CLOUD FORENSICS: INTRODUCTION AND CHALLENGES IN RESEARCH

Prajakta N. Sonone[1], Department of CSE,Government College of Engineering, Amravati, India, sonu11sonone@gmail.com

Pushpanjali Chauragade[2], Department of CSE, Government College of Engineering, Amravati, India. pushpanjalic3@gmail.com

Abstract— Cloud computing is a new model which provides the utility services for shared virtualized resources. It is visualized that in future, Cloud computing can offer everything as a service (EAAS). Cloud Service Provider (CSP) makes infrastructure, platform and software services available over the Internet flexibility at a lower cost. Cloud computing and digital forensics is both developing topics and researching these topics requires an understanding of the main aspects of both cloud computing and digital forensics. Cloud forensics is an approach that attempts to investigate and analyze cloud security threats. It will ensure that attackers will be more cautious to avoid prosecution for their illegal actions. It acts as a deterrent, reducing network crime rate and improving security. The paper aims to provide a better awareness of cloud forensics, understand some of the proposed frameworks and identify the research gaps and challenges. The significance of this work is that it presents the state-of-the-art in cloud forensics, which will be very much useful for security practitioners and researchers.

Keywords- cloud computing, digital forensics,distributed computing,virtualization,cloud forensics, phases, challenges

1. INTRODUCTION

Cloud computing allows computer users to conveniently rent access to fully featured applications, to software development, to computing infrastructure assets such as network-accessible data storage and processing and deployment environments. Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [1]. A style of computing in which massively scalable IT related **capabilities are provided "as a service"** using Internet technologies to multiple external customers [2]. Five essential characteristics of cloud computing are on-demand self-services, broad network access, resource pooling, rapid elasticity, and measured service. Service models of cloud computing offer application software as a service - SaaS, platform (operating system and database) as a service - PaaS, and infrastructure (storage, network and compute) as a service - IaaS. There are four deployment models: public, private, community, and hybrid cloud [1]. The services of cloud are extended over the Internet by the cloud service provider (CSP) who maintains computer systems called data centers (DC). These data centers are placed in clusters and have large storage capacity over storage area networks (SAN). The CSP takes care of the maintenance of infrastructure and software and makes them available to users at a cost. The metering, logging and accounting procedures are handled by the CSP. The clients do not maintain computing facilities instead they have access to the data center over the Internet connectivity on anywhere any time basis [3]. The hypervisor technologies are used to compose virtual machines on demand. These resources are dynamically managed and provisioned over the web services resource framework. The CSP supports distributed identity and trust management, persistence and parallelism of data. The leading service providers at present are Amazon EC2 [4], Google AppEngine [5], Microsoft Azure [6], IBM SmartCloud [7], Ubuntu Enterprise Cloud [8] and many others. They offer computation and data storage services at low cost and can easily be scaled to match the user requirements. Eucalyptus [9] is widely used cloud computing software platform for private IaaS cloud. OpenNebula [10] offers the most feature-rich, flexible solution for the comprehensive and complete management of virtualized data centers. CloudSim is a framework for modeling and simulation of cloud computing infrastructures and services. It supports a virtualization engine and services model to field an application. This application can be tested in a repeatable and controlled environment and be tuned for performance, speed, availability, and other measuring parameters before setting it up on a real cloud computing environment. With the growing popularity of cloud computing, the concerns about security and compliance are also growing. Cloud computing offers tremendous reduction in operation cost. However it has also unfortunately introduced a set of new and unfamiliar risks. Cloud computing gives clients the ability to access computing resources without having to necessarily own the physical infrastructures. Top threats to cloud computing are abuse and nefarious use of cloud services, insecure interfaces and APIs, malicious insiders, shared technology issues, data loss or leakage, account or service hijacking and unknown risk profile. Cloud Forensics involves post attack investigation in order to

know the source of the attack and collect evidence. There are three sources from which evidence can be extracted: the client system, the network layer and cloud service **providers' management server.**

The paper aims to provide a better awareness of cloud forensics by bringing out the differences with cloud security. We also explain the place of cloud forensics under the larger umbrella of digital forensics. We also present a detailed survey of some of the proposed frameworks for cloud forensics. The techniques proposed for some of the phases in cloud forensics are also discussed. The research gaps and challenges are identified and explained.

