Digital Forensics in Cloud Computing

Archit Kapur
Bachelor in Information Technology
Inderprastha Engineering College

ABSTRACT
For years, digital forensics has been used in computer crime investigations. It has changed and progressed in response to technological advances, and is currently entering a new phase as the concept of cloud computing is emerging. Cloud computing is a computing model that enables on-demand network access to a shared pool of configurable computing resources from anywhere in the world (e.g., networks, servers, storage, applications, and services). In the field of digital forensics, the concept of cloud computing has produced unique obstacles. This research paper aims to investigate the digital forensics issues in the model of cloud computing and provide the necessary solutions, guidelines. Traditional digital forensics and cloud computing are also covered in depth.

Keywords
Digital Forensics, Cloud Computing, Security, Cloud Forensics

1. INTRODUCTION
For the past few years, digital forensics has been used in computer crime investigations. It has progressed and changed alongside technological breakthroughs, and is currently entering a new phase with the advent of cloud computing. Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services). It has become a popular, low-cost choice for businesses' IT needs. The nature of cloud computing has produced unique obstacles in the field of digital forensics, despite the various benefits it brings to its clients. The paper focuses on study of forensic difficulties in cloud computing and offers possible solutions, guidelines, and case examples. Traditional digital forensics and cloud computing are also covered in depth.

2. CLOUD COMPUTING

2.1 Definition
Cloud computing is a model for enabling ubiquitous, 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].

This cloud model is composed of five essential characteristics, three service models, and four deployment models.

The characteristics of cloud as mentioned by NIST are: On-demand self-service, broad network access, resource pooling, rapid elasticity and measured service [1].

The various service models are: Saas, Paas, Iaas. Software as a Service (SaaS) is the capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface. Platform as a Service (PaaS) is the capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. Infrastructure as a Service (IaaS) is the capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications [1].

The deployment models are - Private cloud, community cloud, public cloud and hybrid cloud. In private cloud, the cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). In community cloud, the cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). In public cloud, the cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider. In hybrid cloud, the cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds) [1].

2.2 Architecture
A single-site cloud is realized as a datacenter, which consists of compute nodes, switches, network topology, storage nodes connected to the network, front-end for submitting jobs and services. The SLA is responsible for the monitoring of the service contract to ensure its fulfillment in real-time. The actual service contract will also detail criteria associated with any computer forensic investigations, such as jurisdiction and data seizure. The jurisdiction covers the local laws that apply to the service provider and consumer [2].

3. DIGITAL FORENSICS AND CLOUD FORENSICS
The foundations of digital forensics and cloud forensics must be addressed before we can discuss the issues of cloud forensics.

3.1 Digital Forensics
According to [3] digital forensics are the application of scientifically derived and proven methods which aims to preserve, collect, validate, identify, analyze, interpret, document, and provide digital evidence while keeping a recorded trail of evidence for court presentation.

Digital evidence is volatile and fragile, and it can be tampered with if handled incorrectly. Protocols must be followed to ensure that data is not altered during its processing due to its

volatility and fragility. There are four phases involved in the initial handling of digital evidence: identification, collection, acquisition, and preservation. The policies and sequential acts that should be followed to investigate in a manner that ensures the admissibility of collected evidence in a court of law, as well as the tools and other resources needed to conduct the investigation, are all included in a standard operating procedure [4].

3.2 Cloud Forensics

Cloud forensics is the application of digital forensic science in cloud computing environments. Technically, it consists of a hybrid forensic towards the generation of digital evidence. Organizationally it involves interactions among cloud actors (i.e., cloud provider, cloud consumer, cloud broker, cloud carrier, cloud auditor) for the purpose of facilitating both internal and external investigations [5].

Evidence found in cloud forensics must meet the same requirements as traditional evidence, just as it does in digital forensics, and some of the difficulty in cloud forensics stem from meeting these standards.

4. THREATS AND SECURITY IN CLOUD

A survey was conducted by Novell [6] conducted with enterprise respondents, security concerns were raised by nine out of 10 respondents. First category of concern is related to the security of physical and software infrastructure of cloud service providers. The integrity, confidentiality, privacy, and provenance of data in the cloud are the second set of concerns. Strong encryption, transport-level security measures, and requires special can be used to address these difficulties, although some of these mechanisms can be costly to deploy. The third area of concern is user authorization and authentication, which ensures that the user using the cloud is legitimate. With the development of mobile devices with cloud access capabilities, this problem becomes even worse. Sengupta el al [7] emphasizes the need of more flexible cases of identity federation. The fourth and final point of worry is cloud providers' adherence to regulations. There are various regulations such as the Payment Card Industry Data Security Standard (PCI DSS) and various others but the main point is actually following these guidelines and regulations as stated.

