Analysis of Dataset in Private Cloud for Cloud Forensics Using Eucalyptus and Hadoop

Mayur S Patil,
SCET,
MIT Academy of Engineering,
Pune, India
mspatil@mitaoe.ac.in, ram.nath241089@gmail.com

Bharati Ainapure,
Department of Computer Engineering,
Vishwakarma University,
Pune, India.
ainapuressa@gmail.com

Abstract — **At present, in most of the areas of research, development and daily usage, almost all platforms and service providers are dependent on cloud computing. To achieve scalability and sustainability, cloud architecture is considered as a result of technological amalgamation. If the design of the architecture is complex in nature, it could have adverse effects on data recoverability and analysis when the system gets compromised. This type of structure introduces issues like inherent architecture flaws, backdoors, code smelling which could lead to exploitable vulnerabilities for hackers and scammers. Likewise, its complex nature puts limits on forensic investigation methods. For such problems, digital forensic provides a solution. It uses the three-stage approach which consists of evidence collection eye-witnessed by the user, evidence preservation in an unaltered way, offline safeguarding of evidence for feigning its collection. Among its various methods like live forensics, timeline analysis, logging, sandboxing, logging is generalized and easy to use method. The proposed work is about digital forensic technique, log analysis, which is the most effective approach to override investigation issues in the cloud environment. Snort, Network Intrusion Detection System (NIDS), will work as a daemon on Eucalyptus private cloud to monitor and log intrusive attempts of network activities on it. Second, the proposed approach will fix the incapability of Eucalyptus to export logs to the rSyslog server. Third, the generated datasets, irrespective of location and format, would be analyzed by Hadoop, for improved analysis of a system**.

***Keywords** – Cloud Forensics, Dataset, Eucalyptus, Hadoop, Logs, Snort.*

I. INTRODUCTION

Cloud computing is continuously evolving technology which is offering various types of services for different fields. In a controlled environment like hybrid and private cloud platforms, there are high chances of security risks because of mixed resources and infrastructural requirements. Other factors that also impose user data risks are restricted policies, lack of control over infrastructure. Hence, getting dataset used for forensic purpose is also toilsome. For investing both systemand network, cloud forensics concept has introduced to tackle such challenges [1], [2]. Security analysts are more interested in cloud forensic because of issues like software architectural glitches and coalescence of multiple technologies under one platform in a distributed environment. Digital forensics is the collection of techniques in computing the world used to find out the detection and nature of crime with the help of evidence [3]. It applies to one system or group of systems. On an extended level, to maintain artifacts as a result of the procedure in a network, the use of network forensics in such cases becomes essential.

Considering the networking point of view, cloud computing is a network of distributed and virtualized nodes along with scheduled management of resources as per organizational policy [4]. In such a case, strategies need to be mooted about the geographic area of user, authentication policies of the service provider, and consumption of resources from user account. It puts various kinds of access limitations on user data stored on a cloud which is again geo-dispersed. Also, incremental demand for infrastructure results in the mixture of various software and hardware resources makes it difficult for investigators to apply forensics techniques on cloud system architecture [5].

The main issues in cloud forensic are the incidents in which pieces of evidence and their states need to be digitally examined in stipulated time. Another challenge is red tape legislation adds complication for accessing digital data of customers. It’s a paradox situation in which forensic analysts need to access the data which comes under CSP territory but prohibited by legal jurisdiction [6], [7]. Location-based log analysis suffers from a similar kind of problem along with different formats and structures [8]. Problems based on Infrastructure as a Service (IaaS) clouds are allocated logical storage, hosted computation and virtual networks. In such cases, live analysis or cross drive analysis provides some means of reliable solutions [9]. Considering large data stores in storage services, it becomes difficult to get analysis of humongous data within a stipulated time. Using Hadoop map-reduce framework, it is possible to get results of different datasets into a single desired format [10], [11].

Antiforensics is a set of techniques, tools, and methods that can easily make negative impressions on finding out the evidence required for forensic investigations [12]. Because of antiforensics, it becomes a nightmare for investigators even looking for simple sort of data. Approaches include encryption, data wiping, information hiding and so on. The idea behind this is to protect individual privacy and confidentiality but due to their open nature, they are easily available to the attacker and black hat hackers to compromise either individual or corporate systems. While considering the scope of the cloud, the foul play of such powerful tools can make investigation tough.

