Research on Live Forensics in Cloud Environment

Yunting Lei, Yuyin Cui
Key Laboratory of Information Network Security, Ministry of Public Security, People's Republic of China
The Third Research Institute of Ministry of Public Security
Shanghai, China
leiyunting@stars.org.cn, cuiyuyin@stars.org.cn

***Abstract*—Nowadays, with the widely use of cloud computing, the traditional forensics is relatively backward. The new forensic model and skills are needed in the new environment. On the basis of cloud computing, we research on the related technology, infrastructure and legal issues of live forensics in cloud environment and the data sources in cloud forensics are presented in the paper. Our mainly research is focus on providing the basic four processes according to the ISO/IEC 17043 and the latest methodologies of live forensics in cloud environment. At last, we proposed several measures to be further in-depth study in the future.**

Keywords- cloud forensics; live forensics;ISO/IEC 17043

I. GENERAL ISSUES OF LIVE FORENSICS IN CLOUD ENVIRONMENT

Due to the security risks in cloud environment, crime aiming at cloud computing take place frequently which increase the difficulty of digital forensic. The challenges facing by cloud forensics are in the aspects of technology, infrastructure and legal.

A. Technology Issues

The technology issue is mainly in relation to the devices and procedure applied in cloud forensics.

- Data Collection

Data collection in Cloud forensics is to identify, mark, record and access to electronic data from possible data sources in a cloud-computing environment. Across jurisdictions, the data in the cloud is no longer stored in a defined physical node, but there are service providers dynamically provide storage space, which may be located in different regions or countries, thus gathering evidence of a crime becomes quite difficult. Data in the cloud with highly complex and cross-contamination of the data will present challenges to the primitiveness, integrity and effectiveness of the data. Further, according to the volatile of the data in the cloud, the data collection order is important, volatile data need to be collected first.

- Data Analysis

Due to the data features in cloud environment, we need to develop a mechanism to describe the time, spatial extent and correlation analysis of the collected data. The data obtained in the cloud environment must be acquired within a certain range in time or space. In accordance with the requirements of forensics, the semantics of the data and the data itself must have the same level of trust. Data collected in cloud environment, whose formats and descriptions are sometimes specific, and some of which are sensitive transient data found in distributed systems. At this time, access to the context of the data and the system status is crucial important.

B. Infrastructure Issues

In the cloud forensics process, two subjects need to be covered at least, cloud service provider and client, which are important parts of the infrastructure of cloud computing in fact. Cloud service providers and cloud computing applications mostly rely on other cloud service providers and cloud computing applications, thus forming a dependency chain, in which dependencies is dynamic between the service provider and the client. In this case, the forensic investigation largely depends on the number of chains and the complexity of the dependency chain. Any broken links or chain coordination absence of any body will affect forensics and forensic results.

C. Legal Issues

In digital forensics, multi-jurisdiction and multi-tenant brings many legal issues. The situation is even more serious in the cloud forensics. Digital forensics required by law that must not violate the laws of the local jurisdiction and cannot be a threat to the data security of other tenant. There will be a lot of problem needs to be further explored, analyze and solve in the area of digital forensics in cloud environment. As technology continues to progress, researchers will certainly face to more and newer cloud forensics problems.

II. DATA SOURCES IN CLOUD FORENSICS

The first step in digital forensics is data collection, which starts with data sources determination. According to the distribution of data storage in cloud environment, the data sources to be selected are defined as follows.

A. Cloud Computing Developer

In cloud environment, the user's actions are mainly completed online directly, such as a variety of online document editing. Under the premise of the judicial authority, node data can be directly obtained from the cloud computing developing company, base on which the examiners can examine the involved using traces and evidence of illegal behavior under the network environment.

B. Client

Clients mainly store lots of dynamic information, such as memory information, process status, buffer information, documents and other debris. When the suspects take criminal activities in the cloud, and its local client is bound to be left some residual data which examiners can find relevant evidence.

C. Cloud Service Provider

Cloud service providers provide sub-cloud (sub-services)in the cloud providing by cloud computing developer. In some cases， the importance or the quantity of data cannot reach the requirement to investigate cloud computing developer, cloud service provider investigation can be an alternative method. Moreover, cloud service providers coverage is relatively small in terms of geographical, which benefits evidence collection.

III. PROCESSES OF LIVE FORENSICS IN CLOUD ENVIRONMENT

The four basic processes of live forensics are identifying, obtaining and preserving potential digital evidence, as showed in Figure. 1.

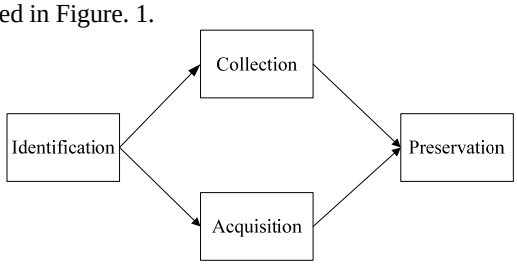

Figure 1. basic processes of live forensics

A. Identification

The forensics process starts with identifying which data may be or may contain potential digital evidence. Formally, identification is the “process involving the search for, recognition and documentation of potential digital evidence” [4].In cloud forensics, the examiners need to identify the type of the cloud services specifically:

SaaS - technical controls or monitors implemented on networks and computers allocated to cloud customers[3] (application level logs like access logs, authorization errors, transaction logs, usage logs, account information, performance issues, data volumes,etc).

PaaS - application specific logs available ideally via an API, patch status, authentication errors, operating system exceptions and warnings, anti-malware software warnings, etc.

IaaS - system level logs, Infrastructure: hypervisor events and logs, raw virtual machine files suspend files capturing unencrypted RAM snapshots, Intrusion detection and firewall events, network events and packet capture, storage logs, backups, etc.

B. Collection and Acquisition

Collection is the process in the digital evidence handling process where devices that may contain digital evidence are removed from their original location to a laboratory or another controlled environment for later acquisition and analysis.[4]

Acquisition is the process of creating a copy of an item of potential digital evidence. Because of the multi-tenant character of cloud infrastructures, examiners usually preferred acquisition over collection to avoid impacts to parties not involved in the case and the gathering of irrelevant information that must be excluded during analysis. However, the specifics of the legal mandate in a particular situation must be followed. It must be emphasized that collection of digital evidence can often only be performed by the CSP and not by the tenant.

Because of the virtual nature of the cloud infrastructure, items normally thought of as physical (hard drives, server memory, etc.) will be logical items (a virtual hard disk file, a file that contains the contents of server memory for a suspended virtual machine, etc.) and acquisition must focus on these logical items rather than the physical containers where they reside.

C. Preservation

Preservation is the protection of the integrity of potential digital evidence. Potential digital evidence and digital devices must be protected from tampering or spoliation. No changes for cloud environments. However, the chain of custody must be preserved as well, which is challenging in multi-geographical and multi-jurisdictional environments.

D. Differences between Cloud Forensics and Traditional Forensics

Cloud computing differs from traditional forensics process flow as follows:

- To determine the purpose of evidence, depending on different case features, there will always be its own evidentiary purposes.
- Confirm the types of cloud services, cloud service types include cloud equipment, cloud software and cloud platforms. Different types of cloud services in forensics face different opportunities and challenges. In the cloud devices, client information is likely to exist in the terminal short; the traditional forensic methods may still apply. However, in the cloud device, you can not live volatile data forensics and access, multi-tenant storage device may contaminate the collected data, different tenants logging may be co-located or scattered in a number of constantly changing device, collecting data may require significant bandwidth and time, data fragmentation and dispersion and ownership issues affect data collection. Software in the cloud, a lot of information can be used for forensic investigations, including application and verification records, most likely in the terminal transient client information, suppliers of terminal equipment including basic access information. Cloud software functionality

may be able to assist in network forensics. However, the software in the cloud, the traditional method is almost impossible to capture, logging, and access depends on the details of the cloud service provider. Cloud software may be provided by different service providers, in addition, the cloud service provider's procedures can be complex, it will be difficult or even impossible to analyze. Mutual isolation of their application process, the system is more proprietary nature. In the cloud platform, the client can obtain evidence for a web server or virtual operating system-related evidence-gathering techniques. However, relying on the cloud service provider logging of the environment, the system will also be more proprietary nature.

- Confirm the underlying data sauce, including cloud computing developer, the client and the cloud service provider as mentioned before.
- Based on the sauces, carry out a specific type of forensic work. For the cloud computing developer, it is needed to collect complete evidence, confirm the tools that used to upload data, confirm the consistency of the time with the cloud service provider. For cloud service providers, first communicate with the cloud service provider to preserve the data and collect logs and evidence from the cloud service provider. For customer service, it is needed to confirm the role of the user, use forensics tools to collect client's activities.

IV. Methodology of Live Forensics In Cloud Environment

Cloud computing platform is characterized as building by a large number of distributed heterogeneous virtual computing resources, these complex structures bring enormous challenges to carry out the work of computer forensics.

Here presents a new cloud computing environment model of cloud computer forensics.[5]Based on the definitions of models of cloud computing, and evidence of strong isolation and integrity of models of cloud computing forensic, a virtual machine image file and can be objects of forensic analysis to achieve the computer forensics process in virtualized environments.

For the analysis of the characteristics of cloud computing environments and the CCFM (cloud computing forensic model) is a new defined cloud computing forensic model. This model through the needs analysis of the process of forensic environment to obtain the scene description and process components partition in cloud environments; through the prove of completeness and constraints of cloud computing model to present that this cloud computing model can be used as the key way of cloud forensic; use the virtual machine image file as the objects of forensic analysis to implement the entire forensics process.

Though the virtualization software layer control in the cloud, use the state transitions, proposed a kind of virtual machine image file migration method. The method can preserve the state of the virtual machine system completely by preservation and reconstruction of the Process Identification of VM when the virtualization software layer migrating, memory mapping, network connection and file system information, and by combining and feedback mechanisms, design data migration system architecture, making the data migration process does not affect the service request of entire cloud computing platform. The migration process can use different strategies over the entire migration, but also can be adjust according to the working status of the migration. By localization image loading, the virtual machine image migrates from a cloud-computing platform to the local forensics environment to analysis, implement computer forensics process in virtualization platform.

V. Further study on Cloud Forensics countermeasures

First, national laws and regulations and the judicial system are needed, as well as the establishment of cooperation with the cloud service provider, such as monitoring of the process in the business, access to system information during the operation and preserved to forensic analysis. Secondly, although the forensic method is still relatively backward, a variety of forensic tools and forensic technology have achieved some progress. Here are some related technologies and methods.

- Service replace is to reproduce the cloud service requests, to restore the unauthorized access that has occurred, and look for possible evidence material.
- Data migration technology is the core technology of the storage system, and is more important in cloud storage. As the prevalence of cloud computing, service providers provide the services that transfer the traditional manual data to cloud-based data backup and recovery. Forensics process may be involved in sampling the monitoring of data migration in order to obtain relevant evidence.
- A comprehensive analysis of intrusion detection system can detect the interactive data between the client and cloud resources, once found the information matching the intrusion pattern library, the system will respond and deal with it. Intrusion detection cannot only make a warning response when danger is detected, but also has a set of security mechanisms to protect client data. It can effectively prevent and reduce internal intrusion and malware harm.

It is very difficult to present the evidence collected from a cloud-computing environment in court, mainly due to the complexity of the data and the data is difficult to be original and integrity. Since the original data differences and omissions, exact reconstruction event is impossible, but use pattern matching and other statistical analysis tools to determine the data set and to infer the actions based on tracking data, such as data fusion technology. Overall, with the popularity of cloud computing, forensics techniques will continue to progress and development, but it is relatively backward now, we need to invest a lot of work.

VI. CONCLUSION

Cloud computing is now rapid development and widely used, but its security issue bring challenges to the forensic. In this paper, on the basis of cloud computing, the related technology, infrastructure and legal issues of live forensics in cloud environment are presented and the data sources in cloud forensics are mentioned. We focus on providing the basic processes and methodologies of live forensics in cloud environment and measures to be further in-depth study.

ACKNOWLEDGEMENT

This work was financially supported by the basic science project of Ministry of public security (project number: 2012GABJC035) and the project of National Development and Reform Commission (project number: [2012]1424).

REFERENCE

[1] Q.Ding andG.Sun, Research on forensics in cloud environment, Netinfo Security, pp36-38

[2] H.Guo, T.Shang and B.Jin,Forensic Investigations in Cloud Environments

[3] MappingtheForensicStandardISO/IEC27037toCloudComputing, Cloud Security Alliance, 2013

[4] ISO 27037, Guidelines for identification, collection, acquisition and preservation of digital evidence, Available athttp://www.iso.org/iso/catalogue_detail?csnumber=44381, 2012 [accessed 22 April 2013].

[5] G.Zhou，Research on Data migration technology for forensics in cloud environment