5. CLOUD FORENSICS CHALLENGES

The continually changing cloud industry contributes to the general issues of cloud forensics. Cloud computing is expected to account for roughly one-third of the IT industry's net new growth. As more people go to the cloud, the aim of cybercrime has moved. Researchers believe that cyber thieves will no longer target users' machines in the future. Their focus will shift to data centers and the cloud. As a result of these phenomena, there is just too much data to evaluate. This is compounded by the fact that there is simply insufficient experience and technology in the field of cloud forensics, making it much more difficult to achieve. Furthermore, the standards and laws governing cloud forensics have yet to be defined.

The major categories of difficulties highlighted by the NIST working group are as follows: Architecture, data collection, analysis, anti-forensics, incident first responders, role management, legal, standards and training [8].

5.1 Identification Stage

Access to evidence in logs: The importance of logs in an inquiry cannot be overstated. The investigators' top objective is gaining access to log files in order to identify an occurrence. Locating logs in cloud environments where data is stored in unknown locations due to system spread is a difficult and time-consuming procedure. It is just partly applicable in IaaS cloud model as it provides the Virtual Machine which behaves almost the same as an Actual Machine [9].

Physical Inaccessibility: Due to the global dispersal of the hardware devices in a cloud environment, locating data is a complex operation. The established digital forensic procedures and tools assume that physical access to the hardware is a fact [10].

Volatile Data: When a Virtual Machine instance in an IaaS service model is shut off or rebooted, the data saved in it is lost. This is due to the loss of critical evidence such registry entries, processes, and temporary internet files. In case an adversary launches an attack on a VM with no persistent storage synchronization, when the attack is completed, the adversary can shut down the Virtual Machine instance leading to a complete loss of volatile data, if no further countermeasures are installed [11].

Client Side Identification: Evidence can be discovered in both the providers' and clients' interfaces. The user agent on the client system is the only application that communicates with the service in the cloud. Hence, in an exhaustive forensic investigation, the evidence data gathered from the browser environment should not be omitted [11].

5.2 Preservation – Collection Stage

Integrity and Stability: In cloud investigations for IaaS, PaaS, and SaaS, evidence integrity preservation and stability are critical. . We must preserve data to acquire evidence in multi-jurisdiction environments, a difficult task, without violating any law. If the integrity is not preserved, then the evidence will not be admissible to the court of law. It is also difficult to maintain the stability of the data because of the transient nature and dedicated description of the data in a cloud [12].

Internal Staffing: From identification to preservation, this issue affects all three service models and all four stages. To conduct an investigation in cloud forensics a number of people must be involved as a team. This team should consist of investigators with technical knowledge, legal advisors and specialized external staff with deep knowledge in new technology and skills [13].

Time Synchronization: Date-time stamps, as digital evidence, are very important in a court of law. Once they can be easily altered, additional verification needs to be obtained; otherwise, investigators cannot ascertain whether the event occurred at a certain time [14].

Multi-jurisdiction: Another challenge for the investigators is obtaining evidence from the three cloud models from separate jurisdictions. Because of the properties of cloud computing, data is typically dispersed around the globe. Thus, it is very difficult, almost impossible, to conduct evidence acquisition when investigators are dealt with different legal systems, where the related laws or regulations may vary by countries. Hence, to determine that someone is the owner of the data from a large number of cloud users distributed globally is an intricate process [15].

5.3 Examination - Analysis Stage

Volume of Data: The amount of data held in cloud data centers is enormous, and it continues to grow on a daily basis.

This has an immediate influence on the information processing required to uncover valuable evidence for the investigation. Appropriate capture and display filters have to be developed and set up in order to make the data volume present in Cloud Infrastructures proccessible. It is very difficult to analyze the VMs directly, even if the CSPs cooperate with investigators, because the VMs for SaaS and PaaS may have a huge storage system and contain many other applications [16].