The diorama of the paper is stated as follows: section 2 deals with a literature survey, section 3 deals with the proposed system, section 4 deals with expected outcomes and section 5 describes conclusions and future work.

II. LITERATURE SURVEY

The 'forensics' term has a wide range of verbose as well as implementations. Considering the opinions, forensics experts and CSPs are having a wide range of differences in their opinions while making decisions about strategies. From forensics expert perspectives, it is a proactive way i.e. the methods to protect the system and analyze it considering future threats and issues. Result of which will help the system to prevent itself from getting compromised or affected in the future from similar kind of attacks. Conversely, a cloud service provider (CSP) prefers a reactive approach i.e. to heal the system after being compromised.

At present, easier access to the internet has contributed to a steep increase in cybercrime activities. Due to its constant growth, it is becoming difficult for even security analyst to pause before acting as intruders are using new techniques and reverse engineering tools to perform penetration testing, doesn't matter how forefront it is [5]. Most of the cases are Distributed Denial of Service (DDoS) attacks on social networks like Facebook, Twitter, Snapchat, etc. which are also examples of Storage as a Service (SaaS) cloud. It's quite challenging for even security experts and architects to find out the exact root cause of the problem through which attack can be triggered.□

In Infrastructure as a Service (IaaS) based private cloud, it is quite possible that vendor-hosted or multi-tenant architecture can introduce user data risk for several reasons. Here, multi-tenant architecture means each customer, as a tenant, will be provided with services using virtualization and web services. Some of the possibilities are malicious intentions of CSP itself which means self-involvement in breaching policies defined for data protection or he is a victim of compromised data by hackers or stolen data by an internal or external source.□

Considering the derivation of characteristics of network forensics, cloud forensics having an inherent network analysis toolkit include in its arsenal. During the investigation, it is possible that along with comprised affected user's data other users' data might get exposed due to the shared nature of resources among multiple users. To this problem, live analysis is helpful but considering its nature of agility.

There are very rare chances for an investigator to get this opportunity. It is nearly as impossible to restrict the monitoring of resources in a cloud environment. This is where traditional methods of digital forensics die on the vine. When users use virtual machines, which are on-demand, they perform their jobs on it and terminates machines after use. In the case of forensic analysis, it is difficult to track records or traces of virtual machine after user exited from it. In this case, the real challenge is having a virtual machine ready for the introspection with the compromised hypervisor. It means if the hosting platform is infected or hacked, it will automatically affect hosted virtual machines as a means of native components of its environment [13] [14]. Other factors such as SLA based cross-border based regulations' norms and conditions prohibit investigators, users, and vendors to proceed further in the investigation. Often, two dependent policies can leverage problem on a different level, which seems to be easy but ambiguous. For example, considering the case of the data protection act, due to the distributed nature of the cloud, it becomes difficult to investigate issues in cloud storage at run time [15]. It is only possible if there is global standardization but not yet possible due to the unanimous decision of CSP [6].

Evidence collection, identification, preservation, analysis, reconstruction, and reporting are the basic steps of traditional forensics procedures. Techniques such as tunneling of VPN, image recording of data evidence can help to make investigation smoother than the traditional approach [8] [9] [16]. Sometimes, it has observed that there exist CSP's policies for the customer, some of them which might include gray shades about privacy [17] [18]. It can impose the risk of jeopardizing user data by improper handling methods that can lead to several questions but again legal and jurisdiction constraints make it difficult to handle the same. In a standardized environment, it becomes difficult to apply resources tagging and allocation for each customer on an individual basis.

Considering networking in a cloud environment, there is a need to conceive data communication between several machines. For investigation purposes, it is essential to look after challenges in network forensics [17] [19]. It consists of the live analysis of network traffic, detecting misuse and anomalies in the network by any means i.e. either false positive or false negative. So in such types of cases, the intrusion detection system (IDS) seems to be useful for analysis and gather information about malicious network activity [20]. Using multilevel IDS along with log exportation tools like rSyslog, it is feasible to imply authentication which can be deployed on a cluster level and snort rule-based levels like confidence, sensitivity, scope in support of anomaly detections [21] [22].

