ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 6, Issue 10, October 2017

Survey on Forensic Investigation in Cloud using VM Snapshots

Nilima Mitragotri[1], Prof M.A.Nirgude[2]

P.G. Student, Department of Computer Science and Engineering, Walchand Institute of Technology, Solapur, Maharashtra, India[1]

Assistant Professor, Department of Information Technology, Walchand Institute of Technology, Solapur, Maharashtra, India[2]

ABSTRACT:Cloud computing systems host most of today's commercial business applications yielding it high revenue which makes it a target of cyber-attacks. This highlights the need for a digital forensic mechanism for the cloud environment. Conventional digital forensics cannot be directly presented as a cloud forensic solution due to the multi-tenancy and virtualization of resources prevalent in the cloud, as it has to address various technical, legal, and organizational challenges typical to the cloud systems. While we do cloud forensics, the data to be inspected are cloud component logs, virtual machine disk images, volatile memory dumps, console logs and network captures. The dynamic nature of cloud computing allows abundant opportunities to enable digital investigations in the cloud environment.

We propose an approach to forensic investigation, using Log Information and VM Snapshots of the malicious activities in the cloud environment. To ensure the security of Log files we aim to apply Encryption algorithms to Log information, which will be helpful for further investigation.

KEYWORDS: Cloud computing, Cloud Service provider, Virtual Machine, Intrusion Detection.

I. INTRODUCTION

Cloud computing is the emerging technology for delivering information technology (IT) services to Internet-connected devices. It abstracts away the physical computer and communication infrastructure, and allows customers to rent, instead of own and maintain, as much compute capacity as needed [1].The cloud service model, although enabled by a number of technological developments, is primarily a business concept, which changes how businesses use and interacts with it.

Business houses migrate to the cloud after verifying the security mechanism of the cloud service provider's infrastructure. Though the security mechanisms employed are good, due to the huge revenue involved in cloud, it is an easy target for hackers, crackers, and other unethical online intruders. If the cyber-crime breaches the installed security system, then forensics can be used to find the evidence and prove the guilty before the court of law. Digital forensics is traditional computer forensic science which involves the process of seizure, acquisition, analysis and reporting the evidence in traditional OS images, USB drives and hard disks [2].

Digital forensics in the cloud environment comprises the stages: Identification, Collection, Examination/ Analysis and Reporting/ Presentation. Identification phase identifies the sources of evidence, Collection phase captures the actual evidence and the data related to it, Examination/Analysis phase examines, analyses the forensic data, Reporting phase is concerned with the presentation of collected evidence to the court of law. The challenge for digital forensics in the cloud is that we cannot seize the physical hardware which runs various applications in the cloud, as they are distributed across various geographical locations [4].

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 6, Issue 10, October 2017

Forensics and cloud: Traditional analytical model of digital forensics has been client-centric, investigator works with physical evidence devices such as storage media or integrated computer devices (e.g., smartphones). On the client (or standalone) device it is easy to identify where the computations are performed and where the results/traces are stored. Therefore, research has focused on discovering and acquiring every little piece of log and Timestamp information [8].

So we propose a system for identifying malicious activity in the cloud by using intrusion detection system in the form of Virtual Machine Snapshots and Log Information.

II. LITERATURE SURVEY

A critical assessment of the work has been done so far on Cloud Forensics to show how the current study related to what has already been done. Numerous companies are nowadays migrating to cloud due to greater economic issues. But for small and medium-sized companies the security of information is the primary concern. For these companies, the best alternative is to use managed service which is also known as outsourced service in which they are provided with the full package of service including antivirus software to security consulting and the alternative model that provides such outsourced security is known as Security as a service (SECaaS). Scientists and researchers together presented their latest ideas and findings on what the real world scenario is and what all efforts are made but it was found that despite being so much research work in the field of cloud forensic there is only a fraction part of the total work that has contributed to the wealth of the society. However, a cloud came into existence in the mid of 90's yet it is not taken up by everyone fully. There have been lots of works before in this field and variety of methods for the forensic analysis of cloud yet there is a huge room for improvement that needs to be carried forward into the research.

