Measuring the Efficiency of Cloud Computing for E-learning Systems

PAUL POCATILU, FELICIAN ALECU, MARIUS VETRICI
Economic Informatics Department
Academy of Economic Studies
6 Piata Romana, Sector 1, Bucharest
ROMANIA
ppaul@ase.ro, alecu.felician@ie.ase.ro, mariusvetrici@softmentor.ro

Abstract: - As with rapid growth of the cloud computing architecture usage, more and more industries move their focus from investing into processing power to renting processing power from a specialized vendor. Education field is no different. E-learning systems usually require many hardware and software resources. There are numerous educational institutions that cannot afford such investments, and cloud computing is the best solution for them. The implementation of a cloud computing e-learning system has its peculiarities and needs a specific approach. This paper measures the positive impact of using cloud computing architectures upon e-learning solutions development. We advance a set of cloud computing efficiency metrics for enhanced e-learning implementation process control. Also, the long term overall efficiency of the cloud computing usage in the field of e-learning system is evaluated.

Key-Words: - Cloud computing, E-learning, Mobile learning, Project management, Paretto Principle

1 Introduction
During the last years, the nature of the Internet was constantly changing from a place used to read web pages to an environment that allows end-users to run software applications. Interactivity and collaboration have become the keywords of the new web content.

There is no doubt the future belongs to the Web 3.0 (also called the intelligent Web) [1]. This new environment supports the creation of a new generation of applications that are able to run on a wide range of hardware devices, like mobile phones or PDAs, while storing their data inside the cloud. The need for education is increasing constantly and the development and the improvement of the e-learning solutions is necessary.

Also, the e-learning systems need to keep the pace with the technology, so the new direction is to use cloud computing.

There are several cloud computing services providers that offer support for educational systems. Among them are Amazon, Google, Yahoo, Microsoft etc.

In [2] are presented the main advantages of using cloud computing in schools.

The following sections focus on cloud computing concepts and the benefits of cloud computing for e-learning solutions. Also, the impact on e-learning solutions based on cloud computing project management is analyzed. This is very important for the development of e-learning solutions based on cloud computing.

2 Cloud Computing
The cloud computing term was derived from the way the Internet is often represented in network diagrams. Due to the fact it involves the existence of data centers that are able to provide services, the cloud can be seen as a unique access point for all the requests coming from the world wide spread clients (see figure 1).

Cloud computing comprises of three layers [5]:
- infrastructure as a service (IaaS)
- platform as a service (PaaS)
- software as a service (SaaS)

Depending on the requirements, the customers can choose one or more services provided.

Hardware devices (such as regular PCs, notebooks, mobile phones, PDAs or any other similar equipments) or software applications (like web browsers, for example Google Chrome) can successfully play the role of a cloud client (see figure 2). The customers are renting or simply accessing the needed processing capacity from the data center using the above mentioned client applications. The quality of the service becomes a crucial factor of the cloud computing success.
Cloud computing is by no means different from grid computing. The later tries to create a virtual processor by joining together a cluster of computers. The aim of a grid computing architecture is to solve large tasks by using the advantage of concurrency and parallelism, while the cloud is focused on collaboration.

Cloud computing becomes very popular because it moves the processing efforts from the local devices to the data center facilities. Therefore, any device, like an Internet connected phone, could be able to solve complex equations by simply passing the specific arguments to a service running at the data center level that will be capable to give back the results in a very short time. In these conditions, the security of data and applications becomes a very major issue.

Cloud computing is widely accepted today due to its key advantages:

- the cost is low or even free in some cases. Also, there are no costs (or very small ones) for hardware upgrades;
- for some applications (like spreadsheets) it can be used even in the offline mode, so when the client goes back online a synchronization process is refreshing the data;
- the strong connection that exists today between the users and their personal computers can be completely broken because a customer can reach the same result by using any Internet connected device having minimum software requirements;
- devices with minimal hardware requirements (mobile phones, for example) could be successfully used as cloud clients;
- in order to become part of the cloud, there is no need to download or install a specific software, only the Internet connection is required;
- the cost of licensing different software packages is moved to the data center level, so there is no need to upgrade the local system when new service packs or patches are released;
- crash recovery is nearly unneeded. If the client computer crashes, there are almost no data lost because everything is stored into the cloud.

