The 8th International Scientific Conference
eLearning and software for Education
Bucharest, April 26-27, 2012
10.5682/2066-026X-12-110

TOWARDS THE SECURITY OF THE E-LEARNING IN THE OPEN SOURCE CLOUDS

Adrian COPIE
West University of Timisoara, Computer Science Department, Blvd. V. Pârvan 4,Timişoara, Romania
E-mail: adrian.copie@info.uvt.ro

***Abstract**: In the last years, Cloud Computing was adopted as a large scale technology in various fields and e-learning is not an exception from this rule. In the meanwhile, many universities have developed their own private clouds in local datacenters, for training and educational purposes, using open source technologies. Together with the information of public interest, stored in the private cloud, there is a lot of confidential data involved in the e-learning process, like exams, students personal data and many more, that must be safely kept inside the boundaries of the private cloud. This paper explores the challenges that face the private cloud based on open source solutions, reviews the general and particular threats existent in such environment, closely related to the e-learning activities and raises some conclusions about the security mechanisms and good practices that must be implemented in order to have a safe e-learning environment.*

***Keywords:** Cloud Computing, Security, E-learning, Open Source*

I. INTRODUCTION

Cloud computing became nowadays a common term and people are more familiar with products like Google Docs, Google AppEngine, Amazon S3, Amazon EC2, Yahoo! Mail, etc, developed by the big companies following the service model, but all of them are part of the public cloud, where anyone has access. The private cloud is something different, tailored to fit specific needs in terms of security and functionality, covering the enterprise level and with limited access. A possible definition of the private cloud is "an environment based on dedicated hardware, under the control of an organization, which is able to run and implement characteristics of cloud computing, like virtualization, layered services over the network and in the same time it applies stricter security policies, latencies, Service Level Agreements (SLAs)" [8]. Usually the provided environment is single-tenant and the consumers of the offered services are called trusted users. The infrastructure that is found inside the datacenters can be very heterogeneous, so the private cloud must support a large variety of operating systems, virtual machines, network switches and routers, it must smoothly integrate with a large range of management tools, must dynamically scale up and down based on the peaks that occurs inside the system. The lessons learned from the public cloud could be applied on a smaller scale to better use and optimize the resources [8]. This niche related to the under-utilization of the universities datacenters is a perfect opportunity for improving the educational process by using specific solutions in environments based on open source private clouds.

The price of the hardware components has declined in the recent years, fact that leads to more acquisitions of computation units inside the educational institutes in order to keep the pace with the new technologies and the new trends in every field of education. Many universities have now their own datacenters used for training purposes and for various internal and external projects. Most of the computers are not always fully loaded, a lot of computation power together with memory and storage

capabilities are wasted, thus, these resources could be used for other purposes. These aspects come together with the general economic context of the last years, when the universities' budgets have dramatically decreased due to the global financial crisis and alternative solutions for reaching the educational purposes had to be found. One approach in optimizing the existent resources and improving the efficiency and effectiveness of the educational processes inside the universities is transforming the learning activities to be service oriented through e-learning paradigm [14].

Most of the existent e-learning systems are implemented as Virtual Learning Environments (VLE) which are electronic platforms used to provide electronic courses, track the students activity in acquiring the necessary knowledge to prepare the exams and they are used even further, in formal education, in corporate trainings to meet various needs [1]. Usually the activities performed by the students are: participation in online courses and online examinations, participation in common projects and sending feedbacks for various courses or seminaries. Teachers, on the other hand, perform activities like : adding content for courses and projects, preparing the tests and examinations, evaluate the exams together with the home works and group or individual projects and at last but not at least they communicate with the students. VLEs are known also as Learning Management Systems (LMS) or Course Management Systems (CMS) and their goal is to simplify the course management activity for a wide area of learning domains. The VLEs are content centric, all the courses are centralized in a single location by the authors (usually teachers) and then they are pushed to the learners regarding their background and learning style. Another complementary approach for the e-learning is given by the new technologies developed in the last years related to the content management over the internet, like wikis, blogs, social networks which allows people to create, cooperate, publish and share their own content organized in a specific way, becoming in the same way producers and consumers of information. This kind of environment is called Personal Learning Environment (PLE).

In order to take advantages from the computation power and storage capacity existing in their own datacenters, the universities can build private clouds using various open-source middleware solutions like Eucalyptus, OpenNebula, Nimbus and many others, which allow flexible configurations of virtual machines to be provisioned dynamically and elastically on-demand [11] . These virtual machines will form the base for hosting the other software layers represented by the e-learning environment. The pros of using such open source solutions consist in lower CapEx with the middleware, better customization that fits specific needs and theoretically better security due the fact that all the components are under educational entity's control. Despite the fact that, in theory, the security could be better managed inside the private clouds, there is a wide range of risks to which the systems built inside the cloud are exposed.