Among seven modes of snort, full mode of NIDS and common authentication policies, there are the possible chances in which exposure of other peoples' data in a group to forensics investigators along with compromised user(s). From the survey, it has discovered the point that there is a notable divergence between problems and techniques to solve them. Hence, log analysis is a forefront choice for the proposed system. As per the study, it is quite effective yet lightweight complementary solutions to forensic techniques to ease the process of investigation. Also, it is one of the handy methods popular among forensics experts [5] [8] [23] [24]. It can be used as an effective means for the preparation of datasets on a large scale. It can consist of the user logged in & logged out attempts, network activities and also explanation of when and why the logs are created [23]. In a cloud environment, it is immensely helpful to detect the errors & problems, whether they belong to Eucalyptus or Snort IDS. Some of the notable log files are cluster logs to detect capability or information exchange issues with clusters, fault logs to know details of known solutions and details error code, debug logs generated from the level of debugging-based logging.

III. Proposed System

The aim behind this paper is neither about finding the vulnerabilities nor cluster formation of the Eucalyptus cloud which is out of scope. It focuses on log analysis in Eucalyptus using Snort and Hadoop. The Eucalyptus v4.4.5 is the latest stable release channel [25]. Administrators and users can easily access and control instances according to their roles using its web service based GUI. Snort is used for intrusion detection. Among its configuration modes which are packet sniffing, packet logging, and network intrusion detection system, NIDS uses promiscuous mode to detect traffic not intend for its MAC address.

Snort running as a daemon will help the system to prepare for self-defense mechanism in such a way that whenever any Linux OS will complete its booting process, along with other daemon processes, snort will also start to run in the background. Generally, snort captures each packet came across its way but when configured for NIDS, it becomes very selective about detection of packet reached to a network port. Hence, using user-defined snort rules in snort.conf file, suspicious packets on the port will be detected and logged in a custom file that can later used by Syslog for further process. In short, using NIDS, snort will effectively log only those packets who try to activate rules in snort.conf and generate alerts about it.

Here, logs or log message is used for summarizing the details of events occurring throughout the operating system at a specific location. Its general format is the date, timestamp, source which can be relegated into assortments like errors, alerts, information, warning, debug, critical. These categories can help to understand the meaning of logs as per its level of importance. Logs can be generated using protocols like SNMP, Reliable Event Logging Protocol (RELP), Syslog. Among which Syslog is most commonly used, easy to configure and flexible for functionality compared to others. So, it is used to manage generated log data. Log data or dataset is the abstract information to point out reason or purpose that why log messages are generated [26]. It will be useful to perform log analysis.

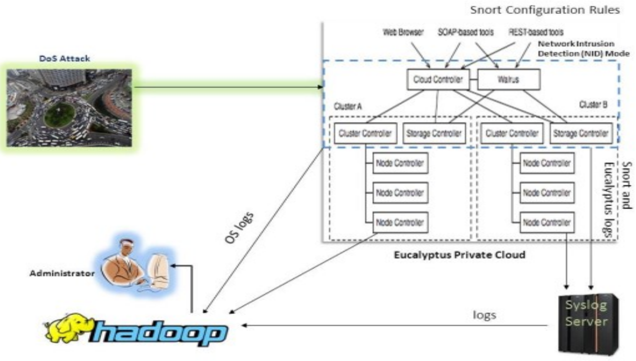

Fig 1: Eucalyptus Dataset Analysis Proposed architecture

The proposed system includes three physical machines; two having CentOS with Eucalyptus Components installed on them; One is having Front end components (Cloud Controller (CLC), Walrus Controller (W), Storage controller (SC), Cluster Controller (CC) while the other one with installation of Node Controller (NC) only. The third machine is a vanilla Linux installation without any of the cloud components.

