Indian Journal of Science and Technology, *Vol 9(17), DOI: 10.17485/ijst/2016/v9i17/93051, May 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Challenges of Digital Forensics in Cloud Computing Environment

Deevi Radha Rani[1]*, Sk. Nazma Sultana[2] and Pasala Lourdu Sravani[1]

[1]Department of CSE, VFSTR University, Vadlamudi – 522213, Andhra Pradesh, India; dradharani29@gmail.com, sravani93.11@gmail.com
[2]Department of IT, VFSTR University, Vadlamudi – 522213, Andhra Pradesh, India; nazma.cs@gmail.com

Abstract

Cloud computing is becoming most promising technology in recent days. It offers the scalable elastic services to many users at a time and it helps to access the resources quickly from cloud service provider. Digital forensics is part of computer forensics. Various challenges of cloud hinder the process of cloud forensics so that no standard framework for cloud forensics can be designed. This paper gathers the challenges and possible solutions. This paper also presents various challenges in every step of cloud forensics with probable solutions that can mitigate those challenges. This paper also proposed a solution for collecting snapshots as evidence in Eucalyptus Cloud and proposed effective framework for cloud forensics.

Keywords: Cloud Computing System, Cloud Forensics, Digital Forensics Process, Eucalyptus, Snapshot

1. Introduction

Cloud computing technology enables convenient, on-demand usage of computing resources with minimal management effort and service provider interaction[1]. The organization can save 37% cost if they outsource data centers to Amazon cloud[1]. Market research media surveys that the cloud computing is expected to grow at an 30% compound Annual Growth Rate (CAGR) and about to reach $271 billion by 2020[2].The virtualization and multi-tenant nature of the cloud gives the better usage of resources and these are main characteristics of cloud computing but these makes the main problems to cloud.

But in every new technology the security takes the best place if that technology is with good security and privacy adoption of that technology is very easy but some technologies like cloud computing, digital forensics and cloud forensics are more useful in today's world but with less security adoption of these technologies take much time. With distributed nature of the cloud the investigators face several challenges and these challenges are different from traditional digital forensics.

Digital forensics is part of computer forensics. The identification, collection, analysis and presentation of the digital evidence is termed as digital forensics process. Applying Digital forensics in cloud environment is called cloud forensics.

[3]proposes a definition for Cloud forensics in the cloud environment technically, organizationally, legally. It consists of a remote, virtual, network, live, large scale thin client, tick client and generates digital evidence, it involves interaction among cloud actors for facilitating investigations, and it implies multi-jurisdictional and multi-tenant situations. This paper discusses the challenges in each step of digital forensics in cloud computing environment.

In this paper section 2 discusses the challenges of cloud forensics, section 3 discusses current solutions of each phase of the forensic process in cloud, section 4 presents the proposed solution. Finally conclusion is given in section 5.

*Author for correspondence

2. Challenges of Cloud Forensics

Figure 1 presents the clauses of digital forensics in cloud. This section presents the challenges in every phase of cloud forensics.

2.1 Identification

The Identification phase mainly defines the purpose and process of Investigation. Identification of crime is the starting step in Digital Investigation Process model. Determining of a malicious activity that happen is simply identification step .The main thing here is how we say that the crime is happen? Traditionally in Digital Forensics the investigators identifies the crime in following ways

- If any Individual made any complaint
- By an anomalies detected by Intrusion Detection System
- At the time of a computer system audit

Identification of crime in cloud is difficult compare to traditional forensics identification. This phase arises in cloud by the complaint of any cloud user or cloud service provider reporting the unauthorized use of cloud recourses.

The intrusion detection system in cloud may identify any anomalies in the virtual machine, in cloud environment one of the virtual machine is monitor all the virtual machines in the cloud that virtual machine can act as Intrusion Detection System.

The evidence is fickle and frail in the context of cloud so e we need to propose new methods to efficiently use existing tools and hence making the evidence to be evaluated and isolated properly.

