Forensics framework for cloud computing☆

M. Edington Alex [a,*], R. Kishore [b]

[a] Department of Information Technology, Rajalakshmi Engineering College, Chennai, India
[b] Department of Electronics & Communication Engineering, SSN college of Engineering, Chennai, India

ABSTRACT

The popularity of cloud computing has been on the rise in recent years, as cloud resources are not only shared by many users but can be allocated on demand. A recent survey reports success of the cyber criminals in using cloud computing technology for fraudulent activities, due to its essential characteristics and the lack of suitable digital forensic techniques for the cloud environment. While mitigating cloud crime, investigators face several challenges and issues dealing with cloud forensics. In this paper, the challenges faced by forensic investigators are highlighted. Most of the research work deals with the identification of challenges in cloud forensics and the proposed solutions reported in literature depends on Cloud Service Provider (CSP) for forensic investigation. The dependence on CSP includes the collection of data for the forensics process and there may be a chance of altering data that affects the entire investigation process. For mitigating the dependency on CSP, a new model for collecting forensic evidence outside the cloud environment is developed.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Cloud forensics

National Institute of Standards and Technology (NIST) [1] defines cloud computing as "Cloud computing forensic science is the application of scientific principles, technological practices and derived and proven methods to reconstruct past cloud computing events. This is done through identification, collection, preservation, examination, and interpretation and reporting of digital evidence". Ruan et al. [2] have defined cloud forensics as "the application of digital forensic science in cloud environments as a subset of network forensics", as shown in Fig. 1. Here, the authors highlight the significance of cloud forensics in three different aspects, namely, technical, organizational and legal. Technical aspects are engaged in forensic tools, mechanisms, and procedures. Organizational aspects incorporate the interaction between cloud actors for forensic investigation. Legal aspects deal with multi-jurisdictional and multi-tenant situations. The authors also identify cloud forensics as an associate of cloud computing and digital forensics.
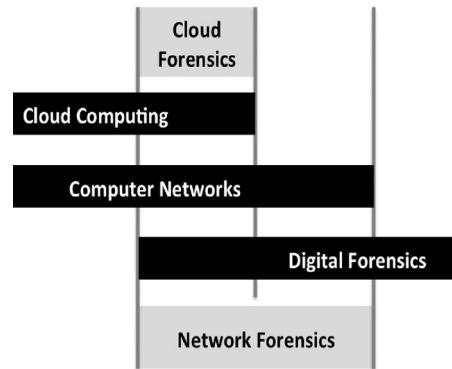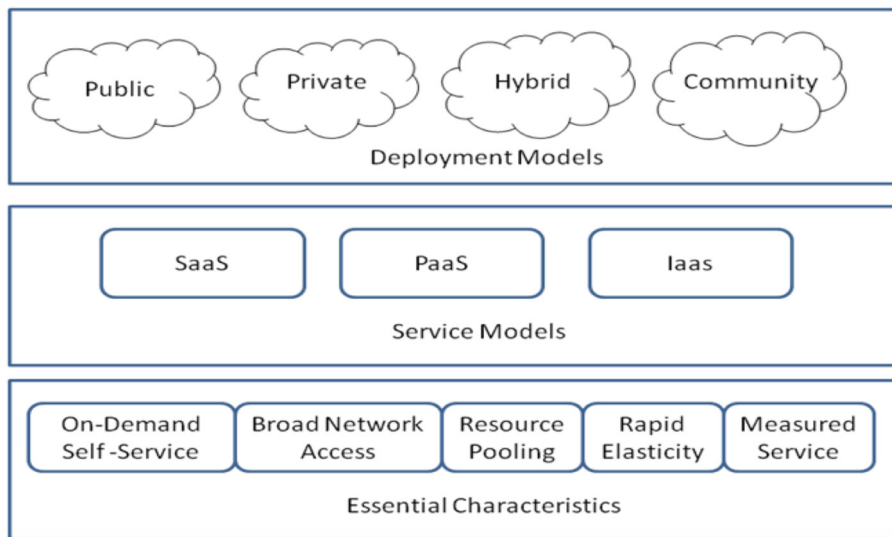

Fig. 1. Cloud forensics [2].

Fig. 2. NIST cloud model [3].

1.2. Cloud computing

The term cloud computing means sharing of computer resources among different users. As per NIST [3] "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". The cloud model consists of five essential characters, namely, on-demand self-service, broad network access, resource pooling, rapid elasticity, measured service, and three service models, namely: Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and four deployment models such as Public Cloud, Private Cloud, Hybrid Cloud, Community Cloud [3] as depicted in Fig. 2.

1.3. Digital forensics

Digital Forensics refers to "an applied science to identify an incident, collection, examination, and analysis of evidence data" [1]. The different phases of digital forensics are:

- Identification: Two major steps are involved in this phase, (i.e.) identification of malicious activity and isolating the evidence towards malicious activity.
- Collection: Evidences related to the malicious activity from different digital media are collected and the integrity of the evidence is maintained.
- Organization: In this phase, the examiner investigates the collected evidence which forms the examination phase and all identified evidence are correlated in the context of the malicious activity.
- Presentation: The investigator produces an organized report to the jury in the context of his investigation towards the case.
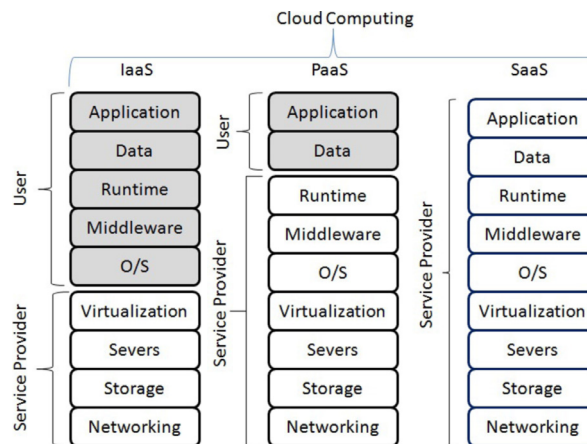
Fig. 3. Access control to the service models [23].