DeeviRadha Rani and Geethakumari G [1] proposed the technique of Forensic investigation of VM using snapshots as evidence that can be shown as a proof in front of the court of law. In that mechanism, software stored and maintained snapshots of running VM selected by the user which acted as good evidence. VM can be created by the user as per his choice from the physical machines that are available. Any cloud software similar to that of Eucalyptus instead of a request of a user takes the snapshots of the machines stores till terminated. Snapshots can be stored only till it reaches the maximum but when once the maximum is reached the snapshots which were taken long before gets deleted. So the huge storage management of snapshots of VM becomes difficult as it affects the performance of the system also, the author proposed a model in which VM was combined with an IDS. This helped to observe the destructive activities being performed between the VMs by thoroughly monitoring it. The basic idea behind this work was to store the log of destructive activities in the form of snapshots using the IDS placed in the system. Simultaneously, the CSP was asked for the logs of the doubtful VM and those logs were collected by the investigator. Investigator then works on those log files to obtain the evidence which can be helpful to the investigator.

BKSP Kumar Raju Alluri and Geethakumari G [2] presented a Model for the self-analysis of VM. They split the entire Introspection into three parts as follows. a) Analysing virtual machines by taking into consideration the swap space where the continuous monitoring of swap space is done. It provides the information about the current process of the VM. b) A self-analysis method for VM instances. In this three models were used, to collect as much accurate data evidence can be collected and reduce the semantic gap. But later, out of these three methods in-band method was proved to be less useful for live forensic as it modified the data at the time of collection phase. c) A Terminated Process based Introspection for Virtual Machines in Cloud Computing. This captured every process that was terminated and later was improvised to capture only the processes that were found doubtful. The proposed method for performing the digital forensic observation in Cloud on VM for introspection which addressed the issues related to the assembling of evidence. For resolving they made use of certain methods of introspection on VM. This work can be useful in current research if incorporated as a part of the investigation process.

Hubert RitzdorfNikolaos, Karapanos Srdjan Capkun proposed [3] Assisted deletion of related Content in ACM, 2014 Hubert and Karapanos have discussed a system which helps the user of that system to diminish the similar and associated files, contents of any project. This system did not affect the user or systems components in any sense as it

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 6, Issue 10, October 2017

was directed embedded with the system of the user itself. It starts functioning from user space and preserves the files along with its metadata. When they executed their work, realized that the resulting accuracy and the overhead were feasible. The results were appropriate to be used for the purpose of deployment. The aim of the system was to aid users by displaying all the associated files of the project to be diminished and it was successful in providing it. Deletion of content using assisted deletion of the content that is related was proposed here. The user was presented with all the associated files to be diminished securely organized manner. This aided user by maintaining the confidentiality of their data. This can help in current research also as it any system is providing facility to delete files this can be integrated.

Mr.DigambarPowar and Dr. G. Geethakumari [4] proposed a technique for Cloud Computing domain and that was named Digital Evidence Detection technique. Some conventional methods have been discussed in their work which has been used as a tool for performing forensic observations and those methods were used to learn and examine the behavior of the digital evidence in a virtualized environment called Cloud. Also, the feasible solutions are shown in which forensic practices can be performed in the virtual environment. Also, authors have introduced the feasible solution in which forensics can be practiced in the virtual environment. This work is a crucial stage as it leads to appropriate data evidence collection and presentation that can be an aid to the forensic investigator.

Mr.Chandrashekhar S. Pawar, Mr.Pankaj R. Patil, Mr.Sujitkumar V. Chaudhari[5] proposed mechanism for Providing Security and Integrity of Data Stored In Cloud Storage.The authors in their research work have tried to propose a solution to lessen the workload and simultaneously provide the integrity and security of the data which is kept on Cloud in a well-organized way. But as the data stored in the cloud is not easily approachable by the users, it becomes difficult to ensure its integrity. So, authors have proposed a technique which once combined with SLA after agreement with CSP and user allows user can test the integrity of data. Also, the author has worked for minimizing the computational overhead. They have performed encryption only for some bits out of the entire block of the file. As a result, at the side of client the overhead has lowered and thus the scheme has been more accepted by the users.

