DIGITAL FORENSIC INVESTIGATION IN AWS CLOUD USING SNAPSHOT

[1]Neeraj, [2]Dr. Rajesh Yadav

[1]M.Tech Scholar, Department of C.F.I.S, G.I.T.A.M, Kablana, Jhajjar, MDU, Rohtak, Haryana,
[2]Associate Professor, Department of C.F.I.S & C.S.E, G.I.T.A.M, Kablana, Jhajjar, MDU, Rohtak, Haryana

Abstract : AWS cloud computing has recently emerged as a technology to allow users to access organization, storage, deployment and software environment based on a pay-for what-they-use model. Outdated digital forensics cannot handle the dynamic and multi-tenant nature of the AWS cloud environment as it has to address various technical, legal, and organizational challenges typical to the AWS cloud systems. The dynamic nature of AWS cloud computing allows abundant opportunities to enable digital investigations in the AWS cloud environment. This paper addresses the challenges of digital forensics in the AWS cloud environment and existing solutions to ease some of the challenges. We propose an efficient approach to forensic investigation in AWS cloud using Virtual Machine (VM) snapshots.

IndexTerms **- AWS cloud Computing System, Digital Forensics, Forensics Challenges and Virtual Machine (VM).**

I. INTRODUCTION

AWS cloud computing has recently emerged as a technology to allow users to access infrastructure, storage, software and deployment environment based on a pay-for-what-they-use model. Digital forensics in remote, ubiquitous provider controlled AWS cloud computing systems is difficult when we compared to outdated digital forensics. Criminal use of AWS cloud computing is an impending possibility as AWS cloud becomes omnipresent. Likewise, the need for digital forensic analysis of AWS cloud computing environment and applications has become customary.

As in the case of outdated computer forensics, digital forensics in the AWS cloud environment also covers the steps: Collection, Identification, Examination/ Analysis and Reporting/ Presentation [1]. Identification phase identifies the sources of evidence, Collection phase captures the actual evidences and related data, Examination/Analysis phase examines and analyses the forensic data, Reporting/ Presentation phase is afraid with the presentation of collected evidences as per the court of law. The technical, legal, and organizational dimensions of AWS cloud forensics are challenging digital investigators to cope up with current developments [2]. The dynamic nature of AWS cloud provides abundant chances to enable digital investigation in AWS cloud environment. The challenge for digital forensics in AWS cloud is that we cannot seize the physical hardware which runs various applications in AWS cloud, as they are distributed across various geographical locations. The rest of paper is organized as trails. In section II we review the challenges and existing solutions for digital forensic investigation in AWS cloud environment. In section III we discuss our proposed digital forensics approach for AWS cloud. Section IV summarizes the paper and explores future research scope of AWS cloud forensics.

II. BACKGROUND AND LITERATURE REVIEW

2.1 Challenges of Digital Forensics in AWS cloud environment

AWS cloud Computing is a technology which orchestrates with virtualization. Every AWS cloud environment would have the administration and management of its services performed by an entity called the AWS cloud Service Provider (CSP). The approach to AWS cloud computing varies with different providers and different service models and deployment models. Thus digital forensics in AWS 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 its partial control over the infrastructure. There are many way for evidences collected for investigation but they vary with service models of provider. In Software as a Service and Platform as a Service (PaaS), logs information are collected as evidence and in IaaS, Virtual Machine (VM) image is taken as evidence. We can get access to the physical devices in private deployment model but not in the public AWS cloud. The extent in which the cloud Service Provider helps to access the sources of evidence is important then there can be two possibilities: first one is the cloud Service Provider is involved in the investigation, second, the cloud Service Provider does not involve. If the CSP does not involve, then the investigators can depend on the law of the land to seek cloud Service Provider's cooperation. If cloud Service Provider agrees to cooperate [3].

The investigator could not rely on the cloud Service Provider for acquiring evidence. If the CSP compromises, then the evidence provided by it may not be unaffected. It may turn out that the data may have been injected by a malicious user.

Many tools exist for performing outdated digital forensics. These tools access the devices physically and perform forensic investigation analysis. Due to inaccessibility to physical devices in AWS cloud, it lacks the tools required for gathering forensic evidences. Using automated forensic tools in the AWS cloud environment is not a feasible option [4]. Outdated digital forensics allows investigators to seize tools to recover the data. In AWS cloud, seizing physical storage is impossible as data of the customers are diverse among various geographical areas across the world. Even to take control over the evidence changes in different service models which makes evidence acquisition is more challenging.

When users request for VM, the associated data will be stored in AWS cloud data centers based on the users' request. Once a user terminates the device access or VM, the associated data will be lost. If virtual device performs malicious activity, he can terminate the VM losing the unpredictable data which makes forensic investigation impossible. Terminating Virtual device does not allow reconstruction of crime scene which would have helped in finding the user responsible for the attack [4]. Birk et al. addresses an issue concerning that the evidence found in AWS cloud computing system is volatile [5].

There is a significant increase of digital devices using the AWS cloud but limited power is given to investigators to obtain the data legally. The Service-level Agreement (SLA) may be not mention the terms and conditions regarding the role of the cloud Service Provider in the investigation and responsibility of the CSP during a crime incident. Also the cloud Service Provider may not have maintained the log files and logging mechanisms which are really useful for identifying malicious activities or malicious data. Collecting these log files for investigation is difficult because the investigator needs to depend on the service provider. Challenges in acquiring log files from the Cloud service provider have been discussed by many researchers [6], [7].

In the AWS cloud environment, sources of evidences can be the client system, the network layer, the virtual AWS cloud instance or servers etc [5]. Evidence data spread across different geo locations under various controls. During evidence collection the AWS cloud instance may get data from one control and store the data in another authority; then the investigator cannot violate the laws to access the evidence and hence chain of custody is difficult to preserve [2]. Maintaining chain of custody in AWS cloud environment is becoming challenging. Existing solutions various solutions exist which can ease few challenges of digital forensics in AWS cloud environment.

2.2 Virtual Machine Introspection

Virtual Machine Monitor (VMM) or a Virtual device running under the VMM analyse the attacked VM when attack is identified. This technique is called Virtual Machine introspection (VMI) and was first introduced by Garfinkel and Rosenblum. Malicious events can be identified by performing Virtual Machine Introspection which is the technique of examining a running VM from either another VM not under examination or from the hypervisor (hyper-V). Proposed hypervisor forensics and presents the possibility of obtaining evidence from hypervisors to perform digital forensics. Live Forensic analysis is also can be done on the targeted device using open-source VMI library and Xen Suite. Virtual Machine Introspection is suggested as the most practical approach to identify the malicious VM. If the intrusion detection system resides on the host, it may be susceptible to attack and if intrusion detection system resides in the network it is more resistant to attack. A virtual machine introspection (VMI) based approach to intrusion detection (ID) is proposed where the Intrusion Detection System (IDS) is outside the host for good attack resistance. In the authors proposed the use of Forensic Virtual Machines (FVM) to analyse memory space of other virtual devices. FVMs are small virtual machines which can monitor other VMs to find symptoms in real time via Virtual Machine Introspection and investigation.

III. Digital Provenance

Digital provenance is a essential feature for forensic investigations which describes the history of a digital object and way to investigate. Proposed the secure provenance scheme which performs digital forensics with trusted evidence in AWS cloud environment. This scheme proves that AWS cloud data evidence is acceptable in court of law. In researchers identified four properties that are crucial for provenance systems and introduced protocols to store data provenance using AWS cloud services. Also provenance is accessible as a layer on top of AWS cloud. Implementing secure provenance in the cloud environment increased the importance of the data on AWS cloud. Focus on enhancing value of AWS cloud data to users using provenance.
In principles of provenance are applied to forensics in AWS cloud and bring in data provenance to track the history and access of AWS cloud objects.

3.1 Isolating AWS cloud Instance

When crime incident happen on AWS cloud, AWS cloud instance and evidence collected from AWS cloud instance need to be isolated for digital investigation. Isolation prevents from the possible corruption of collected evidence. Isolating AWS cloud instance helps to preserves the integrity of the evidence collected from the AWS cloud instance. Introduced new techniques to isolate instances on AWS cloud which are referred in our proposed approach.

3.2 Regeneration of Events

For Acquiring digital evidence is the most widely used mechanism to take snapshots of the events occurred. Snapshots can be restored sequentially by using their time of creation and regenerate the crime incident. Belorkar proposed a new method which is use to regenerate crime events with continuous snapshots. Leading cloud Service Providers like AWS, Openstack are also giving a provision to take snapshots of the AWS cloud events. In AWS the snapshots taken will be stored in the storage to all of the virtual machines 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 to the secondary storage, maintaining huge store of snapshot for each VM event will be difficult,

time consuming, expensive and would degrade the performance. Also the service providers should have a mechanism to segregate and provide mappings as to which snapshot belongs to which virtual device. Through our approach we propose to address this issue and finalise the issue.

3.3 Forensics-as-a-Service (FAAS)

Variables AWS cloud computing system allows everything to be delivered as a service. Dener purposed the technique of Forensic as a Service (FaaS) to utilize AWS cloud resources like storage and processing power. Its demonstrate that the Forensic as a Service is capable and handling large amount of data set efficiently. Dykstra also explored that forensics-as-a-service is one of the methods for acquisition of digital evidence and documents.

3.4 Log model

Logging is a challenging issue in AWS cloud computing systems and becoming prevalent in all service models. Ting presented that some kind of forensics can be made a little easier on AWS cloud if logging ability is improved and proposed a log model that suits for SaaS and PaaS [7]. In SaaS logs can be used locally and synchronously to verify the actions on AWS cloud providing SaaS without interacting with AWS Service Provider. To use the proposed log model to PaaS AWS cloud there is need to depend on AWS cloud service Provider to provide log module to the third-party. The proposed model based on the logging may ease the challenges of the forensics for non-repudiation behaviors in AWS cloud model.

IV. FORENSIC INVESTIGATION USING VIRTUAL MACHINE SNAPSHOTS AS EVIDENCE

AWS service providers provide various types of services to users, few users from specific organization frequently use the same kind of service based on pay-per-what-they-use and these service some providers provide a free trial period with unlimited bandwidth and storage capacity in which gives users get an opportunity to perform malicious activities. Malicious users can steal the sensitive and confidential information from any cloud users which in turn affect the trust of the cloud service provider. AWS cloud necessitates protection from these malicious activities and Provider should have a provision to use either introspection or Intrusion Detection System (IDS) to monitor customer VMs and detect malicious activity or impacted malicious user.

Users can create virtual device of their choice from the available physical machines with same configuration. In spite of users request cloud software like AWS, OpenStack generates snapshots of a running VM continuously and stores it till the VM terminates. Max. no. of snapshots can be saved for a specific VM allotted, if maximum is reached older once are deleted. In a AWS cloud environment snapshots are rich sources of evidence for digital investigation and can be regenerate the events. Storing and managing huge storage of VM snapshots is difficult to maintain. Snapshots can decrease the performance of VM based on how long the snapshots are stored and how much it changed from the time previous snapshots are taken.

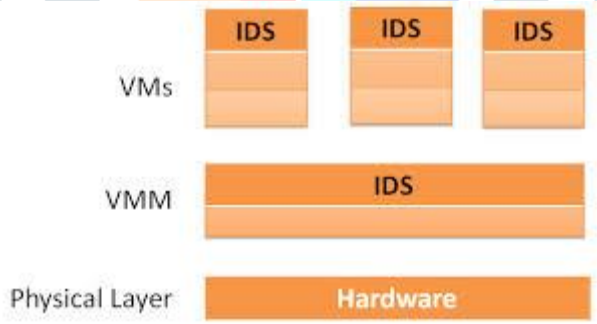

Fig. 1. Incorporating IDS at VMs and VMM

Malicious activities are identified when users of that VM perform any activity like excessive access from location, upload malware to a number of systems in the AWS cloud infrastructure, intense number of downloads and uploads in a short period of time, launch dynamic attack points, cracking passwords, decoding / building web tables or rainbow tables, corruption or deletion of sensitive data, malicious data hosing, altering data, executing botnet commands. Our proposed model incorporates Intrusion Detection System (IDS) on VMs which allows it to monitor itself and on VMM to detect malicious activity between VMs . Fig. 1. shows that Intrusion Detection Systems (IDS) are incorporated in all the VMs and VMM for monitoring malicious activities. Deploying, managing and monitoring the Intrusion Detection System is done by AWS service provider. The idea of the proposed model is that the service provider stores snapshots of a VM whose activities are identified as malicious by an intrusion detection system (ids). Simultaneously the AWS service provider should be requested for log files of the suspected VM and the investigator collects and processes the log files to obtain the evidence and digital data in form of evidence. To collect proper and correct evidence, the suspected VM should be monitored for some more time after it is identified to be performing malicious or any suspicious activities.