Reconstruction: When conducting an inquiry, encrypted data will be useless if the encryption keys cannot be obtained. The evidence also can be compromised if the owner of the data is the only one who can provide the key, or if the key is destroyed. Furthermore, many CSPs are using encryption methods to store clients' data in the cloud. In IaaS, the VM time is under the client's control meaning that all date and times used in logs and other records should be converted to the specific time system. Another solution to overcome the problem is the Network Time Protocol, designed by Mills [17]. The latest protocol RFC 5905 is the most efficient

Encryption: To protect themselves from any activity, many cloud clients in all three service types keep their data in an encrypted way. When conducting an inquiry, encrypted data will be useless if the encryption keys cannot be obtained. The evidence also can be compromised if the owner of the data is the only one who can provide the key, or if the key is destroyed. Wan [18] proposed hierarchical attribute-set-based encryption to achieve scalability, flexibility, and fine-grained access control in cloud computing.

Unification of Log Formats: Analyzing data obtained from service models takes time, especially if we have to deal with and identify a variety of log formats. Unification of log formats in cloud is a difficult operation when we have to access the huge amount of different resources available. Logging in the cloud also has problems such as the decentralization of logs, volatility of logs, multiple tiers and layers, different archiving and retention policies, accessibility of logs, nonexistence of logs, the absence of critical information in logs and incompatible / random log formats [19]. Thorpe [20] stated that the hypervisor system logs can be used to track VM incidences which may later be used to compile potential evidence for a cloud investigation.

6. POSSIBLE SOLUTIONS

Several authors suggested potential solutions to some of the issues in cloud forensics and investigations, despite the fact that many of them acknowledge the urgent need for tools, registration, and procedures that are specifically specialized for cloud forensics.

6.1 Local Machine

Depending on the cloud model, forensic data access varies dramatically. Customers of IaaS, for example, have generally unrestricted access to the data needed for forensic inquiry, but SaaS customers may have limited or no access. As a result, if the victim was a SaaS client, the client's Web browser may be the sole application that communicates with the cloud service.

6.2 Between Cloud and Local Machine

The audit control node at the Internet Service Provider is another place to look for traces of client cloud activity (ISP). Regardless of the cloud's kind or location, the ISP must adhere to the rules and regulations of the nation in which it is based, making the ISP's audit control node in the local area one of the most important targets in the inquiry.

6.3 In the Cloud

Virtual Machine Monitor is a tool that allows you to keep track of your virtual machines (VMM). It is recommended that you take use of the cloud's virtual environment and use virtual introspection to monitor and inspect virtual machines from a VMM. It is also stated that APIs for usage with virtualization, different projects are in the works. PaaS and SaaS, the two cloud models, would add extra layers to account for the platform or service delivered. Layers 5 and 6 are under the administrative control of the cloud customer. The forensic data collecting operation would differ at each layer since each layer requires a different level of confidence to ensure the layer's security and trust; the further down the table you go, the less cumulative trust is necessary. Based on this principle, consider a scenario in which an investigator acquires remote access to the operating system of a guest virtual machine. The investigator can collect evidence from the virtual system, install a forensic programme, and retrieve live evidence remotely, or suspend/terminate the virtual machine and examine it later. The confidence required for this acquisition at this tier is that the guest operating system, hypervisor, host operating system, underlying hardware, and network are free of purposeful and inadvertent tampering, compromise, or error in producing complete and accurate evidence data.

Currently, data is collected through a cloud service provider. They carry out the search, gather the data, and return it to law authorities after receiving a search warrant. Although all of the tools and strategies were successful in producing proof, they all required a high level of faith in the cloud infrastructure at all levels. They offered solutions to address the multi-trust issue, which are detailed below: The cloud management plane, forensics-as-a-service, and contract support. Each of these technologies, or a combination of them, can increase the trustworthiness of remote acquisition by establishing trust at lower cloud layers. The hardware and associated software might evaluate what software is loaded on each computer and verify the health and status of each machine by placing platform modules in each cloud server. Another strategy that can be utilized to help reliable forensics is data collection from the management plane. The management plane's log files, disc images, and packet capture can all be downloaded on demand by the provider, end users, and law enforcement. Because the forensic acquisition takes place within the hypervisor, accessing virtual machine pictures and other data requires just confidence in Layers 3 and below, making it more admissible in court. Cloud companies have already demonstrated their ability to maintain and collect data not only from virtual machines, but also through infrastructure logging methods, packet captures, and billing information. If the company and its infrastructure can be trusted, they can help law enforcement with forensics. The provider can also supply these services to their clients at a low cost and with less effort.

