Digital Forensic Readiness in a Cloud Environment

George Sibiya, Thomas Fogwill
Meraka Insitute
Council of Scientific and Industrial Research (CSIR)
Pretoria, South Africa
{gsibiya, tfogwill}@csir.co.za

H.S. Venter
Department of Computer Science
University of Pretoria
Pretoria, South Africa
hsventer@cs.up.ac.za

Sipho Ngobeni
DPSS
Council of Scientific and Industrial Research (CSIR)
Pretoria, South Africa
sngobeni@csir.co.za

***Abstract*—Although cloud computing is maturing, security issues are still prevalent. Most of the security issues that are in the cloud have existed since the advent of the Internet. These issues are escalated in a cloud environment due to its distributed nature, multi-tenancy and the sensitive and large amount of data that is transmitted over the Internet and hosted by third parties. The security aspect that this paper focuses on concerns digital forensics. The cloud spans over multi-jurisdictions. As such, service providers hosting the data that may be required for digital forensic investigation may be reluctant to comply with foreign law enforcement agencies. Even if they comply, this may be a costly and time-consuming exercise, given the amount of hosted data that belongs to multi-tenants. In this paper we present a forensic readiness model that makes use of a Forensic Service hosted in the cloud. The model is aimed at minimizing costs associated with conducting a digital forensic investigation in a distributed cloud environment. The scope of this paper, however, is limited to examining the impact that a forensic readiness mechanism may have on other hosted cloud services. Preliminary results have shown a negligible effect in performance of cloud services by a having our proposed digital forensic readiness mechanism in place.**

Keywords—Cloud Computing; Digital Forensics; Security as a Service; Forensics as a Service.

I. INTRODUCTION

Cloud computing is a relatively new computing paradigm that presents new research challenges in the information security field [1–3]. The challenges include those that arise when a forensic investigation needs to be performed, as the environment is virtualized and often distributed in nature. The challenge is even higher when no digital forensic readiness mechanisms are put in place before an occurrence of an incident that requires investigation. A lack of forensic readiness mechanisms is common in cloud infrastructures as the cloud itself is still new and digital forensic is also relatively new as a research field.

The remainder of this paper is organized as follows: In Section II we present a brief background that covers digital forensic readiness, digital forensics in cloud environments. These will cover challenges associated with conducting digital forensic investigations in a cloud environment. In Section III we present our digital forensic readiness model that addresses digital forensics in a cloud environment. In Section IV we present a concept evaluation where we demonstrate the applicability of our concept in practice.

In Section V we conclude the paper and also present our future work.

II. BACKGROUND

In this section we present a brief background on digital forensic readiness and the conducting digital forensics in a cloud environment.

Digital forensic readiness is the mechanism that an organization puts in place to enable the continuous collection of information that can be used as evidence in an investigation [4]. Setting up investigation mechanisms after the incident has occurred may contaminate evidence and critical information may be missed. Some electronic investigations may also be abandoned due to an escalation of costs. A digital forensic readiness mechanism reduces costs and also increases the chance of a successful prosecution and convictions [5].

The main essence of digital forensic readiness is to reduce the effort involved in performing a digital forensic investigation. This is done by taking the necessary prior steps so as to be ready for any digital forensic investigation, while maintaining the level of credibility of the digital evidence that is collected [6]. The decrease in effort is the result of the cloud computing environment being in a state of readiness, which reduces the time and cost involved in the subsequent digital forensic investigation process. A cloud computing service provider that is ready in terms of digital forensics can respond to an attack rapidly and efficiently. In general, reducing the time involved in cloud incident response greatly reduces the costs of the entire digital forensic investigation.

III. DIGITAL FORENSIC READINESS MODEL

In this section we present our proposed digital forensic readiness model for a cloud environment. This model makes use of our digital forensic service presented in [8], which in turn makes use of cloud resources to acquire evidence from a cloud environment that may be beyond the jurisdiction of digital forensic investigators. It also uses accessible information to build a case before costly data acquisition process from foreign countries can be executed.

Fig. 1 depicts how security is offered as a service in a cloud environment and how we propose digital forensics as a service to achieve digital forensic readiness in a cloud environment. Three sections are depicted by the model illustrated in Fig. 1. The first section represents client devices or computers that

users use to access cloud services. These client devices can either be other virtual machines within a cloud environment or an ordinary computer or device belonging to a cloud user. Interaction between clients and cloud services is always via a security model that is either implemented by a third party or built in within the cloud services. The second section represents a generic cloud service stack that is comprised of software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS). The third section represents our digital forensic service, which we integrate with the existing cloud model to offer digital forensic services.

In the following sections we expand on security as a service as presented in [2], [3], the cloud model and our proposed digital forensics as a service. Lastly we will present our complete expanded digital forensic readiness model.

A. *Security as a Service (SECaaS)*

There are seven layers in the security model by the Cloud Security Alliance in [3] in which security services can be offered in a cloud environment (see Fig. 2). The figure also depicts devices that cloud service consumers may use to access cloud services. These different layers of security are the physical layer, compute and storage, trusted computing, network, management, information and application layer. The physical layer refers to the physical security at the venue where the physical infrastructure is hosted.

Security services on the compute and storage layer provide security solutions for the hosts and storage devices. These include solutions such as firewalls and encryption. On trusted computing, the security services deal with the APIs that interact directly with the hosts such as hypervisors. Network level security services include firewalls, data packet inspection, intrusion detection and intrusion prevention. At the management level, security services offered include patch management, vulnerability assessment, vulnerability management, identity management and access management. At the information level, the services provided include database activity monitoring, data leak prevention, content monitoring and content filtering. At the application layer, security services include binary analysis, web applications, firewalls, transactional security and binary scanners. Security services may also include among others virus definition updates, competent security expertise, security and administrative tasks [2]. Examples of service providers that offer security services include security services in the cloud. These include (but are not limited to) CloudPassage [9], CipherCloud [10] and CloudLock [11].

In the next section we present the cloud service model and the basic service models in cloud computing.

B. *Cloud Model*

The cloud model offers three service models, i.e. infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS) (see Fig. 3). Depending on the level of service that a cloud service provider may choose to offer, the security services requirement as well as digital forensic services requirements will differ.

Fig. 1. Cloud Forensics Readiness Model

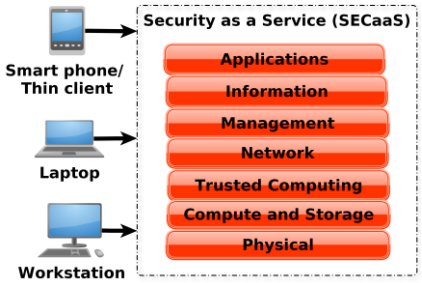

Fig. 2. Security as a Service

The security services requirements of an infrastructure as a service provider may end at the network level. Requirements for a platform as a service provider may end at the information service security level. On the digital forensic service side, digital forensic readiness capability requirements for a SaaS provider may be limited to the application activities and to the platform on which their services are deployed. Even though the IaaS providers have access to PaaS and SaaS hosted on their environment, they may still choose to limit their digital forensic responsibilities and needs to the network level.

In the following section we present services provided by the digital forensic components and indicate how they cater for the digital forensic needs of IaaS, PaaS and SaaS.

C. *Forensics as a Service (FaaS)*

The components of the digital forensic service comprise applications forensics, RAM forensics, network forensics and computer forensics (see Fig. 4). Each of the components is discussed in more detail in the sections to follow.

1) Applications Forensics

Applications that interact with layer 7 of the ISO/OSI model [12]– such as static and dynamic web applications, web clients, web servers, application servers and web services are the major players in an Internet environment and means through which data is exchanged around the world between clients and servers. These applications are joined by a new set of applications in the cloud that offer collaborative environments such as online word processors and Integrated Development Environments (IDEs).The latter applications can therefore be a rich source of data that can be used for investigation purposes. According to [13], "SaaS is the most mature category of cloud services, since it evolved from the application-service-provider model of software hosting". A software offered as a service in a cloud environment is a single instance of an application that is hosted by the provider. Customers access it from a single server and their data and unique configurations are virtually partitioned from other customers. If an instance of a deployed application is created, the cloud platform dynamically determines the server to run the instance [14].

In such a scenario a digital forensic service is used to keep

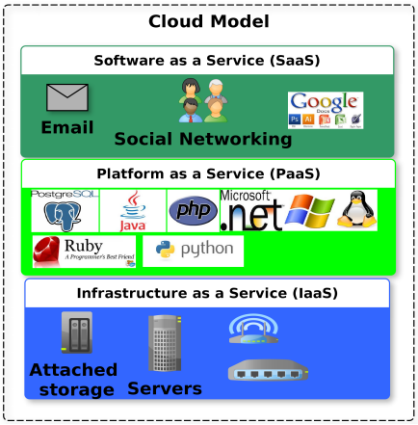

Fig. 3. Cloud Model

track of information so that it can be used for investigation purposes. Log files associated with the running application from those servers are retrieved at intervals determined by the digital forensic service consumer.

2) RAM Forensics

The RAM Forensics component of the digital forensic service is utilized by PaaS providers. The latter are responsible for and have access to application data belonging to the applications hosted as a service, as well as to information regarding their platform. PaaS providers install their platforms and custom applications on virtual machines which they deploy on IaaS [14], [15]. PaaS providers are then responsible for the management of the virtual machines on which their platforms are running. When invoked, the digital forensic service captures snapshots of the running virtual machine that hosts the platform. The snapshots preserve 'state data' for a virtual machine with regard to the exact instant during which the data was captured. The 'state' includes information such as whether the virtual machine was running, shut down or suspended. Data includes disks and all other devices connected to the virtual machine. Such information is critical for investigation purposes and may be costly for PaaS providers if they choose to manage it themselves. Detailed procedures for carrying out RAM forensics are presented in [7].

The digital forensic service can be invoked at any time when needed. A PaaS provider may choose to invoke the service from the time their PaaS is deployed and keep it active for the lifetime of the service. They may also instantiate (or release) it at any time during the lifetime of the service.

A need for a digital forensic investigation may arise at any time and since the moment cannot be predicted, relevant information must be readily available for investigators. Hence, the digital forensic service needs to be instantiated from the very time the PaaS is deployed. In this manner, costs and time involved in the investigation process are reduced.

3) Network Forensics

The network component of the digital forensic service is useful for the providers of both PaaS and IaaS. This is due to the fact that virtual machines for PaaS are connected via virtual networks, while in the case of IaaS, the physical servers and physical devices are connected via physical networks.

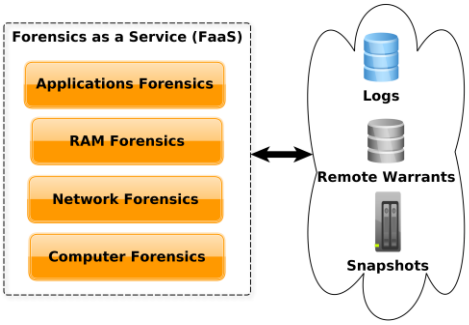

Fig. 4. Forensics as a Service

The virtual networks connecting virtual machines are in turn bridged over the physical networks of IaaS. In both cases, data to be captured is the same. The only difference is that connections in virtual networks occur through virtual devices such as virtual network interface cards and virtual switches. Data captured by the network components of the service include network logs and logs from IDS and IPS devices in the network. Such logs are useful in determining devices and machines that connected or disconnected at any point during the lifetime of the network. Having such data to be managed externally is advantageous as it can still be retrieved even if the network is compromised during an attack or failure and is inaccessible. In [16] we presented detailed procedures on network forensics.

4) Computer Forensics

In this paper we refer to computer forensics as a digital forensics investigation carried out on a physical computer. Since it is an online service, a computer forensics service can be used only to investigate a live system. Data that can be captured while the system is running is the RAM and also includes swap space if it is configured on the system. This component is useful for IaaS providers as they own the physical infrastructure. Cloud computing is designed to handle node failures. As such, nodes can be connected and disconnected from the infrastructure without interrupting the service. This allows traditional digital forensic techniques to be performed – which is beyond the scope of this paper.

In the next section we present the integrated cloud forensic readiness model. We also present the components of the model that interact with each one another.

D. Integrated digital forensic readiness model

In this section we present the complete digital forensic readiness model for a cloud environment, which integrates the components discussed in the previous sections namely security as a service, the cloud model and forensics as a service.

Fig. 5. shows how the different components interact. With SECaaS in place, client devices always interact with cloud service through it. The basic security requirements for users to access electronic resources are authentication and identity management. These functionalities, including more complex ones such as encryption, are handled by a third party as a service. The interaction between the cloud service consumers and the cloud services therefore always occurs via the SECaaS component.

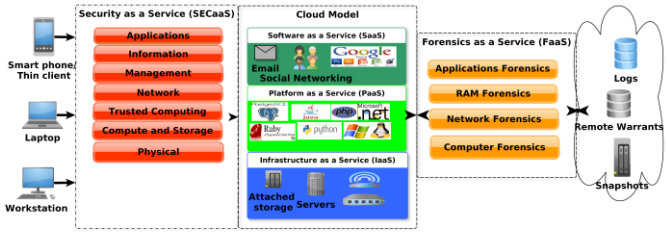

Fig. 5. Complete digital forensic readiness model

The Forensics as a Service (FaaS) model can always be invoked by cloud service providers for their digital forensic readiness model. Since the digital forensic service is offered as a cloud service, it can also be used by cloud service consumers for digital forensic investigations. The digital forensic readiness needs of the SaaS provider, PaaS provider and IaaS provider differ. Depending on their needs, cloud service providers may invoke a relevant FaaS component at any stage of their service provisioning, i.e. on deployment of a cloud service or after detection of an incident.

In the next section we present our experiment that aimed to provide a preliminary evaluation of the concept of our proposed model.

IV. CONCEPT EVALUATION

A use case scenario of our proposed model is presented here. Preliminary tests of our concept were carried out using Nimbula Director 2.0.3 [17] and Splunk [18]. Nimbula Director is cloud infrastructure software that is used to build a private cloud or hybrid cloud. The tests in this experiment were aimed at demonstrating the ability of forensically relevant data to be managed by a third party in a scalable and reliable manner. In respect of scalability we consider the effect introduced by the process of transferring data from the platform that hosts an application to the digital forensic server. These effects include service disruptions (if any) while data is being retrieved for storage in the digital forensic server.

In the next section we present the environmental setup of our experiment. We also present technical specifications of the experimental environment.

A. *Experimental setup*

For purposes of conceptual examination, the experiment was performed in a single cluster cloud infrastructure. Two virtual machines were launched in a Nimbula Director [17] cloud environment with the first virtual machine running a web application (SaaS). The machine was configured to forward logs and other events to a server. The second virtual machine represents FaaS and it runs a Splunk server that continuously pols remote agents for updates on new events. The client machine consumes a cloud service deployed on the first virtual machine. In the next section we present the experimental procedures carried out and the results of our study.

B. *Experimental Procedures*

A human resource management system, OrangeHRM [19] was deployed on the first virtual machine as a cloud service.

We then used Apache Bench [20] to measure the performance of the HR system while the forwarder was running and also while the forwarder was turned off. Apache Bench is an application used to benchmark web servers and web applications. It simulates a typical usage scenario of a web server by flooding it with HTTP requests. Apache Bench was fed with 1 million HTTP requests. The experiment was repeated and each time it was run, the number of concurrent requests to be served was changed with the intervals of 10. These procedures were carried out while the Splunk forwarding agent was down and also while the forwarding agent was running. Apache Bench record elapsed time every time a percentage of the requests is served.

Further tests were conducted on the performance of the HRM application by using Siege, “a multi-threaded http load testing and benchmarking utility”[21]. Siege was configured to send HTTP requests for twelve hours while the forwarder was up. A similar exercise was carried out while the forwarder was down.

The results obtained during these experimental procedures are presented in the next section.

