Chapter 1: Introduction

Security of cloud computing environment and its services and infrastructures is becoming one of the emergency research directions especially digital forensics. Applying digital forensic techniques and principles in the cloud environment which known as ‘Cloud Forensic’ is facing complex obstacles where slight research has done to cover this critical area. There are numerous multifaceted challenges that obfuscate cloud forensics such as crime scene reconstruction, isolating cloud instance, data provenance, evidence segregation, and dynamic nature of cloud computing. This makes the cloud forensic investigation as a tough mission for digital investigators to cope with cybercrimes in the cloud environment in order to reconstruct crime events about an incident that occurred and introduce a digital evidence to a court of law as admissible proof for the committed cloud based-crime. At present, cloud computing services and infrastructures suffer from the absence of the support for digital forensics. Moreover, the current digital forensic techniques, methods and tools cannot contend with the dynamic nature of the cloud environment.

1.1 Overview

Presently, cybercrime has become one of the critical threats to countries, organizations, and individuals around the world because of the problems that are caused breaching security of their digital systems such as computers, servers and storage systems…etc. Therefore, these countries, organizations, and individuals take various strategies and procedures to protect and defend against these threats by securing their systems using firewalls, anti-viruses, intrusion detection, and prevention system and countless other protection mechanisms. Infrequently, these protection mechanisms perhaps not prevent attackers from breaching systems security, so as to, there is a serious need for methods or procedures that are able to discover and trace attackers and criminals who attacked the systems. The science which concerns with discovering and tracing the attackers called digital forensic which is used to extract and analyze digital evidence which collected and acquired from the crime scene to reconstruct the crime events that occurred and provide it as an admissible proof about crime in a court of law.
Today, conventional view to cybercrimes is changed because the appearance of new technologies such as cloud computing that depend on a theory of distributed systems. Cloud computing infrastructures are distributing around the world to provide services for organizations and individuals with a method of on-demand or cost-per-use. Thus, traditional cybercrime techniques are changed to adapt to a dynamic landscape of the cloud computing environment. In the same time, digital investigators have to change and expand their tools and techniques to deal with these the new types of cybercrimes in the cloud.

To understand the prominence of the digital forensic investigation in the cloud, two hypothetical scenarios of cybercrimes that may occur in a cloud environment will be explained. The two scenarios are online banking service and malicious user in the cloud respectively.

1. **Online Banking Service:** a criminal scenario that can be occurred within online banking services is explained as follows: "Alex is a customer on online banking service from Bank XYZ. The banking service using datacenter which installed VMware ESXi server 5 to host virtual machines and the affected virtual guest resource running windows server 2012 R2 Webserver Edition. The customer believes that hackers hacked and exploited an application-level flaw in the guest operating system to steal his information such as name, username, password, credit card information and other critical information about the customer account. The bank has identified the flaw and can patch the system, but is first interested in preserving forensic evidence to support a criminal investigation or civil suit. Technical staff determined that the flaw would not have enabled persistent access to the guest server. The bank has already disabled inbound web access to the virtual guest, but the system needs to be kept up and online to support other critical business services" [1].

2. **Malicious User in the Cloud:** a malicious activity that can be occurred in the cloud by a malicious user is described as shown in Figure 1.1 as follows: "Bucky is a cloud user who runs a Virtual Machine (VM) in a cloud. John is a malicious user. He rented also another VM in the cloud to enable him to access cloud infrastructure. John decided to use Bucky's VM to launch many types of attacks to other VMs running in the cloud to steal their data. One of the attacked VM's owner stores important data
which are stolen by the malicious attacker. Consequently, the owner asked a forensic investigator to investigate the case. The investigator found the Owner's VM records each visiting activity information like IP address. Analyzing the visiting customer records, the investigator found that the attacking has done through Bucky's VM. Eventually, the investigator issued a subpoena to the corresponding cloud provider to provide him the network logs for this particular IP address. On the other hand, there are two scenarios may occur. Firstly, John managed to collude with the cloud provider after the attack. Therefore, while providing the logs to the investigator, the cloud provider supplied a tampered log to the investigator, who had no way to verify the correctness of the logs. Under this circumstance, John will remain undetected. Secondly, Even if the cloud provider was honest, John could terminate her rented VM and leave no trace of the attack. Hence, the cloud provider could not give any useful logs to the investigator. Finally, if the investigator can't find the truth, Bucky will be responsible for the malicious attack. To identify the actual attacker in the above malicious attack scenario, there is a need to introduce new techniques that are able to execute digital forensics procedures successfully in the cloud "[2].