A. Technical Specifications

Each system should have 300 GB HDD, 4 GB RAM and Intel© core 2 duo processor with VT enabled technology each. On the front end, the Snort IDS will be installed on two Eucalyptus machines and Hadoop 2.9.2 on a machine outside of a cloud network. After setting Eucalyptus, Snort needs to be installed on the front end. Now, Snort rules need to be configured with updated rule sets from the official Snort website. Here, if DDoS attack is performed, to detect it Snort rule will look like as follows:

alert tcp $EXTERNAL_NET any ⟶ $192.168.45.2/22 any (msg: “This is a sample rule for DDoS Attacks”; flow: established, to server;)

This simple snort rule gives an idea regarding the working of snort in detecting DDoS in the network. This rule explains that alert should be generated for TCP connection if an attack from an external network irrespective of any port is performed on the specified IP address of port 22 or any other. Then send a message which is in double quote and flow of packets towards the server direction with state of connection specified. Here, whatever activity snort performs, it gets stored at location /var/log/snort and it should be exported to an external server by time for analysis purposes. Due to supportive logging methods, there are chances to miss important and critical alerts about like segments of network connections because of IDS was triggered. Also, due to the lack of detailed logs and summarized results, it becomes quite difficult as well as time-consuming to detect the exact cause of actions in the network.

It will also try to analyze the same for Eucalyptus front end cloud components and export logs to Syslog server. Due to the logging problem discussed earlier, logs will be processed from different locations with Hadoop map-reduce implementation [27], [28]. This proposed system will prove to be helpful in the process of forensic investigation to solve the problem of volatile nature and storage of logs to some extent.

B. Mathematical Conventions

There are three components mainly included; Eucalyptus Front End, NC; and Syslog server. In this scenario,□

$$EF_m \equiv M$$
$$ENC_m \equiv W_1$$
$$S_s \equiv W_2$$

where the triple bar is used for identical equivalence. Notations explained as follows:

EF_m = Eucalyptus Front End Machine
ENC_m = Eucalyptus NC Machine
S_s = Syslog server
M = Master Node
W_1 = Worker Node 1
W_2 =Worker Node 2

When the attacker or a third party perform an attack on the front end, the equation becomes:

$$A_i \rightarrow EF_m$$

where i is the number of machines used for an attack.

Here, the EF_m consists of logs of OS, Snort and Front End Components of Eucalyptus, so the equation becomes:

$$EF_m \supseteq (E_{fel} \cup S_l \cup O_{fl})$$

where E_{fel} = Eucalyptus Front-end logs
S_l = Snort logs
O_{fl} = Operating system front end logs

Similarly, in the case of Eucalyptus Node Controller machine, it becomes:

$$ENC_m \supseteq (E_{ncl} \cup O_{sl})$$

where E_{ncl} = Eucalyptus NC logs

O_{sl} = OS logs of node machine

Here, logs need to be exported from the Eucalyptus front end machine and NC as well. On the other hand, log exporting capability of Eucalyptus needs to be checked as well as affected components of the system on which attack has been performed.

The map-reduce model needs to implement it. When worker nodes will be ready, data present on it can be efficiently used by the map function. Also, the reduce method will help to collect processing input on the master node, so by using a standard map-reduce equation.

For Map function:

$$map(l,n) \rightarrow list(l_r, val_l)$$

where l is a format of input from the n locations
l_r is the output list for map function
val_l is the values after map function

Now, for Reduce function:

$$reduce(l_r, list(l_r, val_l)) \rightarrow list(O_r)$$

where l_r, in this one, input for reduce function
list (val1), a list of all sorted outputs.
list (O_r) is the output of reduce function.

In set theory, the equation will be as follows:

$$M = W_1 \cup W_2$$

C. Modular Data Flow

Data input, in this case, are the packets that are flowing transmitting across the network. Those packets are meant to be analyzed by Snort IDS that should monitor and alert if suspicious content would be found. This record will get exported to the Syslog server and used for map-reduce as per the need of the forensics investigation.

Here, Hadoop comes into the picture. It has the main advantage that using a cluster of local machines, each node acts as a means of terabytes or petabytes of storage with no dedicated RAID machine. Also, in a cloud environment, Hadoop is useful for computational scaling. As a load of data processing increases, the frequency of task distribution among slave nodes increases. In this case, Hadoop analyzes these logs from different locations such as front end, node controller and Syslog server. Namenode stores inode information of all log data and machine outside network acts as datanode resulting in storage of processed logs. This structure will reduce the burden of processing on network and machines as log data need to process already resides on the worker node or datanode.