Some of the main cloud computing disadvantages are the following:

- the Internet connection speed may affect the overall performances;
- on a long term basis, the data center subscription fee may be more expensive than buying the hardware;
- the service quality is crucial and the need of the backups is critical when speaking about data security.

The major players in the field of cloud computing are Google, Microsoft, Amazon, Yahoo and some legacy hardware vendors like IBM and Intel. Cloud Computing applications are mainly
intended to help companies and individuals to stretch resources and work smarter by moving everything to the cloud.

One of the biggest promoters of the cloud computing is Google that already owns a massive computer infrastructure (the cloud) where millions of people are connecting to. Today, the Google cloud can be accessed by Google Apps [6] intended to be software as a service suite dedicated to information sharing and security. Google Apps covers the following three main areas: messaging (Gmail, Calendar and Google Talk), collaboration (Google Docs, Video and Sites) and security (email security, encryption and archiving).

Microsoft is developing a new Windows platform, called Windows Azure, which will be able to run cloud based applications [7].

In 2006, Amazon extended its AWS (Amazon Web Services) suite with a new component called Amazon Elastic Compute Cloud (or EC2), that allows to the users to rent from Amazon processing power to be used to run their own applications [8]. The EC2 users rent out from Amazon virtual machines that can be accessed remotely. The cloud is an elastic one just because the user can start, stop and create the virtual machines through the web service. There are three predefines sizes for the virtual machines that can be rented: small, medium and large, depending on the physical hardware performances.

Grid computing represents a new evolutionary level of distributed computing that tries to create the illusion of a virtual single powerful computer instead of a large collection of individual systems connected together.

The easiest way to use a grid is to remotely run an application on a different computer than the one on it is usually executed. If a computer is busy, the application can be executed on another idle machine from the grid network. The remote machine must meet hardware, software and resource requirements of the application.

We can easily imagine the datacenter in the cloud as being a grid composed by several computers. Grid computing is able to increase the resource usage efficiency because it could be obtained a better balance of resource utilization.

3 Cloud Computing Benefits for E-learning Solutions

Many education institutions do not have the resources and infrastructure needed to run top e-learning solution. This is why Blackboard and Moodle, the biggest players in the field of e-learning software, have now versions of the base applications that are cloud oriented.

E-learning is widely used today on different educational levels: continuous education, company trainings, academic courses, etc.

There are various e-learning solutions from open source to commercial. There are at least two entities involved in an e-learning system: the students and the trainers.

The students' actions within an e-learning platform are:
- Taking online course
- Taking exams
- Sending feedback
- Sending homework, projects.

The trainers involved in e-learning solutions are:
- Dealing with content management
- Preparing tests
- Assessing tests, homework, projects taken by students
- Sending feedback
- Communicating with students (forums).

Each of these actions requires a certain degree of security, depending on the importance and data sensitivity.

Fig. 3 E-learning system

Usually, e-learning systems are developed as distributed applications, but this is not necessary so. The architecture of a distributed e-learning system includes software components, like the client application, an application server and a database server (see figure 3) and the necessary hardware components (client computer, communication
The client hardware could be a mobile device or a desktop computer. The client application can be a simple web browser or a dedicated application.

Even with the current hardware and software limitations, mobile devices are supporting multimedia based applications. Compared with desktop applications, nowadays mobile applications, especially multimedia-based applications, have serious limitations due to the processing power and memory constraints. Due to the fact that the data processing is on the server side, the use of mobile devices for learning is growing fast. Still, the mobile applications need to be optimized to be used for e-learning. In [3], [9] and [18] the m-learning applications characteristics were analyzed.

The e-learning server will use cloud computing, so all the required resources will be adjusted as needed.

