Research on Digital Forensics Based on Private Cloud Computing

Gang Zeng
Dept. of Police Information Technology,
Liaoning Police Academy,
Liaoning, China
dlzenggang@126.com

Abstract :

With development of network and digital devices, traditional digital forensics tools show their drawbacks, and investigators need new forensics tools to deal with enormous digital evidences. Therefore, this paper introduces digital forensics and cloud computing, then lists the advantages of private forensics cloud computing, proposes a structure of Forensics as a Service(FaaS) and the architecture of private forensics clouds, discusses private cloud for security and forensics procedures in FaaS model.

Keywords: *Digital forensics; Cloud Computing; private cloud; security issue*

1. INTRODUCTION

The term digital forensics, defined in wikipedia.org, is a branch of forensic science encompassing the recovery and investigation of material found in digital devices, often in relation to computer crime. It was originally used as a synonym for computer forensics but has been expanded to cover investigation of all devices capable of storing digital data. Therefore, it is called digital forensics, not called computer forensics any more[1].

A digital forensics investigation commonly consists of 4 procedures: identification, evidence collection, analysis, and reporting. Acquisition involves creating an exact sector level duplicate of the media, often using a write blocking device to prevent modification of the original data. Both acquired image and original media are hashed (using SHA-1 or MD5) and the values are compared to verify the copy's accuracy.

During analysis procedure, an investigator recovers evidence material using a number of different methodologies and tools, searching the evidence related to the suspected crime thoroughly. The actual process of analysis can vary between investigations, but common methodologies include conducting keyword searches across the digital media, recovering deleted files and extraction of registry information (to list user accounts, or attach USB devices, for example).

The recovered evidence is analyzed to reconstruct events or actions, and to reach conclusions. When an investigation is completed , the data is presented, usually in the form of a written report in lay persons' terms.

During the forensics procedure, Mass storage device and high speed processing are needed. Enormous digital evidences will require much more processing time and resources to process them. Generally speaking, an image of the evidence is more than 1 TB, while imaging of 1TB data takes 3.5 hours, and the search on 1TB forensic image data takes about 14 hours. The search is being processed at a rate of 20MB/S, and the speed can not be accepted[2][3].

Another major limitation to a forensic investigation is the use of encryption; decryption of the keywords encrypted can't be carried out, for we haven't a powerful computing device.

If we use the technology of cloud computing, we may resolve these problems.

2. Cloud Computing

Cloud computing is defined by the National Institute of Standards and Technology (NIST): it is 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[4].

2.1 new characteristics of Cloud Computing

(1) Security and Reliability

The data saved in cloud is secure, and you don't worry about their security. Many people think that the data in their own computer is secure, but in fact, the data is not secure at all. The data may be lost because of virus or incorrect operation. In cloud computing, many corporations have multiple data center; they back up the data individually. So cloud computing is used as disaster recovery solution.

(2) High Availability

Cloud computing has a shared resources pool, consisting of servers, caches, cloud storage, mass data analysis, cluster management, server virtualization, applications, distributed file systems, distributed databases，etc. The user can share information with other users, make use of mass cloud storage, powerful processing ability, and software, etc.

(3) Low System Requirements in Client

The user can connect to the cloud by a web browser or the dedicated client, and the system requirements in client are very low. Both powerful processing station and maintenance in server are no longer needed.

2.2 Service Models of Cloud Computing

Cloud computing is divided into three classes according to fundamental models:

(1) Infrastructure as a service (IaaS)

IaaS is a basic service model in which cloud providers often offer virtual machines and other resources. The virtual machines are run as guests by a hypervisor, such as Xen or KVM. Management of pools of hypervisors by the cloud operational support system leads to the ability to scale to support a large number of virtual machines. Other resources in IaaS clouds include images in a virtual machine image library, firewalls, load balancers, virtual local area networks (VLANs), and software bundles. IaaS cloud providers supply these resources on demand from their large pools installed in data centers. To deploy their applications, cloud users then install operating system images on the machines as well as their application software.

(2) Platform as a service (PaaS)

In the PaaS model, cloud providers deliver a computing platform typically including operating system, programming language execution environment, database, and web server. Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers. With some PaaS offers, the underlying computer and storage resources scale automatically to match application demand so that cloud user does not have to allocate resources manually.