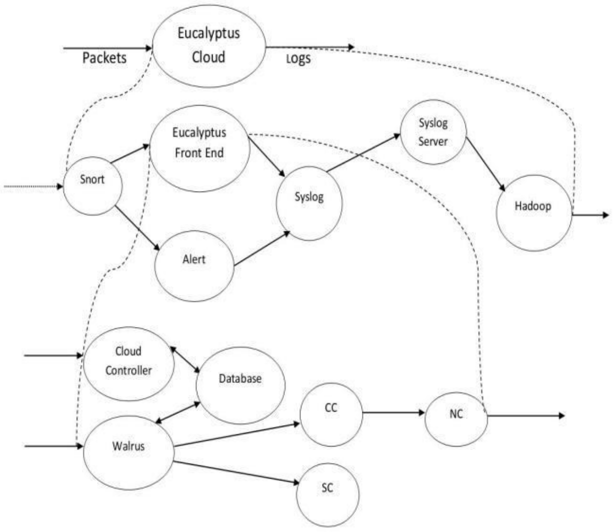

Fig. 2. Data Flow Diagram

Context levels in the data flow diagram are defined in a top-down manner. Level 0 indicates a Eucalyptus cloud with logs and packets. It expands to level 1 with snort for packet capture and alerts send to Syslog including Eucalyptus front end. At the end of this phase, Syslog server data get processed by Hadoop which is the proposed structure. Level 2 indicates the architecture of the Eucalyptus cloud.

D. State Diagram

This simple state diagram shows how data is flowing in the system as well as the implementation of the proposed system.

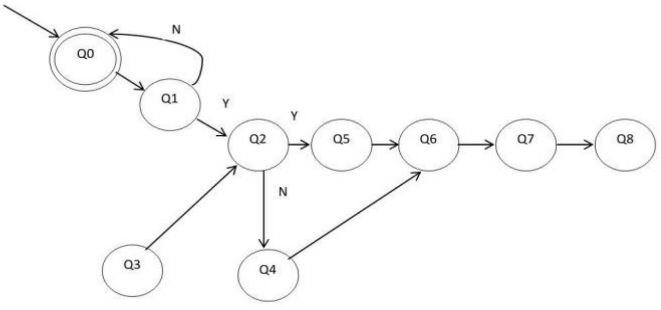

Fig.3. State Diagram of Workflow

Q_0 = User login, Q_1= Credentials verified, Q_2= User account access in Eucalyptus, Q_3 = Snort IDS Monitoring, Q_4 = Attacks performed, Q_5= Logs of Eucalyptus cloud and system components, Q_6 = Attacker event logs due to snort rules, Q_7 = Hadoop NameNode for indexing data, Q_8 = Hadoop DataNode for contents of indexed data.

The user will first log in into Eucalyptus has represented by state Q_0. If invalid, it will return to state $Q_{0;}$ otherwise, it will move to state Q_1 which will verify credentials. After this state, the user will get access to a user account on state Q_2.

Here, attacker state Q_4 is introduced as a means to perform attacks. To verify whether that user is legitimate or outlawed, state Q_3 snort IDS will check these attempts using either community-contributed or user-defined rules. If these logs remain sane, snort logs will not appear on the log server. Otherwise, attack based events will be logged on state Q_6 as a means of snort logs. Meanwhile, logs of Eucalyptus cloud components and Syslog also generated which is represented by state Q_5.

Now, Q_7 & Q_8 will make use of Hadoop for effective processing. Q_7 will collect logs using Hadoop from state Q_5 & Q_6. For result analysis, Q_8 state will store processed multi-input format logs into a unified and understandable format.