Figure 1.1: Hypothetical Scenario for Malicious Purpose in the Cloud.

1.2 Motivations
Recently, cloud computing becomes one of the most significant and popular computing paradigms. Cloud computing is a revolutionary technology that has started changing the manner people live and work. A research by Market Research Media states that the worldwide cloud computing market is projected to rise at a 30% Compound Annual Growth
Rate (CAGR) reaching $270 billion in 2020 [3]. With the growth of cloud and related services, security and privacy in the cloud have become very critical issues in cybersecurity field where hackers can use cloud infrastructure with exceptional bandwidth, storage, and computing power to launch their illegal activities. According to IDC IT Cloud Services User Survey, 74% of IT executives and CIOs referred security as the main reason to prevent their migration to the cloud services model [4].

There are numerous complex challenges such as crime scene reconstruction, isolating cloud instance, data provenance, evidence segregation, and dynamic nature of cloud computing which are making forensic investigation process more challenging task for digital investigators to investigate cloud-based crimes, extract a digital evidence and reconstruct cybercrime events and finally send a forensic report to a court of law as admissible evidence about the committed crime that occurred in the cloud environment. Little work has been done to explore and introduce new theoretical and practical methods in cloud forensic area to support security professionals, cloud investigators and law enforcement agencies to investigate incidents in the cloud in forensically sound and timely fashion manner.

Cloud forensic is challenging at best but can be performed in a fashion consistent with law using new techniques, procedures, and tools. Several researchers have sightseen security issues relating to data stored in clouds which make cloud cannot be used to store data for many sectors such as business, healthcare, banks or national security agencies, which require an audit and regulatory compliance. Though, some researchers have clarified and identified challenges and opportunities along with starting design and develop techniques and tools to easily accomplish cybercrimes investigation in cloud computing environment. In addition to there are little case studies of cloud investigations of cybercrimes to illustrate and explain the issues and to educate practitioners in this domain.

The major challenges of cloud forensics originate from the very characteristics with which the cloud computing platform is identified. A number of cloud forensics challenges that obfuscate cloud forensic are such as each cloud server contains documents from many users. Hence, it is not possible to seize servers from a data center without violating the privacy of various other customers, the storage system is no longer local. Therefore, even with a subpoena, law enforcement agents cannot seize the suspect’s machine and get access
to the suspect’s documents and although the data belonging to a particular suspect is recognized, detaching it from other consumers’ data is problematic. Moreover, except the cloud provider’s word, there is usually no evidence that links a given data file to a specific suspect. In addition, there are more challenges which will discuss in detail in the next chapter.

In the other side, scores of opportunities are available through utilizing massive cloud resources for supporting cloud forensic such as Forensic as a Service (FaaS) which introduce one powerful option for digital investigators through utilizing the vast capabilities of cloud computing. This makes digital forensics as an “on-demand” service for allowing for as much storage and processor power as needed to conduct an investigation. In this case, forensic servers will reside on the cloud side, offline, until the need arises for them. Documents could be backed up to the cloud for investigators to use without having to disrupt normal business. Naturally, the cloud resources could be used for sorting, searching, and hashing the evidence data. There are many benefits of conducting forensics on a cloud environment which such as Improve log indexing and searches, reduce service downtime, reduce forensic image verification time, reduce evidence acquisition time, decrease time to access protected documents, reduce evidence transfer time, and virtually unlimited log storage [5].

This thesis identifies some shortcomings and complicates in cybercrimes investigation in the cloud through studying technical issues and offering particular proposed solutions. Cloud computing poses numerous challenges that are specific to the forensic investigation. From these challenges which addressed in this thesis are as follows: Firstly, secure cloud storage as well as understand and identify data remnants that are available after accessing, using and conducting Box cloud storage as a case study and the need for a method to the forensic investigation of this cloud storage services in forensically sound manner. Secondly, forensic analysis of large volume of log files to reconstruct a timeline of cybercrime events in the cloud environment. Thirdly, investigation of cybercrimes in private cloud environment. Fourthly, gather and analyze digital evidence from VMware clustered ESXi hypervisors. Fifthly, offers cloud services with ensuring that Cloud Service Providers (CSPs) support the digital investigation process (i.e. building cloud architecture support forensics). Sixthly, a need for design and guidance for building a cloud forensic laboratory to investigate
cybercrimes as well as provide a learning environment for law enforcement officers, digital investigators researchers, and students through using colossal cloud capabilities.