E-learning systems can use benefit from cloud computing using:
- Infrastructure: use an e-learning solution on the provider's infrastructure
- Platform: use and develop an e-learning solution based on the provider's development interface
- Services: use the e-learning solution given by the provider.

A very big concern is related to the data security because both the software and the data are located on remote servers that can crash or disappear without any additional warnings.

Even if it seems not very reasonable, the cloud computing provides some major security benefits for individuals and companies that are using/developing e-learning solutions, like the following:
- improved improbability – it is almost impossible for any interested person (thief) to determine where is located the machine that stores some wanted data (tests, exam questions, results) or to find out which is the physical component he needs to steal in order to get a digital asset;
- virtualization – makes possible the rapid replacement of a compromised cloud located server without major costs or damages. It is very easy to create a clone of a virtual machine so the cloud downtime is expected to be reduced substantially;
- centralized data storage – losing a cloud client is no longer a major incident while the main part of the applications and data is stored into the cloud so a new client can be connected very fast. Imagine what is happening today if a laptop that stores the examination questions is stolen;
- monitoring of data access becomes easier in view of the fact that only one place should be supervised, not thousands of computers belonging to a university, for example. Also, the security changes can be easily tested and implemented since the cloud represents a unique entry point for all the clients.

Another important benefit is related to costs. If the e-learning services are used for a relative short time (several weeks, a quarter, a semester), the savings are very important.

4 Project Management Challenges

This section deals with the impact of employing cloud computing architectures in the field of e-learning software systems development. We analyze the individual influence of the cloud computing model characteristics on the project development process.

4.1 Definitions

A project is “a temporary endeavor undertaken to create a unique product, service, or result” [10]. In [11] we find that “a software development project is a temporary endeavor undertaken to create a unique piece of software.” Considering the two former definitions, we state that a cloud computing based development project is a temporary endeavor undertaken to create a unique system (both hardware and software) that will be running on a cloud computing architecture. High quality cloud computing based development projects deliver the required product within scope, on time and within budget. It is the project manager’s duty to skillfully balance the competing demands for project quality, project duration and cost of resources in order to be able to deliver the system as planned.

4.2 Project management for cloud computing system development

At the top level view, cloud computing based development of e-learning systems follows the same pattern as any other software development project.
Namely, cloud computing development project management comprises of the following process groups (see figure 4):

- Project initiation
- Project planning
- Project execution
- Project monitoring and controlling
- Project closing

Subsequent sections of this article will go into great details of the project management process groups. More than this, a set of e-learning cloud computing performance metrics will be advanced for each of the following process groups.

4.2.1 Project initiation process group

The initiation of an e-learning system development project using cloud computing architecture comprises of developing the project charter and the development of the preliminary project scope statement.

The project charter represents the document that formally authorizes the development project and endows project manager with the authority to employ organizational resources to project activities.

Preliminary project scope statement defines what needs to be accomplished, i.e. the functional specifications of the future e-learning system and the project objectives that have to be met.

At this stage special care should be given to the strong correlation between the project objectives and project scope. It is important to quantitatively measure the efficiency of the project objectives as they will have a crucial impact upon the efficiency of the future e-learning cloud computing system.

In order to assess the fitness of the proposed objectives, the If-then analysis from the LogFramework analysis [12] will be employed. The fitness \( f \) of the objective \( O_i \) is defined function of the project scope \( S \) as:

\[
f_{O_i}(S) = \begin{cases} 
1, & \text{if the statement "If } O_i \text{ then } S \text{" is true} \\
0, & \text{if the statement "If } O_i \text{ then } S \text{" is false}
\end{cases}
\]  

when \( f_{O_i}(S) \) equals 1 this means the completion of objective \( O_i \) leads to the accomplishment of project scope \( S \).

In order for the project scope \( S \) to be considered completely defined one must have:

\[
\sum_{i=1}^{N} f_{O_i} = N \quad (2)
\]

which means, all the stated objectives \( O_i \) must lead to the accomplishment of the project scope \( S \).