(3) Software as a service (SaaS)

In this model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients. The cloud users do not manage the cloud

infrastructure and platform on which the application is running. This eliminates the need to install and run the application on the cloud user's own computers, which simplifies maintenance and support. What makes a cloud application different from other applications is its elasticity. This can be achieved by cloning tasks onto multiple virtual machines at run-time to meet the changing work demand. Load balancers distribute the work over the set of virtual machines. This process is inconspicuous to the cloud user who sees only a single access point. To accommodate a large number of cloud users, cloud applications can be multitenant, that is, any machine serves more than one cloud user organization. Commonly it refers to special types of cloud based on application software with a similar naming convention: desktop as a service, business process as a service, Test Environment as a service, communication as a service.

2.3 Deployment Models of Cloud Computing

(1) Public cloud: Public cloud applications, storage, and other resources are made available to the general public by a service provider. These services are free or offered on a pay-per-use model. Generally, public cloud service providers like Amazon AWS, Microsoft and Google own and operate the infrastructure and offer access only via Internet (direct connectivity is not offered).

(2) Community cloud: Community cloud shares infrastructure between several organizations from a specific community with common concerns (security, compliance, jurisdiction, etc.), whether it is managed internally or by a third-party and hosted internally or externally. The costs will be shared among fewer users than a public cloud (but more than a private cloud), so only some of the cost savings potential of cloud computing are realized.

(3) Private cloud: Private cloud is cloud infrastructure operated solely for a single organization, whether managed internally or by a third-party and hosted internally or externally

(4) Hybrid cloud: The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability.

3. Conversion of Traditional Digital Forensics into Cloud Computing Service

3.1 Limitations of the Traditional Digital Forensics

(1) limitation of forensic model: the digital forensics is usually after the case, then the digital evidence is processed, and forensics is static. Investigators produce an image from digital devices, then analyze evidence thoroughly by using tools such as Encase or FTK, trying their best to recover the original scene. In this model some evidences can not be acquired. As a result, dynamic forensics is often needed, but dynamic forensics is not secure: it maybe modify the original evidence, or it may be harmful to the system.

(2) limitation of hardware performance: Traditional digital forensics runs on a single platform, and its performance depends on the hardware performance (CPU, memory, storage, etc.) of the single platform. Traditional digital forensics tools are designed based on mono-tasking. This means only one task can be processed when many digital evidences are going to be processed [6].

(3) more maintenances of traditional digital forensic tools: The Errors of traditional digital forensic tools are often reported, the users must download, reinstall the software, and modify the environment of the system[5].

(4) limitation of the position of the forensic laboratory: Some digital evidence need to be processed immediately, but forensic tools must be used in laboratory, it is a long time for transfering the evidence to laboratory[5].

3.2 Advantages of Digital Forensics Based on Cloud Computing

(1) Real-time dynamic forensics

In cloud computing the digital evidence is dynamic, which consists of local host's evidence and network evidence. Local host's evidence include information in memory, information in caches, read/write the files, and all these evidences are very important in the chain of evidence. Network evidences include network protocol, the size of digital package, the position of computing entity, port, connecting time, etc. The information can not be acquired in static forensic model[7].

(2) resources integration

All computing resources are managed as a whole in cloud computing, and the investigators use the resources in the shared resources pool when they need them, they can carry out distributed, parallel analysis tasks[8].

(3) no limitation to position of the client.

The investigators can use the forensic application at any node accessing forensic cloud computing in any place.

(4) low requirement for the client

No need to install forensic tools, the users can access the cloud by browser or dedicated client software, which avoids the reconstruction of forensic system.

(5) one or multiple tasks together

The users can carry out one or multiple tasks using all the resources in the forensic cloud in the following mode: one person, one task; multiple persons, multiple tasks; one person, multiple tasks.

3.3 Public Cloud or Private Cloud

According to the definition of NIST, cloud computing has four development models: private cloud, public cloud, community cloud, and Hybrid cloud as well. The discussion was held among the scholars, some scholar support public cloud, others support private cloud. What's the difference between them? Let’s answer these questions below.

