Cloud Technology and the Challenges for Forensics Investigators

Alex Roney MATHEW and Jamal Abdullah AL-ZAHLI

Department of Information Technology, College of Applied Sciences, Ministry of Higher Education,
P.O Box 135, P.C, 311, Sohar, Sultanate of Oman

Keywords: Cloud services, Cloud forensics, Cross-border and multi-jurisdictional challenges.

Abstract. Cloud computing is a recent technology that offers a cheap and almost limitless computing power and storage space for data which can be leveraged to commit either new or old crimes and host related traces. This paper aims to review and discuss the challenges for digital forensics in cloud that we cannot seize the physical hardware which runs various applications in cloud. Cloud forensics is difficult because there are challenges with multi-tenant hosting, synchronization problems and techniques for segregating the data in the logs. Cloud forensic is the process of identifying, preserving, analyzing and presenting digital evidence in a manner that is legally acceptable. The complexity of forensics and investigations in a cloud environment is broadly related to the various challenges which are discussed in the paper. We came to know that the main challenge in cloud forensic is data acquisition. It is important to know exactly where the data is located and actually acquiring it. If Cloud Service Providers practice to preserve volatile data, the loss of important artifacts, which could be potentially crucial evidence, can be made avoided.

Introduction

Cloud computing has recently emerged as a technology to allow users to access infrastructure, storage, and software and deployment environment based on a pay-for-what-they-use model. Cloud computing has been gaining popularity with businesses and end users in recent years. A certain level of hype and inconsistent definition has led to some confusion about what cloud computing is, and what services it can provide. Along with general confusion, some concerns have been raised about the security of cloud environments. As seen with traditional computing, a growing concern for security leads to consideration of incident response and eventually digital forensic investigation capabilities. The Cloud offers a cheap and almost limitless computing power and storage space for data which can be leveraged to commit either new or old crimes and host related traces. Conversely, the Cloud can help forensic examiners to find clues better and earlier than traditional analysis applications, thanks to its dramatically improved evidence processing capabilities. From one side, the pervasive availability of cheap cloud computing services for data storage, either as a persistence layer to applications or as a personal store for documents and pictures, is remarkably increasing the chance that cloud platforms potentially host evidence of criminal activity. When this happens, collecting data in a way that is able to resist to legal and technical vetting may reveal itself very tricky, because forensic tools targeted to cloud infrastructures are still in their infancy and issues concerning jurisdiction may apply. Relevant data may indeed be fragmented in countless shards, possibly available for a very limited timeframe and residing in more than one country.

Digital forensics in remote, ubiquitous provider-controlled cloud computing systems is difficult when compared to traditional digital forensics. Criminal use of cloud computing is an impending possibility as cloud becomes omnipresent. Likewise, the need for digital forensic analysis of cloud computing environment and applications has become customary. As in the case of traditional computer forensics, digital forensics in the cloud environment also comprises the stages: Identification, Collection, Examination/Analysis and Reporting/Presentation. 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 concerned with the presentation of collected evidence to the court of law. The technical, legal, and organizational dimensions of cloud forensics are challenging for digital investigators to cope up...
with current developments. The dynamic nature of cloud provides abundant chances to enable
digital investigation in cloud environment. The challenge for digital forensics in cloud is that we
cannot seize the physical hardware which runs various applications in cloud, as they are distributed
across various geographical locations.

Cloud Forensics

To investigate the crimes involving clouds, investigators have to carry out a digital forensic
investigation in the cloud environment. This particular branch of forensic has become known as
Cloud Forensics. Digital Forensics (DF), as defined by McKemmish, is the “process of
identifying, preserving, analyzing and presenting digital evidence in a manner that is legally
acceptable”. “The discipline that combines elements of law and computer science to collect and
analyze data from computer systems, networks, wireless communications, and storage devices in a
way that is admissible as evidence in a court of law”.

The aim of a forensic investigation is to identify and preserve the evidence, extract the
information, document every process, and analyze the extracted information to find answers with
respect to the 5Ws (Why, When, Where, What, and Who).

Wolfe defines computer forensics as “a methodical series of techniques and procedures for
gathering evidence, from computing equipment and various storage devices and digital media that
can be presented in a court of law in a coherent and meaningful format”. In computer forensics,
maintaining the integrity of the information and strict chain of custody for the data is mandatory.
Several other researchers define computer forensic as the procedure of examining computer system
to determine potential legal evidence.