C. *Results*

In this section we present and describe the data collected from our experiment. In our tests, the number of concurrent requests was varied and the results obtained are depicted in Fig. 6. In Fig. 6, the curves show an almost unaffected performance of the HR application deployed in the same environment that forwards data to a digital forensic server. However, a closer examination on both ends of the curves reveals that when a forensic forwarder is down, requests are completed earlier than when the forwarder is running.

TABLE I show results from a t-test paired two samples for means analysis of the obtained results with the level of significance, $\alpha = 0.05$. It shows analysis of completion times per request percentage served and analysis for completion time per concurrent requests.

From the analysis table, TABLE I it can be observed that the test statistic (t Stat) is lesser than critical statistic values (t Critical) on both one-tailed test and two-tailed tests. The one-tail and the two-tail tests probabilities (P) are both greater than the level of significance in both tables. These observations indicate that there is no significant difference on the performance of the HRM application between when the forwarding agent is up and when the forwarding agent is down. The Pearson’s correlation coefficient also shows a very high positive correlation between the performance of the HRM application between when the forwarding agent is running and when the forwarding agent is down.

In TABLE II we present the results that were obtained when using Siege while the forwarder was up and down respectively.

While testing the performance of the HRM application where Siege was used, only two values indicated towards a comparatively bad performance of the application, i.e. concurrency and longest transaction time.

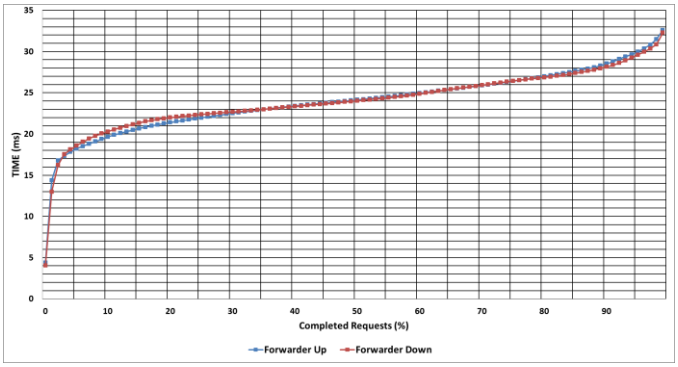

Fig. 6. HRM application performance analysis

TABLE I. Average completion time percentage served analysis

	Completion time / % served	**Completion time/ Concurrent Requests**
Pearson Correlation	0.996233046	0.998564779
t Stat	-0.958124909	-0.135193085
P(T<=t) one-tail	0.170166904	0.447717275
t Critical one-tail	1.660391156	1.833112933
P(T<=t) two-tail	0.340333808	0.89543455
t Critical two-tail	1.984216952	2.262157163

TABLE II. Siege Evaluation

	Forwarder Up	**Forwarder Down**
Transactions	1285788	1285610
Availability	100%	100%
Elapsed time	43201.9 secs	43201.82 secs
Data Transferred	633.34 MB	633.26 MB
Response time	0 sec	0 sec
Transaction rate	29.76 trans/sec	29.76 trans/sec
Throughput	0.01 MB/sec	0.01 MB/sec
Concurrency	0.12	0.11
Successful transactions	642903	642812
Failed transactions	0	0
Longest transaction	4.72 secs	3.04 secs
Shortest transaction	0 sec	0 sec

The difference between the longest transaction times was 1.46 seconds. Both values were beyond the 2 seconds recommended in [22], but when the forwarding agent was running, the delay time was higher.

V. CONCLUSION AND FUTURE WORK

Cloud computing provides many benefits and although it is maturing, security still remains a concern. Cloud services are still subject to compromise and cloud customer data is always at risk. This emphasizes the need for forensic investigations to be carried out whenever incidences occur. Conducting a digital investigation in a cloud environment is challenging and costly, due to the distributed nature and multi-tenancy of the cloud. A digital forensic readiness mechanism is therefore required to conduct a cost-effective digital investigation in a cloud environment.

In this paper we presented a digital forensic service that can be used by cloud service providers as a digital forensic readiness mechanism. The service can be used by cloud service providers to manage data that can be used for digital forensic investigations. Open source tools were used to evaluate the concept of our proposed model. Nevertheless, the preliminary test of this model indicates that the inclusion of a digital forensic service agent on a cloud service host has a negligible effect on the performance of cloud services.