In the field of e-learning system implementation using cloud computing, \( S \) might be “implement an e-learning system for the 100 students of the PhD program with a maximum initial investment of 50,000Euro”.

4.2.2 Project planning process group

The very nature of a cloud computing business model and of its technical architecture makes the planning of a cloud computing based project different than any other IT development project. Considering the cloud computing infrastructure will be rented from the service provider, the project manager’s focus moves from choosing the right technology to choosing the right vendor. Instead of concentrating on the computing power of the architecture and the costs of scaling up the e-learning system, the project manager will be looking at such parameters as service availability, data security, backup and contingency plans the cloud computing vendor offers, etc.

The availability of the cloud computing based e-leaning system can be calculated as:

\[
A = \frac{UT}{31536000} \times 100\% \quad (3)
\]

where:

- \( A \) – the availability of the system during a year.
- \( A \) is measured in percents.
- \( UT \) – the total uptime of the e-learning system, measured in seconds.
- 31536000 – represent the total number of seconds during a year.
A service availability $A$ of more than 99% is considered a highly available e-learning system.

The data security e-learning efficiency metrics is defined as:

$$DS = \frac{DA}{TNA} \times 100\% \quad (4)$$

where:
- $DS$ – the security degree of the e-learning system, measured in percents;
- $DA$ – the number of successfully denied attacks upon the e-learning system during a year;
- $TNA$ – the total number of denied attacks upon the e-learning system during a year.

The more business oriented and less technical nature of the cloud computing based project activities can be immediately observed in the project scope planning, project work breakdown structure, activity identification and sequencing, activity duration estimation and schedule development. This is because the service provider takes on this initial burden and then offers on-demand virtualized processing power. For project activity duration estimation, one or several of the described techniques [13], [14] and [19] can be successfully employed.

Cost estimation and cost budgeting project processes are highly biased because of the cloud computing architecture. The service supplier absorbs up-front costs and spreads the costs over a longer period and over several cloud computing customers. Thus, the initial capital expenditure of the project is converted to ongoing operational expenditure of the e-learning system maintenance. Table 1 shows the average fees for cloud computing services.

Google App Engine includes a free quota. After this quota is exceeded, the rates from table 1 apply [15].

Amazon EC2 services are charged based on the required resources (small, medium or large) [4].

Microsoft Azzure services are not yet commercially available [6].

Using cloud computing instead of investments in datacenters (hardware and software licenses) will result in a shift from capital expenditure (CapEx) to operational expenditure (OpEx).

Quality planning project process involves creating test plans for the future e-learning system. The tests should take into consideration both intrinsic system features and performance testing on the cloud computing architecture.

Human resource planning entails allocating both legacy software development personnel and cloud computing engineers that are aware of the peculiarities of these platforms.

Communications planning means deciding what project processes and tools will be used for timely and appropriate generation, collection, distribution, storage and retrieval of project information. More exactly, this means setting up the reports that have to be generated, their content and frequency. Also, this implies setting up a bug tracking system for recording all the issues that arise during e-learning system development.

The ratio of bugs to the total number of features developed is called bug-feature ration:

$$BFR = \frac{\text{Bugs}}{\text{Features}} \quad (5)$$

where:
- $BFR$ – the ratio between the number of bugs (defects) discovered and the number of features developed. This metric show how many bugs are there for every developed feature;
- Bugs – the number of defects found in the system;
- Features – the number of features developed according to the project plan.

Risk management planning project processes need special attention when working with cloud computing infrastructure. It is true that the cloud computing maintenance burden resides solely within the responsibilities of the provider. Even though there are service level agreement items in the contract with the vendor, still Gartner analysts advises us [16] to discuss the following items with our future cloud computing vendor:

- Privileged user access, means asking who has specialized access to data and what are the procedures regarding hiring and management of such administrators.
- Regulatory compliance, means asking whether the vendor is willing to undergo external audits and/or security certifications.
- Data location, does the provider allow for any
control over the location of data.