- Who use the forensic cloud computing?
- Where are the digital evidences stored?
- How is the data being protected?
- Who access or view the data?
- How is the facility secured both physically and digitally?
- How often do they conduct routine checks and audit access to restricted areas by the system administrators?
- Which model of cloud-source is used?
- What industry standards do your services met?[9]

To answer these questions, it is necessary to explain the meaning of cloud computing security including privacy protection, safeguarding the security, whether it be national in case of government, or internal in the case of companies. The aim of digital forensics is to prevent computer crimes, and this task in many countries is carried out by police and prosecutor, or a third party who are usually civil servants of government.

After evidences being acquired, the digital evidence is often stored in a safety place which can not be modified, but remained in original states. The sensitive data can not be exposed to unnecessary risks. Security issue has been widely discussed among industries and governments. In contrast to the public cloud, the private cloud is secure, for the private cloud provides the exact location on which the data is being stored. The private clouds are managed by the organization itself or outsourced. Specific users tend to use private clouds behind their own firewalls.

The selection of the cloud-sourcing model is very important. Private government cloud computing is a part of cloud computing, and the private government cloud for digital forensics has two forms according to deployment model: private government cloud and private commercially-hosted cloud. Each of them has a specific security level. Private government cloud is established by government for digital forensics, running in the government dedicated intranet. We consider that private commercially-hosted cloud is suited to government, and private government cloud can be outsourced and managed by commercial companies.

3.4 Architecture of Private Forensic Clouds

Private forensic cloud usually runs in government dedicated intranet. The dedicated intranet has been set up in many countries. It has different name in different countries, and the private dedicated intranet is called Golden Shield network in China.

The Golden Shield network is one of the achievements of Golden Shield projection. The network is secure for its physical isolation from other networks, which covers the whole of China where it is very suitable to build a private forensic cloud. We design the architecture of the forensic cloud computing which has three layers according to builders: national cloud, provincial (state) cloud, civic cloud. The Ministry of Public Security builds national cloud. It is responsible for the formulation of security strategy, task scheduling and resource allocation, and development of application. The Public Security Departments set up provincial cloud, and the public security bureaux set up civic cloud. The structure is conducive to load balancing and distribution of work. Users can access the private forensic cloud anytime, anywhere. The architecture of private forensics cloud is shown in “Fig 1”.

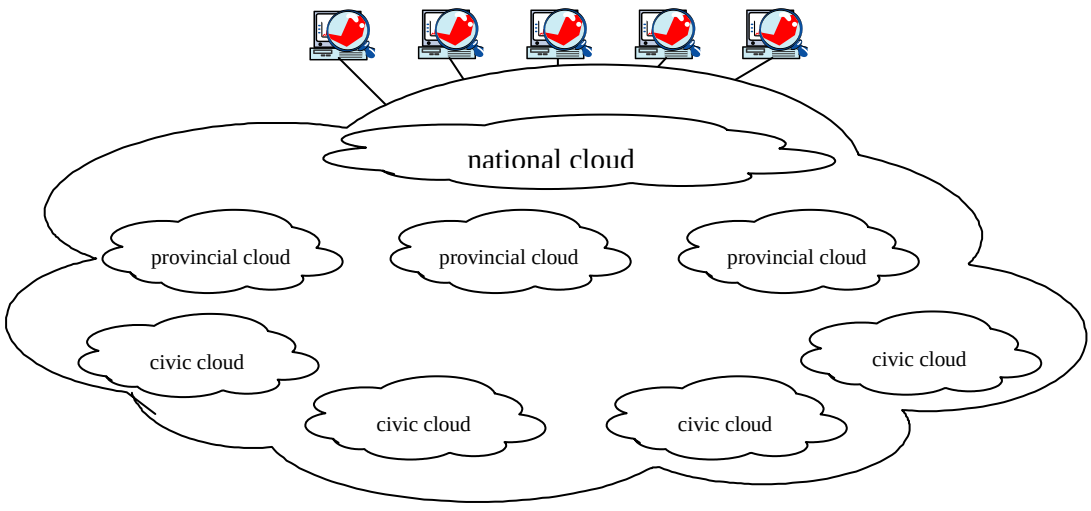

Figure 1. construction of private forensics cloud

4. Forensic Cloud Computing

Forensic cloud computing may be mainly set up in two model: IaaS and SaaS. PaaS model is rarely used, because of the limitation of research and development capacity. IaaS is a simple model relatively in which investigators can use the traditional forensic tools. SaaS is the development direction and tendency in the future.

