Digital Forensics in the Cloud:
The State of the Art

George SIBIYA\textsuperscript{1}, Hein S. VENTER\textsuperscript{2}, Thomas FOGWILL\textsuperscript{1}
\textsuperscript{1}Meraka Institute, CSIR, Pretoria, SA
Emails: \{gsibiya, tfogwill\}@csir.co.za
\textsuperscript{2}Department of Computer Science University of Pretoria, Pretoria, RSA
Email: hsventer@cs.up.ac.za

Abstract: The advent of cloud computing has brought new challenges to digital forensics. To address these challenges, new approaches in conducting digital forensic are required. In this paper, challenges that are faced by digital forensic investigator when faced with cloud based incident scenes are presented. The presented challenges are obtained from survey articles that explore outstanding and future challenges in digital forensics in general. In this paper we zoom in into cloud forensics as it is the main focus of the paper. Based on the challenges brought to light by the considered survey articles, we present requirements that should be met by digital forensic systems that aim to investigate cloud environments. Existing architectures and implementations of digital forensic systems are evaluated based on these requirements. Through this evaluation, gaps that are left out by the evaluated architectures are brought to light.

Keywords: Cloud computing, Digital forensics, Architecture, Digital forensic process, survey

1. Introduction

The continuous developments in technology have ushered in cloud computing in the last few years. Cloud computing has unique characteristics that require changes in practices in which digital forensics has been conducted. The unique characteristics include its virtualised, volatile, multi-tenancy and multi-jurisdictional distributed nature. These characteristics make it difficult to conduct an investigation in the cloud using conventional processes that include evidence identification, evidence collection, evidence analysis, reporting and presentation \cite{1}. Furthermore, conventional digital forensic tools that often required pulling the plug (powering off incident scene) or to attach digital forensic devices on the incident scene physical host are not applicable in the cloud as incident scenes are not always physically accessible. These challenges require new specialised processes and tools for cloud environments.

Research efforts have been on going in addressing the issues of standard processes and tools. In this paper we explore research endeavours that provide digital forensic systems and/or implement standardised digital forensic processes. The research endeavours in this paper are evaluated based on criteria deduced from survey articles.

The paper is organised as follows: Section 2 is dedicated to the methodology used to sample survey and cloud forensic system research articles that are considered in this paper. The section also covers how the criteria used to evaluate the cloud forensic systems are deduced from the survey articles; In Section 3 we present requirements for systems that are aimed at investigating cloud environments as raised by other researchers; In Section 4, existing digital forensic systems are analysed including the extent to which they satisfy the
requirements presented in Section 3; In Section 5, we present a discussion on the presented analysis of architectures and in Section 6 we conclude the paper.

2. Methodology

The survey articles that were used in this research were obtained by using the following search phrases: “cloud forensics, a survey”; “cloud forensics, a state of the art”; “digital forensics in the cloud, a survey” and “digital forensics in the cloud, state of the art” from IEEE Explore [2], ACM [3], Elsevier [4] and Springer Link [5] digital libraries. The digital forensics systems were obtained by searching the same digital libraries with the following phrases: “digital forensic system”; “cloud forensic system”; “digital forensic lab”. Search results are further filtered by publication year where articles that are older than five years are excluded from the results. The digital forensic challenges identified in the survey articles are used to identify features required to be on a digital forensic system in order to address the corresponding challenges.

In the next section the survey articles including deduced digital forensic system requirements are presented.

3. Cloud Forensic Systems Design requirements

In this section we present summaries of surveys on digital forensic systems and later, the deduced digital forensic systems requirements.

3.1 Survey Literature Summary

In this section we present prevalent issues associated with conducting digital forensics in the cloud. This was done by taking into consideration, the surveys conducted by other researchers concerning issues of digital forensics. The research works considered are research work by Fahdi et al. [6], Henry et al. [7], Quick and Choo [8], Wazid et al. [9], Ruan et al. [10] and Mishra et al. [11]. The criteria used to select the research works was based on them being survey articles that are not older than five years (not published after 2011) and are exploring digital forensic issues and more specifically issues with digital forensics in cloud computing. These are survey articles that discuss challenges that faced by digital forensics investigators.

Fahdi et al. [6] selected respondents to their surveys from digital forensic practitioners and from digital forensic researchers. In Fahdi et al. [6], top six limitations of digital forensics as rated by digital forensic practitioners in their order are Volume of data, Legal Aspect, Time, Tool Capacity, Visualisation and Automation of Forensic Analysis. On the other hand, the top six limitations according to researchers are Time, Volume of Data, Automation of Forensic Analysis, Tool Capacity, Visualisation, Forensic Training and Legal Aspect. Further in Fahdi et al., Cloud computing is highly rated among the technologies that are a cause for concern for digital forensics both currently and in future. Surveys by Wazid et al. and Simou et al. [12] also agree with Fahdi et al. [6], on the amount of data forming one of the challenges in investigating cloud environments.

In Henry et al. [7], results have shown that investigations on cloud based systems and services only form a fraction of all investigations carried out by investigators. This is a reflection of the challenges attributed with cloud based investigations meaning that investigators are reluctant or unable to extend an investigation to the cloud. These challenges force investigations to end on client devices owned by perpetrators and not to extend to cloud based data. In Henry et al., the lack of standard and tools constitutes most of the challenges associated with investigating cloud environments according to respondents. The standards and tools challenge is followed by lack of skills and certifications on the side of investigators then followed by legal issues of ownership and privacy. The respondents
also rate imaging incident scenes the highest in complicating digital forensic investigations in virtualised environments. This is followed by disk acquisition, legal processes, live response, monitoring/scanning for events and evidence analysis.

Quick and Choo [8] address the issue of the large volumes of data that investigators are often faced with. In Quick and Choo it is pointed out that by searching for evidence in a large volume of data, chances of missing crucial evidence data is high and this also increases time required to conduct an investigation. It is also conceded that tools that can automate some investigation tasks are required.

Surveys by Ruan et al. [10] and Simou et al. [12] agree on collaborations among multi-jurisdictions as a challenge. Other challenges that emerged from Wazid et al. [9] include diverse digital media types, anonymity of IP addresses which may be difficult to trace and anti-forensic techniques such as encryption. Encryption is also cited in Simou et al. [12] as a challenge. Challenges identified by Ruan et al. [10] include increase in mobile devices (this can be linked to diverse digital media types in Wazid et al.), interdependency among cloud service providers (challenging when they all have to be involved in an investigation), and inadequate control over cloud data by a customer. Diverse log formats are identified as being a challenge by Mishra et al. in [11] and Simou et al. in [12].

From the challenges presented in this section, it is clear that there cannot be a one size fits all solution to the challenges. It is in our view, however, that digital forensic standards tailored for the cloud in combination with collaborative environments can play a role in addressing these challenges. In the next section we present digital forensic system requirements that emanate from the challenges presented in this section. As it will be observed, the requirements lean towards digital forensic process standards implementation and collaborative environments.

3.2 Deduced Digital forensic system requirements

Based on the analysis of the surveys, we now summarise and propose cloud forensic systems requirements. The requirements presented are not standardised as of yet but are aimed at inspiring research towards the direction. The proposed cloud forensic system requirements are as follows:

1) **Live Cloud Forensic Investigation** - With the ability to analyse a live cloud based system in a standardised manner would help eliminate the need to image data and real time crucial evidence can be obtained from the incident scene. Live forensic response has also emerged as a one of the digital forensic challenges in the survey by Henry et al. in [7].

2) **Implements a standardised digital forensic process** – Fahdi et al. [6] and Quick and Choo [8] investigations both revealed the importance of automation during a digital forensic process. It is in our view that standards which were also found to be lacking by Henry et al. [7] can play a role as they can help in specifying tasks that can be automated in a digital forensic process. Automating some of the tasks can result into an efficient digital forensic system. An efficient system can lead into lesser time spent in an investigation and time is regarded as one of the issues in Fahdi et al. [6] and Quick and Choo [8].

   It is in our view also that standards can solve this challenge as standards. Standards can assist investigators with processes and procedures that need to be carried out while investigating cloud based incidences. In this way, standards can therefore also help facilitate cooperation from cloud service providers and multi-jurisdictional investigators.

3) **Aid and Lead Investigator through a Standard digital forensic Process** - There are a number of digital forensic processes proposed by researchers some of which are not
standardised. There are however emerging efforts towards standardisation of the digital forensic processes such as the ISO/IEC27037 [13] and ISO/IEC27043 [14]. Given the number of existing ones, it would be impractical for an investigator to know them all even after training. A digital forensic system that can lead an investigator through a chosen standard process would help can contribute in efficiency as an investigator would not always need to consult with documentation of the chosen standard every time an investigation is carried out.

With this requirement satisfied, a digital forensic system can assist investigators go through a complex digital forensic process efficiently.

4) **Cloud Based (i.e. runs on the cloud)** - The volume of data that has to be dealt with in an investigation emerged as one of the challenges in Fahdi et al. [6] and Henry et al. [7] surveys. Deploying a digital forensic system on the cloud would not only help address the challenge of dealing with large data by utilising unlimited processing and storage resources provided by the cloud but, the self-service and collaborative environment in the cloud would help enable digital forensic investigations to commence timorously and be executed efficiently.

5) **Aimed at investigating cloud environments** - Existing digital forensic tools and processes are either meant for general purposes or specific for other electronic environment but the cloud. Though the tools can be used in combination to investigate the cloud, there is need for a specialised investigation platforms meant for the cloud. A specialised platform would ensure that all unique aspects of the cloud are adequately covered during a digital forensic investigation.

6) **Digital Forensics Readiness** - Monitoring or scanning for events in a cloud has emerged as one of the challenges in the Survey by Henry et al. [7]. We regard monitoring as digital forensic readiness as it is mechanism that is always on standby to trigger an investigation and also makes data available to initiate the investigation. In the case of the cloud, digital forensic readiness is an additional mechanism incorporated in cloud services that makes digital forensic to be readily available for investigation. Maintaining digital forensic readiness mechanisms come at a cost on the cloud service provider side or on the customer side if the service provider transfers the costs to the customer.

A driving force of businesses is the delivery of their product and digital forensic readiness does not always add value to a business. Businesses therefore do not always invest on digital forensic readiness on their cloud services. If a cloud forensic system would offer this functionality and be deployed as a cloud service, the readiness feature can be utilised by cloud service providers as and when needed. Digital forensic readiness therefore needs to be used as a criterion to evaluate a digital forensic system.

In the next section we present an evaluation of existing cloud forensic system architectures. The evaluation will be based on the requirements presented in this section as criteria.

4. **Cloud forensics architectures and their evaluation**

Digital forensics as a service has been proposed by many researchers in the last 4 years. In this section we discuss some of the research works that are aimed at investigating cloud environments and/or deployed on the cloud. The research works are then evaluated based on the requirements presented in Section 3.

4.1 **Evaluation Criteria**

We evaluate the suitability of a digital forensic system for live forensics based on its ability to interact with a live system and gather evidence from it in a standardised manner. In cloud
environments it is not possible to power a virtual machines or a physical server for imaging as they may be hosting essential services that belong to fellow tenants. Evaluation on whether a digital forensic system implements a standardised digital forensic process is based on whether it has defined digital forensic processes and procedures, the defined processes are standardised (i.e based on published standards), and whether the processes are incorporated in the digital forensic system implementation.

The ability to lead an investigator through standardised digital forensic processes is based on the system’s ability to guide an investigator through investigation tasks and automatically renders the tasks to investigators in their standard sequence. Some of the investigation tasks need to be automated in the digital forensic system. By the way, automation of investigation tasks emerged as one of the challenges that face digital forensic investigators in the survey by Fahdi et al. in [6].

Any web application or service can be developed and deployed in a cloud environment without necessarily implementing the features found in cloud services. To evaluate a digital forensic system on its deployment in the cloud we check whether the system offers a collaborative environment, investigators have access to the same version of the service and that the digital forensic system has resources that scale on demand.