1.3 Thesis Statement

The thesis statement is adopted as follows:

“Cloud Forensic cannot be performed as computer and network forensic in traditional IT systems due to distributed and dynamic nature of cloud computing but can be performed in a manner that is consistent with law by developing and introducing practical digital forensic procedures, approaches and techniques to help Cloud Investigators for investigation of cloud-based crimes in legal and easy manner. For instance, cloud forensic storage require understand and identify data remnants that are available after accessing, using and conducting client systems, moreover analyse large generated log data from cloud services and infrastructures in order to reconstruct timeline of crime events as well as investigate private cloud environment plus investigate crimes in clustered VMware hypervisors along with the need for building cloud architecture support digital forensics besides utilizing immense cloud capabilities to provide forensic as a service for countries, organizations and individuals”.

This statement reflects several points. First, cloud storage services need more understanding from a digital forensic point of view to facilitate investigation of crimes related to cloud storage services. Second, Cloud infrastructures and services generating an immense amount of log data which need high and fast processing platforms to reconstruct timeline about cloud-based events that may assist in tracking criminals in an effective manner. Third, VMware provides hypervisors for deploying and serving virtual machines. A cluster of ESXi hypervisors can be subject for criminals so that there was little work to provide tools for extract and acquire digital evidence from them. Fourth, One of lack in cloud systems is it is no support digital forensics therefore, there is a serious need to build a cloud forensic model for being ready for further investigation of cloud-based crimes. Finally, utilizing the huge capabilities of storage and computing resources in building cloud-based forensics laboratories for training students, law enforcement officers as well as building a central laboratory at a country level and other countries to enhance collaboration between their forensics agencies in fighting and investigation severe cyber-attacks.
1.4 Thesis Objectives
The central theme of this research work concerns on applying digital forensic techniques in cloud computing environment for the investigation of cybercrimes. Applying digital forensic techniques and principles in the cloud environment facing complex challenges and obstacles which are required to be resolved to perform a convenient cloud forensics. Thus, this thesis concerns with developing proficient digital forensic techniques for investigation of cybercrimes in cloud computing environment in forensically sound and timely manner. The aim of this research is to study and resolve the following issues:

- Explore and identify challenges and opportunities for performing digital forensics investigation in cloud computing environment.
- Secure cloud storage with understand and identify data remnants that are available after accessing, using and conducting Box cloud storage as a case study and provide a method to the forensic investigation of this cloud storage services in forensically sound fashion.
- Perform forensic analysis on a large volume of log data to reconstruct a timeline of cybercrime events in the cloud.
- Specify a method to perform a forensic investigation of cybercrimes in private cloud environment.
- Develop an application to help in performing forensic investigation of VMware clustered ESXi hypervisors.
- Design and develop cloud computing architecture model to support digital forensics.
- Design a cloud forensic laboratory to investigate cybercrimes as well as provide a learning environment for law enforcement officers, researchers, and students.

1.5 Thesis Contributions
This thesis concerns with developing proficient digital forensic techniques for investigation of cybercrimes in cloud computing environment in forensically sound and timely manner. It is introduced decent seven different research contributions in the field of cloud forensics. They can be summarized as follows:
A literature study is done to explore and identify challenges and opportunities for performing digital forensics investigation in the cloud computing environment which is known as ‘cloud forensic’. The identification of cloud forensic challenges and opportunities helped us to accomplish and complete this research.

A cloud forensic approach based on data integrity checking and digital forensic is proposed for securing cloud storage as well as assisting and helping digital investigators for performing automatic digital forensics for cloud storage services such as Box cloud storage service.

A forensic analysis approach is presented using Apache Hadoop and Apache Spark to analysis large size of cloud log data and extracts knowledge from them which can assist digital investigators during investigating cloud-based crimes that occurred through a particular time.