- Data segregation means making sure that data encryption is available at all stages. Also this implies making sure that the encryption schemes are designed and tested by experienced professionals.
- Recovery, means discussing what happens to data in the case of a disaster, and whether the vendor offers complete restoration. If so, how long does that process will take.
- Investigative support means making sure that the vendor has the ability to investigate any inappropriate or illegal activity.
- Long-term viability, what happens to data if the company goes out of business. Also this implies making sure that the date will be returned in an appropriate format.
- Data availability means asking what are the procedures of moving the data onto a different environment, should the vendor decide to do that.

4.2.3 Project execution process group

Executing the project means directing and managing the project development and performing ongoing quality assurance. For a successful cloud computing based e-learning system to be developed, legacy software development techniques can be successfully employed. That is source control software, build scripts for building the deployment package and automated tests for regression testing. Project criticality analysis techniques [17] can be easily employed in order to enhance the project development quality.

The criticality index of a task represents the probability that this task will be on the critical path:

\[ TC = \frac{\sum_{i=1}^{N} TC_i}{N} \]  

where:
- \( TC \) – task criticality, a number between 0 and 1 inclusively.
- \( TC_i \) – equals 1 if task is on critical path at iteration \( i \) and 0 otherwise.
- \( N \) – the total number of Monte Carlo simulations.

The closer to 1 is TC for a given task, the higher the probability that that task will be on the critical path. The closer to 0 is a task’s TC, the higher the probability that the task will not reside on the critical path. The higher the TC of a task, the higher is the importance to manage the duration of that task in order to avoid project delays.

The sensitivity index of a task represents the correlation between task duration and the overall project duration. In practice, the sensitivity index SI is calculated as the Spearman’s Rank Correlation between task duration and project duration:

\[ SI = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \]  

where:
- \( SI \) – sensitivity index of a task;
- \( d_i = x_i - y_i \) – the difference between the ranks of the corresponding values \( x_i \) (task duration) and \( y_i \) (project duration);
- \( n \) – the number of simulations performed.

The sensitivity index SI values lie between -1 and 1. In the field of project duration estimation, a SI less than 0 has no sense because the project duration cannot be shorter as long as the task duration goes longer. So the only meaningful values are between 0 and 1 inclusively. The greater the SI of a task, the higher is the correlation between task duration and the overall project duration.

The cruciality index CI represents the product of the two indexes calculated above and shows the importance to manage the duration-uncertainty of an activity:

\[ CRUI = CI \times SI \]  

where:
- \( CRUI \) – the cruciality index of a task;
- \( CI \) – criticality index of a task;
- \( SI \) – sensitivity index of a task.

The CRUI metric has no unit of measure but its significance lies in its ability to rank project tasks according to the descending order of the importance to manage the uncertainty of an activity. The higher the CRUI of a task, the more attention the task needs from the manager of the project regarding timely execution of the task.

Project execution during e-learning system implementation can benefit from such performance metrics used in the field of automated software testing. Test success rate is defined as:

\[ TS = \frac{TP}{TNT} \times 100\% \]  

where:
TS – test success rate, measured as a percentage; TP – test cases that passed with success; TNT – total number of tests.

The closer the TS test success rate to 100%, the higher the quality of the e-learning system. For most of the e-learning systems the acceptance criteria is a TS rate of 100%.

In [22] are presented metrics related to the cost effective software testing process.

4.2.4 Project monitoring and controlling process group
The monitoring and controlling processes are performed during the entire project lifecycle in order to take preventive and corrective actions so as to meet the established project performance goals. Continuous monitoring gives the project management team insights regarding project health and identifies any areas that might need special attention.

The cloud computing based e-learning systems are no different than other software development projects. More exactly, monitoring and controlling processes are concerned with:

- Assessing project current performance.
- Comparing planned and actual project performance.
- Analyzing, tracking and monitoring identified project risks.
- Provide accurate information regarding project status report.
- Provide updated project costs and schedule information.
- Manage ongoing feature change requests.