Finally, to evaluate if a digital forensic system is aimed at investigating cloud environments, we check if it is designed and implemented with remote access to cloud based incident scene capability. To evaluate digital forensic readiness in a forensic system we check if the system adopts a proactive approach to digital forensic investigations.

4.2 Evaluated Architectures

In this section we present the evaluation of existing digital forensics systems based on the criteria presented in Section 3.2.

The literature discussed in this section is summarised in Table I. In Table I, the sign ✓, means that the feature is supported by the digital forensic system. The mark, ✗, means that the feature is not supported by the digital forensic system. The sign, —, indicates that no information could be found on the article concerning the feature. The literature represented in the table are summarised below.

In FaaS by Wen et al. [15] is concerned with scalable evidence data processing using the cloud. They use capabilities of the cloud in analysing large evidence data.

Sang [16] proposes an approach whereby a Software-as-a-Service provider and a Platform-as-a-Service provider also provides log modules that would be executed on the client side in addition to services that they offer. This approach partially removes cloud service providers from the equation since they may sometime not be willing or have no capacity to participate in an investigation.

Forensic OpenStack Tools (FROST) [17] involves modification of cloud management software in this case OpenStack to integrate digital forensic capabilities. In digital forensic terms, this is referred to as digital forensic readiness. Shende [18] presents a Cloud Forensic-as-a-Service (FRaaS) which was also published by the author as a book chapter in (Dykstra, 2012). The cloud based digital forensic service deployment concept presented in the system is required as it was also proposed by many other researchers in the last few years including: Lee and Un [23]; Marty, [35]; Ruan et al. [36] and Wen et al. [15].
## Table 1 Cloud Forensic Systems Evaluation

<table>
<thead>
<tr>
<th>Live Cloud Forensic Investigation</th>
<th>Implements a standardised digital forensic process</th>
<th>Aid and Lead Investigator through a Standard digital forensic Process</th>
<th>Cloud Based (i.e. runs on the cloud)</th>
<th>Aimed at investigating Cloud environments</th>
<th>Digital Forensics Readiness</th>
<th>Proactive Forensic Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieve data from live system</td>
<td>Has defined forensic processes and procedures</td>
<td>System renders tasks to investigator based on their standard sequence</td>
<td>Collaborating investigators have access to same Service Instance</td>
<td>Hardware resources can scale on demand</td>
<td>Designed with remote access capability to incident scene</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Processes Standardised</td>
<td>In implementation</td>
<td>Investigators run similar service version at all times</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Wen et al. [15]                  | X                                               | X                                                               | ✓                                                                   | ✓                                                         | X                           | X                           |
| Sang [16]                        | X                                               | X                                                               | ✓                                                                   | ✓                                                         | ✓                           | ✓                           |
| Dykstra and Sherman [17]         | ✓                                               | X                                                               | ✓                                                                   | ✓                                                         | ✓                           | ✓                           |
| Shende [18]                      | X                                               | –                                                               | ✓                                                                   | ✓                                                         | X                           | ✓                           |
| Lee et al. [19]                  | ✓                                               | X                                                               | ✓                                                                   | ✓                                                         | X                           | ✓                           |
| Deepak et al. [20]               | X                                               | X                                                               | X                                                                   | ✓                                                         | ✓                           | X                           |
| Zeng [21]                        | –                                               | ✓                                                               | –                                                                   | –                                                         | –                           | –                           |
| Lee et al. [22]                  | X                                               | ✓                                                               | X                                                                   | ✓                                                         | X                           | ➜                           |
| Lee and Un [23]                  | X                                               | X                                                               | X                                                                   | ✓                                                         | X                           | X                           |
| Baar et al. [24]                 | X                                               | ✓                                                               | X                                                                   | –                                                         | X                           | X                           |
| Ting and Yang [25]               | ✓                                               | X                                                               | X                                                                   | ✓                                                         | ✓                           | X                           |
| Thorpe et al. [26]               | X                                               | X                                                               | X                                                                   | –                                                         | –                           | –                           |
| Shirkherdkar and Patil [27]      | ✓                                               | X                                                               | X                                                                   | –                                                         | –                           | ✓                           |
| Shields, Frieder and Maloof [28] | –                                               | X                                                               | X                                                                   | X                                                         | X                           | X                           |
| Reddy and Venter [29]            | X                                               | X                                                               | X                                                                   | X                                                         | X                           | X                           |
| Li and Du [30]                   | –                                               | ✓                                                               | X                                                                   | –                                                         | –                           | –                           |
| Yan [31]                         | –                                               | X                                                               | X                                                                   | –                                                         | –                           | –                           |
| Belorkar and Geethakumari [32]   | ✓                                               | X                                                               | X                                                                   | –                                                         | –                           | –                           |
| Zawoad and Hasan [33]            | X                                               | X                                                               | X                                                                   | –                                                         | –                           | –                           |
| Reichert, Richards and Yoshigoe [34] | ✓                                           | X                                                               | X                                                                   | –                                                         | –                           | –                           |