The Private cloud is able to support both VLE and PLE approaches due to its high level of customization of the software stack. As a result of the activities performed by teachers and students, sensitive and confidential data is generated: the grades from the examination processes, the results obtained as a consequence of researches in various projects, personal data, exams subjects, etc. To protect this data, various levels and techniques of security inside the private cloud must be put in practice.

II. PRIVATE CLOUD SECURITY CONSIDERATIONS

The private cloud is a collection of various components, hardware and software, that cooperate together to provide to the end user the image of a whole system with limited resources, unlike the public cloud which offers the illusion of unlimited resources. All the components, concepts and relations used to define the private cloud form an *ontology* [9] that is better illustrated in Figure 1 which is also depicted in [9]. One of the private cloud important characteristics is the management performed locally, entirely by the IT department of the entities that owns the system. For a secure operation of the whole system, a wide range of policies, technologies and controls used to protect data, applications and associated infrastructure [6] are involved and the responsibility for their implementation and compliance is entirely moved to the local staff.

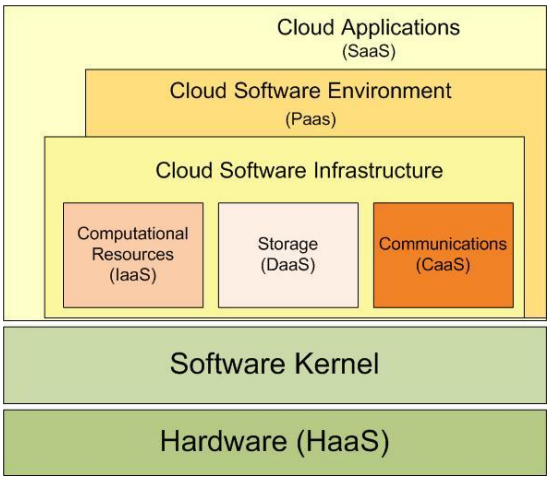

Figure 1. Cloud Computing Ontology [9]

A private cloud can basically provide many types of services, structured on various layers, starting at the lowest level, with *Hardware as a Service* (HaaS), where physical machines are made available to the end users. Over this layer, through specialized software applications called hypervisors, virtualized resources are offered, in terms of configurable virtual machines and the service is called *Infrastructure as a Service* (IaaS). Other services provided at this level are: *Data storage as a Service* (DaaS) which made possible data persistence in specialized storage configurations and *Communication as a Service* (CaaS) through various types of message queues. More sophisticated software components installed over the virtualized infrastructure allow a higher level of abstraction and offer an environment for the execution of special applications that take advantages of the dynamically character of the private cloud, this layer being called *Platform as a Service* (PaaS). The highest level in the private cloud ontology stack consists of specialized single-tenant software applications typically accessed by the end users through a thin client, entitled *Software as a Service* (SaaS).

The e-learning systems installed into an educational entity benefit from almost all these service levels: they use virtual machines in order to install the software stack for running applications, obtained from the Infrastructure level, some applications need a specific platform to be executed, which is the Platform level responsibility, and finally some e-learning components are implemented as software services hosted by the private cloud itself. Due to this layered structure and the way in which all the components interact, every level presents various security risks that must be understand and mitigate, especially those related directly to data, as the most valuable asset inside the e-learning system. No matter the service offered by the private cloud, some requirements must be satisfied in order to preserve data security: *confidentiality*, *integrity* and *availability* [6].

- Confidentiality ensures that the data which resides in the private cloud cannot be accessed by unauthorized persons. In fact, confidentiality has two components: *authentication*, which is about who can access the data and *authorization* which controls if the legitimate person performs legitimate operations against the data, once it was already authenticated. For the e-learning systems implemented in the private cloud this is the case, for example, of the students which have access to various courses or learning content but they are not allowed to see the subjects of the future examinations;
- Integrity is about keeping the data unaltered in the cloud. This could be achieved through some well established techniques like *Message Authentication Code* (MAC) and *Digital Signature* (DS), based on symmetric keys (MAC) and asymmetric keys (DS), which appends a checksum at the end of the data. Like all the check mechanisms, the stored checksum is compared in the reading phase with an ad-hoc checksum computed based on the read chunk of data;
- Availability is about keeping data accessible all the time for the authorized parties. There are various threats targeted against availability: Distributed Denial of Service (DDoS) or Cloud Service Platform (CSP) availability.

