Cloud Forensics Solutions: A Review

Stavros Simou[1], Christos Kalloniatis[1], Evangelia Kavakli[1], and Stefanos Gritzalis[2]

[1] Cultural Informatics Laboratory, Department of Cultural Technology and Communication, University of the Aegean, University Hill, GR 81100 Mytilene, Greece
{SSimou,chkallon}@aegean.gr, kavakli@ct.aegean.gr
[2] Information and Communication Systems Security Laboratory, Department of Information and Communications Systems Engineering, University of the Aegean, GR 83200, Samos, Greece
sgritz@aegean.gr

Abstract. Cloud computing technology attracted many Internet users and organizations the past few years and has become one of the hottest topics in IT. However, due to the newly appeared threats and challenges arisen in cloud computing, current methodologies and techniques are not designed for assisting the respective forensic processes in cloud environments. Challenges and issues introduced, require new solutions in cloud forensics. To date, the research conducted in this area concerns mostly the identification of the major challenges in cloud forensics. This paper focuses on the identification of the available technical solutions addressed in the respective literature that have an applicability on cloud computing. Furthermore it matches the identified solutions with the respective challenges already mentioned in the respective literature. Specifically, it summarizes the methods and the proposed solutions used to conduct an investigation, in comparison to the respective cloud challenges and finally it highlights the open problems in the area of cloud forensics.

Keywords: Cloud Computing, Cloud Forensics, Cloud Forensics Challenges, Cloud Forensics Solutions, Review.

1 Introduction

In recent years, the traditional computer technology has been transformed into new forms of services dictated by Internet changes. Cloud computing has dominated our world giving a different perspective and new horizons opened to companies and organizations with virtualized services providing, mostly, flexibility, elasticity and on-demand service. Complimentary to the above, the cost reduction along with the benefit of eliminating training and maintenance made cloud very popular. However, the development of this new technology attracted a number of people to carry out criminal activities leaving almost no evidence behind.

Due to the distributed and virtualized environment, cloud forensics is facing a great deal of challenges in comparison to traditional forensics. The ability to conduct a proper cloud forensic investigation depends on the methods and tools used to acquire digital evidence. Unfortunately, current methods and tools do not meet the standards

L. Iliadis, M. Papazoglou, and K. Pohl (Eds.): CAiSE 2014 Workshops, LNBIP 178, pp. 299–309, 2014.

on cloud computing [1], [2], [3], [4]. Many authors have dealt with these challenges and proposed new solutions and frameworks to overcome the problems.

Within this work, we summarize the challenges and issues raised in cloud computing environments regarding the cloud forensics process. We, then, explore the cloud forensic solutions addressed to the challenges and a summary of these findings is presented based on a detailed literature review. Specifically the paper is organized as follows. In section 2 a table summarizing the challenges identified in cloud forensics is presented. Section 3 presents the current cloud forensic solutions addressed based on the identified challenges. Finally, section 4 concludes the paper by presenting a categorization of the results and raises future research issues derived from this review.

Since cloud forensics is a newly developed research area our main and primary fo-cus was to conduct a thorough analysis of the respective literature in order to present an analytic review of the existing solutions regarding the challenges in the respective field. It is not in the scope of this paper to cover the quality of the audit control. The starting point of the review was some certain scientific papers on cloud forensic issues and we broadened our review to the related work referenced. After studying these papers our review broadened to less related academic reports and papers from the field of cloud computing.

2 Cloud Forensic Challenges

Although cloud computing has been introduced and used in the market enough years, the cloud forensics is still at its infancy. There are still many open issues and challenges to be explored. Past years many researchers have tried to identify the challenges and the work produced by some of them is accurate and well documented. The cloud forensics challenges identified from the review conducted in the respective area can be categorized into identification stage, preservation and collection stage, examination and analysis stage, and presentation stage. Table 1 summarizes the challenges identified in the three service models [5]:

Table 1. Cloud Forensics Challenges

Cloud Forensic Challenges / Stage	Applicable to		
	IaaS	PaaS	SaaS
Identification			
Access to evidence in logs	partly	√	√
Physical inaccessibility	√	√	√
Volatile data	√	X	X
Client side identification	√	X	√
Dependence on CSP - Trust	√	√	√
Service Level Agreement (SLA)	√	√	√
Preservation - Collection			
Integrity and stability	√	√	√
Privacy	X	√	√

Table 1. (*Continued.*)

Time synchronization	√	√	√
Internal Staffing	√	√	√
Chain of custody	√	√	√
Imaging	X	√	√
Bandwidth limitation	√	X	X
Multi-jurisdiction - collaboration	√	√	√
Multi-tenancy	√	√	√
Examination – Analysis			
Lack of forensic tools	√	√	√
Volume of data	X	√	√
Encryption	√	√	√
Reconstruction	√	√	√
Unification of log formats	√	√	√
Identity	√	√	√
Presentation			
Complexity of testimony	√	√	√
Documentation	√	√	√
Uncategorised			
Compliance issues	√	√	√

Some may consider that certain of the above challenges are not part of the cloud forensics process/stage as they regard them input to the forensic process itself. Still, most of the researchers identified in the review [1], [2], [3], [4], argue that all of the challenges referred in Table 1 must be considered as challenges.

3 Current Solutions

After summarizing the cloud forensics challenges this section presents all possible solutions addressing clarified challenges, identified from an analytical review conducted in the respective area. In the following section identified solutions are presented categorized per challenge.

3.1 Access to Evidence in Logs

This challenge is the most important one in cloud forensics and is referred from every researcher that deals with the respective field. Many of them have come up with solutions such as Sang [6], who proposed a log-based model which can help to reduce the complexity of forensic for non-repudiation of behaviors on cloud. He proposes that we should keep another log locally and synchronously, so we can use it to check the activities on SaaS cloud without the CSP's interference. The local log module will use information such as unique id and timestamp on the log record locally. HASH code will be also used to detect modification on the log files. In PaaS, the CSPs should

supply a log module on PaaS to the third-party in order to create a customized log module, for both of the consumer side and the cloud side.

In PaaS, since the customers have full control on their application over a prepared API, system states and specific application logs can be extracted. Birk et al. [1] proposed a logging mechanism which automatically sign and encrypt the log information before its transfer to a central logging server under the control of the customer. This mechanism will prevent potential eavesdroppers from being able to view and alter log data information on the way to the logging server.

Solving the cloud logging problems Marty [7] proposed a log management architecture that involves three steps: enable logging on all infrastructure components to collect logs from, setup and configure log transport and finally tune log sources to make sure we get the right type of logs and the right details collected. He states that every log entry should log what happened, when it happened, who triggered the event and why it happened. According to this, the minimum fields need to be present in every log record are: Timestamp, Application, User, Session ID, Severity, Reason and Categorization. He also recommends an application on how log entries should be structured. At the end an application logging infrastructure at SaaS company was implemented using application components such as Django, JavaScript, Apache, MySQL, Operating system and Java Backend. Zawoad et al. [2] mentioned that although the advantages to this approach are several, the specific work does not provide any solution about logging network usage, file metadata, process usage and many other evidence, which are important for forensic investigation in IaaS and PaaS.

Damshenas et al. [8] suggested a solution in PaaS, to prepare an API to extract relevant status data of the system, limited by the data related to the client only. In SaaS, depends on the interface, he proposed to implement the feature to check the basic logs and status of the client's usage. The above features should provide read-only access only and demands for specific log and system status manager running as a cloud service.

According to Zafarullah et al. [9] logging standards should be developed, which ensure generation and retention of logs and a log management system that collects and correlates logs. A cloud computing environment was setup using Eucalyptus. Using Snort, Syslog and Log Analyzer (e.g. Sawmill) Eucalyptus behavior was monitored and all internal and external interaction of Eucalyptus was logged. Observing the log entries were generated by the Eucalyptus, not only the attacker's IP address was recorded, but also details on number of http requests along with timestamps, http requests/responses and fingerprinted attacker's OS & web browser were provided.

