processing at low level- If the database management system possess a low level language (the data is processed at registration), it should not be used to avoid restrictions of integrity.

## 7. The MySQL's characteristics

MySQL as we have stated is a management system relational database and implements SQL (Structured Query Language), which is the standard query language of relational databases. We choose MySQL because it is open source, is the most widely used and can handle very large databases.

Table 1. MySQL's characteristics

MySQL's characteristics	
<i>1.High Performance</i>	MySQL can run on a variety of operating systems such as: Linux, Windows, Mac OS X, Solaris. Meanwhile, he can work with millions of tranzactions.
<i>2.Easy to use</i>	MySQL is easy to conFigure and administered and has a high performance.
<i>3.Scalability</i>	MySQL is open source and it can be easily adapted by the user for his own requirements. It also manages very large databases.
<i>4.Speed</i>	Implements a variety of customer interfaces such as: Java, PHP, Python, C++, C, Perl and provides a great

	data processing speed.
5. Good security	MySQL provides increased security through the implementation of encrypted passwords and ensure data protection through specific mechanisms.

MySQL has a high performance. He can run on a variety of operating systems such as: Linux, Windows, Mac OS X, Solaris. Meanwhile, he can work with millions of transactions.

MySQL is easy to configure and administered and it has a high performance.

MySQL is open source and it can be easily adapted by the user for his own requirements. It also manages very large databases.

Also it implements a variety of customer interfaces such as: Java, PHP, Python, C++, C, Perl and provides a great data processing speed. MySQL provides increased security through the implementation of encrypted passwords and ensure data protection through specific mechanisms.

## 8. Conclusions

We choose working with MySQL technology because has many advantages: high performance, easy to use, scalability, speed, increased security.

## References:

[1]. A. O. Afolabi, A. O. Ajayi, "Performance Evaluation of Database Management System (A case study of INTERBASE and MySQL)", *Journal of Engineering and Applied Sciences* 3 (2), Medwell Journals, 2008, ISSN: 1816-949X, pp. 155-160;

[2]. K. Changsu, T. Wang, S. Namchul, K. Ki-Soo, "An empirical study of customers' perceptions of security and trust in e-payment systems", *Electronic Commerce Research and Applications* 9 (2010), pp. 84-95;

[3]. K.E. Corey, M.I. Wilson, "e-Business and e-Commerce", *International Encyclopedia of Human Geography*, Elsevier, Oxford, 2009, ISBN 978-0-08-044910-4, pp. 285-290;

[4]. K. Wikstrom, K. Artto, J. Kujala, J. Soderlund, "Business models in project business", *International Journal of Project Management*, Volume 28, Issue 8, IRNOP 2009 in Berlin, December 2010, ISSN 0263-7863, pp. 832-841;

[5]. M. Turle, "Data security: Past, present and future", *Computer Law & Security Report*, Vol. 25, Issue 1, 2009, ISSN 0267-3649, pp. 51-58;



**Andreea IONESCU** graduated from the Faculty of Cybernetics, Statistics and Economic Informatics of the Academy of Economic Studies in 2008(Bachelor's degree) and Master of E-business in 2010. She is university assistant in computer science at The Faculty of Economic Sciences at Hyperion University of Bucharest. She is currently a PhD candidate at Institute of Doctoral Schools at Academy of Economic Studies from Bucharest, in Economic Informatics Field. Her interest includes: e-commerce and project management.