The rest of the paper is organized as follows. Challenges faced in cloud forensics are discussed in Section 2. The various relevant solutions in the literature are highlighted in Section 3. The limitation in current solutions and the proposed solution for mitigating those challenges are discussed in Section 4. Experimental setup for the proposed work and the inferences are depicted in Sections 5 and 6 respectively. Finally, conclusions are discussed in Section 7.

2. Challenges in cloud forensics

2.1. Data acquisition

This is the fundamental and vital step in the forensic procedure. Any flaw in this phase is passed on to the successive phases, resulting in transformation of the course of the investigation process. In digital forensics, investigators get hold of the affected computer (digital equipment) and carry out the investigation process by applying forensic principles in search of evidence towards malicious activity, to ensure no alteration in the evidence. But, in cloud forensics, grabbing the equipment is infeasible owing to the multi-tenancy and remote nature of cloud computing. Birk [4] indicates that the evidence may be in three different states in cloud, namely, at rest, in motion, and in execution. This will complicate the data acquisition compared to traditional forensics. Some of the challenges faced by the investigators towards data acquisition in a cloud environment are as follows.

2.1.1. Physical inaccessibility
Evidences are scattered and saved in different locations due to the significant characteristics of cloud. This leads to inaccessibility towards the collection of data and affects the data acquisition process.

2.1.2. Less control in cloud
In cloud, both users and investigators have restricted access, unlike digital forensics seizing of digital equipment is uncertain in the cloud. This complicates the data acquisition process in the cloud environment. The access control to the cloud varies by service models as shown in Fig. 3.

Only logs related to the application can be accessed by the investigator in the SaaS and PaaS models. In PaaS, application can be built by customers for getting certain additional forensics information compared to SaaS model, which has very limited access. Customers can move up to the operating system level in the IaaS model. IaaS has more privileges compared to the other two models. Even though access control is available for various levels in the cloud, forensics investigators have to anticipate Cloud Service Provider (CSP) for collecting data. J. Dykstra and A. Sherman quote the data acquisition problem by a hypothetical case study of child pornography [5]. In this case study, the authors have addressed the warrant issue, i.e. the location must be specified in the warrant, but the data is scattered and stored in various locations in the cloud. A cloud server cannot be seized by the investigator despite reaching the location due to its multi-tenant nature.

2.1.3. Volatile data
Virtual Machines (VM) are used by service providers for provisioning their customers. In this VM, volatile data like registry entries or temporary internet files will be lost if it is not synchronized with storage devices like Amazon S3, i.e. all information in VM is erased when VM gets restarted or shutdown [6].

2.1.4. Trust issue

Another serious problem is the dependence on the third party for collecting evidence in the cloud [8]. This issue is pointed out in a child Pornography case after a search warrant was issued. The investigator needs an internal staff to assist him in collecting data. Sometimes this person may be from the same CSP or may not be a certified investigator and this may affect the integrity of data to be produced in law.

2.1.5. Multi-tenancy

In cloud computing, different clients share individual resources. While acquiring evidence from cloud**,** two issues are addressed by the investigator. To start with, he has to prove that the extracted data is not mingled with other's data and has to maintain the integrity of the other user's data.

2.2. Logging

Analysis of logs is the first step in digital forensics. The logs may be process logs, application logs, system logs or network logs. These are the key for the investigation process, but getting this log data from the cloud is a crucial one. Several challenges that are recognized while obtaining logs are as follows [7–9].

2.2.1. Decentralization

In cloud, the logs are spread all over the network. Due to this phenomenon, the gathering of logs from various sources becomes difficult for cloud investigators.

2.2.2. The volatility of logs

Virtual Machines are used by CSPs for providing service to their customers. In the case of VMs, the volatile data like temporary internet files, registry data is completely lost once VM gets restarted or shutdown.

2.2.3. Accessibility of logs

There is no procedure or method for accessing logs in distinct places and the logs are used for troubleshooting, debugging, etc.

2.3. Dependence on CSP

Logs are collected and stored at CSP premises requiring the need for the investigators and the users to depend on CSPs for accessing network logs and server logs. In this point, CSP may tamper logs.

2.4. Chain of custody

Chain of custody is "the chronology of the ownership, custody or location of a historical object, document or group of documents" [10]. This is one of the most significant issues in forensic investigation, clearly indicating when and how the evidence was collected, analyzed, organized and presented in court [11]. Application of this procedure in digital forensics is easier than that of cloud forensics since seizing of equipment is possible in digital forensics. In the case of cloud forensics, this is not applicable because of its multi-jurisdictional laws and procedures. Hence a chain of custody produces many challenges [12,13] in cloud forensics. In a hypothetical case study of compromised cloud-based website, J. Dykstra and A. Sherman have highlighted the access available for multiple users to evidence. So, investigators have to depend on CSP for acquiring the chain of custody [5]. Birk et al. queried the reliability of hypervisor for a chain of custody [4].

2.5. Crime scene reconstruction

Crime scene reconstruction is infeasible in a cloud environment as data in VM gets erased completely when VM gets power off or rebooted [14].

2.6. Cross border law

Data centers afforded by cloud providers are distributed worldwide, so the cross-border law is an important issue in cloud forensics. The investigation process should be carried under the laws in the specific jury, whereas the measures for preserving data and chain of custody differ according to the jury and the entire investigation process will be affected by the cross-border law.

2.7. Law presentation

Presentation under jury is the final step in both digital and cloud forensics. In a cloud environment, thousands of VMs run in cloud data centers and hundreds of users are accessing simultaneously. This creates a serious challenge in cloud forensics than digital forensics [14].

3. Related works

This section highlights, the various solutions discussed by researchers for mitigating the challenges in cloud forensics are highlighted.

Dykstra et al. [5] have projected the use of cloud management plane in IaaS model. In this model cloud users and investigators have to trust the management plane to obtain data for investigation. Management plane in cloud premises makes the investigation process complex, but it mitigates the dependence of CSP.