The system by Lee [19] is a deployment of a digital forensic service in a scalable resources platform such as the cloud. Lee goes further and offers the capability to investigate an incident remotely. The approach by Lee however is based on a physical...
accessibility of the incident scene so that investigation components can be attached to the incident scene or evidence device and physical accessibility is not always possible in the cloud. The system by Deepak et al. [20] is a cloud service which utilises cloud resources to analyse videos in order to detect illegal content. Its aim is not to investigate cloud based incident scenes which in this paper is viewed as having to be a requirement.

The technologies proposed in Zeng [21], Lee et al. [22], Lee and Un [23], Baar et al. [24] and Ting and Yang [25] implement cloud based services to carry out computing expensive tasks including data analysis and performing searches on big evidence data. The technologies can be referred to as “cloud based forensic labs” designed to process data obtained from any storage media. There is, however, a need for system designs with a goal to conduct an investigation by interacting with a live cloud based incident scene in a standardised manner.

Thorpe [26] and Shirkherdkar and Patil [27] are digital forensic readiness frameworks for cloud computing. In addition to digital forensic readiness, there is a need to also focus on investigating an incident as and when it occurs regardless of whether forensic readiness mechanisms were in place or not. Thorpe presents a detailed framework on how digital forensic readiness mechanisms can be incorporated in a cloud service stack. Shirkherdkar and Patil [27] present an approach where communications between the incident scene and a potential attacker can be monitored.

Other research works included in the evaluation are: Shields, Frieder and Maloof [28]; Reddy and Venter [29]; Li and Du [30]; Yan [31]; Belorkar and Geethakumari [32]; Zawoad and Hasan [33] and Reichert, Richards and Yoshigoe [34]. The aspects or requirements that these research works address are as shown in Table I.

5. Discussion
Analysis of the evaluation in Table I and the discussion of the literature, it can be deduced that progress is being made on building digital forensic systems capable of being deployed on cloud environments. More work still needs to be done on digital forensic systems aimed for the cloud as well as on systems that support digital forensic readiness. A larger amount of work still had be done on standardising digital forensic processes as this aspect is the least supported in Table I. Live forensics as well still requires more attention from researchers and implementers of digital forensic systems.

6. Conclusion and Future Work
In this paper we have presented a criteria that can be used in evaluating architectures and systems that are aimed at investigating cloud environments. Published literature was analysed based on this criteria. Results from the analysis revealed that more work still needs to be done in standardising digital forensic process and more specifically, for cloud environments.

Since the European Union has restrictions on data migration to countries that are perceived to have weaker privacy policies [37], this can become a challenge when a multijurisdictional investigation has to be conducted which often encountered in cloud based investigations. On the other hand, the use of Internet and eventually, cloud computing is growing at an alarming rate in Africa. This presents an opportunity for researchers to develop digital forensic standards that can facilitate collaborations between African and European cloud service providers and multijurisdictional investigators.

Digital forensic standards tailored for cloud environments were found to be receiving less attention from researchers as can be seen in Table 1. There is therefore a need for more research endeavours in this direction.
References