Along with the data generated as a result of the learning processes, the other building blocks of the private cloud, together with the e-learning components are also vulnerable to various attacks or leaks and the next paragraph will discuss some of the most common threats and risks that the educational entities must face in order to maintain a secure and operational learning environment.

III. PRIVATE CLOUD SECURITY RISKS

The private cloud scaffolding is built usually behind firewalls, so the threats coming from the Internet are lowered enough, the main risks concerning the security are coming from the inside. However, the risks palette of the inside threats is still wide, some of them are relate with the infrastructure, some with the personnel using the applications and some with the applications themselves. Further we will detail some of the most important ones:

- *Hardware loss or theft*. If the datacenters are not well physically secured, hardware theft could occur and together with the hardware, important data could disappear. The employees or students carrying USB flash drives or external hard drives with confidential or critical data could be risky and prone to severe damages if their storage devices land in wrong hands;
- *Dynamic character of the VMs*. In the private cloud all resources are centrally pooled and then delivered on demand, based on algorithms that assure the scalability of the resources up and down, correlated with the users needs. The virtual machines are created and destroyed at any time and more than this, the hypervisors are now able to support the VMs migration, which means that the virtual machines are no more bound to specific hardware, they could migrate to other physical server in order to execute in its environment, or even they could leave the current network and be hosted in another network;
- *Control unbounded from the hardware*. Due to the dynamic character of the virtual machines in terms of migration, it is not possible to implement traditional security rules related to MAC addresses and IPs;
- *Users policies*. VLEs and PLEs hosted in the private cloud must face with a wide range of users with specific access rights and security policies;
- *Security of the interface APIs.* This is a potential risk for the e-learning system and its data located in the private cloud due to the fact that the APIs represent the connection between the applications and the virtual infrastructure. They are built to protect the services against accidental and malicious attempts to circumvent the security policies in the private cloud and if they have security breaches they could affect the entire system security;
- *Hijacking of the accounts.* This could happen from inside the private cloud through phishing, fraud or exploitation of the software security vulnerabilities that could lead to credentials and password thefts and further to unauthorized access or manipulation of confidential data, monitoring of the activities or even identity thefts;
- *Data leaks*. The educational entities tend to have a relaxed level concerning the security policies. They may have various procedures defined and assigned responsibilities, but they are not rigorously enforced in the daily activity and could lead to losses in intellectual property and data;
- *Weak authentication*. Relaxed authentication system in terms of password strength and rotation is a major risk for accessing the network and get in touch confidential data;
- *Virus threats*. The private cloud has its own risks in what concerns the virus and malware threats, the antivirus updates being the responsibility of the local IT team and must be done on regular basis and all the virtual machines templates must have the up-to-date AV programs;
- *Improper data backups*. Having old backups or worst, don't have backups at all, is a major risk in case of disasters because of the irreparable data loss. Unlike the public cloud where

data backup is the responsibility of the provider, this solution enforce the local backups management;

- *Failures recovery policies*. Power outages or hardware failures could be very damaging in a private cloud environment, on one hand due to their physical impact on the existing running hardware, on the other hand due to the effects resulted from the interruption of the learning process.

The experience gained in the commercial private cloud environments reveals that all the issues enumerated above must be seriously taken into account and they must be addressed individually using specific methods.

IV. RECOMMENDATIONS TO IMPROVE THE PRIVATE CLOUD SECURITY

In the previous chapter we emphasized some of the most relevant threats and risks that an e-learning environment based on the open source private cloud must face. It is important to know that despite the wide palette of potential security issues, these could be mitigated, addressed or even avoided by implementing and obeying various policies and good practices as follow:

- *Secure the hardware.* In order to counter thefts of the hardware located in the educational institutes datacenters, this should be locked in dedicated rooms and security policies regarding the access must be designed and followed;
- *VM security.* Due to the dynamic character of the virtual machines that could migrate from the initial physical host to another physical host or even to another network, the hardware specific restriction could not be used to implement various security policies, so these defense techniques must be designed at the VM level, in the installed software stack;
- *Define pragmatic users policies.* Efficient security policies must be set in order to have a better control over the entire users database. Various activities related to content producing and consuming should be virtually set in separate sub-networks with separate security policies, users and risks;
- *Use the most recent APIs.* This is an effective approach since there are chances to be less prone to security holes. E-learning system implemented in the private cloud rely on very complex software and hardware scaffoldings. Many software layers are used and exposed as interfaces to communicate each other and security issues could be present at any level;
- *Enforce strict security policies.* Enforce strict rules regarding the strength of the passwords, password rotation, disable the unused accounts, audit the sensitive resources;
- *Enforce the use of anti-virus application.* This must be performed on all the eligible machines from the private cloud;
- *Do regular backups*. The backup for the entire e-learning system which consists in VLE or PLE components together with the related data is the responsibility of the local IT team and must be performed on regular basis and the backups must be rotated and some copies kept in a safe place, preferably not in the same building with the operational cloud. Sometimes the public cloud could be considered in order to keep a backup copy or even a mirror of the data from the private cloud, if this will not break any data confidentiality law or if the data are not restricted by law to reside in a specified geographically area;
- *Protect the hardware.* The critical hardware from the datacenters must be protected with uninterruptible power supplies and surge protections in order to avoid as much as possible the failures related to the power interruptions or current spikes.

The list is not exhaustive and for each configuration, according to its concrete problems, specific measures must be applied.

V. CONCLUSIONS

The e-learning process will represent the cornerstone for any respectable university. The information related to any educational field is growing tremendously and a highly collaborative and distributed system is mandatory in order to organize and manage all this data. On the other hand, the interest of the new generations of students is mainly focused on the acquisition of information using e-learning platforms, according to Gartner [15]. The private clouds are cost effective solutions in order to benefit from the unused computation power and data storage from local datacenters in universities, even if the hardware is not always the last generation, together with better customization of the software and a closer and effective security of the whole system. However, due to the system complexity and the large number of components involved, a wide range of security risks threats the integrity of the services implemented in the cloud. Knowing them and being aware of their existence allow us to address them individually and take the appropriate actions to mitigate their effects. We have emphasized in this paper some security considerations related to the private cloud and also we have reviewed the most important security risks that an e-learning system implemented in such an environment must face. The proposed list of recommendations could be a starting point to improve the security and mitigate the inherent risks in an e-learning environment that works in a private cloud.

References

[1] Mohammed Al-Zoube, 2009. E-Learning on the Cloud, International Arab Journal of e-Technology, vol. 1, No. 2

[2] Paul Pocatilu, Felician Alecu, Marius Vetrici, 2010. Measuring the Efficiency of Cloud Computing for the E-larning Systems, WSEAS Transactions on Computers, Volume 9 Issue 1

[3] Pardeep Sharma, Sandeep K. Sood, Sumeet Kaur, 2011. Security Issues in Cloud Computing, Communications in Computer and Information Science Volume 169, pp 36-45

[4] Yun Bai, Sean Policarpio, 2011. On Cloud Computing Security, Communications in Computer and Information Science, Volume 162, Part 4, 388-396

[5] Amit Sangroya, Saurabh Kumar, Jaideep Dhok, Vasudeva Varma, 2010. Towards Analyzing Data Security Risks in Cloud Computing Environments, Information Systems Journal

[6] Puneet Jai Jaur, Sakshi Kaushal, Security Concerns in Cloud Computing, Communications in Computer and Information Science Volume 169, 2011, pp 103-112

[7] Niall Sclatter, 2010. eLearning in the Cloud, International Journal of Virtual and Personal Learning Environments, Volume 1, Issue 1

[8] Krishnan Subramanian, 2011. Private Clouds, Whitepaper sponsored by Trend Micro Inc.

[9] Youseff, L., Butrico, M., Da Silva, D, 2008. Toward a unified Ontology in Cloud Computing, Grid Computing Environments Workshop

[10] Yu Bai, Sean Policarpio, 2011. On Cloud Computing Security. Communication in Computer and Information Science, Volume 162

[11] Peter Sempolinski, Douglas Thain, 2010. A Comparison and Critique of Eucalyptus, OpenNebula and Nimbus, Second International Conference on Cloud Computing Technology and Science

[12] Cloud Security Alliance, 2010. Top Threats to Cloud Computing v1.0

[13] Cloud Security Alliance, 2010, Security Guidance for Critical Areas of Focus in Cloud Computing v2.1s

[14] Marinela Mircea, Anca Ioana Andreescu, 2011, Using Cloud Computing in Higher Education: A Strategy to Improve Agility in the Current Financial Crisis, Communications of the IBIMA, Vol. 2011, ID 875547

[15] Kathy Harris, 2002, E-Learning: An Application Whose Time has Come, url: http://www.gartner.com/id=375574