Brik et al. [6] have recommended the use of application programming interface (API) to enable access log information to customers by read-only API, and the customer can provide information for the forensic investigation. Trusted third party issue was solved by this solution since customers are directly involved in continuous synchronization. But the dependency of CSP still exists and the authors have also suggested the encryption of logs before sending to the API for defending external breaches.

Marty [9] has suggested some guidelines for gathering logs which are used for forensics investigation. The author proposed a logging framework to solve logging issues that help in developing business oriented log framework which will be used by various IT professionals. SaaS model gets benefited by this framework.

Dykstra et al. [15] have offered a six layer trust model for mitigating the dependency of CSP and to preserve the trust. In this model, investigators have to trust resultant layers alone. IaaS model gets benefited from this model. Trust in individual layers and PaaS model is not discussed.

Wolthusen [16] suggest interacting evidence presentation and visualization mechanism for mitigating the dependence on CSP by granting confidence and trust. This model depends on CSP for getting access to the collected data as it is operated within the cloud and works for IaaS and PaaS models.

Zafarullah et al. [17] have proposed a solution inside cloud premises in IaaS model for getting OS logs and security logs. In this solution, distributed denial of service (DDoS) attack is launched in an eucalyptus environment. The service type and attacking machine IP are identified by logs in the eucalyptus. Dependence on CSP still exists since the solution is implemented inside the cloud premises. The authors conclude that the CSPs have to adopt a new mechanism for taking cloud forensics to a new extent.

Biggs et al. [18] have presented a universal law called global unity solution for cloud forensics investigation. This solution facilitates the investigation process, which solves Cross-Border law issue. For implementing this solution, all cloud providers have to adopt global unity solution.

Shah et al. [19] have highlighted the challenges and a possible malicious activity in cloud computing. The authors propose a three-layer architecture for cloud forensics. In this approach, the investigator has to depend on CSP for getting data.

Khorshed et al. [20] have highlighted the major threats in cloud computing and proposed a Support Vector Machine (SVM) technique performance based on kernels. They have created an attack set and compared it with other convenient machine learning techniques. Possible threats in the cloud are identified by this model, but predefined attacks are alone detected by this method.

Hale [21] has highlighted digital artifacts and procedures for forensic investigation that have to be followed by the investigator in Amazon cloud. Dependence on CSP exists for collecting evidence for the investigation.

Ruen et al. [22] have suggested the necessary things to be focused on cloud computing after conducting a survey over 257 respondents. Definition of cloud computing, cloud forensics, the significance of cloud, challenges, and opportunities of cloud forensics and research directions are included in this survey. Forensics as a Service is suggested by 55% of respondents and recommended by 87% of respondents.

4. Proposed solution

All previously identified solutions can be implemented only in cloud premises and investigators must depend on CSP for collecting forensic data for investigation. To overcome these limitations the proposed solution is implemented outside the cloud premises.

The proposed solution addresses the data collection issues discussed in literature by introducing a centralized forensic server and a forensic layer called forensic monitoring plane (FMP) outside the cloud Infrastructure, after obtaining permission from the international telecommunication union (ITU). So, the investigators need not depend on the CSP for collecting data.

The proposed model [24] for cloud forensics is shown in Fig. 4.1, where forensic monitoring plane (FMP) and forensic server are introduced for enhancing cloud forensics. The forensics tool such as forensic toolkit (FTK) analyzer, E-Detection running at the top of the FMP will monitor entire inbound and outbound connections in a cloud environment and the monitored data are forensically imaged (i.e.) bit by bit stream encryption and is stored in separate forensics server which is located in cybercrime premises. The forensic tool also monitors the actions of cloud service models. The tool automatically acquire**s** a forensic image of a current state in cloud service models which include VM and stores it in separate forensics server.