2. CLOUD SECURITY

Cloud Security has to be understood before attempting to examine the intricacies of Cloud forensics. We discuss the important issues of Cloud Security in terms of architecture and the attacks on cloud services.

2.1 Cloud Architecture

Architectures are useful for understanding how various recommendations come together to provide a complete solution. A whitepaper, jointly developed by VMware and Savvis suggests enterprises interested in cloud computing to consider the reference architecture. The following security components are to be placed.

1) Security profile must be defined at each level.

2) CSP must keep the infrastructure and software behind a demilitarized zone (DMZ).

3) Operating Systems and Virtualization must be handled behind the DMZ on a CSP.

4) CSP must handle resource provisioning by separating and isolating VM resources.

5) Network Security must be provided through router ACLs, perimeter firewall or web application security.

6) CSP must provide access paths to physical servers that have permissions for the desired functionality.

7) Security authentication, authorization and auditing (AAA) must be provided by the CSP.

2.2. Attacks on Cloud Services:

The following are some of the attacks in the cloud Computing environment:

1. Wrapping Attack: A SOAP (Simple Object Access Protocol) message is generated when a user makes a request from his Virtual Machine to the browser. The request is directed to the web server. A wrapping attack is done by duplication of the user account and password in the log-in phase so that the SOAP messages that are exchanged during the setup phase between the Web browser and server are affected by the attackers

2. Malware-Injection Attack: The attacker creates a normal operation, like deleteUser, and embeds in it another command, such as setAdminRight. When the user request is passed to the server, it discloses a user account to the attacker rather executing the command to delete an user account.

3.Flooding Attack: Attacker generates bogus or malicious data, which could be resource requests or some type of code to be run in the application of a legitimate user, **engaging the server's CPU, memory and all other de**vices to compute the malware requests. The servers finally end up reaching their maximum capacity, and thereby offload to another server, which results in flooding.

4. Browser Attack: It is committed by sabotaging the signature and encryption during the translation of SOAP messages in between the web browser and web server, causing the browser to consider an adversary as a legitimate user and process all requests communicating with web server.

5. Insecure Interfaces and APIs: Cloud computing service providers expose a set of software interfaces or APIs that customers use to manage and interact with cloud services. Reliance on a weak set of interfaces can expose an organization to a variety of security issues related to confidentiality, availability, and password integrity.

6. Malicious administrators: Cloud computing as a process is governed, managed, and maintained by site administrators. By default, they hold the key to managing all the data, files, and privileged company resources. As revenge, or for other reasons, administrators may end up spreading, or allowing privileged information to leak.

7. Data Stealing: System administrators stealing any volume of data without leaving a trace is one of the biggest overlooked security holes in virtualized data centers. Three simple steps are login as an administrator on the hypervisor, create a replica of a virtual machine and mount the disk image onto the hypervisor and lastly delete the original copy.

8. Data Leakage*:* Data leakage is the movement of data from one customer to another. The data leakage problem comes when a customer deletes their drive and then a new customer creates a new drive. The areas on the physical disks used for the old and new drives can overlap. Its therefore possible for the new customer to try and image off previously written data from other customers.

3. CLOUD FORENSICS

Forensics is a formal and proven approach to the gathering of evidence and processing of a crime scene. The conclusions are used in the court of law.

3.1 Digital Forensics

Digital forensics is defined in DFRWS 2011 as "the

use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence derived from digital sources for the purpose of facilitation or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions Shown to be disruptive to planned **operations."**

4.2 Computer Forensics

Computer forensics is related to the forensics of computer components and their content. The field of computer forensics attempts to narrow the search for evidence to the computer itself, the content on the computer and devices attached to the computer. The processes of the model are to find useable evidence immediately, identify victims at acute risk, guide the ongoing investigation, identify **potential charges and accurately assess the offender's** danger to society.