2.2 Challenges

i Accessing the evidence in logs: Distributed nature of the cloud make the identification of data difficult. The availability of log files depends on the servicing model of cloud. In SaaS, PaaS identification is more difficult because of limited access, identification is better in IaaS but not full access. Many researchers are find number of tools and procedures to identify the digital evidence, but cloud is volatile in nature investigators need to access the logs to identify the crime unfortunately many of the researchers are focused on identification of digital evidence only. Some solutions are purposed by many researchers[4].

ii Volatile data: Cloud is volatile in nature, volatile data means once the device is turned off all the data will erased similarly in cloud once the VM is turned off all data will lost unless the is stored at somewhere. RAM might contain valuable evidence including username, passwords and encryption keys. Due to the increase in the size of RAM and the increase in the use of data encryption, live data forensics is becoming increasingly[5].

iii Lack of control on the system Cloud is an on demand network access to a shared pool of resources and the resources are virtual in nature, exactly the physical location of the resources are never known to any cloud user. Only the CSP knows the physical location of the resources, the cloud investigator and the cloud user didn't get any control on the real system and it poses number of obstacles to the investigator when they carry out evidence acquisition[4]. Indeed, consumers have varied and limited access and control at all levels within the cloud environment and have no knowledge where their data are physically located[4].

iv Lack of customer awareness: In cloud all is under the control of CSP and the cloud user have little interaction sometimes no interaction with the CSP. A lack of CSP transparency along with little international regulation leads to loss of important terms regarding forensic investigations in the Service Level Agreement (SLA). This issue is applicable to all three service models[6].

Figure 1. Phases of cloud forensics.

2.3 Evidence Collection and Preservation

Evidence collection collects the evidence from identified sources of evidence. Collected evidence need to be preserved. Preserving data is maintaining data integrity original data is not to be changed till investigation completes. In traditional system the investigation process starts by seizing the hard disk of the system and taking the bit wise copy of the same maintaining integrity of the system. But in cloud, it is practically impossible because the evidence is untouchable and it is volatile in nature.

So the investigators and the researches need the better preservation methods. Some of the methods are proposed and are discussed later in this paper.

Challenges

i Data integrity: The investigators need to maintain the integrity of the evidence to preserve the integrity of the original data for cloud investigator its very difficult[7]. Data integrity is the difficult part in entire process of cloud forensics because the original data need not be changed up to the evidence is submitted in front of law[8]. To maintain the integrity of the evidence a piece of incident related information is listed in chain of custody register which included how, where and by whom the evidence was collected[9]. The evidence is valueless in front of law if the integrity of the data was missed[10]. Number of users are involved in the investigation process due to this the errors may occur in the preserving phase[10] says that data integrity and preserving is very difficult and challenging phase for the cloud investigator.

ii Cloud Instance Isolation: When crime event happen on cloud, cloud instance and evidence collected from cloud instance need to be isolated for digital investigation. Isolation prevents from possible corruption and contamination of collected evidence. Isolating cloud instance helps to preserves the integrity of the evidence collected from the cloud instance. [11] introduced new techniques to isolate instances on a cloud which are referred in our proposed approach.

iii Digital Provenance: It is an essential feature for forensic investigations which describes the history of a digital object. The secure provenance scheme[12] was proposed which performs digital forensics with trusted evidence in cloud environment. This scheme proves that cloud data evidence is acceptable in court of law.

iv Chain of custody: In the traditional investigation process the investigators need to establish and maintain the chain of custody. Chain of custody is the documentation of the gathered evidence, that how the evidence is collected by whom and when, and how the evidence is preserved and by whom. The investigator needed to maintain the proper chain of custody before it documenting. APCO gives the specific guidelines for documenting the evidence and maintaining the chain of custody. In traditional digital forensics the chain of custody starts: when the investigator took the physical device like hard disks into custody.

2.4 Examination and Analysis

In the Digital Imaging Process (DIP) model once the data is collected and preserved various examination techniques and several software tools are available to aid the investigators. FTK (Forensic Tool Kit) and Encase are widely used commercial forensic tool suites; another Open source tool is Sleuth tool kit. These all tools are used to perform filtering and pattern matching for searching the content or files or file types. By using these tools one can recover the data deleted or modified. In entire analysis phase the evidence need to be evaluated. The generated report supports the evidence help to regenerate the crime event. It is also possible to correlate evidence with cloud users. The evidence generated in the analysis phase is validated to compare with the alternative sources of evidence to confirm that the evidence is not altered. The examination and analysis phase of cloud forensics is similar to digital forensics examination and analysis phase.

Challenges

i Lack of available cloud forensic tools: Cloud is new technology cloud forensics is not known to even some regular cloud users also. Cloud forensics is thrust area of cloud, at present no specific tools for cloud forensics most of the cloud investigators are uses the digital forensics and network forensics tools together in cloud., but these are not enough cloud forensics is different from digital and network forensics at some point of investigation these tools are not sufficient in cloud. Many cloud researchers are start their research in cloud forensics and some tools are introduced but we need better than that tools.

ii Evidence correlation across multiple sources: In cloud one resource is shared by number of cloud users. Evidence also spread across multiple resources which bring in various problems for investigators.

iii Crime-scene reconstruction: Crime scene reconstruction is the crucial part in forensics process. Reconstruction of crime scene in cloud forensics is difficult and sometimes may be impossible to reconstruct the crime event if the VM terminates after committing of malicious activity.