Hence all the actions, including network traffic in the specific cloud are forensically imaged whenever an event occurs, or the request processed in the cloud is acquired and is again encrypted and stored in the forensic server to enable reduction in the trust amount on CSP. The forensically imaged data is unaltered since bit by bit stream imaging is done during the
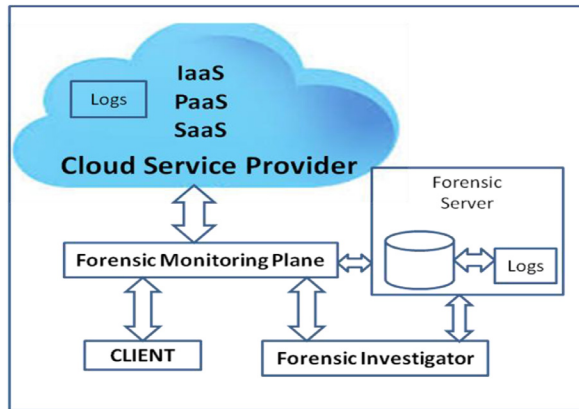
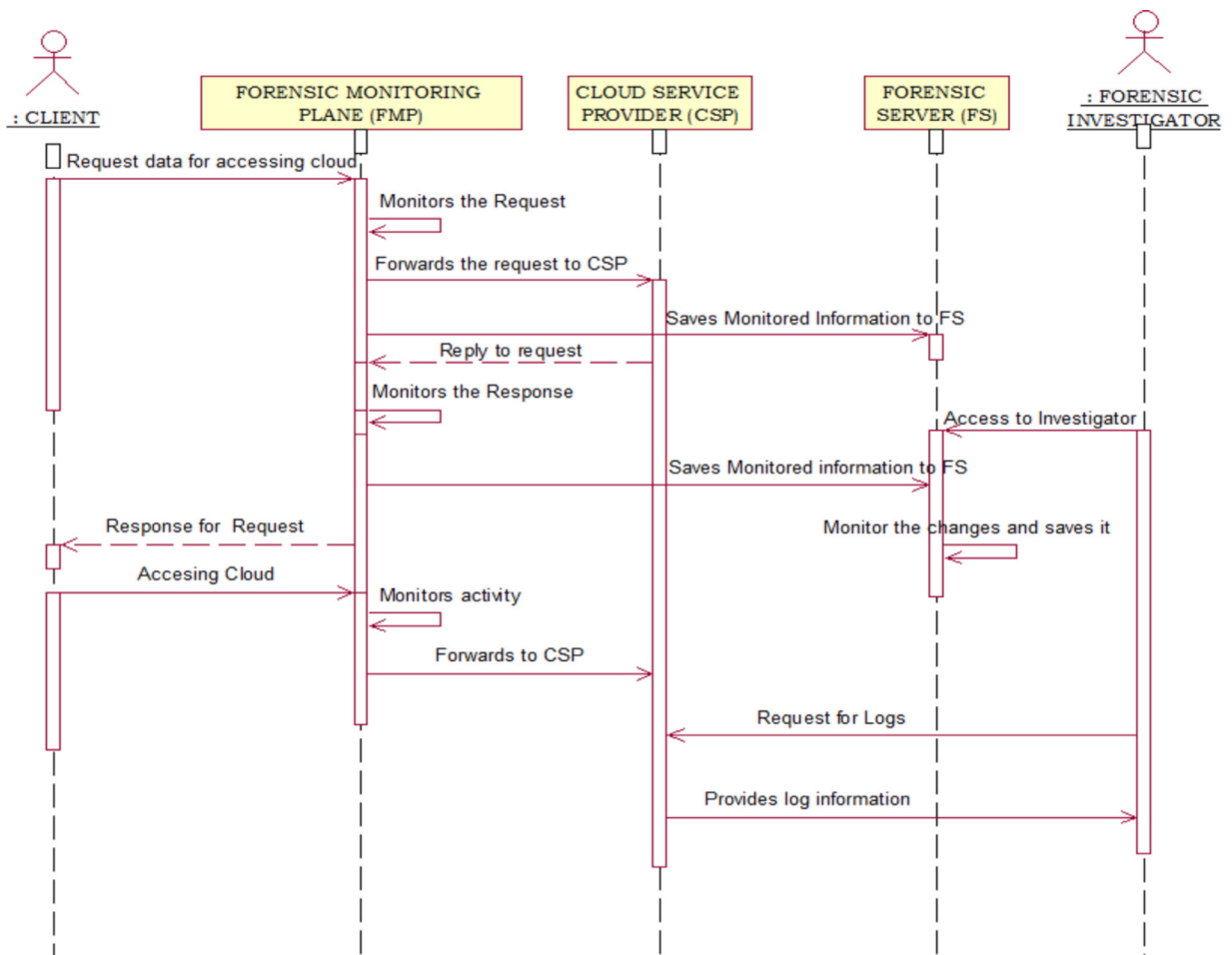**Fig. 4.1.** Proposed model for cloud forensics.

Fig. 4.2. Sequence diagram for proposed model.

forensic image process. The captured forensic images are not raw data and can be processed only through forensics tools. Network logs are also acquired from adjacent network devices (routers) and are imaged in forensics server that will give high proof for finding out the attacker.

In the case of any malicious activity, the investigator can directly login to the forensic server with their user credentials and can acquire forensic data within a time frame of the event. Meanwhile, upon suspicion, the investigator can request data from CSP and can verify it with the data obtained from the forensic server. Forensic tools are running in the forensic

server, and it takes the forensic image of the forensic server in the case of any unexpected event since there is also a chance of a suspect logging as a forensic investigator and tamper the data. The sequence of operation in **our** proposed model is depicted in the Fig. 4.2 by a sequence diagram for better understanding.

Steps involved in the proposed model:

- The client initiates the request to the cloud service provider.
- The request and response are intercepted by FMP monitoring tool which forwards the request to the server and the response to the client, and at the same time, it forensically images the request and saves it in the forensic server.
- Forensic investigator logs into the forensic server for analyzing the evidence collected.
- Actions in the forensic server also get forensically imaged and saved in the forensic server. If the investigator suspects the CSP, he initiates the request for evidence sources from CSP and compares with one another, and hence the integrity of the collected data also gets verified.

From the proposed solution, it is evident that dependence on CSP for acquiring data is reduced. The logging challenge is also reduced by storing log files separately in a centralized manner. The new proposed forensic model will take cloud forensics to the next level.

5. Experimental setup

A virtual prototype was created for this research work in the lab with four systems which include cloud server (own cloud), forensic management plane (FMP), agent system (i.e.) compromised system in that network and master system (Original Attacker) which is shown in Fig. 5 here FMP is validated by initiating distributed denial of service (DDoS) attack and verified that all necessary information's are captured in CMP. The DDoS attack setup is assumed to be working as a Master-Handler system which includes an agent system in the network. The attacker places an HTTP DDoS attack code on the agent and commands the agents through a remote connection, for launching an attack on the victim cloud server. However the scale of the prototype is small, it does bring the essential architecture of DDoS attack. We used our attack launch code as well as DDoS tools such as sprut [25] to bring down the server.
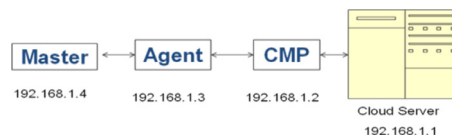

Fig. 5. Prototype setup.

Fig. 6.1. Server logs- owncloud.

6. Results and discussions

Forensics procedure starts with tracing back from server logs in the cloud server. The attack time is the main clue for tracing the attacker, but server logs are located within the cloud service provider and the proposed model collects the entire evidence outside the cloud environment. So, a remote log analyzer is used for capturing logs from the server and packet sniffer is used to collect information in FMP. During verification of data collected in FMP, server logs identify similar IP addresses at the time of the attack. This is shown in Fig. 6.1.

Further analysis on the captured packet shows HTTP packets flooded towards the cloud server and PsExec service was executed as shown in Fig. 6.2. Tracing agent from the victim is carried out by some existing IP traceback mechanism. The focus is to traceback to the origin of the attack.