The ongoing performance of the project can be successfully tracked using Earned Value Management technique. See figure 5.

\[ EV = \sum_{\text{completed}} \text{PV} \] (10)

where:
- EV – current earned value of the project, i.e. the sum of the planned value of the completed elements of the project;
- PV – planned value of various project elements. Only completed project elements will be taken into consideration for this formula calculation.

PV from figure 5 represents the planned (budgeted) project value earning as the e-learning system implementation advances. PV will be calculated on a per case basis, but generally it will be a function of the implemented features:

\[ PV = f(\text{features}) \] (11)

AC from figure 5 represents the actual cost of the work performed. AC is calculated as the sum of all the individual costs incurred:

\[ AC = \sum C_i \] (12)

where:
- AC – actual cost of the work performed;
- C_i – the cost of executing the feature i.

Earned value management method is a valuable tool for monitoring project progress and for anticipating and mitigating any problems the project may sustain.

4.2.5 Project closing process group
Project closing involves finalizing all project activities and performing the acceptance and delivery of the final e-learning system. During this stage the project scope is checked against the initial objectives, the e-learning system installation and maintenance is documented, the acceptance testing of the final product is performed and the formal closing of the project is executed.

5 Using Pareto Principle for quality control
The name of the Pareto Principle (also known as the 80/20 rule or Pareto’s Principle of Unequal Distribution) was suggested by the Joseph Juran (1904-2008), an American management consultant that was born in Romania (Braila).
He revealed the work of the 19th century Italian professor of political economy Vilfredo Pareto that discovered that the top 20% of any country’s population accounts for more or less 80% of its total income.

Initially, Pareto noticed the distribution for Italy but right after he extended the analysis on several other countries obtaining very similar results. Of course, the principle can be virtually applied in any area, like domestic behavior, for example – we can easily notice that we eat 20% of favorite food for 80% of time or, similarly, we spend 80% of time doing the most frequent 20% of activities.

Mathematically speaking, there is nothing special about the proportion of 80/20 but many real systems come across a ratio very closed by the Pareto’s distribution.

In the modern economy, the principle was quickly extended to quality control [20], stating that most defects in production are the result of a small percentage of the causes of all defects. This is generally defined as “the vital few and the trivial many” or “the vital few and the useful many”.

Starting from the Pareto Principle, Juran lately introduced the concept of CWQM (Company Wide Quality Management) [21] that was based on three pillars (also known as Juran Trilogy):

- quality planning – is focused on identifying the customers together with their needs that should be satisfied;
- quality control – follows the process of producing goods and services that meet the previously identified needs;
- quality improvement – the efforts performed to constantly improve the previous processes.

In our days, the principle is still very applied in a variety of areas, especially the quality control field. For example, in the IT industry it is considered that 80% of users are actually using only 20% of the features.

In the recent years, the principle was mostly applied on errors rather than features because it was observed that 80% of errors are generated by the 20% of the detected bugs, so a small proportion causes most of the errors (figure 6).

This is a winning strategy applied by the big software companies that are fighting with the bugs since the customer satisfaction can be quickly improved by fixing only a small amount of the reported errors.

This principle can be successfully applied for e-learning projects based on cloud computing.

6 Conclusions and Future Work

The development of e-learning solution cannot ignore the cloud computing trends.

There are many benefits from using the cloud computing for e-learning systems. Also, there are some disadvantages that have to be taken into account.

Using cloud computing for e-learning solutions influences the way the e-learning software projects are managed. There are specific tasks that deal with finding providers for cloud computing, depending on the requirements (infrastructure, platform or services). Also, the cost and risk management influences the way the e-learning solutions based on cloud computing are managed.

A metrics system has been developed in order to measure the efficiency of cloud computing based e-learning solutions.

Also, the Pareto Principle is still a strong mechanism constantly used in quality control of projects from various areas, including the IT field.

The Academy of Economic Studies from Bucharest uses a e-learning solution based on Moodle and it has its own datacenter that can be in the future a platform for cloud computing.

Acknowledgements

This paper presents some results of the research project IDEI 2673: Project management methodologies for the development of mobile applications in the educational system, financed within the framework of IDEI research program.
References: