SPECIAL ISSUE PAPER

A framework for cloud forensics evidence collection and analysis using security information and event management

Muhammad Irfan¹, Haider Abbas^{1,2}, Yunchuan Sun^{3*}, Anam Sajid⁴ and Maruf Pasha⁵

- ¹ National University of Sciences and Technology, Islamabad, Pakistan
- ² King Saud University, Riyadh, Saudi Arabia
- ³ Business School, Beijing Normal University, Beijing, 100875, China
- ⁴ Shaheed Zulfiqar Ali Bhutto Institute of Science and Technology (SZABIST), Karachi, Pakistan
- ⁵ Department of Information Technology, Bahauddin Zakariya University, Multan, Pakistan

ABSTRACT

A primary feature of cloud computing is the provision of a variety of transparent services with efficient resource utilization. However, there are concerns with cloud computing in terms of the user's data privacy and security, especially in evidence collection for forensics analysis, because the tangible resources and hardware are out of reach for users who own the data. This paper presents a framework using security information and event management (SIEM), to address the issue of efficient evidence collection for crime investigation, such as that for cloud forensics, with respect to the cloud service provider. Indeed, evidence could be shared with cloud users when required. SIEM can be considered as a major player in terms of evidence collection in a virtualized environment. The proposed mechanism using SIEM focuses on passive attacks and provides a solution from the cloud administrator or service provider's point of view. The proposed framework can help in performing detailed cloud forensics in terms of efficient evidence building for crime investigations. Copyright © 2016 John Wiley & Sons, Ltd.

KEYWORDS

cloud computing; cloud forensics; security information and event management; virtual infrastructure

*Correspondence

Yunchuan Sun, Business School, Beijing Normal University, Beijing 100875, China. E-mail: yunch@bnu.edu.cn

1. INTRODUCTION

Virtualization and cloud computing are two most recent topics in the field of information technology that offer transparent services and change the way in which these services are created, outsourced, managed, and performed. Companies and institutions can benefit through reduced costs and increased efficiency. Meanwhile, virtualization and cloud computing open the doors to operational and security challenges. Nowadays, cloud forensics is an important issue in cloud and virtual environment in terms of crime investigation. Cloud services are still evolving, and it is the perfect time to address cloud forensics in an effort to prevent and fight malicious and illegal activities [1].

The size of the cloud is an important factor with respect to security. It is difficult for cloud administrators to protect the cloud environment from all threats owing to a rapid increase in the number of threats. Such security challenges lead to security incidents in even small cloud environments, despite the fact that they are being administered by large teams. The large cloud environment is more complex, and carrying out forensics is thus a difficult task. There is an urgent need to develop a novel framework that is able to support effective forensics in the cloud environment as well as handle cloud security challenges by using a proactive approach.

This study presents an effective framework for the virtual-cloud environment that adopts visualization to show a single centralized location and thus make visible all security events of a virtual-cloud environment, hence allowing us to easily follow the activities of cyber criminals, to reproduce crude information identified with respective incident, and to execute proactive work that shelters our virtual-cloud environment from prevailing advanced and sophisticated security threats [2,3]. It is necessary to converge different domains—namely, digital forensics, cloud computing, information security, and event management and monitoring [4]—in producing an

end product that is capable of providing a proactive defense against modern security challenges.

As compared with the existing work, this present paper proposes an approach that is more efficient and takes less time in performing cloud forensic investigations. It begins by providing a background knowledge on all three domains, followed by their convergence and related work. Further, the proposed framework is presented in detail, and the deployment for this framework is presented, which also describes the lab setup. The outcomes for the proposed study are then presented in results and discussion section, ending with conclusion and future directions.

2. BACKGROUND

2.1. Cloud computing

Cloud computing is considered a new computing paradigm originating from grid computing, distributed computing, parallel computing, virtualization technology, and utility computing. It has advanced characteristics, such as largescale computation, data storage, virtualization, high expansibility, high reliability, and low-price services [5]. In the cloud environment, deployment models involve private, public, hybrid, community, and distributed cloud deployments. Similarly, there are multiple service models available, for example, infrastructure as a service, software as a service, platform as a service, desktop as a service, database as a service, information as a service, storage as a service, and security as a service. It is necessary to evaluate the choice of deployment and service model in the virtualcloud environment; for example, if someone wants to publish a web-based application over the Internet via cloud, then he or she would choose a model of public cloud deployment with software as a service. Similarly, if an organization wants to deploy vulnerability assessment solutions for its internal application security testing, then it would choose a model of private cloud deployment with security as a service. The main advantage is that cloud users have to choose the respective model and service type only, and the cloud administrators or cloud service providers (CSPs) then perform the remaining actions.

2.2. Cloud forensics

Cloud computing solutions are rapidly emerging. Organizations are changing their culture to adopt virtual-cloud-based solutions, as the Internet offers countless possibilities. Features and facilities, such as high availability, high speed, scalability, online storage, and elasticity, are readily available on demand in no time. Thus, many organizations have adopted the virtual-cloud infrastructure. However, there are many challenges with respect to cloud forensics. Although cloud environments have been improving with time, CSPs do not have clearly defined standard procedures for handling security incidents. Additionally, cloud users are unable to take effective measures at their end.

This is because of the decentralized nature of data processing in the cloud environment. Instead of using traditional approaches for cloud forensics, a new approach is presented in the proposed framework for the effective handling of challenges.

Downtime in cloud environments is disastrous for businesses as they can lead to millions of dollars in losses per minute. A proactive approach is thus required to deal with the evolving threats in a virtual-cloud environment. In the case of a security incident, a timely incident response is required, and forensic analysis should ensure that future unforeseen incidents are avoided. In the cloud environment, normally at the customer end, whenever a security breach occurs, a forensic analyst is responsible for analyzing the possible malicious activities from the information available and for generating a report from the information available. If an attack is found from the CSP end or from the network end, then dependencies may cause trouble. This is because the cloud or network service providers usually have to be requested to provide logs for analysis and, unfortunately, network service providers usually do not share this information. We therefore need a mechanism at the service provider's end that is able to track all the logs and characterize or segregate all the artifacts at a central location for forensic analysis. This will play a vital role in providing a better incidence response against evolving and sophisticated security threats in the virtual-cloud environment.

2.3. Information security and event management

Logs play a vital role in the field of information security, especially in incident analysis. There are numerous benefits of log management systems, and the return on investment of log management systems is appreciable. Such logs are useful in analysis as they provide detailed information about the activities being performed. The management and monitoring of logs are not easy tasks, as they require considerable time and resources for analysis and, in the case of a large environment, the analysis of logs for multiple appliances and their correlation is difficult. In the case of a security breach or incident, finding the root cause is time consuming. It is sometimes impossible to find the root cause, as in the case of very large environments. Therefore, to provide effective and efficient monitoring at a centralized location, security information and event management (SIEM) solutions are used.

Security information and event management solutions are basically configured in a way that they acquire logs from different appliances of an infrastructure at a central location, called the SIEM database. All logs in the database are then normalized, such that the logs can be interpreted. Finally, the normalized logs are correlated using correlation directives or a correlation engine, and events (or alarms) are generated, on this basis. The solutions also offer a dashboard for viewing all correlated data in a single place, which allows us to view runtime statistics, to make

risk calculations, to categorize events, and to create reports. Multiple SIEM solutions are currently available, including International Business Machines QRadar [6], AlienVault Unified Security Management (USM) [7], Hewlett Packard Arc Sight [8], Solar wind log and event manager [9], and Splunk by Splunk Enterprise [10]. The choice of SIEM depends on the feature set and the budget of an organization, as each solution has its pros and cons.

3. RELATED WORK

Cloud computing has revolutionized the way we process, store, and transmit digital information and data [11]. This advancement has shifted traditional networks, servers, and endpoints towards cloud-based virtual infrastructure, and has resulted in serious challenges in terms of digital forensic investigations within these environments. The National Institute of Standards and Technology has analyzed cloud computing forensic science challenges and categorized these challenges into nine major groups, to facilitate detailed analysis and understanding. As per a National Institute of Standards and Technology draft [12], these challenges relate to architecture, data collection, analysis, anti-forensics, incident first responders, role management, legal issues, standards, and training. Each of these has further sub-categories; the details of which are given in the literature [12]. Digital forensics frameworks and standards are needed to address these challenges.

A number of research projects are in progress with the aim to overcome cloud forensic challenges. A framework has been presented for the automated detection of anomalies in a cloud environment [13]. Architecture-level changes are performed, and a module for cloud forensics, which is capable of learning malicious activities, is embedded in the management layer of the cloud infrastructure. Similarly, a mechanism that uses the Struts and Hadoop distributed file system for forensic data collection and rendering in a cloud environment has been proposed and implemented [14]. Virtual machine disk images and logs are collected using a pull model, by an investigator, and network captures are pushed periodically to the Struts and Hadoop distributed file system, which is later used for forensic analysis.

A hypervisor-based approach has been used for thread monitoring and forensic analysis [15], and provides an option for virtual machine introspection, through a hypervisor (a virtual machine manager), for the monitoring of virtual machines and their related activities. This model operates at a layer between the hardware and virtual environment. It is capable of effective, uninterruptable, and undetectable monitoring and analysis. An effective strategy has been proposed [16] for forensic analysis in the referenced model and focuses on forensic analysis with respect to the CSP. However, no practical approach has been presented for the referenced model.

An architecture has been presented that uses cloudbased forensic tools for digital forensic investigation and eliminates hardware dependency by providing an online solution to forensic experts [3,17]. In this architecture, forensic tools such as sorting, indexing, data recovery, bookmarking, and hex viewer tools are accessed on local systems for the digital investigation of the evidence collected. In a cloud environment, to preserve the confidentiality and integrity of activity logs that are present on virtual machines, a secure logging service model has been proposed [18]. The logs are then accessed by cloud forensic investigators via a secure application programming interface. The accuracy of logs presented in this model will assist effective forensic investigations.

The cloud environment is dynamic in nature, and a snapshot-based approach for in-depth analysis has been discussed [19]. According to this approach, whenever the intrusion detection system of a cloud environment detects an anomaly, it identifies the suspected virtual machine and notifies the CSP. The CSP immediately takes a snapshot of the suspected virtual machine, isolates it from the network, and stores it in permanent storage. The forensic investigator then requests and collects the logs of the forensics investigation and collects possible evidence. This collected information is then analyzed in depth and reported to the concerned party.

Sun, *et al.* [20,21] presented an event-linked network model that is able to query and organize big data. In this model, events are considered as primary units in organizing the data, whereas links represent the association among the events. This model is effective and efficient for cloud or virtual-environment analysis, as a huge quantity of data is involved, such as in the case of internet service provider with SIEM solution having a huge quantity of data at centralized locations. Considering previous feasibility analysis [22], the present paper proposes a framework that will help in addressing the challenges of cloud forensics.

4. FRAMEWORK DESCRIPTION

4.1. Roadmap and infrastructure

The solution proposed in this study combines elements from several distinct technological areas. The aim was to develop a framework for digital forensics in a virtual-cloud environment, using an information security and event management system. The feasibility analysis of the framework has already been presented [22], and the framework proactively assists in dealing with modern cyber security challenges. Within this framework, centralized information of threats is gathered via a SIEM solution and is tuned to extract information needed for forensic analysis, which helps in performing a correlated analysis and forensic investigations of huge virtual-cloud environment data. The framework also offers an option of defining customizable rules and correlation directives for filtering out specific data targeted for analysis. Another benefit is realtime monitoring and analysis, on a single dashboard, which is vital to active monitoring.



In initial testing [22], AlienVault USM was used as a SIEM solution. Although AlienVault USM provides initial results, owing to its limitations and lack of in-depth analysis, it does not consider detailed forensics or further attacks. To overcome these limitations, Rivest, Shamir, and Adelman (RSA) Envision was adopted as the SIEM solution. RSA Envision is able to gather log event data from various event sources within an enterprise-level environment and then to consolidate the data for a single logical location. As a result, data are easily viewed, tracked, analyzed, reported, and stored securely. The Internet Protocol Database of RSA Envision provides an architecture that automatically collects and protects data from network devices and event sources. It is able to independently monitor network and security events, to generate alerts for possible breaches, to conduct analysis, and to issue reports on network performance. RSA Envision comprises three components: an application, a collector, and a database. The application supports interactive users and runs the suite of analysis tools, the collector captures incoming events, and the database manages the access and retrieval of captured events.

To begin at a level higher than the previous testing, a virtual infrastructure is deployed using VMware technology. This virtual infrastructure is based on blade servers, and a storage area network is used for storage. The VMware ESXi 5.5 operating system is used, and the underlying servers are Hewlett Packard ProLiant BL460c Gen8 Blade Servers with 16 central processing units operating at 2.988 GHz with 64 GB of random access memory on each. For virtual networking, a vSphere standard virtual switch, which communicates between the virtual machines, is used. The storage used for data stores and virtual machines is an EMC VMAX/Clarion storage area network. This virtual infrastructure includes all those machines and servers that an enterprise already has, such as an active directory, exchange servers, web servers, Linux-based servers, and network and security appliances. Although the main purpose of this infrastructure is to cover passive attacks, some active attacks are also covered; hence, the virtual-CSP offers a defense mechanism. Moreover, the SIEM solution and attacker machine exist within the same virtual infrastructure, as shown in Figure 1.

4.2. Digital forensics process

Digital forensics is a process that tracks an information security incident and effectively discovers what, when, and how an incident occurred. Currently, most organizations are business oriented and are reluctant to think about security-related issues [23]. However, the number and diversity of security threats and attacks have been increasing. Attackers are becoming more adept, leaving little evidence of their hacking attempts. It is thus necessary to take measures to prevent intrusions. If, for example, zero-day vulnerability is exploited, then it is necessary to find how this event occurred and how to reduce its effect. At this stage, the forensics process and investigation play an

important role, by searching for clues to determine the criticality and collect in-depth information about the security breach. Through the forensics process, we can easily track the step-by-step occurrence of security incidents. The forensics process also assists in determining incident severity, reconstruction, and review, and in identifying the root cause behind the breach, thus allowing preventive measures to be taken.

Although it is possible to use a generic digital forensics process, it is better to use a customized process to obtain efficient and effective output. This customization depends on the environment in which the forensic investigation is expected to be performed. In the literature, a six-step forensic process is defined for forensic collection in the virtualcloud environment [24]. The forensic collection first obtains an organization's administrator credentials and then connects to the environment and collects the available event logs. Thereafter, it collects organizational metadata, virtual data center metadata, virtual apps, virtual machine metadata, and virtual machine data. This forensic process is helpful in conducting forensic investigations for the cloud environment. Similarly, in another study [25], a cloud forensic investigation process was presented from the CSP's point of view. In that process, the first step is the identification of the artifacts to be collected and analyzed. The type of artifacts depends on the type of incident reported. Once artifacts are identified, the next step is the acquisition and preservation of data. This is followed by the analysis of acquired data for the identification of malicious activity, where the search space is narrowed and data are filtered, correlated, and extracted. Finally, evidence is identified from the analysis phase, and results are presented. Keeping in mind the digital forensic process and work presented in the cited studies [24] and [25], the forensics process adopted in the present work involves the following steps.

- Digital forensics data collection includes data from logs, applications, vulnerabilities, flows, events, and other machines.
- (2) Digital forensics data generation includes the monitoring of file integrity, processes, registry, network connections, application IDs, and in-depth packet inspections.
- (3) Real-time processing includes the classification of data, extraction of metadata, time normalization, context infusion, risk prioritization, indexing, and persistence.
- (4) Machine analysis includes advance correlation, pattern recognition, whitelist profiling, and behavioral and statistical baselines.
- (5) Digital forensic analytics involves the filtering of data via a search, visualization, and pivot or drill down analysis.
- (6) Actionable intelligence involves risk prioritization, report gathering, and real-time dashboard views.
- (7) Incident response consists of a smart response, design of workflows, and case management.

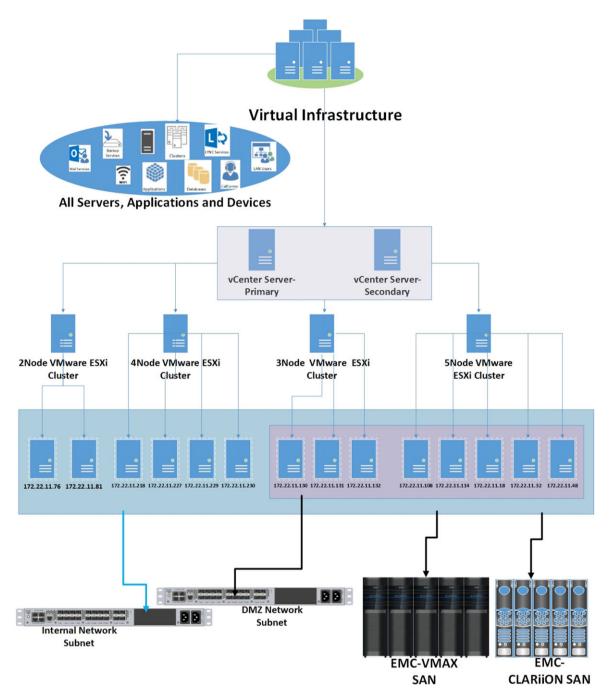


Figure 1. Virtual infrastructure.

4.3. Theory/calculation

The framework follows a three-stage process. The first stage is the input phase, the primary purpose of which is the collection of data for forensic analysis. In the second stage, the collected data are forensically analyzed. In the third stage, output is obtained, which is used to determine and carry out a suitable action. The following briefly describes these three stages.

4.3.1. Input stage

Digital forensic data are collected and generated. In digital forensic data collection, the logs, applications, vulnerabilities, flows, events, and other machine data are collected at a central location, whereas in digital forensic data generation, the host and network level forensics is performed for the monitoring of processes, file integrity, registry, network connections, layer-7 flows, application IDs, and packet capturing for deep packet inspection.



Security Comm. Networks (2016) © 2016 John Wiley & Sons, Ltd.

4.3.2. Analysis stage

Once all data are collected at a centralized location, they are analyzed. The analysis first requires real-time data processing, which includes metadata extraction, data classification, time normalization, context infusion, indexing, and risk prioritization. Machine analysis is then performed, including whitelist profiling, advance correlation, pattern recognition, and the setting of behavioral or statistical baselines. Finally, data are filtered according to a search and visualized as required.

4.3.3. Output stage

The output stage involves actionable intelligence and the incident response. The term actionable intelligence covers the risk prioritization of alerts, real-time data viewing on a dashboard, and the gathering of reports, whereas the incident response relates to the provision of a smart response, the design of workflows, and overall case management.

As a result of the three stages, we obtain an incidence response and forensic investigation for the case under consideration. The results section provides outputs obtained in response to several practical attacks for the abovementioned framework.

5. DEPLOYMENT, ATTACK CLASSIFICATION, AND RESULTS

This section covers all attacks that were performed during testing of the proposed solution. Results obtained using the SIEM solution are presented along with countermeasures. The steps are based on the discussions in the previous sections. Devices, such as routers, firewalls, servers, and similar devices, are configured to forward their logs to the SIEM or its sensor via syslog for central monitoring. Detailed findings are presented as follows.

5.1. Brute-force attack

A brute-force attack was performed using the utility Hydra. The target server is a Linux-based server having the

Internet Protocol (IP) address 172.16.13.19, and an attack is made from a Windows-based server having the IP address 172.16.11.75. Figure 2 shows Hydra attacking the targeted system.

In the highlighted text of the figure, also given next, the "-l" switch is for the username, and "-P" is for the list of passwords tried, whereas "ssh" is the targeted protocol and 172.16.13.19 is the targeted system.

 C:\Users\Irfan\Desktop\hydra-7.5-windows\hydra-7.5 \hydra > hydra -1 irfan -P passwords.txt ssh:// 172.16.13.19

Once an attack is performed, the SIEM solution immediately identifies it as a malicious activity and generates an alert that notifies the respective administrator, so that he or she may respond immediately to the incident. An alert could be sent via email or cell phone text message or both; however, we only consider email alerts in this study. Figure 3 shows the alert generated in response to the detection of a brute-force attack.

As shown in Figure 3, the alert states from where the attack was made and the target of the attack, namely, the source and destination IP addresses. The count is the number of attempts made in terms of the brute-force attack, and the message text describes the attack by stating that an authentication failure has been detected in the system. Therefore, useful information is received by the incident analyst, which allows him or her to take effective measures quickly to prevent the ongoing attack and to take necessary countermeasures against future attacks.

Countermeasures against a brute-force attack include account locks after a specific number of failed login attempts, blacklisting an IP that generates a huge number of failed login requests, allowing logins from specific IPs, assigning unique login uniform resource locators to a specific block of users, delaying the login console after specific failures, placing an account in a locked mode with limited capabilities, and using Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) and multifactor authentication mechanisms.

```
Syntax: hydra [[[-] LOGIN:-L FILE] [-p PASS:-P FILE]] : [-C FILE]] [-e nsr] [-o FILE] [-t TASKS] [-M FILE [-T TASKS]] [-w TIM E] [-w ITME] [-s PORT] [-x MIN:MAX:CMARSET] [-suvb46] [service://server[:PORT][/OPT]]

Options:

-1 LOGIN or -L FILE login with LOGIN name, or load several logins from FILE
-p PASS or -P FILE try password PASS, or load several passwords from FILE
-p PASS or -P FILE try password PASS, or load several passwords from FILE
-t FILE colon separated "login:pass" format, instead of -L-/P options
-t TASKS run TASKS number of connects in parallel, one entry per line
-t TASKS run TASKS number of connects in parallel (per host, default: 16)
-t more command line options (complete help)
server the target server (use either this OR the -M option)
service module usage details
-h more command line options (complete help)
server the target server (use either this OR the -M option)
service services corack (see below for supported protocols)
OPT some service modules support additional input (-U for module help)
Supported services: asterisk cisco cisco-enable cus ftp ftps http[s]-(head!get) http[s]-(get!post)-form http-proxy urlenum icq inap[s] irc Idap2[s] Idap3[-(cranidigest) add][s] messql mycql nntp oracle-listener oracle-sid pcanywhere penfs po p3[s] rdp rexec rlogin rsh sip snb sntp[s] snb sntp[s
```

Figure 2. Brute-force attack using Hydra.

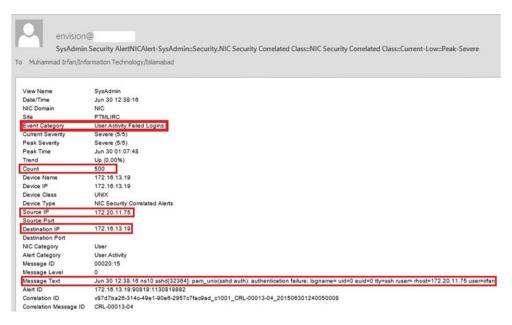


Figure 3. Alert of a Brute-force attack detection.

5.2. Untrusted device gaining authorized access

Mobile devices are among the most vulnerable devices to attackers, mostly because they access unsecured external networks rather than organizational networks with security parameters installed. When a mobile device that belongs to an organization accesses non-organizational networks, the domain user account might be hacked and later used by an untrusted mobile device to gain authorized access to the domain.

Such authorized access may pose threats to an organization, as the account could be misused. In the deployment of the proposed environment in the present study, when such an attack took place, the SIEM solution detected and issued an alert, specifying that an untrusted device (iPhone) had been found using the credentials of a trusted account to gain organizational email services (Figure 4).

In this case, the device used by the attacker was an iPhone with the IP address 172.56.29.162, whereas the target of the attack had the IP address 172.16.12.6, which is that of the Microsoft Exchange Server. The compromised user account was nedal.mustafa, which is the domain user, and the attacker is able to attack other services, as he has authorized access. The generated alert was thus triggered, and the SIEM administrator was notified via email. As a result, the SIEM administrator was able to take necessary actions against the attack.

An organization's business-critical services could be disrupted by such attacks, possibly leading to a loss in

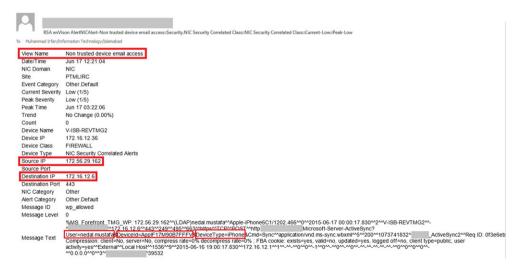


Figure 4. Alert of an untrusted device detection.



business, reputation, data, and trust. By implementing proper controls, such as network-based access control, multifactor authentication, network segregation, mandatory access control, and unique ID-based device hiding, these attacks can be mitigated.

5.3. Log tampering

A log plays an important role in the monitoring and analysis of events or activities of any system, device, or application. The log helps clarify whether everything is working as desired. If the integrity of logs is not maintained, then an anomaly cannot be easily traced.

In the deployed environment, there is a syslog server with the IP address 172.16.15.174, which is responsible for getting logs from various servers, ESXi hosts, and network devices. An attempt to tamper with the log was made with the intention of falsifying the records. This attempt was instantly detected by the SIEM solution, and an alert was triggered, as shown in Figure 5.

The alert shows that the target IP address of the syslog server is 172.16.15.174 and the message shows the tampering of logs. Once this alert is sent to the SIEM administrator, necessary steps would then be taken by the administrator to prevent against the tampering of logs in the future. The administrator can inspect other system logs to identify who logged into the system to conduct the malicious activity.

Countermeasures against log tampering include proper log management, integrity checks, input validation before log writing, validation of all log data before outputting, and log encryption techniques.

5.4. Malicious content detection

Although there are many types of malicious content, the term generally refers to documents or programs that are infected with viruses, web sites that attempt to infect a computer with a virus, or web sites that attempt to solicit sensitive personal information [26].

In the deployment environment of the proposed solutions, whenever a user accesses malicious content or a malicious activity is detected at a network or endpoint device, an alert is generated, and an endpoint protection solution blocks or deletes the malicious content. Figure 6 shows the alert triggered once the malicious content is detected.

Here, the alert shows the time stamp, infected asset, type of risk or malicious content detected, and path where it was found. It also shows the action performed in response (i.e., the risk is blocked, quarantined, or deleted). Figure 6 shows that a system having the IP address 172.17.131.30 was compromised by a malicious executable activity using malware named PUA RegCleanPro. It was quarantined via endpoint protection.

The SIEM solution presented in this study is capable of generating reports for a specific time period with details of all the infected and cleaned systems, which will be of benefit in performing a consolidated analysis. Figure 7 shows a similar report generated via RSA Envision, the SIEM solution.

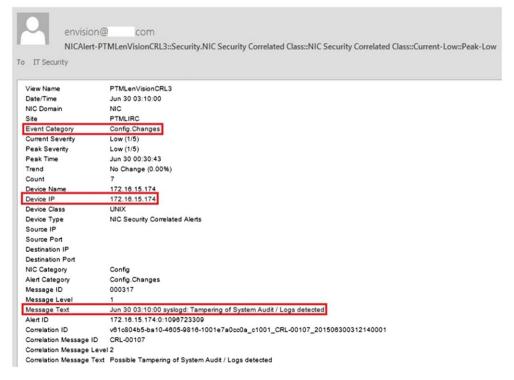


Figure 5. Alert of log tampering detection.

detected a new risk on one or more of the client computers. For more information about the event that triggered this notification, see the Reports page, Quick Reports tab. Select the Risk report type and run the "New Risks Detected in the Network" report.