7. CONCLUSIONS

Although cloud forensics can make use of digital forensic tools and technologies, they have limitations. By removing navigation data after surfing, the examination of the local machine and web browser database for traces of client-cloud interaction can be simply cancelled. Anyone with a basic understanding of browser history and database may delete all traces of their cloud interactions. The collection from the management plane method is appealing since it is user-driven, but it necessitates management plane confidence. A natural choice is forensic support as a service, which is already

available from one source. As long as each cloud provider's forensic preparedness can be assessed by a third-party auditor, which might include internal staffing requirements, recording of user activity, and security measure installation, forensic as a service can be a huge help to law enforcement in investigations. Another critical matter in the sector is the training and recruitment of technical employees for forensic investigations in order to address the shortage of experienced cloud forensic personnel.

8. REFERENCES

[1] NIST, "The NIST Definition of Cloud Computing," September 2011. [Online]. Available: https://csrc.nist.gov/publications/detail/sp/800-145/final.

[2] C. W. T. B. Denis Reilly, "Cloud Computing: Pros and Cons for Computer Forensic Investigations," March 2011. [Online]. Available: https://infonomics-society.org/wp-content/uploads/ijmip/published-papers/volume-1-2011/Cloud-Computing-Pros-Cons-for-Computer-Forensic-Investigations.pdf.

[3] "Digital Forensic Research Workshop, "A Roadmap for Digital Forensic Research"," 2001. [Online]. Available: https://www.oreilly.com/library/view/practical-forensic-imaging/9781492018049/xhtml/footnote.xhtml.

[4] "ISO/IEC 27037," 2012. [Online]. Available: https://www.iso.org/standard/44381.html.

[5] "NIST Cloud Computing Forensic Science Challenges," August 2020. [Online]. Available: https://csrc.nist.gov/publications/detail/nistir/8006/final.

[6] Novell, "Survey Finds Mixed Feelings On Cloud Computing Among IT Pros," [Online]. Available: https://www.netiq.com/docrep/documents/0h82v8pmt1/SurveyReport_CloudComputing_en.pdf.

[7] "Cloud Computing Security--Trends and Research Directions," July 2011. [Online]. Available: https://ieeexplore.ieee.org/document/6012787?signout=success.

[8] "NIST Cloud Computing Forensic Science Challenges," [Online]. Available: https://csrc.nist.gov/csrc/media/publications/nistir/8006/draft/documents/draft_nistir_8006.pdf.

[9] M. Damshenas, A. Dehghantanha, R. Mahmoud and S. b. Shamsuddin, "Forensics investigation challenges in cloud computing environments," July 2012. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6246092.

[10] S. T. Rainer Poisel, "Discussion on the Challenges and Opportunities of Cloud Forensics," 2012. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-32498-7_45.

[11] D. Birk and C. Wegener, "Technical Issues of Forensic Investigations in Cloud Computing Environments," March 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6159124.

[12] Y. D. P. Q. J. D. Guangxuan Chen, "Suggestions to digital forensics in Cloud computing ERA," 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6418812.

[13] C. J. K. T. C. M. Ruan K, "Cloud forensics: an overview. In Advances in Digital Forensics VII, 7th IFIP WG 11.9 International Conference on Digital Forensics," [Online].

[14] D.-Y. Kao, "Cybercrime investigation countermeasure using created-accessed-modified model in cloud computing environments," January 2016. [Online]. Available: https://dl.acm.org/doi/10.1007/s11227-015-1516-7.

[15] H. F. Sibiya G Venter, "Digital forensic framework for a cloud environment," May 2012. [Online]. Available: https://researchspace.csir.co.za/dspace/handle/10204/5890.

[16] K.-K. R. DarrenQuick, "Impacts of increasing volume of digital forensic data: A survey and future research challenges," December 2014. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S1742287614001066.

[17] U. D. D. Mills, "Network Time Protocol Version 4: Protocol and Algorithms Specification," June 2010. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc5905.

[18] J. L. R. H. D. Zhiguo Wan, "HASBE: A Hierarchical Attribute-Based Solution for Flexible and Scalable Access Control in Cloud Computing," April 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6046132.

[19] R. Marty, "Cloud application logging for forensics," March 2011. [Online]. Available: https://dl.acm.org/doi/10.1145/1982185.1982226.

[20] R. Marty, "Cloud application logging for forensics," March 2011. [Online]. Available: https://dl.acm.org/doi/10.1145/1982185.1982226.