Saibharath S and Geethakumari G [6] have implemented a data collection and rendering mechanism for cloud through Hadoop file system using struts 2.0 MVC framework integrated with Hadoop and cloud, a web software tool has been successfully implemented to do cloud forensics. Pre-processing of the evidence files has performed through the log and VM disk drives clustering. It helps in minimizing the time of the forensic investigation. Using the correlation function between drives helps the investigators to perform cross-drive analysis. As future work, a soft clustering model with definite search strategy can be designed.

Curtis Jackson1, Rajeev Agrawal2, Jessie Walker3, William Grosky4 [7] proposed virtual environment for testing utilizing Proxmox, an open source virtualization management tool, and KVM, a virtualized environment. Initially, they have captured data from VM. In phase 2 investigation of the Virtual Machine Monitor has been done, which creates, observes and manages the virtual machine. In the last phase of this project, they have designed scenarios of cyber-attacks and data loss. These enacted scenarios are logged by the Virtual Machine Monitor. Comparing the datasets from the initial attack scenarios to the Virtual Machine Monitor's data, they have identified and verified the activity in the Virtual Machine Monitor's dataset. This data has been used to enhance our understanding of VMI how it does and should interact with the hypervisor. Using this dataset from the first phase, they have been able to identify activities for threats and normal activities.

Ting Sang[8] stated a log-based model for the cloud environment. The log-based model can help to reduce the complexity of forensic for nonrepudiation of behaviours on the cloud. However, it is totally not enough for the other kinds of digital forensics. What makes matters worse is that, till now, there are still no guidelines or standards for the cloud security. Most of the times, we modified the guidelines of traditional digital forensics to suit for cloud computing environment independently.

FilipoSharevski [9] presented an initial effort to describe the potential privacy implication that might arise during the cloud forensic investigation. Additionally, he has approached every dimension of the cloud investigation process with a

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 6, Issue 10, October 2017

set of preliminary recommendations that can greatly contribute to the formal definition of privacy requirements in the extremely complex cloud environment. The work presented here is generic in nature and can be easily extended to cover privacy aspects in different cloud service or deployment models against various types of cybercrimes, cyber-attacks or incidents that may have potential impact on the cloud entities' privacy, this work is an important step that not just contributes to the improvement of the cloud forensic process, but to the overall evolution of the digital forensic science.

Zhen Chen, Fuye Han, Junwei Cao, Xin Jiang, and ShuoChen[10] proposed cloud computing systems to explore the large volume of collected data from CNSMS to track the attacking events. An IaaS cloud platform has been constructed with Eucalyptus and existing cloud platforms such as Amazon EC2 and S3 were used for comparison purposes. Phishing attack forensic analysis as a practical case has presented and the required computing and storage resource has evaluated by using real trace data. The result shows that the proposed scheme is practical and can be generalized to forensic analysis of other network attacks in the future. This work has been supported by the National Key Basic Research and Development (973) Program of China.

All the existing approach to cloud computing varies with different providers and different service models and deployment models. Thus digital forensics in cloud varies according to the service and deployment models. In Infrastructure as a Service (IaaS) model, digital forensics is affable as compared with Software as a Service (SaaS) and Platform as a Service (PaaS) model. This is because of its limited control over infrastructure. The types of evidence collected for investigation vary with service models. In SaaS and PaaS, log information is collected as evidence and in IaaS, VM image is taken as evidence. We can get access to the physical devices in private deployment model but not in the public cloud. The work presented in this survey takes due care of the data which is kept on the cloud as it not only provides the integrity check but also security for the data as well. This lets us test the integrity at the moment of retrieving the stored data from Cloud.

Current methodologies of cloud computing have generated too many snapshots to prevent the malicious activities from attackers. Sometime attacker not only generates the malicious queries from GUI but also from command line interface. The existing approaches were unable to find some command line attacks at runtime. Some approaches generate snapshot in very high quantity even when the user follows the normal activities, so it leads to high time complexity issues for the investigator for manual verification. The next challenging work of this system to eliminate the high time complexity using on-demand snapshot generation when malicious activity has generated.