Upon successfully locating the agent system, the use of DDoS attack is identified. When the search is carried out on the basis of the time of attack in the event logs, PsExec application is found running on the agent that correlates with an attack time. This is shown in Fig.6.3. The event logs in agents do not disclose any explicit information about the master-handler.

Event logs reveal only the name of the process related to attack time and are insufficient for tracing master. Further analysis on agent using forensics toolkit analyzer (FTK) reveals the file which uses PsExec service, and it is highlighted in the Fig. 6.4.
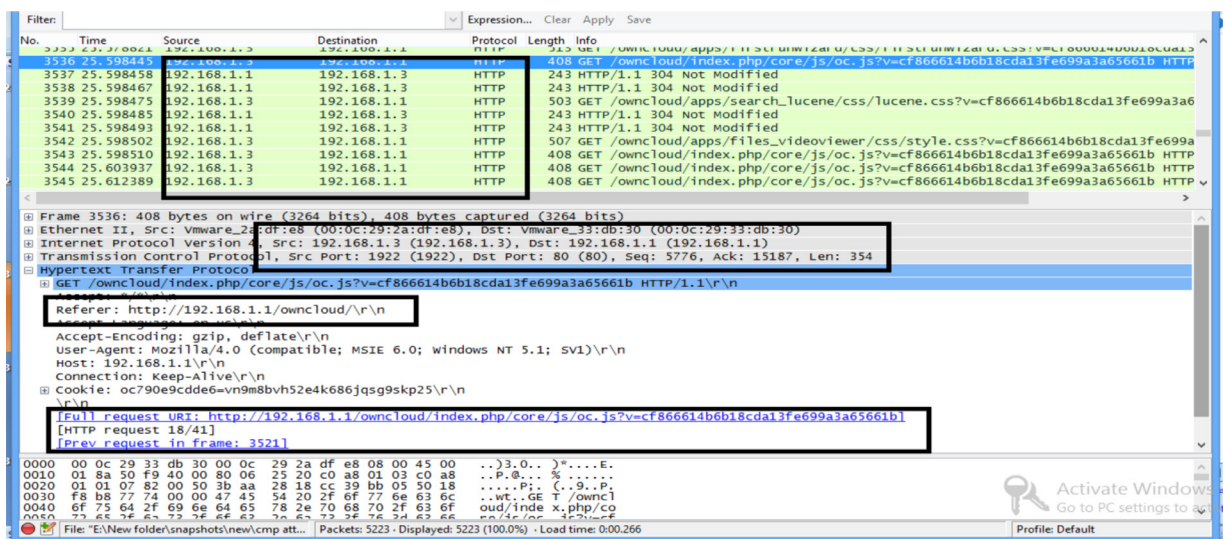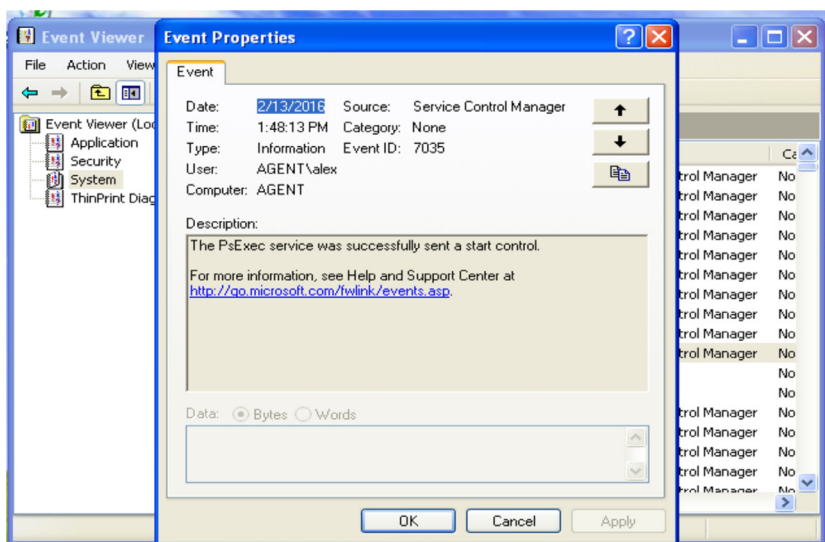

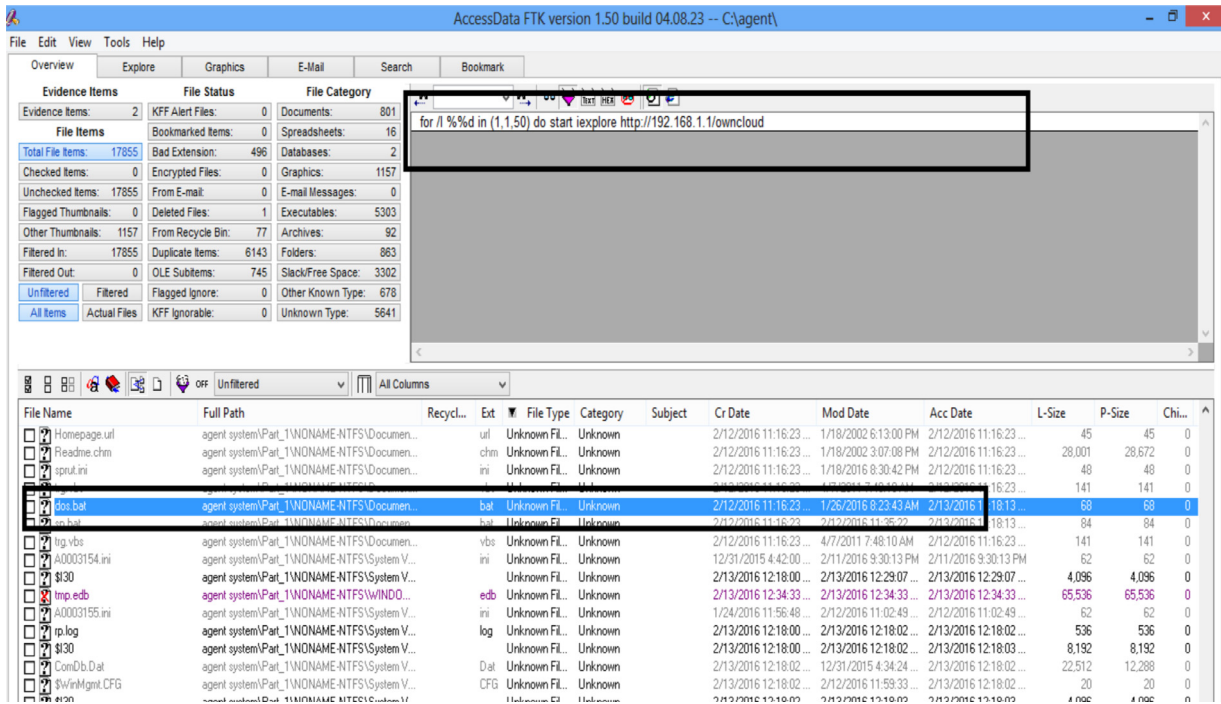**Fig. 6.2.** FMP-data.