A digital forensic approach is provided for investigation of cybercrimes in a private cloud environment. The proposed approach can help digital investigators and practitioners in acquisition and collection of digital evidence from the private cloud infrastructures especially virtual machine which is considered the core element of virtualized cloud systems. In addition to, present a forensic methodology called ‘FMRCA’ for investigating and analyzing virtual machine and its snapshots for helping in the reconstruction of criminal activities which done using a virtual machine.

An efficient forensic acquisition application called Hypervisor Forensic Acquisition Application (HFAA) is proposed for data acquisition from clustered VMware ESXi servers. This application can provide an initial step and guide to digital investigators and researchers to develop new applications and tools for real-time forensic acquisition and analysis of virtualized environments like cloud environment. With some additional development and evaluation through adding new features to the proposed application, this can become a good tool for using in cloud forensic community.

A new Cloud Forensics Investigation Model (CFIM) to investigate cybercrimes in a cloud computing environment is introduced. The proposed system is a smart system
that is able to take a snapshot of the state of running virtual machine in virtual datacenter and send to Trusted Center Server (TCS) which monitoring the status of the VMs as well as store snapshots of them to send to Forensic Server (FS) up to require for performing forensics process. The proposed model supports a concept of Forensics as a Service (FaaS) that provide various benefits of conducting digital forensics through using Forensic Server on the cloud side.

- Provide the idea of designing and establishing a Cloud Forensic Laboratory (CFL) to face the sophisticated yet advanced level of cybercrimes taking place currently in classical IT and cloud systems. The proposed system will benefit from enormous storage and processing capabilities which providing by cloud computing to perform acquisition, extracting, analysis, examination and reporting for large size of digital evidence for both of crimes in IT systems and cloud The proposed system will help the digital investigators and practitioners to perform the digital investigation process in forensically sound and timely fashion manner.

1.6 Thesis Organization

This thesis is organized as follows:

**Chapter 2:** provides basic concepts of cloud computing, digital forensics, cybercrimes and cloud forensics in addition to literature review about challenges and opportunities of performing digital forensic techniques in cloud computing environment.

**Chapter 3:** provides a cloud forensic approach based on data integrity checking for assisting and helping digital investigators for performing automatic digital forensics for cloud storage services such as Box cloud storage service.

**Chapter 4:** provides a forensic approach using Hadoop ecosystems and Apache Spark to analysis large size of log data and extract knowledge from them which can assist the digital investigators during investigating of cloud-based crimes that occurred through a particular time.

**Chapter 5:** provides a forensic approach for investigation of cybercrimes in a private cloud environment. In addition to, introduce a forensic methodology called ‘FMRCA’ for investigating and analyzing virtual machine and its
snapshots for helping in the reconstruction of criminal activities which done using a virtual machine.

Chapter 6: provides an efficient forensic application called Hypervisor Forensic Acquisition Application (HFAA) for extracting digital evidence from clustered VMware ESXi hypervisors and helping digital investigators and experts in performing the digital investigation process where the virtual data center is used in the cloud computing environment.

Chapter 7: provides a new Cloud Forensic Investigation Model (CFIM) to investigate of cybercrimes in the cloud. The proposed system is an intelligent system that able to take a snapshot periodically for the state of each Virtual Machine (VM) which runs in the cloud and send it automatically to Trusted Center Server (TCS) that works as a datastore for VMs. This TCS server is responsible for monitoring and recording the status of the VMs and send it to Forensic Server (FS) upon request from Cloud Service Provider (CSP) to deal with the FS.

Chapter 8: provides the idea of designing and establishing a Cloud Forensic Laboratory (CFL) which is based on cloud computing platform for facing the sophisticated yet advanced level of cybercrimes taking place currently comprises of various layers requiring levels of analysis consequently in classical IT and cloud systems.

Chapter 9: gives conclusions and directions for future work on the subject.

1.7 Chapter Summary

This introductory chapter introduced an overview about cybercrimes in cloud environment, two hypothetical scenarios of cybercrimes that may occur in a cloud environment to understand the prominence of the digital forensic investigation in the cloud, a brief about research problem and motivation to working in cloud forensic area as well as the research objectives and the proposed contributions. Finally, the thesis organization is provided.