4.3 Network Forensics

The Network forensics is the science that deals with capture, recording, and analysis of network traffic for detecting intrusions and investigating them. The concept of Network forensics deals with data found across a network connection mostly ingress and egress traffic from one host to another. Network forensics tries to analyze traffic data logged through firewalls or intrusion detection systems or at network devices like routers and switches [15]. Network forensics involves monitoring network traffic and determining if there is an anomaly in the traffic and ascertaining whether it indicates an attack. If an attack is detected, then the nature of the attack is also determined.

Network forensic techniques enable investigators to track back the attackers. The ultimate goal is to provide sufficient evidence to allow the perpetrator to be prosecuted. Steps in network forensics are preparation, detection, incident response, collection, preservation, examination, analysis, investigation, presentation.

4.4 Cloud Forensics

Cloud forensics is digital forensics applied on cloud environment. Cloud forensics is a subset of network forensics as a cloud runs on a network and consists of network equipment. Cloud forensics also entails computer Forensics as a cloud consists of nodes that are computers. Cloud forensics ties computer and network forensics together. Network forensics deals with forensic investigations of networks. Cloud computing is based on broad network access and follows the main phases of network forensics with techniques tailored to cloud computing environments. Network Forensics deals with data found across a network connection mostly ingress and egress traffic from one host to another. Network forensics involves monitoring network traffic and techniques enable investigators to track back the attackers. Cloud forensics involves gathering information from cloud environment for the purpose of investigation, while Cloud security is a term used to protect a cloud environment from intrusions and maintaining privacy [15].

4. CONCEPTS IN CLOUD FORENSICS

In addition to the potential source of evidences, there are several concepts such as cloud crime types and where to perform investigation in/on cloud that can significantly assist researchers in this field.

4.1 Crime Type

Similar to computer crimes any crime conducted using the cloud either as an object, subject or tool is considered a cloud computing crime [10] [11]. Cloud computing is an object when the CSP has been as a crime target e.g. Distributed Denial of Service (DDoS) attacks. It is a subject when the cloud environment is used to conduct crime e.g. identity theft, as discussed in section 1, the Google case [2]. Finally, cloud is considered as a tool when one cloud service is used to attack another service **provider's** network e.g. dark cloud.

4.2 Performing Investigations In the Cloud

Building a case based on evidence located in the cloud is considered as an "In" cloud investigation. With current digital forensics methodologies, organizations must be aware of a CSP's incident response strategy, including incident identification, notification and incident recovery. Snapshots provide an image of the system at a specific point in time. It can be considered as a rich source of evidence for services provided either based on virtualization or distributed systems. However, given the current approaches of taking the snapshot, its reliability and soundness for forensics purposes needs to be investigated. As a proactive measure, cloud users should check the availability of their virtual environment snapshots offline, together with the periods when these snapshots are performed. For example, the Amazon Elastic Block Store (EBS) Boot Volume provides storage services on a block level along with Elastic Compute Cloud (EC2). The EBS provides snapshots of the user storage. In the case of cyber-attacks, a snapshot can later be analyzed offline without tampering with the original storage and disturbance to the course of business [15]. As stated in, when examiners have to access 'live' systems, capturing volatile data will result in changes to the target system. From our point of view, having both consistent snapshot of a running system and maintaining an audit trail of the examiners actions should minimizes the chances of error.