Fig. 6.3. Agent – event logs.
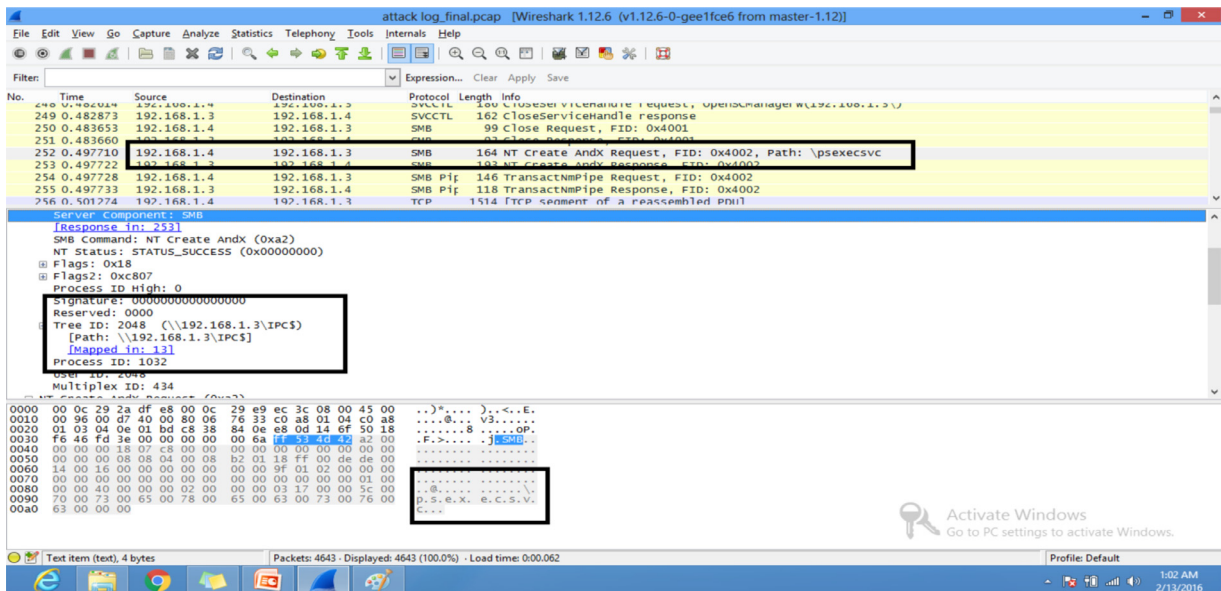
Fig. 6.4. Agent FTK.

Fig. 6.5. FMP – master to agent (attack initiation).

The earlier investigation indicates that PsExec service was started during the attack time as shown in Fig. 6.3, but an in-depth analysis on agent system using forensic tool kit (FTK) confirms that agent system is used for attacking the server and is controlled remotely as shown in Fig. 6.4.