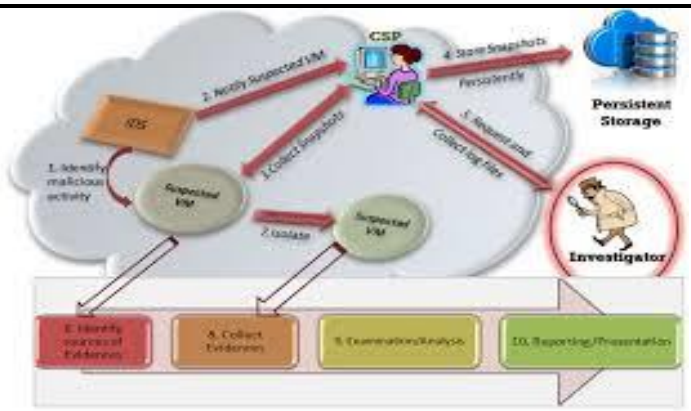

Fig. 2. A digital forensics using virtual device snapshots as digital evidences

In this approach investigator collect the logs file from the service provider and analysis the logs files of different devices to reach a valuable result.

The more time the suspected virtual device is monitored more then it can be sure of the possibility of malicious behavior. Once the investigator identifies the source of evidence and select the device, the suspicious VM is moved to other nodes to preserve confidentiality, integrity and authenticity of other VMs parameter. By moving or isolating, VM evidence can be protect from impurity and damaging. Delportetal introduced a new techniques to isolate VM instances on the cloud to be investigated for forensic. After that isolating the suspected device, the investigators can collect the evidence or data. After all this the evidence can be analysed using forensic tools and presented it to court of law. Fig. 2. Shows the approach to perform forensic investigation using snapshots as evidence.

Our experimental set up comprises AWS private cloud where 3 VM instances are created namely m1.small, m1.large, m. xlarge. Every virtual instance is attached with 70GiB of volume space. Snapshots of all instances are taken when change in source data is identified and are stored in a storage device. Snapshot image size would be same as the volume size of the device. Suppose m1.small instance has 3 snapshots, the total snapshot of image size is 210GiB and the total time taken to acquire each image is approx. 4.19hrs and image verification time is approx. 0.21hrs. In the normal scenarios for 3 instances with 3 snapshots it would require 630GiB of space in walrus and time taken is approx. 39.6hrs. The space for snapshots and time to acquire the snapshot would be reduced exponentially if the impacted device is identified and those virtual device snapshots are taken and stored persistently.

V. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a novel approach to enable digital forensics in the AWS cloud environment with respect to performance by taking virtual machine snapshot as evidence or digital data. The approach incorporates intrusion detection system (IDS) in VM and VMM to identify the malicious VM and improves the AWS cloud performance in terms of size and time by storing snapshots of malicious VM. The proposed approach takes snapshots from suspected VMs and stored in persistent storage device, hence improves the performance of AWS cloud.

Our future work will be implement the proposed approach with multiple VMs and its snapshots. Also, we plan to explore the implications of acquisition of data and evidence from AWS cloud VMs and develop a framework for digital forensics in AWS cloud IaaS.

REFERENCES

[1] B. Martini and K.K. R. Choo, "An integrated conceptual digital forensic framework for cloud computing," for Digital Investigation, vol. 9, no. 2..

[2] K. Ruan, J. Carthy, T. Kechadi, and M. Crosbie, "A cloud Forensics," Advances in Digital Forensics VII, vol. 361, no. A Advances in Information and Communication Technology

[3] J. Dykstra and A. Sherman, "Understanding issues in AWS cloud forensics: Two hypothetical case studies," Journal of Network Forensics, vol. b, no. 3, pp. 19–31, 2011.

[4] D. Reilly, C. Wren , "A cloud computing: Pros and cons for computer forensic investigations," International Journal Multimedia and Image Processing (IJMIP),
vol. 1, no. 1, pp. 26–34, March 2011.

[5] D. Birk and C. Wegener, "Technical issues of forensic investigations in AWS cloud computing environments," Systematic Approaches to Digital Forensic Engineering, 2011.

[6] R. Marty, "AWS cloud application logging for forensics," in proceedings of the 2011 ACM Symposium on Applied Computing. ACM, 2011, pp. 178–184.

[7] Ting, S.: A Log Based Approach to Make Digital Forensics Easier on AWS cloud Computing. In: 2013 Third International Conference on Intelligent System Design and Engineering Applications (ISDEA). IEEE (2013).