IV. EXPECTED OUTCOMES AND DISCUSSIONS

The expected outcomes are when any kind of service-based or packet-based attacks will be performed on Eucalyptus, Snort must be able to monitor and detect the same. After the monitoring of attack, logs need to be gathered from Eucalyptus Cloud and Snort IDS. These logs need to be exported on a machine outside of the cloud network, in case the whole cloud network is compromised. These logs can be configured and managed using Hadoop for efficient data processing using the map & reduce function for collecting and processing respectively. Due to different extensions and structure of logs, Hadoop will help to parse those heterogeneous and multi-input format logs in one format for further process.

V. CONCLUSION AND FUTURE WORK

In cloud forensic, continuous evolution with nippy adoptions of systemarchitectures increases security challenges [29]. There are still ambiguous issues such as evidence access period (due to the volatile nature of data), live forensics (presence of evidence at investigation time) and cross drive analysis (multiple hard drives analysis mainly obstructed by legal as well as jurisdictional issues). This proposed method of log analysis provides a trusted solution to some extent. Considering the dynamic nature of datasets, the exportation process to outside network server seems to be helpful and effective for analysis purposes. Also for analysis of datasets, flume can be used for solving location and format problems of logs. Flume will be mainly useful if this implementation is for a distributed system to get processed data on a centralized system.

The scope of the paper can be further extended, from a forensic point of view, by the phenomenon of Event Regeneration. In detail, by taking a snapshot of Virtual Network Environment (VNE), maintaining minimum downtime, ensuring the attack stage and whole events of attack get recorded. Analysis of attacks could be done by fuzzy clustering techniques. With the help of it, a scope can extend towards the development of Forensics-as-a-Service (FoaaS) [30] which can provide built-in support for a cloud platform to investigate issues at its earliest considering volatile the nature of data in the cloud.

REFERENCES

[1] M. Uphoff, M. Wander, T. Weis, and M. Waltereit, "SecureCloud: An Encrypted, Scalable Storage for Cloud Forensics," *Proc. - 17th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. 12th IEEE Int. Conf. Big Data Sci. Eng. Trust. 2018*, pp. 1934–1941, 2018.

[2] D. Reilly, C. Wren, and T. Berry, "Cloud Computing: Pros and Cons for Computer Forensic Investigations," *Int. J. Multimed. Image Process.*, vol. 1, no. 1/2, pp. 26–34, 2011.

[3] Z. Chen *et al.*, "Digital forensics for network, Internet, and cloud computing: a forensic evidence guide for moving targets and data," *Proc. - 17th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. 12th IEEE Int. Conf. Big Data Sci. Eng. Trust. 2018*, vol. 1, no. 12, pp. 110–116, 2018.

[4] D. Kumar, "REVIEW ON TASK SCHEDULING IN UBIQUITOUS CLOUDS," *J. ISMAC*, vol. 1, no. 01, 2019.

[5] T. V Lillard, *Digital forensics for network, Internet, and cloud computing: a forensic evidence guide for moving targets and data*. Syngress Publishing, 2010.

[6] Zafarullah, F. Anwar, and Z. Anwar, "Digital forensics for Eucalyptus," *Proc. - 2011 9th Int. Conf. Front. Inf. Technol. FIT 2011*, pp. 110–116, 2011.

[7] B. C. Sekhar, "Access Control for Cloud Forensics through Secure Logging Services According to Guide to Computer Security Log Management," *2017 Int. Conf. Energy, Commun. Data Anal. Soft Comput.*, pp. 3527–3532, 2017.

[8] M. Damshenas, A. Dehghantanha, R. Mahmoud, and S. bin Shamsuddin, "Forensics investigation challenges in cloud computing environments," in *Proceedings Title: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec)*, 2012, pp. 190–194.

[9] Z. Chen *et al.*, "Electronic Evidence Service Research in Cloud Computing Environment," *Proc. - 17th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. 12th IEEE Int. Conf. Big Data Sci. Eng. Trust. 2018*, pp. 332–338, 2018.

[10] V. Amiry, S. Z. Rad, M. K. Akbari, and M. S. Javan, "Implementing hadoop platform on eucalyptus cloud infrastructure," in *2012 Seventh International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, 2012, pp. 74–78.

[11] V. Bhosale, A. Thakar, C. Pandit, A. Deshpande, and H. Khanuja, "Hadoop in Action: Building a Generic Log Analyzing System," in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2018, pp. 1–7.