Further analysis of network log collected in FMP indicates that PsExec service was started remotely by remote user as highlighted in the Fig. 6.5. Analysis on CMP data clearly indicates that the source of the remote process is not the agent system but the attack code is launched remotely from the agent. Even though the investigation has collected enough evidence against the agent and master of the attack, it reveals little information about the master system. On further analysis of FMP network logs, the master placing attack codes in the agent system can be identified as depicted in Fig. 6.6.
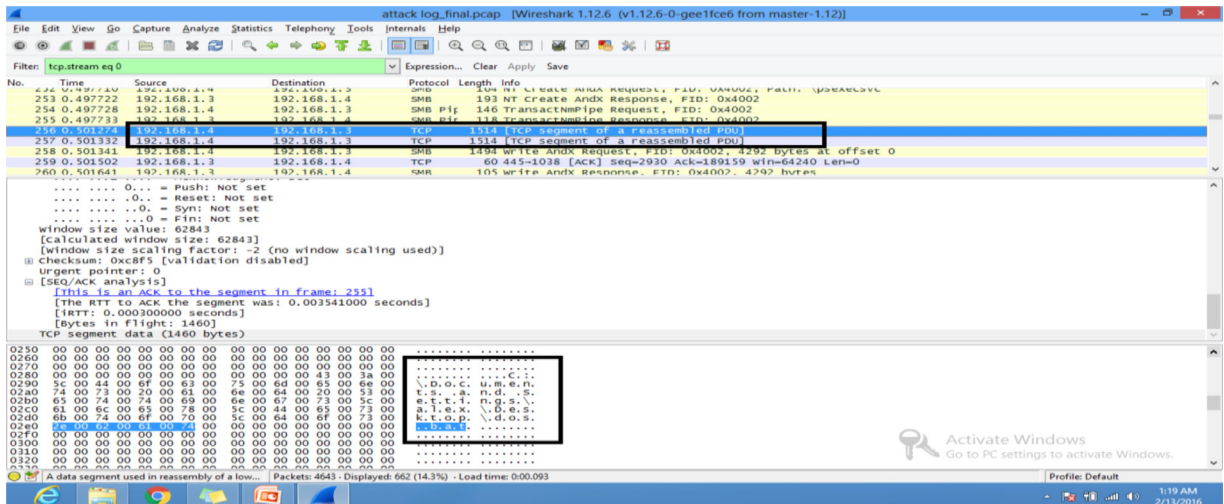
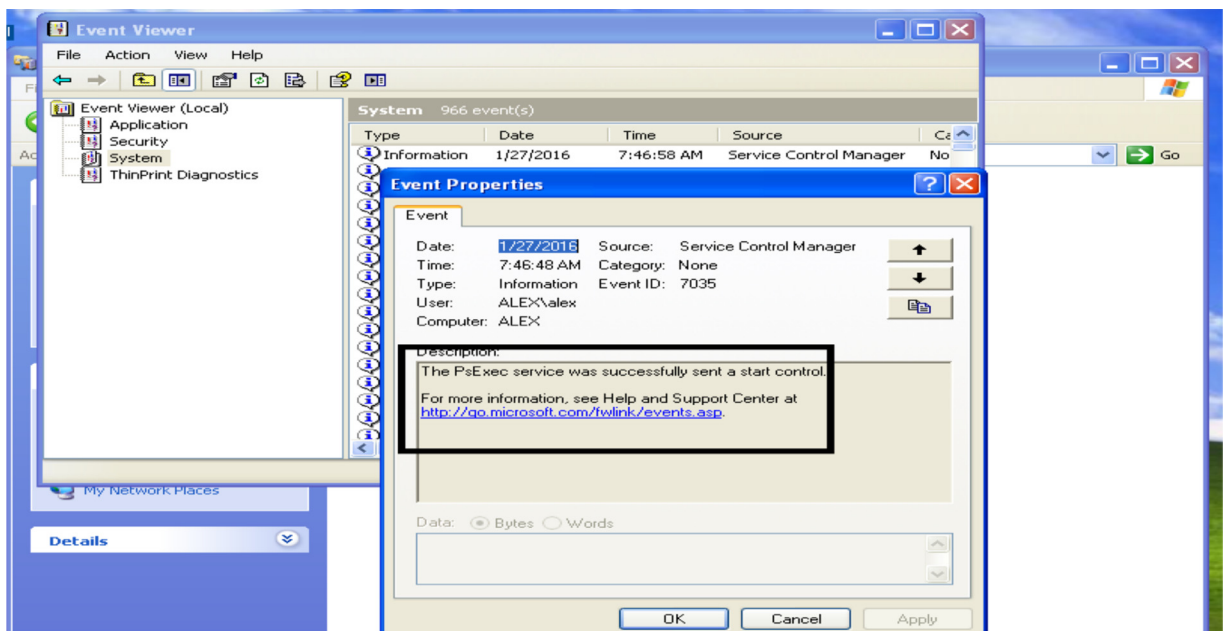**Fig. 6.6.** Master placing attack (dos.bat) file.

Fig. 6.7. Event logs – master.

For creating additional evidence against the master, the investigator moves into event logs in the master as done in the agent system and it indicates that PsExec service is started which correlates with attack scenario as shown in Fig. 6.7.

Further analysis on the master system using FTK analyzer reveals the attack code used for initializing the agent code that attacks the server. This creates additional evidence against the master, leading to the investigator's conclusion that the corresponding master is the original attacker. This is depicted in Fig. 6.8.

7. Conclusions and future work

The need for cloud forensics is on the rise, because of its rapid growth in cloud computing and due to the possibility of cloud-related crime occurring in the digital world. There are many challenges in cloud forensics and only a few researchers have addressed these challenges. In this paper, the challenges faced in cloud forensics and corresponding solutions addressed by the researchers have been highlighted in depth. A new model for mitigating the challenges in cloud forensics has been proposed and validated with DDoS attack to check whether the proposed FMP collects all necessary information related to fraudulent activities required for forensics analysis. In future, the entire attack scenario will be modeled inside the cloud
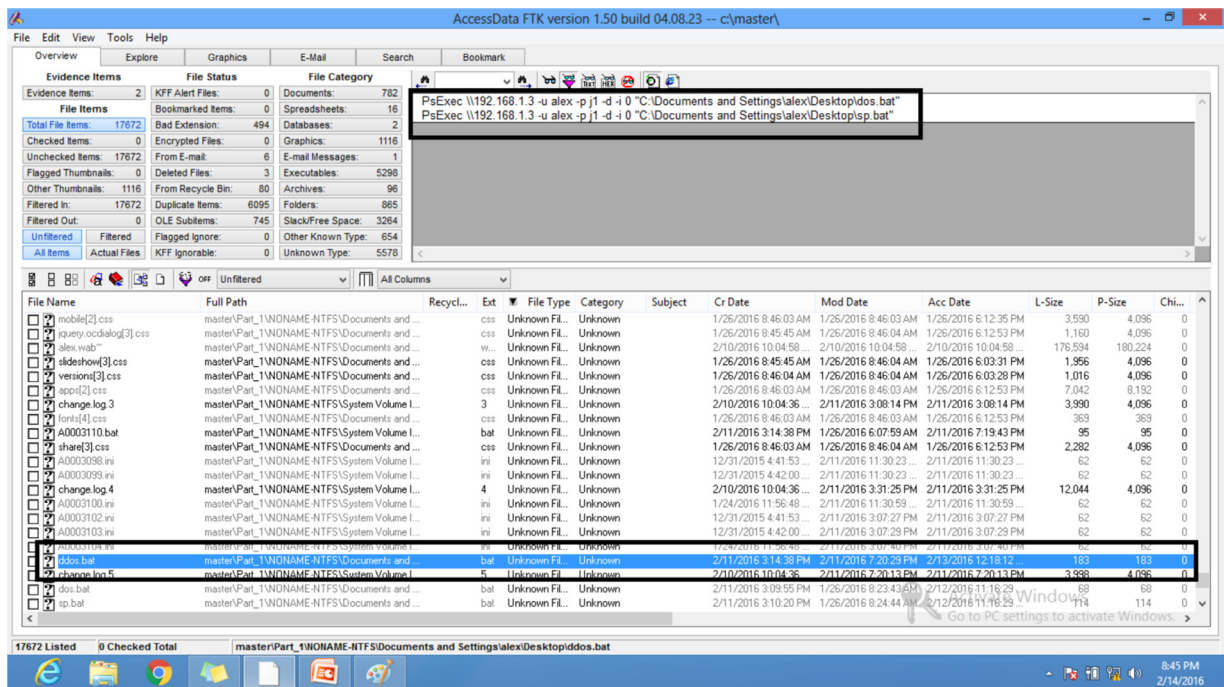
Fig. 6.8. FTK analysis - master.