Based on this above research we are proposing a system to implement the proposed approach with multiple VMs. In summary, the work presented in this paper has been built on previous research to explore how the security of data stored on cloud relates to people's trust. While earlier work focused on data storage impacts people, we focus on its impact on the worldwide acceptance of cloud.

III. PROPOSED SYSTEM

There are too many systems which are used for attack detection and forensic IDS in the cloud environment. The traditional digital forensic process undergoes the following steps which can be incorporated in cloud forensics considering its different service and deployment models.

Identification: Identifies the sources of evidence, Collection phase captures the actual evidence and related data.
Collection: Collection is nothing but the physical acquisition of investigation related data
Examination: After collecting the desired large amount of data with the help of CSPs, these are to be processed through a combination of manual and automated processes too. The main motto of examining is to extract and assess data of the particular interest of the classified incident scene. The integrity must be preserved through this entire process.

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 6, Issue 10, October 2017

Analysis: All the relevant data are analyzed using suitable and legally justified techniques so that the proper suspected hosts or data can be identified through this investigation procedure. Investigators must be able to meet up with all queries those are raised during the presentation of the analyzed report to the court.
Reporting & Presentation: These are the final stages of any investigation process. The report must be comprised of all the details of this investigation process (explanation against what, why and how). The detail report is to be presented to the jurisdiction section with authenticity and accuracy without tampering the evidence which is the most crucial part of the investigation. For acquiring digital evidence the most widely used mechanism is to take snapshots of the events occurred. Snapshots can be restored sequentially using their time of creation to regenerate the crime incident.

Leading Cloud Service Providers like Eucalyptus are also giving a provision to take snapshots of the cloud events. In Eucalyptus the snapshots taken will be stored in the walrus component. It was noticed that the size of the snapshot will be the same as that of the original. Even though snapshots can be stored in the secondary storage, maintaining the huge store of the snapshot for each VM event will be difficult, time-consuming, and expensive and would degrade the performance. Also, the CSPs should have a mechanism to segregate and provide mappings as to which snapshot belongs to which VM. Through our approach, we propose to address this issue.

A.Cloud Service Provider: In our daily routine we use this cloud service without our notice like web-based email service, watching movies through the internet, editing documents, storing pictures etc. uses cloud computing on the back-end. Using such cloud technology we can design and create new applications, store and recover data, hosting the websites etc.Generally, cloud computing services are categorized into three types.

We will use amazon Web service as a cloud service provider.Amazon Elastic Compute Cloud (Amazon EC2) provides scalable computing capacity in the Amazon Web Services (AWS) cloud. Using Amazon EC2 eliminates your need to invest in hardware up front, so you can develop and deploy applications faster. You can use Amazon EC2 to launch as many or as few virtual servers as you need, configure security and networking, and manage storage. Amazon EC2 enables you to scale up or down to handle changes in requirements or spikes in popularity, reducing your need to forecast traffic.

B. VirtualMachine: In computing, a virtual machine (VM) is an emulation of a computer system. Virtual machines are based on computer architectures and provide the functionality of a physical computer. Their implementations may involve specialized hardware, software, or a combination.

C. Virtual Machine Monitor: VMM is the primary software behind virtualization environments and implementations.VMM facilitates the creation of VMs, each with separate operating systems (OS) and applications. VMM manages the backend operation of these VMs by allocating the necessary computing, memory, storage and other input/output (I/O) resources.

VMM also provides a centralized interface for managing the entire operation, status and availability of VMs that are installed over a single host or spread across different and interconnected hosts.

D. Intrusion Detection Systems:
Intrusion detection systems (IDS) are used to alert users of an attack. These systems monitor a specific system in order to gather and analyse information to determine whether there is a security threat. These security threats can range from something as being as someone mistyping a password to something malevolent as a Denial of Service attack.
In order to maximize the efficiency of checking for security threats, IDSs are specialized to handle specific components. For instance, there are network IDSs which specialize in monitoring network traffic for malicious activity [12].

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 6, Issue 10, October 2017

E. Snapshots:
A "snapshot" is capable of capturing a machine in its current state. They are typically used so that a user can easily back up current configurations of their IaaS machine before attempting to install software or reconfiguring software that may break the machine. This snapshot capability is offered by many hypervisors and CSPs.