4.1 IaaS Model

In IaaS model, the customer uses the virtual machine provided by the CSP to install his own system on it. The system can be used like any other physical computer with a few limitations. Therefore, a lot of investigators use the virtual machine for evidence acquisition, making image of digital evidence from crime scene. The image can be used by forensic tools, such as Encase, FTK, etc. Log data and other evidence information in the image could be analyzed. Log data information include currently-logged users, open ports, running processes, system and registry information etc.

At the same time, investigators can load and run the image in a virtual machine. Snapshots provide a powerful way to freeze specific status of the virtual machine. Therefore, virtual instances can be still running, which leads to the case of live investigations, or can be turned to lead to static image analysis. It provides much more information that could be used as forensic evidence than the PaaS and SaaS models do.

In IaaS model, the cloud computing significantly reduces time for data acquisition, data copying, transfering, and analyzing, for parallelization of data processing. Digital forensic cloud computing decreases evidence acquisition time and onsite service downtime.

A dedicated forensic server can be created and kept offline until it is needed. A copy of VM can be easily distributed as new sources of evidence.

Copies between two co-resident VMs (residing in the same physical server) are made very fast. Forensic image verification is reduced if a Cloud Application generates cryptographic checksum or hash.

Cryptanalysis of data is the most complex and time-consuming operation in digital forensics, requiring sufficient CPU size and memory capacity. Typically, the computing power of cloud environment far exceeds capabilities of portable or stationary forensic computers, thus reducing the time needed for cryptanalysis and discovery of passwords[10].

4.2 FaaS Model

The Virtualization is the core technology of IaaS model, which provides storage, processing power, memory, and management of recourses for customers. The digital forensics in IaaS model decreases the forensic time by using the technology, but the application of evidence acquisition, analysis, reporting is out of date, even traditional forensics tools. The new technology for digital forensics must be developed to deal with these drawbacks of traditional forensics tools. We propose a model named private digital forensics cloud computing working in SaaS model, and we call it as FaaS(Forensic as a Service). It is flexible, elastic and dynamic platform, owning unlimited storage and processing power.

The forensic cloud consists of control center, data processing center, storage system. Control center supports Authentication of the users and task scheduling. Data processing center supports Case management, Evidence Collection, Evidence Processing/Analysis, Reporting and Visualization. Documents, analysis result, evidences data, logs are stored in the

cloud storages. The Architecture of FaaS is shown in fig. 2.

Users of FaaS consist of the role, manager, investigator and analyst who can use all the devices to access the integrated platform for evidence collection, analysis and storage through the resources pool anywhere, anytime. Forensics tasks are distributed by control center, application servers to achieve optimal resource utilization for parallelization of data processing, multi-tenant model.

Digital forensics can be carried out as follows in FaaS model:

(1) Preparation: After identified by fingerprint or other biological features, the investigator can create a new case. The information of the case is recorded in case storage of the storage system.

(2) Evidence collection: In the crime scene, the investigators use the collection service of FaaS to make a forensic image of a disk or other medias, then store the image in evidence storage.

(3) Data processing/analysis: the processing/analysis service of FaaS is used for static forensics and dynamic forensics. The investigators make two copies of the original image, then a static forensic on one copy is carried out, inclusive of key word search, decryption, etc. then the other copy of the image is loaded in a virtual machine, and a dynamic forensic is carried out in the virtual machine instance. The result of analysis is stored in result storage of the storage system.

(4) Reporting/Presentation: reporting service of FaaS is used to produce a report, and the report is stored in report storage, then presented in court.

During all the procedures, logs are recorded for reliability. Logs are stored in log storage of the storage system. Of course, legal issues of digital forensics are still very weak, and we must do more work to improve it.