environment for checking whether the proposed FMP collects all necessary information related to fraudulent activities. On completion, other modules in the proposed solution will be implemented.

References

[1] Kent K, Chevalier S, Grance T, Dang H. Guide to integrating forensic techniques into incident response. NIST Special Publication; August 2006. p. 800–86.
[2] Ruan K, Carthy J, Kechadi T, Crosbie M. Cloud forensics. IFIP advances in information and communication technology advances in digital forensics January 2011;vol. 361:35–46.
[3] Mell PM, Grance T. SP 800-145. The NIST definition of cloud computing. Gaithersburg, MD: National Institute of Standards & Technology; September 2011.
[4] Birk D. Technical challenges of forensic investigations in cloud computing environments. In: Workshop on cryptography and security in clouds, March; 2011. p. 1–6.
[5] Dykstra J, Sherman A. Understanding issues in cloud forensics: two hypothetical case studies. J Network Forensics 2011;b(3):19–31.
[6] Birk D, Wegener C. Technical challenges of forensics investigation in cloud computing environment. In: Proceedings of the 6th international workshop on systematic approaches to digital forensic engineering (SADFE); May 2011. p. 1–10.
[7] Guo H, Jing B. Forensic investigations in cloud environments. In: International conference on computer science and information processing (CSIP); 2012. p. 248–51.
[8] Ludwig Slusky MD, Partow-Navid P. Cloud computing and computer forensics for business applications. J Technol Res July 2012;3:1.
[9] Marty R. Cloud application logging for forensics. In: Proceedings of the 2011 ACM symposium on applied computing. ACM; 2011. p. 178–84.
[10] http://en.wikipedia.org/wiki/Chain_of_custody (accessed on 12/5/2014).
[11] Vacca JR. Computer forensics: computer crime scene investigation. Charles River Media, Inc.; 2002.
[12] Taylor M, Haggerty J, Gresty D, Hegarty R. Digital evidence in cloud computing systems. Comput Law Secur Rev 2010;26(3):304–8.
[13] Grispos G, Storer T, Glisson WB. Calm before the storm: the challenges of cloud. In: Emerging digital forensics applications for crime detection, prevention, and security, vol.4; 2013. p. 28–48.
[14] Reilly D, Wren C, Berry T. Cloud computing: pros and cons for computer forensic investigations. Int J Multimedia Image Process 2011;1(March (1)):26–34.
[15] Dykstra J, Sherman AT. Acquiring forensic evidence from infrastructure-as-a-service cloud computing: exploring and evaluating tools, trust, and techniques. Digital Invest 2012;9(Supplement):S90–8.
[16] Wolthusen S. Overcast: forensic discovery in cloud environments. In: Proceedings of the fifth international conference on IT security incident management and IT forensics (IMF). IEEE; 2009. p. 3–9.
[17] Zafarullah, Anwar F, Anwar Z. Digital forensics for eucalyptus. In: Frontiers of information technology (FIT). IEEE; 2011. p. 110–16.
[18] Biggs S, Vidalis S. Cloud computing: the impact on digital forensic investigations. In: Proceedings of the international conference for internet technology and secured transactions, ICITST. IEEE; 2009. p. 1–6.
[19] Shah JJ, Malik LG. An approach towards digital forensic framework for cloud. In: IEEE international advance computing conference (IACC); 2014. p. 798–801.
[20] Khorshed MT, Ali ABM, Wasimi SA. A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing. Future Gen Comput Syst 2012;28(June (6)):833–51.
[21] Hale JS. Amazon cloud drive forensic analysis. Digital Invest 2013;10(3):259–65.
[22] Ruan K, Carthy J, Kechadi T, Baggili I. Cloud forensics definitions and critical criteria for cloud forensic capability: an overview of survey results. Digital Invest 2013;10:34–43.

[23] Almulla S, Iraqi Y, Jones A. Cloud forensics: a research perspective. In: Innovations in information technology (IIT), 2013 9th international conference on, 17-19 March; 2013 66,71..

[24] Alex ME, Rajendiran K. Forensic model for cloud computing: an overview. In: IEEE international conference on wireless communications, signal processing and networking (WiSPNET) 23-25 March; 2016. p. 1334–8.

[25] http://ihackers.co/sprut-dos-tool-dos-attack-tool/ (accessed on 12/5/2014).

Edington Alex. M, graduated from Francis Xavier Engineering College in Information Technology. He obtained his Master degree in Computer and Communication from SSN College of Engineering, Chennai, during the year 2011. At present he is working as Assistant Professor at Rajalakshmi Engineering College, Chennai. Research interests include Cloud forensics, digital forensics, cloud computing, network security, digital forensics and cryptography.

Kishore Rajendiran, graduated from Madras University, in Electronics and Communication Engineering. He obtained his Master degree in Communication Systems from Pondicherry Engineering College and Ph.D. from Anna University, Chennai. At present he is working as Associate Professor in the Department of ECE, SSN College of Engineering, Chennai. He has 15 years of teaching experience. His research interest includes security issues, Cloud computing.