From the above definitions, we can say that computer forensics is comprised of four main
processes:

- Identification: Identification process is comprised of two main steps: identification of an
  incident and identification of the evidence, which will be required to prove the incident.
- Collection: In the collection process, an investigator extracts the digital evidence from
different types of media e.g., hard disk, cell phone, e-mail, and many more. Additionally, he
  needs to preserve the integrity of the evidence.
- Organization: There are two main steps in organization process: examination and analysis of
  the digital evidence. In the examination phase, an investigator extracts and inspects the data
  and their characteristics. In the analysis phase, he interprets and correlates the available data
to come to a conclusion, which can prove or disprove civil, administrative, or criminal
allegations.
- Presentation: In this process, an investigator makes an organized report to state his findings
  about the case. This report should be appropriate enough to present to the jury.

Cloud forensics procedures will vary according to the service and deployment model of cloud
computing. For SaaS and PaaS, we have very limited control over process or network monitoring.
Whereas, we can gain more control in IaaS and can deploy some forensic friendly logging
mechanism. The first three steps of computer forensics will vary for different services and
deployment models. For example, the collection procedure of SaaS and IaaS will not be same. For
SaaS, we solely depend on the CSP to get the application log, while in IaaS, we can acquire the
Virtual machine instance from the customer and can enter into examination and analysis phase. On
the other hand, in the private deployment model, we have physical access to the digital evidence,
but we merely can get physical access to the public deployment model.
Investigation Activities

According to ISO/IEC 27037 and 27042, there are seven main activities in a forensic investigation. The first two activities focus on readiness, before an incident happens; the rest are carried out after the incident happens. Figure 2 was taken from ISO/IEC 27041 and represents the activities before and after an incident has been identified.

- Plan: A scenario-based planning approach to the investigators needs is recommended. The idea is to plan scenarios that investigators might face.
- Prepare: Forensic practitioners should put all essential services and efforts in place in order to support future cases. This includes preparing tools, techniques and safeguards.
- Respond: This is when the incident has happened and the forensic practitioners start determining the scope of the event like what the situation is, the nature of the case and its details. This step is important because it helps determining the characteristics of the incident and defining the best approach to carry out the investigation.
- Identify: In this step the investigators start gathering information about the specific event or incident. Notes describing the systems to be analyzed, their network position and general configurations may be taken at this stage.
- Collect: This third step, after the incident has been identified, aims to maximize the collection of evidence as well as minimizing the impact to the victim. Recording of the scene is also included on this step.
- Acquire: The most important task here is to maintain the integrity of the evidence and provide assurance that the evidence has not been changed. This is carried out by maintaining a chain of custody of all evidence, ensuring that they have been collected and protected by legally acceptable processes.
- Preserve: Isolation, securing and preservation of the original evidence is comprised in this step. The main aim is to prevent any cross-contamination.
- Understand: In this step, investigators need to determine the significance of reconstructed data and draw conclusions.
- Report: Here a summary, explanation of findings and conclusions are reported. The reports should be written such that they are legally admissible. In addition, a 3rd forensic investigation team should reach the same conclusions following the investigation steps in the report.
- Close: In the last step, practitioners need to ensure evidence is returned to rightful owner or securely stored if needed.
Challenges for Cloud Forensics Investigators

There are challenges in applying existing digital forensic practices in investigating issues in cloud networks. Most tools currently used for digital forensic investigations are largely intended for offline investigations with the assumption that the storage media under investigation is within the control of an investigator. Limited tools and methodologies that can assist in extraction and analysis of potential evidence which are acceptable in legal proceedings, significantly dependent on the service models or deployment model opted on a cloud infrastructure and the way a cloud service provider is managing those models. Non-availability of expert advice and inadequate oversight, right from the initial stages of planning a migration to a cloud infrastructure, can expose a user to legal or compliance issues later.

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 complexity of forensics and investigations in a cloud environment is broadly related to the following challenges:

1. 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 find number of tools and procedures to identify the digital evidence, but cloud is volatile in nature, so 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.

2. 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 data is stored at somewhere. RAM might contain valuable evidence including user-name, 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.

3. 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 cloud service provider 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. 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. Lack of customer awareness: In cloud all is under the control of cloud service provider (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.

5. Data integrity: The investigators need to maintain the integrity of the evidence to preserve the integrity of the original data for cloud investigator, it’s very difficult. 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. 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. The evidence is valueless in front of law if the integrity of the data was missed. Number of users are involved in the investigation process due to this the errors may occur in the preserving phase says that data integrity and preserving is very difficult and challenging phase for the cloud investigator.

6. 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.

7. Authentication & 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.

8. 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.