[12] S. Garfinkel, "Anti-forensics: Techniques, detection and

countermeasures," in *2nd International Conference on i-Warfare and Security*, 2007, vol. 20087, pp. 77–84.

[13] H. Guo, B. Jin, and T. Shang, "Forensic investigations in cloud environments," in *2012 International Conference on Computer Science and Information Processing (CSIP)*, 2012, pp. 248–251.

[14] D. Birk and C. Wegener, "Technical issues of forensic investigations in cloud computing environments," in *2011 Sixth IEEE International Workshop on Systematic Approaches to Digital Forensic Engineering*, 2011, pp. 1–10.

[15] Y.-Y. Teing, A. Dehghantanha, K.-K. R. Choo, Z. Muda, and M. T. Abdullah, "Greening Cloud-Enabled Big Data Storage Forensics: Syncany as a Case Study," *IEEE Trans. Sustain. Comput.*, vol. 4, no. 2, pp. 204–216, 2017.

[16] A. Huseinovic and S. Mrdovic, "Comparison of computer forensics investigation models for cloud environment," *2018 41st Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2018 - Proc.*, pp. 850–853, 2018.

[17] S. Ishihara and T. Akiyama, "A Tuning Method of a Monitoring System for Network Forensics in Cloud Environment," *Proc. - Int. Comput. Softw. Appl. Conf.*, vol. 1, pp. 951–954, 2018.

[18] D. Zou *et al.*, "A multigranularity forensics and analysis method on privacy leakage in cloud environment," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1484–1494, 2019.

[19] B. Hay, K. Nance, and M. Bishop, "Storm clouds rising: security challenges for IaaS cloud computing," in *2011 44th Hawaii International Conference on System Sciences*, 2011, pp. 1–7.

[20] M. S.Patil and B. Ainapure, "Intrusion Detection by Forensic Method in Private Cloud using Eucalyptus," *Int. J. Comput. Appl.*, vol. 85, no. 12, pp. 50–60, 2014.

[21] J.-H. Lee, M.-W. Park, J.-H. Eom, and T.-M. Chung, "Multi-level intrusion detection system and log management in cloud computing," in *13th International Conference on Advanced Communication Technology (ICACT2011)*, 2011, pp. 552–555.

[22] D. K. Anguraj and S. Smys, "Trust-based intrusion detection and clustering approach for wireless body area networks," *Wirel. Pers. Commun.*, vol. 104, no. 1, pp. 1–20, 2019.

[23] A. Waqar, A. Raza, and H. Abbas, "User privacy issues in eucalyptus: A private cloud computing environment," in *2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*, 2011, pp. 927–932.

[24] G. Sibiya, H. S. Venter, S. Ngobeni, and T. Fogwill, "Guidelines for procedures of a harmonized digital forensic process in network forensics," in *2012 Information Security for South Africa*, 2012, pp. 1–7.

[25] M. S. Patil, "A Survey on Opensource Private Cloud Platforms," *IJCST*, vol. 3, no. 4, 2012.

[26] A. Chuvakin, K. Schmidt, and C. Phillips, *Logging and log management: the authoritative guide to understanding the concepts surrounding logging and log management*. Newnes, 2012.

[27] Y. Hui and L. Zesong, "Research on Real-time Analysis and Hybrid Encryption of Big Data," in *2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, 2019, pp. 52–55.

[28] W. Xia, Y. Li, T. Jia, and Z. Wu, "BugIdentifier: An Approach to Identifying Bugs via Log Mining for Accelerating Bug Reporting Stage," in *2019 IEEE 19th International Conference on Software Quality, Reliability and Security (QRS)*, 2019, pp. 167–175.

[29] S. Sridhar and S. Smys, "A Survey on Cloud Security Issues and Challenges with Possible Measures," in *International Conference on Inventive Research in Engineering and Technology*, 2016, vol. 4.

[30] A. Roy, S. Midya, K. Majumder, and S. Phadikar, "Forensics-as-a-service for mobile cloud environment," *Proc. - 2018 4th IEEE Int. Conf. Res. Comput. Intell. Commun. Networks, ICRCICN 2018*, pp. 6–11, 2019.