F. Log-based Approach:
In the digital world, a log is a regular or systematic record of actions that object has taken or statuses that object have been. It is the most common component that is used in digital forensics.

Figure1. Shows the incorporation of Intrusion detection system in Virtual Machine and Virtual Machine Monitor. IDS are used to start the forensic process and start the gathering of data.VMM analyses the attacked VM when an attack is identified. This technique is called Virtual Machine Introspection (VMI). Snapshots and Log Files are used to capture the virtual machines under inspection and these snapshots and log files can be provided to Investigator as Evidences.

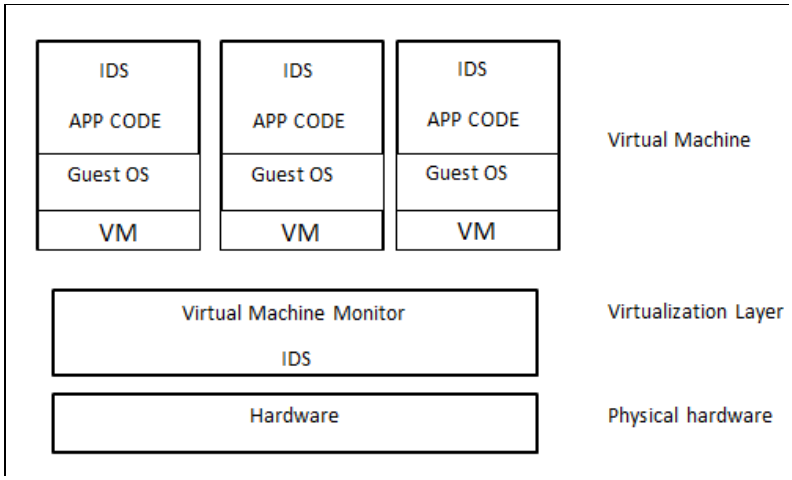

Figure 1: Placement of Intrusion Detection System in Virtual Environment

The Methodology of proposed system:

1. The proposed systemwill incorporate Intrusion Detection System on multiple VMs which will allow it to monitor itself and to detect malicious activity between VMs.
2. Users will make a request to create a VM by logging onto the Amazon AWS cloud portal after the VM is created and ready to use, the user will gain total control of VM as shown in Figure 2.
3. Malicious activities will be identified by IDS when users of that VM perform any activity like dos attacks, crashing server, cracking passwords, wrong OTP attack, and SQL Injection attack.
4. IDS will identify the suspected VM after it is identified to performing malicious activities, to collect proper and correct evidence, the suspected VM will be monitored for some more time.