As part of future work, further evaluations of the concept need to be carried out. These will include the evaluation of the concept on a distributed cloud infrastructure with a digital forensic server that supports the retrieval of raw files Splunk, used in these experiments is not capable of retrieving them. Further investigations will also include secure transmission of evidence between agents and the server.

REFERENCES

[1] S. Zimmerman and D. Glavach, “Cyber Forensics in the Cloud,” *IAnewsletter*, vol. 14, no. 1, pp. 4–7, 2011.

[2] M. Rouse, “Security as a Service,” 2010. [Online]. Available: http://searchsecurity.techtarget.com/definition/Security-as-a-Service.

[3] A. Reed, C. Rezek, and P. Simmonds, “SECURITY GUIDANCE FOR CRITICAL AREAS OF FOCUS IN CLOUD COMPUTING V3.0,” 2011.

[4] A. Mouhtaropoulos, M. Grobler, and C.-T. Li, “Digital Forensic Readiness: An Insight into Governmental and Academic Initiatives,” *2011 European Intelligence and Security Informatics Conference*, pp. 191–196, Sep. 2011.

[5] T. Grobler and B. Louwrens, “Digital Forensic Readiness as a Component of Information Security Best Practice,” in *New Approaches for Security, Privacy and Trust in Complex Environments, Proceedings of the IFIP TC-11 22nd International Information Security Conference (SEC 2007), 14-16 May 2007, Sandton, South Africa*, 2007, vol. 232, pp. 13–24.

[6] B. Endicott-Popovsky, D. Frincke, and C. Taylor, “A Theoretical Framework for Organizational Network Forensic Readiness,” *Journal of Computers*, vol. 2, no. 3, pp. 1–11, May 2007.

[7] G. Sibiya, H. S. Venter, and T. Fogwill, “Procedures for a Harmonized Digital Forensic Process in Live Forensics,” in *SATNAC*, 2012.

[8] G. Sibiya, H. S. Venter, and T. Fogwill, “Digital Forensic Framework for a Cloud Environment,” in *IST-Africa 2012*, 2012, pp. 1–8.

[9] “Cloud Passage.” [Online]. Available: http://www.cloudpassage.com/.

[10] “Cipher Cloud.” [Online]. Available: http://www.ciphercloud.com/.

[11] “Cloud Lock.” [Online]. Available: http://www.cloudlock.com/.

[12] R. L. Miller, “The OSI Model: An Overview.” SANS, 2001.

[13] C. Spence, J. Devoys, and S. Chahal, “Architecrutre Software as a Service for the Enterprise.” 2009.

[14] “Deploying SaaS Apps.” [Online]. Available: http://docs.apprenda.com /printpdf/book/export/html/222. [Accessed: 03-Jul-2012].

[15] D. Mueller, “Deploying Private PaaS on CloudStack with Stackato,” *ActiveState*, 2012. [Online]. Available: http://www.activestate.com/ blog/2012/05/deploying-private-paas-cloudstack-stackato.

[16] G. Sibiya, H. S. Venter, S. Ngobeni, and T. Fogwill, “Guidelines for Procedures of a Harmonised Digital Forensic Process in Network Forensics,” in *ISSA*, 2012.

[17] “Nimbula Director.” [Online]. Available: http://nimbula.com/.

[18] “Splunk.” [Online]. Available: http://www.splunk.com/?r=header.

[19] “OrangeHRM - Open Source HR Management.” [Online]. Available: http://www.orangehrm.com/.

[20] “Apache HTTP server benchmarking tool.” [Online]. Available: http://httpd.apache.org/docs/2.2/programs/ab.html.

[21] J. Fulmer, “Siege.” [Online]. Available: http://linux.die.net/man/1/siege.

[22] F. F.-H. Nah, “A study on tolerable waiting time: how long are Web users willing to wait?,” *Behaviour & Information Technology*, vol. 23, no. 3, pp. 153–163, May 2004.