3.2 Volatile Data

To overcome the problem of volatile data, live investigation has been used as an alternative approach to dead acquisition. Grispos et al. [3] mentioned that the specific approach enables investigators to gather data that might otherwise be lost if a computer is powered down. On the other side it may increase the amount of information an investigator is able to extract. To address this challenge, Damshenas et al. [8]

proposed the cost to be globalized between CSPs to offer persistent storage device for client's data.

To prevent loss of volatile data, Birk et al. [1] suggested frequent data synchronization between the VM and the persistent storage or a non-cloud based storage. According to Zawoad et al. [2] this solution does not provide any guideline about the procedure and he proposed two possible ways of continuous synchronization. CSPs can provide a continuous synchronization API to customers and CSPs can integrate the synchronization mechanism with every VM and preserve the data within their infrastructure.

3.3 Multi-jurisdiction – Distribution – Collaboration

New regulations have to be developed in order to solve the cross border legislation issue. Biggs et al. [10] proposed an international legislation that will police the internet and cloud computing specifically. Global unity must be established so the investigations on cloud environment to be fast and successful. Grispos et al. [3] suggested that a partial solution to different jurisdictions could be the CSPs to have trained and qualified personnel to perform forensic investigations when needed. According to Ruan et al. [4] and Sibiya et al. [11] international laws should be developed to secure that forensic activities will not breach any laws or regulations under any jurisdiction.

3.4 Client Side Identification

To identify evidence on client's side, Damshenas et al. [8] suggested designing and implementing an application to log all potential evidence on the client's machine. However, they did not provide any methodology about the application and the procedure.

3.5 Dependence on CSP - Trust

In cloud environments, customers have to depend completely on the CSPs, which affect the trust relationship between them. The lack of transparency and trust between CSP's and customers is an issue that Haeberlen [12] was dealt with the accountable cloud. He suggested a basic primitive called AUDIT that an accountable cloud could provide. The idea is that the cloud, records its actions in a tamper evident log, customers can audit the log and check for faults and finally they can use log to construct evidence that a fault has (or not) occurred. When an auditor detects a fault, it can obtain evidence of the fault that can be verified independently by a third party. A TrustCloud framework proposed by Ko et al. [13], which consists of five layers of accountability: System, Data, Workflow, Policies, Laws & Regulations layers. To increase accountability detective approaches used rather than preventive.

Nurmi et al. [14] presented Eucalyptus, an open-source software framework for cloud computing that implements IaaS, which is the answer to the trust relationship between CSPs and customers. A model showing the layers of trust has been introduced by Dykstra et al. [15]. In IaaS, six layers have been established and more layers would have added in the other two cloud models. Each layer requires a different amount of confidence. The further down the stack, the less cumulative trust is required.

3.6 Service Level Agreement

SLAs should include important terms regarding cloud forensic investigations. According to Ruan et al. [4] SLAs should include: Service provided, techniques supported, access granted by the CSP to the customer, trust boundaries, roles and responsibilities between the CSP and the cloud customer, security issues in a multi-jurisdictional environment in terms of legal regulations, confidentiality of customer data, and privacy policies and security issues in a multi-tenant environment in terms of legal regulations, confidentiality of customer data and privacy policies. A well-written SLA between CSP and customer should include the client's privacy policies Damshenas et al. [8].

Biggs et al. [10] proposed SLA's to be robust in order to be effective in combating cybercrime. For example illegal activities such as DDOS etc. should test cloud vendors' systems and procedures and return useful feedback to assist forensic procedures. To overcome the SLA's issue with different and multiple relationships Birk et al. [1] suggested a trusted third-party to audit the security measures provided by the CSP. Finally SLAs' violation is another problem in which Haeberlen [12] proposed the trusted timestamping. Timing information must be added to a tamper-evident log in order to detect the violations.

3.7 Integrity and Stability – Privacy and Multi-Tenancy

To validate the integrity of the evidence Zawoad et al. [2] suggested a digital signature on the collected evidence should be generated and then the signature should be checked. Hegarty et al. [16] developed and implemented a distributed signature detection framework that enables forensic analysis of storage platforms. Based on the meta-data driven data storage model and provenance integrity, in SaaS, Shi et al. [17] presented a multi-tenancy model where the data storage security issue should be mapped as a series of integrity issues of data chunks. To ensure the primitiveness and integrity of the evidence Yan [18] proposed a new cybercrime forensic framework to image the relative records and files absolutely.

Juels et al. [19] explored proofs of retrievability (PORs) in which a prover (i.e. back-up service) can produce a concise proof that a verifier (client) can retrieve a file in its entirety. PORs method and cryptographic techniques can help users to ensure the privacy and integrity of files they retrieve. To preserve the integrity of the data Birk et al. [1] proposed the Trusted Platform Module (TPM) to assure the integrity of a platform. This standard allows a secure storage and detects changes to previous configurations. Damshenas et al. [8] suggested all the issues concerning clients' privacy data should be included in an SLA contract.

3.8 Time Synchronization – Reconstruction

To solve the time zones' problem Damshenas et al. [8] suggested a specific time system (i.e. GMT) to be used on all entities of the cloud, as it brings the benefit of having a logical time pattern. In IaaS, the VM time is under the client's control meaning that all date and times used in logs and other records should be converted to the specific time system.

3.9 Internal Staffing

It is hard to find the right people to work as a team in order to be involved in a cloud investigation. Ruan et al. [4] proposed a solution that involves internal staffing, CSP-customer collaboration and external assistance with specific roles. Individuals of the team must be trained on, law regulations, new methodologies, specialized tools and techniques. According to Chen et al. [20] an investigator should possess the abilities of professional forensics skills such programming, networking etc., co-operating, communicating and negotiating with CSPs and understanding laws and regulations.

3.10 Chain of Custody

Grispos et al. [3] suggested trained and qualified personnel in forensic investigations should be hired by CSPs. When an investigation arises the personnel should begin the chain of custody process which will be passed onto the investigation party. According to Ruan et al. [4] organizational policies and legally binded SLAs need to be written, in which, communications and collaborations regarding forensic activities through the chain of CSPs and customers dependencies should be clearly stated.

3.11 Imaging

To overcome the issue of acquiring forensic image Damshenas et al. [8] proposed to generate a track record of all clients' activities Later on, to generate a forensic image of specific clients all it requires is to check the track record of the client and then copy bit-by-bit stream of all the area the client has accessed to.

3.12 Forensic Tools

Most of the researchers acknowledge that tools need to be developed to identify, collect and analyze forensic data. Juels et al. [19] developed Proofs Of Retrievability (PORs) tool for semi-trusted online archives which guarantees the privacy and the integrity of files. In IaaS, Dykstra et al. [15] recommended the appropriate forensic tool for acquiring cloud-based data is the management plane. This is a web-based point and click interface to manage and monitor the infrastructure. They concluded that it offers the most attractive balance of speed and control with trust option. EnCase and Accessdata FTK tools were also used to acquire evidence and the results were successful, but authors do not recommend them because too much trust is required. Recently, Dykstra et al. [21] designed and implemented a management plane forensic toolkit in a private instantiation of the OpenStack cloud platform (IaaS), which is called Forensic Open-Stack Tools (FROST) – It consists of three new forensic tools and it provides trustworthy forensic acquisition of virtual disks, API logs, and guest firewall logs.

3.13 Volume of Data

A solution to the challenge is to use the public clouds to store the evidence but this method arise new issues from a legal and technical perspective Grispos et al. [3]. The other solution is the adoption of triaging techniques. New methods should be

developed to allow only partial recovery of data and they should be according to accepted forensic principles.

3.14 Complexity of Testimony

Wolthusen [22] suggested of using interactive presentation and virtualization environments which allow the exploration of data sets in such a way that a focus on relevant data is possible without engendering the risk of leading questions and investigations.

3.15 Documentation

The documentation of the investigation according to Wolthusen [22] must be presented in a way pointing: possible gaps in the data sets, uncertainties about the semantics and interpretation of data and the limitations of the collection mechanisms alongside the actual data.