9. 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.

10. 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.

11. Cross-border and multi-jurisdictional issues: Cross-border and multi-jurisdictional challenges are common to almost any cloud deployment model. However in cases where the host machines and data centres of the CSPs are located around the world, more legal issues are likely to emerge considering the time constraints that characterize the investigations. Specifically, the experts have underlined the following:

a. Cross-national legislations and collaboration mechanisms or channels between law enforcements, CERTs and public authorities that can facilitate the exchange of data for the forensic investigations are not well developed.

b. Formal requests issued during cross-border investigations to allow access to data stored remotely are conditional to the local legislation. Access is granted only after preliminary analysis realized by the responsible authorities of the countries in which the data resides.

c. There is no specific regulation governing the CSP obligations in terms of standard operating procedures, the standard actions the CSP has to take in case of investigations to which request the CSP has to comply, the timing for that (e.g. immediately, in a week, in a month), and finally the instructions on how data logs should be kept, managed and stored in order to provide evidence.

d. There are currently no agreements among cloud providers, law enforcement and customers to collaborate on investigations when necessary.

e. For law enforcement dealing with international investigations, legal access to data is still one of the most relevant challenges. A number of different issues might occur: the concept of locality of the law enforcement mandate (national or regional), the national legislations which do not allow
some categories of information to be transferred abroad, privacy aspects related to sensitive information, and finally the presence of consulting authorities or institutions (as centres of expertise).

f. International cyber policies and laws must progress to help and solve the issues surrounding multi-jurisdiction investigations.

On the other hand, for the private sector, the cross-border and multi-jurisdictional issues can be managed in a better way, as the access and exchange of data for investigations is performed exclusively by the CSP that is the owner of the data.

**Solutions and Future Directions**

Cloud forensics investigation has immense challenges due to its abstract and black box architecture. Since it is a new technology in its budding ground, researchers are quite interested to expose this new technology in a full-fledged mode. We have addressed challenges occurs in cloud forensic investigations in the previous section of this paper. In the cloud architecture, consumers are served by several Cloud Service Providers according to their demands. Since cloud is a distributed platform, it is quite evident that consumers are supposed to get cloud services from several CSPs rather than any singleton provider. All the CSPs basically reside within a network or any sub-networks of those networks. In this respect when a malicious activity is reported, all those inter and intra related hosts must be taken under the consideration of the investigation process\(^1\). In this scenario if an investigator is supposed to examine all the IP addresses, it becomes a menace. To come out of this particular overhead there must be some automated system which will identify all the probable malicious hosts with the help of the previous history. By training the system based on the log of the malicious hosts, an investigator can identify all the malicious hosts whenever a set of hosts, with the same incident and attributes are to be examined by calculating the probability of being malicious. This reduces the investigation time and costs in a major rate.

It is quite obvious that if the CSPs cannot ensure the providing of the information, consumers will not leave the complete computation control over cloud. At the same time it is particularly true from the cloud forensics investigators’ aspect too. Since investigators have to collect all the evidences from a distributed architecture of cloud, therefore there is always a question of evidence integration verification properly and accurately due to much dependence on CSP. Therefore a mechanism should be developed which should have trustworthy and effective framework for the investigator so that they won’t be misguided at the time of evidence collection and integration.

In cloud environment, investigators are supposed to analyze a large volume of collected evidences using all cloud service models. Cloud computing software platform which is free and open source supports in forensic acquisitions.

**Conclusion**

Cloud computing has a number of benefits, such as high availability, potentially lower cost, and potentially improved security. However, cloud computing also has a number of associated risks. Some of these risks have been inherited from traditional computing models, while the cloud business model introduces others. As more businesses and end users move their data and processing to cloud environments, these environments will increasingly become the target, or even the originator, of malicious attacks. For this reason, digital forensic investigation methods must be tested against cloud computing environments with the help of CSPs to ensure digital evidence is available, and integrity can be maintained and verified. As cloud computing blurs geographical borders for businesses, end users and even criminals, law too must begin to look to a more open, global system; a system where everyone, law enforcement, private sector, academia, and even the public can play a role to detect, research and reduce digital crime. Cybercrime affects every Internet connected country, and without effective international collaboration, laws, and efficient communication channels that support international collaboration, cybercrime will continue unchecked. We came to know that the main challenge in cloud forensic is of data acquisition. It is
important to know exactly where the data is located and actually acquiring it. If Cloud Service Providers practice to preserve volatile data, the loss of important artifacts, which could be potentially crucial evidence, can be made avoided.

References