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 6, Issue 10, October 2017

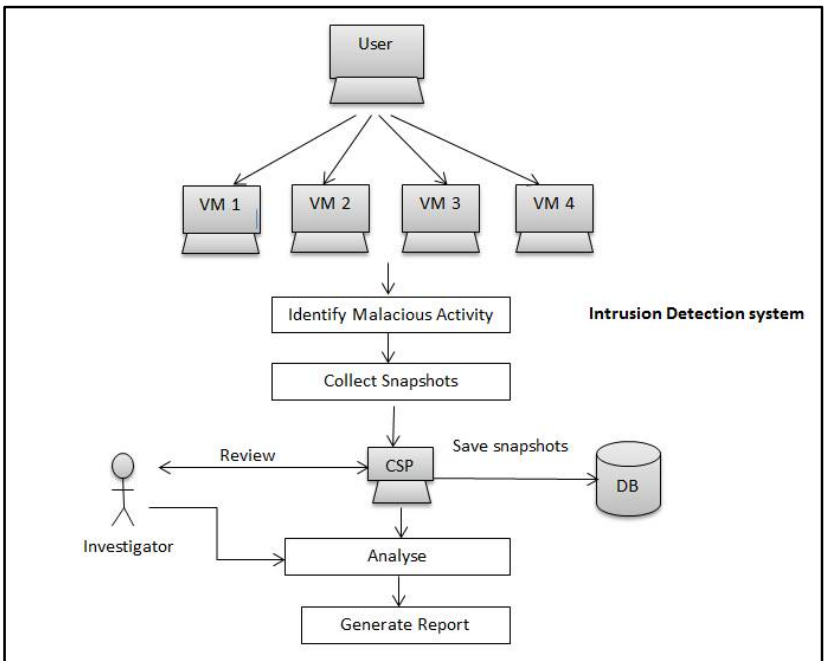

Figure 2: Proposed System

5. Simultaneously the CSP will request for log files of the suspected VM and the investigator will collect and processes the log files to obtain the evidence.
6. Once the investigator identifies the sources of evidence, the Ip-address of attacker VM will get blocked.
7. Thus the system will provide prevention mechanism to preserve confidentiality, integrity, and authenticity of other VMs; also VM evidence will be protected from contamination and tampering by providing encryption to log files.

IV. CONCLUSION

All the existing approaches for forensic investigations generate snapshot in very high quantity for all activities including normal user activities. Thus, time required for investigator is high in case of manual verification. Also maintaining huge storage of snapshots for each VM event will be time-consuming and expensive degrading the performance. The challengeis to eliminate the high time complexity using on-demand snapshot generation when malicious activity has generated.

In this paper, we have proposed the system to enable digital forensics in the cloud environment with respect to the performance by taking VM snapshots as evidence.We propose IDS system which will be residing on VM to identify the malicious activities by storing snapshots of malicious VM in persistent storageand improve the cloud performance in terms of size and time.

REFERENCES

[1] DeeviRadha Rani, G. Geethakumari "An Efficient Approach to Forensic Investigation in Cloud using VM Snapshots" International Conference on Pervasive Computing (ICPC), 2015.
[2] BKSP Kumar RajuAlluri, Geethakumari G"A Digital Forensic Model for Introspection of Virtual Machines in Cloud Computing" IEEE, 2015.
[3] Hubert Ritzdorf, NikolaosKarapanos, SrdjanCapkun "Assisted Deletion of Related Content"ACM, 2014.

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 6, Issue 10, October 2017

[4] Mr. DigambarPowar, Dr. G. Geethakumari "Digital Evidence Detection in Virtual Environment for Cloud Computing" ACM, 2012.
[5] Mr. Chandrashekhar S. Pawar, Mr. Pankaj R. Patil, Mr. Sujitkumar V. Chaudhari "Providing Security and Integrity for Data Stored In Cloud Storage" ICICES, 2014.
[6] Saibharath S, Geethakumari G "Cloud Forensics: Evidence Collection and Preliminary Analysis" IEEE, 2015
[7] Curtis Jackson, Rajeev Agrawal, Jessie Walker, William Grosky "Scenario-based Design for a Cloud Forensics Portal" IEEE, 2015.
[8] Ting Sang, "A Log-based Approach to Make Digital Forensics Easier on Cloud Computing"CPS,2013
[9] FilipoSharevski "Digital Forensic Investigation in Cloud Computing Environment: Impact on Privacy"
[10] Zhen Chen, Fuye Han, Junwei Cao, Xin Jiang, and Shuo Chen"CloudComputing-BasedForensicAnalysisforCollaborativeNetwork SecurityManagementSystem"
[11] AmitKumawat, Cloud Service Models,http://www.cmswire.com/cms/information-management/cloud- service-models - --iaas -saas-paas-how-microsoft-office-365-azure-fit-in-021672.php
[12] Cloud Tweaks, Cloud deployment Models,http://cloudtweaks.com/2012/07/4-primary-cloud-deployment-models/
[13] Amazon EC2 instances deletion in Cloud, https://aws.amazon.com/choosing-a-cloud-platform/?sc_channel=PS&sc_campaign=acquisition_IN&sc_publisher=google&sc_medium=cloud_computing_b&sc_content=sitelink&sc_detail=%2Bamazon%20%2Bclouds&sc_category=cloud_computing&sc_segment=choosing_a_cloud_platform&sc_matchtype=b&sc_country=IN&s_kwcid=AL!4422!3!92346737581!b!!g!!%2Bamazon%20%2Bclouds&ef_id=WKf9NAAAADDF7BAD:20170224152021:s
[14] Openstack, OpenStack command-line interface cheat sheet,http://docs.openstack.org/user-guide/cli_cheat_sheet.html