3.16 Compliance Issues

According to Birk et al. [1] recommended customers should check their compliance requirements and CSPs services to find out which CSP matches customers' needs. On the other hand CSP should offer as much transparency as possible. Finally a Third Party Auditor could be used acting as a trustee between the customer and the CSP.

4 Discussion

Cloud computing undeniable offers various benefits to the users. However, there are plenty of issues that need to be resolved in order to conduct a proper investigation regarding cloud forensics. In the previous section, we have mentioned several solutions addressed to the challenges proposed by researchers. The problem is that most of them have not been tested in real conditions (i.e. only recently, Dykstra implemented FROST, the first dedicated collection of forensic tools). The above mentioned findings regarding the available solutions for every identified cloud forensics challenge are summarized in table 2. This table summarizes the cloud forensic challenges identified from the review conducted in the respective area and all the researchers' proposed solutions addressed to challenges.

As we can see in table 2 there are some challenges missing and some have been combined with others. This is due to the absence of a solution for a respective challenge or that a solution can satisfy more than one challenge. Even though, forensic investigation in cloud environments has moved forward the past years, there are still open issues to explore. Dependence on CSP is still required in various issues, such as access to log files and trust relationship. Most of the problems rely on the CSPs' point of view. Absence of international standards and regulations cannot establish the global unity which can help to cross the boundaries in multi-jurisdiction and collaboration challenge.

Table 2. Cloud Forensics Solutions

Cloud Forensic Challenges	Solution	Related Work
Access to evidence in logs	Log-based model	[6]
	Logging mechanism	[1]
	Log management architecture	[7]
	Status data extraction and checking	[8]
	Eucalyptus framework	[9]
Volatile data	Live investigation	[3]
	Cost globalization between CSPs	[8]
	Data synchronization	[1]
	Continous synchronization API	[2]
Multi-jurisdiction - collaboration	International legislations and global unity	[10]
	Trained and qualified personnel	[3]
	International laws	[4]
		[11]
Client side identification	Log application	[8]
Dependence on CSP - Trust	Accountable cloud	[12]
	TrustCloud framework	[13]
	Eucalyptus framework	[14]
	Layers of trust model	[15]
Service Level Agreement (SLA)	Well and clear-written terms	[8]
		[4]
	Robust SLAs	[10]
	External auditors	[1]
	Trusted timestamping	[12]
Integrity & stability - Privacy & multi-tenancy	Digital signature	[2]
	Distributed signature detection framework	[16]
	Multi-tenancy model	[17]
	Cybercrime forensic framework	[18]
	Proofs Of Retrievability (PORs)	[19]
	Trusted Platform Module	[1]
	SLA contracts	[8]
Time synchronization - Reconstruction	Unified/specific time system	[8]
Internal Staffing	Team collaboration with wide range of skills	[4]
		[20]
Chain of custody	Trained and qualified personnel	[3]
	Organizational policies and SLAs	[4]

Table 2. (*Continued.*)

Imaging	Track record generator	[8]
Lack of forensic tools	Proofs Of Retrievability (PORs)	[19]
	Management plane	[15]
	Forensic Open-Stack Tools (FROST)	[21]
Volume of data	Public cloud storage	[3]
	Triaging techniques	
Complexity of testimony	Interactive presentation	[22]
Documentation	Targeted/pointed presentation	[22]
Compliance issues	Survey	[1]
	Transparency	
	Third Party Auditor (TPA)	

Unification of log formats is another issue which needs to be solved. All the evidence need to be presented in a court of law in such a way that the jury could understand the complexity of the non-standard data sets. Depending on the volume of data, bandwidth limitation is another issue that needs to be solved, when the time, is a crucial factor to an ongoing investigation. The identity of the user who has been engaged in a criminal act is also an unanswered case.