2.5 Presentation

The gathered evidence in the digital investigation process is needed to be submitted in the court of law to prove the crime. For that the investigator submits a report with

summarized investigation process and explained conclusion. At the end of investigation the investigator need to present a report and it must be useful for cross- examination. The result report should be used by an organization to improve their security policy and must be documented for future investigation[8].

3. Current Solutions in Cloud

3.1 Identification

Accessing the evidence in logs: The cloud researcher Zaferullah proposed some simple and standard logging methods termed as log management system for generation and holding of logs for long time with that collect and correlates logs[13]. This solution was assessing in eucalyptus cloud environment. Eucalyptus is a LINUX based open source cloud tool. It is a set of virtualization technologies with in the single cloud to support the resources that are already virtualized. The snort, system logs and log analyzer are some log monitoring and analyzing tools are to monitor the eucalyptus behavior and logging. This logging information is useful to identify VMs controlled by single eucalyptus, time, attacker's IP address, type of browser used, number of HTTP requests and content. The set of results proved that if CSP may give the better log information Cloud forensics may go advance[14].

Another logging model is proposed by Sang. It was well in SaaS and PaaS only. This model does not require CSP support as it mainly focuses on cloud users. It improved the efficiency as well as it reduces the verification time. In PaaS 3rd party member supplies the logging information to cloud user and CSP[9].

The Marty devised frame work that defines the recovering logging information at the time of investigation in standard manner define when, where and what to log[15]. The frame work is simple to understand, but volatile data contain the real evidence it does not work with those data.

A frame work is encrypted logging model here the logging data is gathered and sends to the central logging server which control everything[16]. This mechanism prevents potential eves dropping who changes or views the content at the time of transmission. This frame work also suggests that CSP could provide read only permission to the logs like networks, process and access logs to get the wanted information from cloud service models.

Volatile data: A solution was proposed by[4]. The solution aims to provide persistent storage to clients for storing their data. The volume attached to client, can be used for data recovery and data safety[17]. This makes collection of evidence easier. CSP should provide this kind of service to the clients and this process should be standardized. But this method is confined to small and medium organizations because of cost issue.

Wegner and Brik suggest a frame work as a solution to volatile data problems[16]. This framework suggest that synchronization of volatile data and persistent storage is required. But this frame work is not providing practical implementation and procedures.

3.2 Collection and Preservation

A Trust Platform Module (TPM) is proposed to preserve the integrity and confidentiality of the data in the cloud[16,18]. While running the virtual instance, trusted log files, and the trusted deletion of data to customers provides the integrity. However TPM is not trusted as it allows modifying the process running without it being detected by the TPM[10].

Multifactor authentication methods[19] and cryptographic tunneling protocols such as a virtual private network (VPN) are used to authorize the client and ensure the confidentiality and integrity of the evidence[9]. Researchers are proposed encryption mechanism[20] because security is main issue in cloud, but it increases the burden on investigator by increasing the complexity of investigation process. But some advantages for investigators are public key Infrastructure (PKI) can be used to trace the victim; it also states that the SLA should contain users privacy data [4]. A frame work was proposed by Yan it says that image the relative records and files completely[21] for future purpose.

3.3 Examination and Analysis

Lack of cloud forensic tools: Black Hat developed an open source software and is launched in 2011 termed as OWADE (Offline Windows Analysis and Data Extraction). The software finds the website viewed by cloud user and also extract the information stored in the cloud, it renovate the Internet actions and explore for the online identities. But it works only for Windows XP drives this version is still in under development[22].

Encase and FTK are commercial forensics tools[18] these are digital forensics tool and are used for cloud

forensics also but Dykstra and some other cloud researchers suggests using of these tools are risk because of less security. FORST is cloud management plane developed by [23]. It acquires the evidence from API logs, virtual disks and guest fire wall logs. FROST is the first forensics tool built into IaaS model. It operates on cloud management plane but will not interact with the OS inside the guest VMs.

Crime-scene reconstruction: A method was developed by [21] that allows the investigators to replay the event of attack. It restores the system to the state of before attack by using snapshot.

4. Proposed Solution

In cloud among all the VMs one of the VM is selected as monitor VM which monitors all the VMs, if any VM is suspected as malicious VM snapshot of that suspected VM is taken and is stored in persistent storage not in volatile storage regularly all snapshots are stored in volatile storage once the VM terminate we loss snapshots it makes the investigation process more difficult ,for this reason we stores the snapshots in persistent storage but storing of all snapshots in persistent storage is difficult because the size of snapshot is equal to the volume of VM as this reason only the suspected VM snapshot is placed in persistent storage. We implemented this frame work in Eucalyptus private cloud as follows.

At present day, Eucalyptus is most widely used to setup private Infrastructure-as-a-Service cloud. It consists of i) Cloud Controller which is an entry point for end-users and administrators ii) Node Controller on all physical machines offering VM instances and controls VM instances iii) Cluster Controller to manage all Node Controllers iv) Storage Controller for storing users data and VM images v) Walrus allows users to store their persistent data[20]. Eucalyptus cloud uses Xen hypervisor, Dom0 is designated as Privileged VM and DomU as Guest VM. Privileged VM is not offered to any users, it is assigned to monitor the malicious activities and users. The volatile storage is assigned initially to all VM instances. If the user request for more space then volume is attached to VM instance. Snapshots cannot be created for volatile storage but snapshots are created from volumes which stores persistent data.

Create volume of size 10GB in specified availability zone: #euca-create-volume –z eucalyptus –s 10

Then storage controller creates a image file in /var/lib/eucalyptus/volumes

```
#ls –lsrth/var/lib/eucalyptus/volumes/
Total 1.5M
1.5M –rw –r –r -1 root root 10.1G Jul 15 10:15 vol-149D46EA
```

If the volume is not attached to an instance mount the volume file to a loopback and enable Logical Volume (LV)

```
losetup/var/lib/eucaplytus/volumes/vol-149D46EA
lvchange –ay/dev/<vg>/euca-vol-149D46EA
```

Logical volume management (LVM) is a mechanism for managing storage space. In LVM there are 4 layers Physical Volume, Volume Groups, Logical Volumes, File Systems on top of each other. Dom0 is assigned to monitor the other VMs for detecting malicious events. If crime event is notified, then CSP stores all the snapshots to persistent storage ie walrus. For taking snapshots.

Create new temporary file for taking snapshots of VG
$EUCALYPTUS/var/lib/eucalyptus/volumes/vol-149D46EA
/dev/loop10
Create PV of temporary files loopback
pvcreate /dev/loop10
Extend source volume VG
vgextend <vg> /dev/loop10
Take LVM snapshot of LV
lvcreate –snapshot –n lv-snap-123
Copy snapshot LV to file in $EUCALYPTUS/var/lib/eucalyptus/volume/snap-123
ddif=**/dev/vgX/lv-snap-123 of=/var/lib/eucalyptus/volumes/snap-123**
Remove snapshot LV
lvremove –f /dev/<vg>/lv-snap-123
Reduce VG
vgreduce <vg> /dev/loop10
Remove PV
pvremove /dev/loop10
Remove loopback
losetup –d /dev/loop10
Deactivate LV
lvchange –an /dev/<vg>/euca-vol-149D46EA
Remove volume's loopback
losetup –d <X>

Now entire snapshot is transferred to Walrus. The investigators can use Euca2ool to collect snapshot from

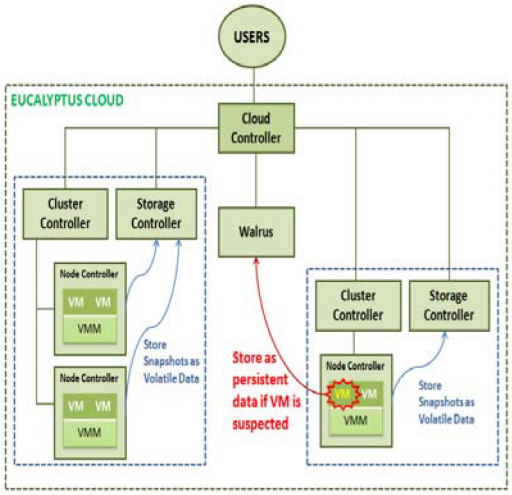

Figure 2. Proposed solution in eucalyptus cloud.

walrus for investigation. Euca2ools are command line tools that manage volumes and snapshots in Eucalyptus cloud[25].The collected snapshots can be used to regenerate the events from the volume.

5. Conclusion and Future Work

Various challenges of cloud computing environment hinder the process of cloud forensics. There is no standard framework for digital forensics in cloud computing environment. To have a standard framework, there is a need to gather challenges and possible solutions. This paper presents various challenges in every step of cloud forensics with probable solutions that can mitigate those challenges.

It is not possible to take the single snapshot of the suspected VM as evidence. The snapshot stored in persistent storage will not help to reconstruct the crime scene before crime happens, for this CSP need to maintain a table contents like snapshot VM id, at what time that snapshot taken and volume of that snapshot. Cloud faced a situation that it never get an opportunity for full acquisition it always an incomplete evidence.