4.3 Performing Investigation On the Cloud

Unlike conducting investigations "in" the cloud, using cloud computing resources to improve the investigation

process can be considered a silver lining. Computer and **network forensics can be provided as "on-demand"**

services, where investigators will have as much storage and computing power as they need. Recently, Dell has provided a forensics-as-a-Service solution. Dell applies the process of digital forensics and then utilizes the datacenters capability to image seized devices on site[15]. Eventually, the examiners should be able to create a coherent timeline of events. Seizing the service provider's devices such as servers will not only affect business continuity and potentially violate legitimate user privacy, but it is also impractical to image Petabytes of information and analyze it[15]. An Access Data Forensics Tool Kit (FTK) performance test report stated that to process a 120GB hard drive using top-of-the-line workstations would require around 5.5 hours. Similarly, to analyze 2 TB of hard drive would require around 85 hours. It is reasonable to comment that analyzing digital evidence is extremely time consuming and that the larger the storage capacity, the greater the time required.

5. CONCLUSION AND FUTURE WORK

The low cost of services provided in cloud computing has pushed many users to adopt cloud based services. At the same time, there is an increasing need for forensically based cloud computing services. We have discussed the impact of enabling technologies such as virtualization and distributed computing in providing forensically ready cloud computing. It can be achieved by enhancing current virtualization and distributed computing methodologies. Given the heterogeneity of the cloud environment, investigators must identify a set of guidelines which can help throughout the investigation. As Future work and to better address the challenges of cloud forensics, a comprehensive real life scenario will be constructed that covers different aspects and supports it through case studies. Also, we will develop a framework that will support the production of forensically sound evidence.

REFERENCES

[1] L. Badger, T. Grance, R. Patt-Corner, and J. Voas, "Draft cloud computing synopsis and recommendations," 2011.

[2] Gartner, (2010), "Gartner Says Cloud Computing Will Be As Influential As E-business," [Online]. Available: http://www.gartner.com/it/page.jsp?id=707508, [October 15, 2012]

[3] S. Ahmed and M. Y. A. Raja, "Tackling cloud security issues and forensics model," in *High-Capacity Optical Networks and Enabling Technologies (HONET)*, Islamabad, Pakistan, 2010, pp. 190-195.

[4] Amazaon, (2010), "Amazon Elastic Compute Cloud http://www.amazaon.com/ec2, [Oct. 15, 2012]

[5] Google, (2008), "Google App Engine," Available: http://appengine.google.com, [Oct. 15, 2012]

[6] Microsoft, (2012), "Windows Azure," [Online]. Available: http://www.windowsazure.com, [Oct. 15, 2012]

[7] IBM, (2011), "IBMSmartCloud," [Online]. Available: http://www.ibm.com/smartcloud, [Oct. 15, 2012]

[8] Ubuntu, (2012), "Ubuntu Enterprise Cloud," [Online]. Available: http://www.ubuntu.com/cloud, [Oct. 15, 2012]

[9] Eucalyptus, (2012), "Eucalyptus Cloud," [Online]. Available: http://www.eucalyptus.com/eucalyptus-cloud, [Oct. 15, 2012]

[10] OpenNebula, (2012), "OpenNebula 3.6," [Online]. Available: http://www.opennebula.org/, [Oct. 15, 2012]

[11] CLOUDS, (2011), "CloudSim," [Online]. Available: http://www.cloudbus.org/cloudsim/, [Oct. 15, 2012]

[12] R. Buyya, R. Ranjan, and R. N. Calheiros, "Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities," in *International Conference on High Performance Computing & Simulation, 2009. HPCS'09.* , Melbourne, VIC, Australia 2009, pp. 1-11.

[13] D. Reilly, C. Wern, T. Berry, "Cloud computing: forensic challenges for law enforcement," in *Proc. of the International Conference for Internet Technology and Secured Transactions (ICITST)*, 2010.

[14] S. Ahmed and M. Raja, "Takling cloud security issues and forensics model," in Proc. of the *High-Capacity Optical Networks and Enabling Technologies (HONET)*, 2010.

..

[15] International Symposium on Cloud and Services Computing Cloud Forensics,2012 : State-of-the-Art and Research Challenges By Anand Kumar Mishra, Priya Matta, Emmanuel S. Pilli and R. C. Joshi Department of Computer Science & Engineering Graphic Era University Dehradun, India