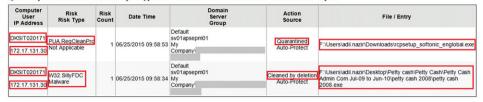


Figure 6. Alert for malicious content detection.

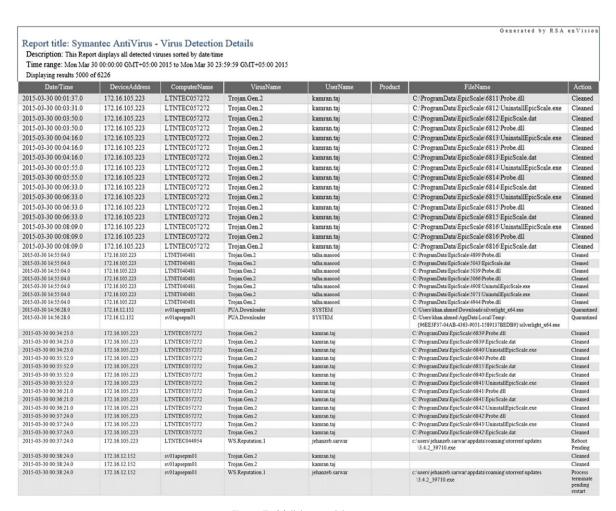


Figure 7. Malicious activity report.

5.5. Data leakage detection and prevention

Data leakage is considered an unauthorized transmission of data (or information), from within an organization to an external destination or recipient, where the leak of information may be via electronic or physical means. Although data leakage, also known as information leakage, is often intentional or malicious, it can also be unintentional or inadvertent unauthorized transmission [27].

For the deployment, an initial policy is made for the detection and prevention of customer data records (CDRs). A CDR contains the customer's private information and comprehensive details of all records. Once data leakage is detected, an alert is triggered for further investigation. The



present investigation analyzed whether the user was authorized to send the data outside. For strict compliance, according to defined policy, no user is able to send data outside the organization, even in the case of an authorized user, prior to investigation or approval. Figure 8 shows the triggered alert.

The policy is implemented for data that are sent outside via emails or any other media. In the alert, as shown in the first row, the user shameel.munir@domain.com is trying to send to aamir.moong@gmail.com an email that contains "call history of customer", which is data restricted to within the organization. This violates the CDR policy, and the transmission of data is blocked until further investigation. The alert also states the sender and receiver of the data so that it is easy to identify the culprits.

5.6. Denial of service attack

A denial of service (DOS) attack is a dangerous attack. In the deployment environment of the proposed systems, a DOS attack was made on a web server using the tool DoSHTTP and was detected by the SIEM solution.

The source of the attack was a Windows server with the IP address 172.16.105.72, and the target was an Internet Information Service web server with the IP address 172.16.24.32. Figure 9(a) and (b) shows the attacker side whereas Figure 10 shows the alert triggered by RSA Envision.

Figure 9(a) shows the continuous flooding of packets for a web server with the IP address 172.16.24.32, leading to unresponsive behavior and a DOS for legitimate users. The attack report shown in Figure 9(b) gives the number of requests generated for the server and the number of responses received from the server. It also shows that the server was taken over by a DOS state, resulting in a 99.99% loss of requests.

Figure 10 shows the alert triggered by RSA Envision when a DOS attack is made on a web server. The alert shows that the attacker's IP address was 172.16.105.72 and the targeted port was port 80. This alert is triggered when there is an abnormal change in traffic on the web server; according to the alert, there was a 79.14% change relative to the normal trend. Therefore, once the attack is detected, the SIEM administrator can easily backlist the identified IP address to block such requests.

Figure 11 shows that the web server with the IP address 172.16.24.32 is not accessible at the time of a DOS attack, hence making it unavailable to legitimate users.

Although DOS and distributed DOS attacks are difficult to prevent, possible countermeasures are the implementation of strict access control lists at the switch (firewall or router) level, load balancing, using an intrusion detection system or intrusion prevention system, implementing a DOS defense system, rate limiting, black holing, and using clean pipes.

5.7. Man-in-the-middle attack

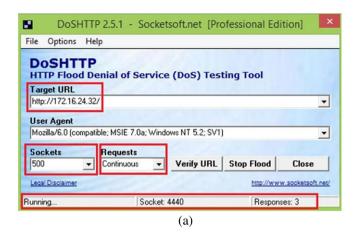
A man-in-the-middle (MITM) attack is also considered dangerous. Such an attack can intercept or alter the communication between two parties. In the deployment environment of the proposed solutions, address resolution protocol (ARP) poisoning is performed for a MITM attack using the Cain & Abel tool. The targeted systems in this attack have the IP addresses 172.16.16.149 and 172.16.16.1, whereas the IP address of the attacker's system is 172.16.16.2. Figure 12 shows the Cain & Abel interface used to poison the ARP entries of the targeted systems.

In this way, all traffic is routed in a manner that the attacker's system can sniff all the data being communicated between two parties. Figure 13 shows the wireshark output that was taken from the attacker's system, showing the successful sniffing of data.



Figure 8. Alert for data loss/leakage detection.

Security Comm. Networks (2016) © 2016 John Wiley & Sons, Ltd. DOI: 10.1002/sec



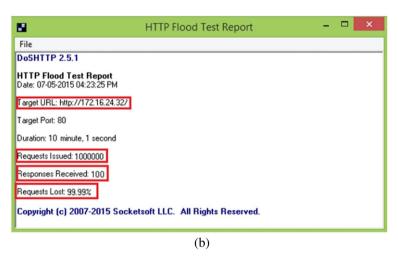


Figure 9. (a) DOS attack using DosHTTP. (b) DosHTTP attack report.

Figure 14 shows how an alert was generated when this attack was detected by the SIEM solution. The alert contains the device IP address 172.16.16.149 and the source IP address 172.16.16.2.

Countermeasures against a MITM attack are using static ARP entries, implementing network access control solutions, implementing endpoint protection solutions, implementing ARP security or dynamic ARP inspection, and using several other tools to watch ARP entries.

5.8. Monitoring network flows

Network traffic can be monitored and analyzed in multiple ways. For a quick incidence response and forensic analysis, real-time network flows can be monitored and operate according to defined policies. It is also possible to store identified data for in-depth analysis and as forensic evidence.

In the present deployment, RSA Envision network flows and real-time data are monitored and analyzed. This includes the monitoring and analysis of data from the wireless local area network controller, secure sockets layer/virtual private network, VMware ESX/ESXi, failed login attempts, user logons during non-working hours, highest Internet bandwidth consumers, most targeted systems, and most browsed web sites. Figure 15 presents all network flows for which the real-time logs of a wireless local area network controller highlight the receiving of an invalid packet request, which may be a possible attack attempt. In this manner, we can check for possible attack attempts in real time by monitoring network flows and logs.