References

1. Dominik, B., Wegener, C.: Technical issues of forensic investigations in cloud computing environments. In: IEEE Sixth International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE) 2011. IEEE (2011)
2. Shams, Z., Hasan, R.: Cloud Forensics: A Meta-Study of Challenges, Approaches, and Open Problems. arXiv preprint arXiv:1302.6312 (2013)
3. George, G., Storer, T., Glisson, W.B.: Calm Before the Storm: The Challenges of Cloud. Emerging Digital Forensics Applications for Crime Detection, Prevention, and Security, p. 211 (2013)
4. Ruan, K., Carthy, J., Kechadi, T., Crosbie, M.: Cloud forensics: An overview. Advances in Digital Forensics 7, 35–49 (2011)
5. Simou, S., Kalloniatis, C., Kavakli, E., Gritzalis, S.: Cloud Forensics: Identifying the Major Issues and Challenges. In: Jarke, M., Mylopoulos, J., Quix, C. (eds.) CAiSE 2014 26th International Conference on Advanced Information Systems Engineering. LNCS, Springer, Heidelberg (June 2014)
6. Ting, S.: A Log Based Approach to Make Digital Forensics Easier on Cloud Computing. In: 2013 Third International Conference on Intelligent System Design and Engineering Applications (ISDEA). IEEE (2013)
7. Marty, R.: Cloud application logging for forensics. In: Proceedings of the 2011 ACM Symposium on Applied Computing. ACM (2011)
8. Mohsen, D., Dehghantanha, A., Mahmoud, R., Shamsuddin, S.B.: Forensics investigation challenges in cloud computing environments. In: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec). IEEE (2012)

9. Zafarullah, Z.,Anwar, F., Anwar, Z.: Digital forensics for eucalyptus. In: Frontiers of Information Technology (FIT). IEEE (2011)
10. Biggs, S., Vidalis, S.: Cloud computing: The impact on digital forensic investigations. In: International Conference for Internet Technology and Secured Transactions, ICITST 2009. IEEE (2009)
11. George, S., Venter, H.S., Fogwill, T.: Digital forensic framework for a cloud environment (2012)
12. Haeberlen, A.: A case for the accountable cloud. ACM SIGOPS Operating Systems Review 44(2), 52–57 (2010)
13. Ko, R.K., et al.: TrustCloud: A framework for accountability and trust in cloud computing. In: 2011 IEEE World Congress on Services (SERVICES). IEEE (2011)
14. Nurmi, D., et al.: The eucalyptus open-source cloud-computing system. In: 9th IEEE/ACM International Symposium on Cluster Computing and the Grid, CCGRID 2009. IEEE (2009)
15. Dykstra, J., Sherman, A.T.: Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques. Digital Investigation 9, S90–S98 (2012)
16. Hegarty, R., et al.: Forensic analysis of distributed data in a service oriented computing platform. In: Proceedings of the 10th Annual Postgraduate Symposium on the Convergence of Telecommunications, Networking & Broadcasting, PG Net (2009)
17. Shi, Y., Zhang, K., Li, Q.: A new data integrity verification mechanism for SaaS. In: Wang, F.L., Gong, Z., Luo, X., Lei, J. (eds.) WISM 2010. LNCS, vol. 6318, pp. 236–243. Springer, Heidelberg (2010)
18. Yan, C.: Cybercrime forensic system in cloud computing. In: 2011 International Conference on Image Analysis and Signal Processing (IASP). IEEE (2011)
19. Juels, A., Kaliski Jr., B.S.: PORs: Proofs of retrievability for large files. In: Proceedings of the 14th ACM Conference on Computer and Communications Security. ACM (2007)
20. Chen, G., Du, Y., Qin, P., Du, J.: Suggestions to digital forensics in Cloud computing ERA. In: 2012 3rd IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC). IEEE (2012)
21. Dykstra, J., Sherman, A.T.: Design and implementation of FROST: Digital forensic tools for the OpenStack cloud computing platform. Digital Investigation 10, S87-S95 (2013)
22. Wolthusen, S.D.: Overcast: Forensic discovery in cloud environments. In: Fifth International Conference on IT Security Incident Management and IT Forensics, IMF 2009. IEEE (2009)