A State-of-the-Art-Review of Cloud Forensics by Sameera Almulla, Youssef Iraqi, and Andrew Jones from The Journal of Digital Forensics, Security and Law is available under a Creative Commons 4.0 International license. © 2006-2015 Association of Digital Forensics, Security and Law. UMGC has modified this work and it is available under the original license.
2. OVERVIEW OF CLOUD COMPUTING AND DIGITAL FORENSICS

2.1 Cloud Computing

Cloud computing has a number of definitions. The National Institute of Standards and Technology (NIST) defines cloud computing as:

A model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction (NIST, 2014b).

The uniqueness of the cloud environment is a result of its characteristics. In order to analyse and study the opportunities and challenges presented by digital forensics in cloud computing, it is essential to understand these characteristics.

2.1.1 Characteristics

As identified by NIST (NIST, 2014b) there are five characteristics for cloud computing, namely, on demand self-service, ubiquitous network access, resource pooling, rapid elasticity, and pay-per-use business model. Acquiring digital evidence of a suspect whose data is stored in a shared, multi-tenant and elastic pool of resources may result in privacy violation of legitimate users data. Hence, an enhanced digital forensics model is required to collect forensically sound evidence for cloud based incidents. In order to measure the severity of the impact of cloud computing on digital forensics, there must be a clear understanding of the cloud computing service and deployment models (NIST, 2014b).

2.1.2 Service and Deployment Models

Service models are categorised by the type of computing resources provided to the end users.

- Software as a Service (SaaS), applications are delivered as a service over the Internet e.g., Google Mail.
- Platform as a Service (PaaS), the development platform is provided as a service e.g., Microsoft Azure.
- Infrastructure as a Service (IaaS), the server(s), storage and hardware are delivered as a service e.g., Amazon Simple Storage Service (S3).
Deployment models There are four deployment models for cloud computing, namely public, private, community, and hybrid cloud (NIST, 2014b). The deployment models differ based on the users control on the computing resources and their location. For example, public cloud is owned by the Cloud Service Providers (CSPs) and its infrastructure is located within CSP premises. However, in the case of private cloud, the computing infrastructure is owned by the user (e.g., a federal or private company) and it is located within its premises. In each deployment model, cloud services can be provided as SaaS, PaaS and IaaS. In a hybrid cloud, the infrastructure consists of a combination of the private and public models. In the community cloud model, computing resources (server or network) are shared between several organizations of similar interests, needs and requirements.

2.2 Digital Forensics

NIST (2014a) defined digital forensics as

The application of science to the identification, collection, examination and analysis of data while preserving the integrity of the information and maintaining a strict chain of custody of data.

There is no single process model that can be followed to obtain evidence. As stated by Ruan (2013), there are about four standards for the forensics process, namely Digital Forensics Research Workshop (DFRW), National Institute of Justice (NIJ), National Institute of Standards and Technology (NIST) and Integrated Digital Investigation Process (IDIP). The main difference between the first three and IDIP is the integration of digital forensics process with the physical forensics. Next is a discussion of each process in terms of its definition and how the presence of cloud computing might impact each procedure.

2.2.1 Process

Cloud computing is a dynamic service oriented technology. It creates challenges to the applicability of existing digital forensics procedures. The NIJ process model encapsulates the DFRWS and NIST process models as well as the IDIP. Hence, in our review, the previous works are categorised based on the NIJ forensics process as follows:

Identification Determining the type of crime, software and hardware used by the suspect and possible evidence locations. In a cloud computing environment, identifying the digital forensic requirements to conduct a sound investigation is considered to be the main building block in the process of identification.

Preservation Ensuring evidence integrity by preserving the integrity of the original data. However, in a cloud environment, the challenge is how to preserve the data and then determining whether the existing approaches of measuring data integrity (e.g., using hash functions) are applicable or not.

Collection Extracting the exact bit-by-bit image of the required data. Most of the literature reviews emphasise that collecting the whole target environment might not be possible in the cloud environment. This is due to the fact that the infrastructure is outsourced and owned by the CSP. Also, the variations of cloud service models present a whole new set of challenges on evidence collection.

Examination Studying the collected data and its attributes. Current computer forensics practices examine well-structured storage e.g., hard disks; however, in cloud computing a significant proportion of the target data may be held in memory/network dumps and/or log files.

Analysis An in-depth systematic evidence search is performed on suspect owned devices in two ways: live and/or static systems analysis. To perform the analysis, many tools and applications such as EnCase (EnCase, 2014) and Forensic Tool Kit (FTK) (FTK, 2014) can be used to aid the investigators. In a cloud environment, the analyst must consider the dependencies of a cloud based application either on the service provided within the CSP bound-
In the case where a complete chain of custody is not possible, investigators need to be able to perform analysis on the partial resources in hand (Almulla, Iraqi, & Jones, 2013; Dykstra & Sherman, 2011). Also, there is the need for a digital forensics tool capable of acquiring and analysing cloud-based cases e.g., FROST (Dykstra & Sherman, 2013).

Presentation

The findings will be presented to either the management of an organization or a court of law.

2.2.2 Types

Based on the target evidence media, there are two main types of digital forensics;

**Static forensic** This is the process of obtaining a bit-by-bit copy of powered off digital media. Attaching the media to the forensic machine via a write blocker (before starting the imaging process) preserves the integrity of the original data. In spite of its strengths, there are several limitations of static forensics such as the failure to capture information stored in the Random Access Memory (RAM), which may include encryption keys and network related data.

**Live forensic** This is a process of collecting volatile network/user related evidence.

Live forensics is becoming increasingly important due to the increase in the size of RAM and the increase in the use of data encryption. RAM might contain valuable credentials such as usernames, passwords and encryption keys.

Another important aspect of live forensics is to define whether the logged on account is in a real or virtual environment. The latter requires further analysis such as the imaging of both the real operating environment and any virtual machines located on the real system.

The main issue with live forensics is that, by virtue of its transient and temporal nature, it will not normally be possible to reproduce the results and as a consequence the reliability of the produced evidence may be questionable.

Most of the literature considers live forensics to be the best fit for cloud forensics. In fact, it is strongly dependent on the cloud service model, CSP cooperation and the target suspect category (client or CSP). In the case of IaaS e.g., Amazon EBS (AmazonEBS, 2014), static forensics may be a more suitable choice. On the other hand, in PaaS or SaaS, live forensics will best suit the situation. However, the feasibility of data access for investigators in the case of the SaaS and PaaS model will differ significantly based on whether the CSP was the victim and based on its Service Level Agreement (SLA) support to digital forensics for incidents.
REFERENCES


Chung, H., Park, J., Lee, S., & Kang, C. (2012). Digital forensic investigation of cloud stor-
This work is licensed under a Creative Commons Attribution 4.0 International License.
(GC Wkshps) (pp. 775–780).