Figure 16 shows the real-time network flow of secure sockets layer/virtual private network connections through which we can easily monitor who is accessing a network, when, and from where. We can also know if a failed attempt was made.

Similarly, Figure 17 shows all real-time logs related to the virtual infrastructure, which provides detailed information about all sessions.

Figure 18 shows a report giving details about failed login attempts. With the report, we can easily analyze the lockout accounts, especially the accounts on which a brute-force attack was made. This report clearly describes the targeted account, failed login attempts, and time. We



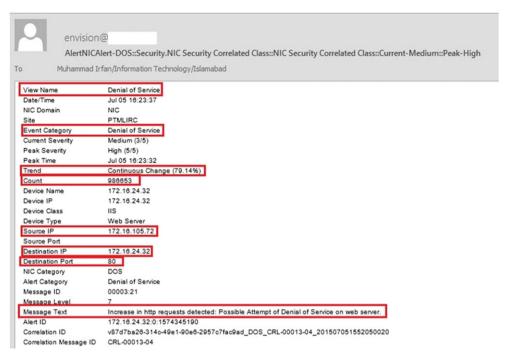


Figure 10. Alert of DOS attack detection.



Figure 11. Inaccessible web server.

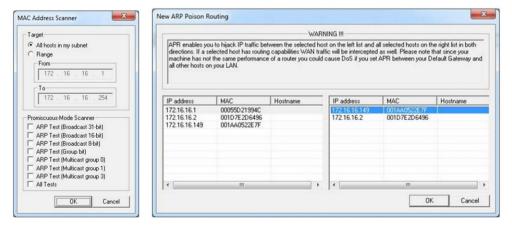


Figure 12. ARP poisoning using the Cain & Abel tool.

| No. | Time | Source | Destination | Protocol | Info |
|-----|-----------|-------------------|-------------------|----------|---|
| 323 | 28.711649 | 00:21:6a:5b:7d:4a | 00:1a:a0:52:2e:7f | ARP | Who has 172.16.16.149? Tell 172.16.16.1 |
| 324 | 28.711854 | 00:21:6a:5b:7d:4a | 00:05:5d:21:99:4c | ARP | Who has 172.16.16.1? Tell 172.16.16.149 |
| 325 | 28.711950 | 00:21:6a:5b:7d:4a | 00:1a:a0:52:2e:7f | ARP | 172.16.16.1 is at 00:21:6a:5b:7d:4a |
| 326 | 28.712037 | 00:21:6a:5b:7d:4a | 00:05:5d:21:99:4c | ARP | 172.16.16.149 is at 00:21:6a:5b:7d:4a |
| 327 | 28.712408 | 00:1a:a0:52:2e:7f | 00:21:6a:5b:7d:4a | ARP | 172.16.16.1 is at 00:1a:a0:52:2e:7f |
| 328 | 28.713480 | 00:05:5d:21:99:4c | 00:21:6a:5b:7d:4a | ARP | 172.16.16.1 is at 00:05:5d:21:99:4c |

Figure 13. Sniffed data.

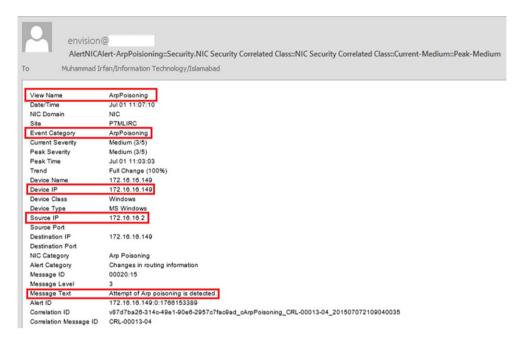


Figure 14. Alert of MITM attack detection.

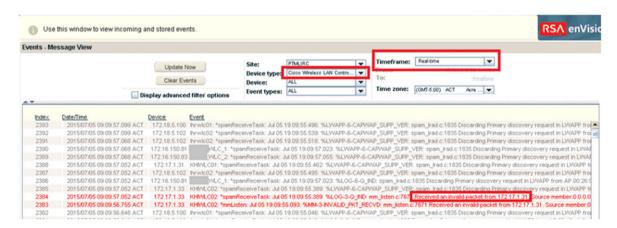


Figure 15. Real-time wireless LAN controller monitoring.

can easily customize this report to include other details such as the source of the attack or system from where false login attempts are made.

Figure 19 shows the systems and accounts that were logged in during non-working hours. From here, we can easily identify any malicious activity during non-office

hours or identify the culprits who are authorized but are performing activity in non-official hours.

Figure 20 shows the report of top Internet users, from which we can easily identify who are using most of the network bandwidth. In this way, any internal user who is creating massive network traffic can easily be traced.



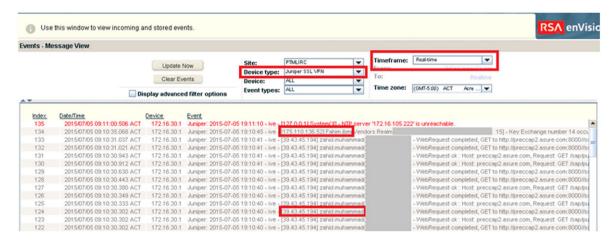


Figure 16. Real-time SSL/VPN monitoring.

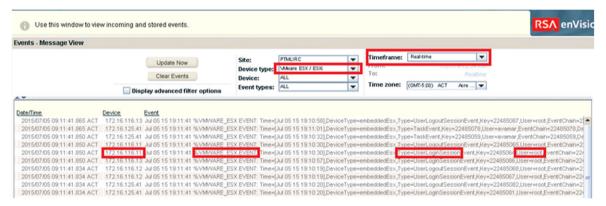


Figure 17. Real-time virtual infrastructure monitoring.

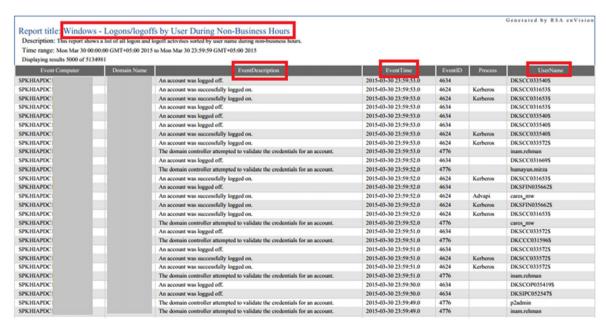


Figure 18. Report of failed login attempts.

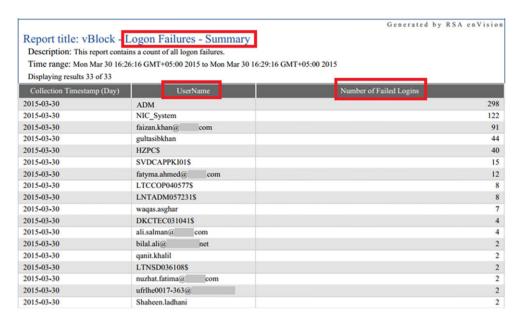


Figure 19. Report of logged-in accounts in non-office hours.