6. References

1. Zargari S, Benford D. Cloud forensics: concepts, issues, and challenges. 2012 Third International Conference on Emerging Intelligent Data and Web Technologies; 2012. IEEE. pp. 236–43.
2. WordPress Used as Cloud Cover in New APT Attacks. Available from http://www.darkreading.com/attacks-breaches/dropbox-wordpress-used-as-cloud-cover-in-new-apt-attacks/d/d-id/1 140098. Accessed on 2015 Nov 25.
3. Dzombeta S, Stantchev V, Colomo-palacios R, Brandis K, Haufe K. Governance of Cloud Computing Services for the Life Sciences. IEEE Computer Society. 2014.
4. Alqahtany S, Clarke N, Furnell S. Christoph Reich2A forensic 10acquisition and analysis system for IaaS.
5. Almulla S, Iraqi Y, Jones A. A state-of-the-art review of cloud. 2014 ADFSL. 2014; 9:7–28.
6. Taylor M, Haggerty J, Gresty D, Lamb D. Forensic investigation of cloud computing systems. Netw Secur. 2011; 4–10.
7. Dykstra J, Sherman AT. Acquiring forensic evidence from infrastructure-as-a-service cloud computing: exploring and evaluating tools, trust, and techniques. Digit Investigate. 2012; 9:S90–8.
8. Reilly D, Wren C, Berry T. Cloud computing?: pros and cons for computer forensic investigations. Int. J. Multimed. Image Process. 2011; 1:26–34.
9. Damshenas M, Dehghantanha A, Mahmoud R, Shamsuddin S. Forensics investigation challenges in cloud computing environments. cyber security. 2012 International Conference on Cyber Warfare and Digital Forensic (CyberSec); 2012; Kuala Lumpur. pp. 190–4.
10. Zawoad S, Hasan R. Digital Forensics in the Cloud. 2013.
11. Yan C. Cybercrime forensic system in cloud computing. Proceedings of 2011 International Conference on Image Analysis and Signal Processing, IASP 2011. pp. 612–3.
12. Zawoad S, Hasan R. Cloud Forensics: A Meta-Study of Challenges, Approaches, and Open Problems. 2013. pp. 1–15.
13. Zaferullah Z, Anwar F, Anwar Z. Digital forensics for eucalyptus. 2011 Frontiers of Information Technology; 2011; Islamabad. pp. 110–6.
14. Wolski R. Available from https://www.usenix.org/conference/lisa09/eucalyptusopen-source-infrastructure-cloud-computing. 16/01/2016.
15. Marty R. Cloud application logging for forensics. Proceedings of the 2011 ACM Symposium on Applied Computing—SAC '11; 2011. pp. 178.
16. Birk D, Wegener C. Technical issues of forensic investigations in cloud computing environments. 2011 Sixth IEEE International Workshop on Systematic Approaches to Digital Forensic Engineering; 2011; Okland. pp. 1–10.
17. Kirubakaramoorthi R, Arivazhagan D, Helen D. Survey on Encryption Techniques used to Secure Cloud Storage System. Indian Journal of Science and Technology. 2015; 8(36):1–7.
18. Delport W, Olivier MS, Kohn M. Isolating a cloud instance for a digital forensic. ISSA. 2011.
19. Senthil KP, Kamal ARNB. Optimal Integrity Policy for Encrypted Data in Secure Storage using Cloud Computing. Indian Journal of Science and Technology. 2016; 9(8):1–10.

20. Kirubakaramoorthi R, Arivazhagan D, Helen D. Survey on Encryption Techniques used to Secure Cloud Storage System. Indian Journal of Science and Technology. 2015; 8(36):1–7.
21. Geethakumari G, Belorkar A. Regenerating cloud attack scenarios using LVM2 based system snapshots for forensic analysis. Int J Cloud Comput Serv Sci. 2012; 1:134–41.
22. The Eucalyptus Open-Source Private Cloud. Available from http://www.cloudbook.net/resources/stories/the-eucalyptus-open-source-private-cloud.
23. Dykstra J, Sherman AT. Design and implementation of FROST: digital forensic tools for the openstack cloud computing platform. Digit. Investig. 2013; S87–95.
24. Zawoad S, Hasan R. I Have the Proof?: Providing Proofs of Past Data Possession in Cloud Forensics. 2012.
25. Sabout N, Parthiban L. SecAuthn: Provably Secure Multi-Factor Authentication for the Cloud Computing Systems. Indian Journal of Science and Technology. 2016; 9(9):1–18.