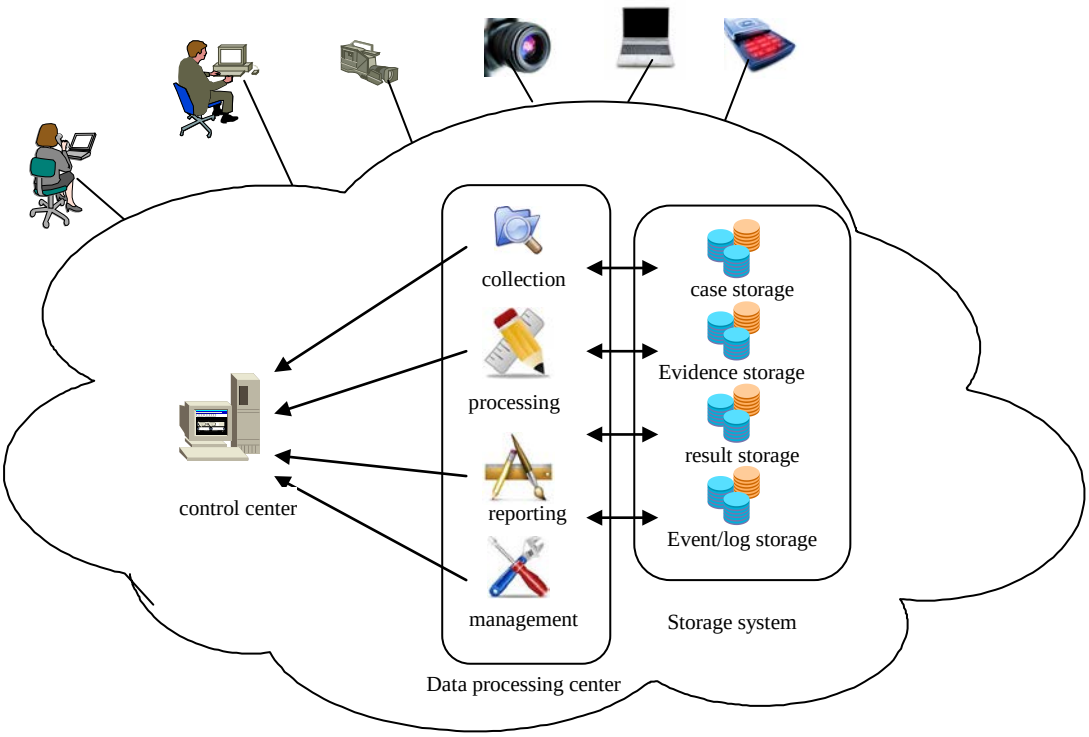

Figure 2. The Architecture of FaaS

5 SUMMARY

In this paper, we introduced digital forensics and cloud computing first, then listed the advantages of private forensics cloud computing, discussed emphatically private cloud for security, proposed the architecture of private forensics clouds. Finally we proposed forensics procedures of Forensics as a Service(FaaS).

As a further research, many issues have to be studied, for example, patterns of use, degrees of reliability, privacy and other security, etc.

REFERENCES

[1] digital forensics. Retrieved July 6, 2012, from http://en.wikipedia.org/wiki/Digital_forensics.

[2] S.L, Garfinkel, Digital forensics research : The Next 10 years, Digit. Investig.(2010), doi: 40.4016 /j.diin.2010.06.009

[3] Jooyoung Lee, Sungkyong Un, Dowon Hong, "New Paradigm of Digital forensics: Forensic Cloud" Digital forensic Technology Workshop, pp.101-107, 2010. 8

[4] NIST. DRAFT Cloud Computing Synopsis and Recommendations. NIST Special Publication 800-146.

[5] Bon Min Koo, Tae Rim Lee, Hun Kim. A Study on Digital Forensic Software as a Service on Cloud Computing. International Conference on Internet & Cloud Computing Technology (ICICCT) – Singapore 2012.

[6] Accenture. Cloudrise: Rewards and Risks at the Dawn of Cloud Computing. High Performance Institute. 2011.

[7] ZANG Jun, MAI Yong-hao. Cloud Computing Environment Simulation Computer Forensics. Netinfo Security. 2011(10):7-9,12.

[8] DAI Shi-jian, SHEN Yi-wen. Cloud Computing Mode Network Crime Detection and Forensics Research. Netinfo Security. 2011(04):32-34.

[9] Dener Didone, Ruy J. G. B. de Queiroz. "Forensic as a Service - FaaS". http://dx.doi.org/ 10.5769 /C2011024

[10] Ludwig Slusky, Parviz Partow-Navid, Mohit Doshi. Cloud computing and computer forensics for business applications. Journal of Technology Research.