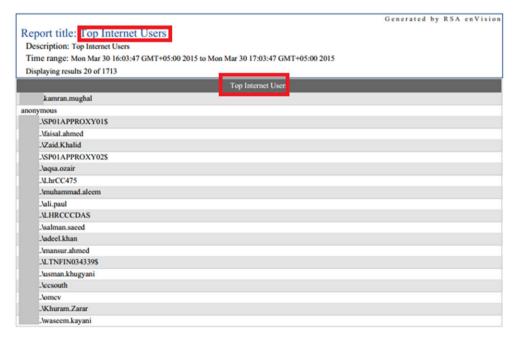


Figure 20. Report on top Internet users.

The report presented in Figure 21 gives the most targeted systems that are infected by malicious links, software, or an executable. It also provides details about the infected user, the rule applied, the source and destination ports, and the source address. In this way, the nature of attacks can easily be monitored, and users made aware about threats and avoidance strategies.

The report in Figure 22 gives the web sites most browsed. From this, we can easily analyze that web sites are most surfed by users.

6. DISCUSSION AND ANALYSIS

This paper extended previous research to provide results that are more detailed and precise. The old SIEM solution was replaced with RSA Envision, as it provides more detailed results and covers a broad range of attacks. The deployment environment is also extended, and the virtual infrastructure at an enterprise level is used for testing and analyzing the results. This virtual infrastructure includes all types of virtual instances for an enterprise, such as



| | Generated by RSA enVision | | | | | |
|------------------|---------------------------|--------------------------------------|-----------------|---|------------|-----------------|
| Report title | | | | | | |
| Description: | | | | | | |
| 0 | | 05:00 2015 to Tue Mar 31 10:59:59 GM | T+05:00 2015 | | | |
| Displaying resul | ts 20 of 17170 | | | | | |
| HostName | UserName | RuleName | count(HostName) | SourceAddress | SourcePort | DestinationPort |
| sv01apsepm01 | abdullah.abbas | Block all other IP traffic and log | 831 | 172.20.10.243 | 3 | |
| sv01apsepm01 | Administrator | Block all other IP traffic and log | 526 | 0.0.0.0 | 546 | 547 |
| sv01apsepm01 | kamran.muhammad | Block all other IP traffic and log | 511 | 0.0.0.0 | 546 | 547 |
| sv01apscpm01 | | Block all other IP traffic and log | 502 | 0.0.0.0 | 546 | 547 |
| sv01apsepm01 | kamil.hussain | Block all other IP traffic and log | 422 | 0.0.0.0 | 547 | 546 |
| sv01apsepm01 | kamil.hussain | Block all other IP traffic and log | 422 | ff02:0000:0000:0000:0000:0000:0001:0002 | 547 | 546 |
| sv01apsepm01 | faisal.ayub | Block all other IP traffic and log | 360 | ff02:0000:0000:0000:0000:0000:0001:0002 | 547 | 546 |
| sv01apsepm01 | siraj.alam | Block all other IP traffic and log | 354 | 172.20.9.201 | 3 | 3 |
| sv01apsepm01 | nasir.shafi | Block all other IP traffic and log | 346 | 0.0.0.0 | 547 | 546 |
| sv01apsepm01 | qummer.afreen | Block all other IP traffic and log | 338 | ff02:0000:0000:0000:0000:0000:0001:0002 | 547 | 546 |
| sv01apsep02 | imran.syed | Block all other IP traffic and log | 338 | ff02:0000:0000:0000:0000:0000:0001:0002 | 547 | 546 |
| sv01apscpm01 | wajahat.bashir | Block all other IP traffic and log | 324 | 0.0.0.0 | 547 | 546 |
| sv01apscpm01 | wajahat.bashir | Block all other IP traffic and log | 303 | ff02:0000:0000:0000:0000:0000:0001:0002 | 547 | 546 |
| sv01apsep02 | muddassir.hussain | Block all other IP traffic and log | 303 | ff02:0000:0000:0000:0000:0000:0001:0002 | 547 | 546 |
| sv01apsepm01 | muddassir.hussain | Block all other IP traffic and log | 292 | 0.0.0.0 | 547 | 546 |
| sv01apsep02 | kamran.muhammad | Block all other IP traffic and log | 289 | fe80:0000:0000:0000:a4b8:55d6:0d80:2594 | 546 | 547 |
| sv01apsepm01 | k.suleman | Allow IGMP traffic | 283 | 224.0.0.252 | 0 | |
| sv01apsep02 | kamran.muhammad | Block all other IP traffic and log | 283 | fe80:0000:0000:0000:e119:f8dd:ade4:df23 | 546 | 547 |
| sv01apsepm01 | zeb.sidra | Block all other IP traffic and log | 281 | ff02:0000:0000:0000:0000:0000:0001:0002 | 547 | 546 |
| sv01apsepm01 | sarfraz.khan | Block all other IP traffic and log | 276 | 192.168.1.2 | 3 | 3 |

Figure 21. Most targeted system with respect to malicious activities.

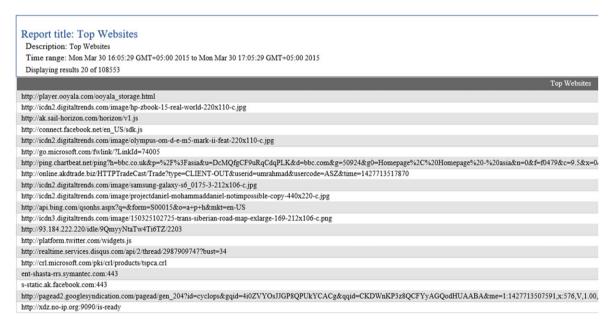


Figure 22. Top browsed web sites.

domain controllers, web servers, mail servers, load balancers, configuration managers, and network and security appliances.

Several attacks were performed for this virtual infrastructure, and SIEM solutions detected all attacks. Alerts are configured and generated in response to each detected attack to inform the SIEM administrator, so that appropriate actions could be taken. Additionally, the real-time monitoring of live activities is possible. The SIEM solution collects data at centralized locations and provides effective correlation for forensic investigation. It can help in finding the root cause of security incidents, and it is thus easy to

design proactive strategies that will counter upcoming security threats and challenges.

Several existing intrusion detection systems are able to achieve this purpose; however, they may not correlate all security events at a centralized point. They may detect intrusions at a particular node or between certain channels, whereas the SIEM solution is able to receive events from everywhere and is capable of pushing these events to a centralized location, therefore allowing correlation and a better identification of the attacks. As we know, effective monitoring, in-depth packet inspection, and log collections are complex actions, and their correlation requires much

computation. This must be considered when choosing the deployment. Except for this limitation, all the results and findings reveal that the proposed framework is feasible and that it can be applied for detection and forensics analysis of more sophisticated attacks, in any sort of virtual-cloud infrastructure, even at a high CSP level.

7. CONCLUSIONS AND FUTURE WORK

Cloud forensics constitutes a difficult challenge when it comes to preparing evidence, because the cloud is used by or shared with cloud users. Cloud providers require a systematic way of collecting evidence and traces that could be later used for tracking malicious activity and regulatory purposes. This study presented a framework, along with deployment details, for accomplishing this task using minimal resources of SIEM. This will provide a baseline for cloud providers to systematically collect and prepare evidence for cloud forensics. In future work, we will develop a more comprehensive framework for cloud users, and we will streamline the best possible evidence for forensics purposes. In doing so, it will be possible to have a cloud in which partial information is shared by CSPs and users.

ACKNOWLEDGEMENTS

The authors extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding of this research through Research Group Project RG-1435-048. This research is also partially sponsored by the National Natural Science Foundation of China (61371185) and the China Postdoctoral Science Foundation (2015M571231). The authors thank the National University of Sciences and Technology, Islamabad, Pakistan, for its support during the research.

REFERENCES

- Cloud forensics. Available from: http://www.techstagram.com/2013/03/20/cloud-forensics-importance/. Retrieved on Nov 3, 2014.
- Yu L, Cai Z. Dynamic scaling of virtualized networks with bandwidth guarantees in cloud datacenters. The 35th Annual IEEE International Conference on Computer Communications (INFOCOM 2016) 2016.
- Abbas H, Mahmoodzadeh QM, Khan FA, Pasha M. Identifying an OpenID anti-phishing scheme for cyberspace. Security and Communication Networks 2014; 9(6):481–491.
- Gao J, Li J, Cai Z, Gao H. Composite event coverage in wireless sensor networks with heterogeneous sensors.
 The 34th Annual IEEE International Conference on

- Computer Communications (INFOCOM 2015) 2015;217-225. DOI 10.1109/INFOCOM.2015.7218385.
- Liu W. Research on cloud computing security problem and strategy. Consumer Electronics Communications and Networks (CECNet) 2012:1216–1219. doi:10.1109/ CECNet.2012.6202020.
- IBM security QRadar SIEM. Available from: http:// www-03.ibm.com/software/products/en/qradar-siem. Retrieved on Jan 10, 2014.
- AlienVault. Available from: https://www.alienvault. com/. Retrieved on Jan 10, 2014.
- Security information and event management. Available from: http://www8.hp.com/us/en/software-solutions/siem-security-information-event-management/.
 Retrieved on Jan 10,2014.
- Log and event manager. Available from: http://www.solarwinds.com/log-event-manager.aspx. Retrieved on Jan 10, 2014.
- Splunk. Available from: http://www.splunk.com/. Retrieved on Jan 10, 2014.
- Yu L, Chen L, Cai Z, Shen H, Liang Y, Pan Y. Stochastic load balancing for virtual resource management in datacenters. *IEEE Transactions on Cloud Computing* 2016. doi:10.1109/TCC.2016.2525984.
- NIST cloud computing forensic science challenges. Available from: http://csrc.nist.gov/publications/ drafts/nistir-8006/draft_nistir_8006.pdf. Retrieved on Oct 20, 2015.
- Patrascu A, Velciu M-A, Patriciu VV. Cloud computing digital forensics framework for automated anomalies detection. Applied Computational Intelligence and Informatics (SACI), IEEE 10th Jubilee International Symposium 2015;505-510. DOI 10.1109/SACI.2015.7208257.
- Saibharath S, Geethakumari G. Cloud forensics: evidence collection and preliminary analysis. IEEE International Advance Computing Conference (IACC) 2015; 464-467. DOI 10.1109/IADCC.2015.7154751.
- Jackson C, Agrawal R, Walker J, Grosky W. Scenariobased design for a cloud forensics portal. IEEE International Symposium Technologies for Homeland Security (HST) 2015; 1-6. DOI 10.1109/THS.2015.7225260.
- Meera G, Kumar Raju Alluri BKSP, Powar D, Geethakumari G. A strategy for enabling forensic investigation in cloud IaaS. IEEE International Conference Electrical, Computer and Communication Technologies (ICECCT) 2015; 1-5. DOI 10.1109/ ICECCT.2015.7226103.
- Mohite MP, Ardhapurkar SB. Design and implementation of a cloud based computer forensic tool. Fifth International Conference Communication Systems and Network Technologies (CSNT) 2015; 1005-1009. DOI 10.1109/CSNT.2015.180.



Zawoad S, Dutta A, Hasan R. Towards building forensics enabled cloud through secure logging-as-a-service.
 IEEE Transactions on Dependable and Secure Computing 2015; 1-1. DOI 10.1109/TDSC.2015.2482484

- Rani DR, Geethakumari G. An efficient approach to forensic investigation in cloud using VM snapshots. International Conference Pervasive Computing (ICPC) 2015;1-5. DOI 10.1109/PERVASIVE.2015.7087206.
- 20. Sun Y, Yan H, Zhang J, Xia Y, Wang S, Bie R and Tian Y. Organizing and querying the big sensing data with event-linked network in the Internet of things. *International Journal of Distributed Sensor Networks* 2014;vol. 2014: 11 pages. doi:10.1155/2014/218521
- Sun Y, Jara A. An extensible and active semantic model of information organizing for the Internet of things. *Personal and Ubiquitous Computing*, 2014;
 18(8): 1821-1833. DOI:10.1007/s00779-014-0786-z.
- Irfan M, Abbas H, Iqbal W. Feasibility analysis for incorporating/deploying SIEM for forensics evidence collection in cloud environment. Computer and Information Science (ICIS), IEEE/ACIS 14th International Conference 2015;15-21. DOI 10.1109/ ICIS.2015.7166563

- 23. Abbas H, Magnusson C, Yngstrom L, Hemani A. Addressing dynamic issues in information security management. *Info Mngmnt & Comp Security* 2011; **19** (1):5–24.
- Martini B, Choo K-KR. Remote programmatic vCloud forensics: a six-step collection process and a proof of concept. IEEE 13th International Conference Trust, Security and Privacy in Computing and Communications (TrustCom) 2014;935-942. DOI 10.1109/ TrustCom.2014.124.
- Meera G, Kumar Raju Alluri BKSP, Powar D, Geethakumari G. A strategy for enabling forensic investigation in cloud IaaS. IEEE International Conference Electrical, Computer and Communication Technologies (ICECCT), 2015;1-5, DOI 10.1109/ICECCT.2015.7226103.
- Malicious content filtering. Available from: http://it. emory.edu/security/malicious_content.html. Retrieved on August 15, 2015.
- Data leakage threats and mitigation. Available from: https://www.sans.org/reading-room/whitepapers/ awareness/data-leakage-threats-mitigation-1931. Retrieved on August, 2015."