Design and Implementation of a Cloud Digital Forensic Laboratory

Yi-Hsiung Ting * Chung-Huang Yang †

Abstract: With the continuous increase of the network bandwidth and the quality improvement of the connection, cloud computing provides a stable network environment and an innovative thinking platform. Through the high-speed and stable network platform, applications can respond to dynamic demand and require the cloud resources quickly. Digital forensics including obtain, preserve, analyze, and document digital evidence in a court of law. Analyzing digital evidence needs a fundamental network infrastructure and large capacity computing power, so we design and implement a digital forensic laboratory (DFL) based on cloud computing platform. The proposed system generates forensic report automatically and it not only provides a centralized storage for digital evidence but also performs multiple forensic tools for analyzing evidence. We use the proposed DFL to support the compute and storage needs.

Keywords: Cloud Computing, Digital Evidence, Computer Forensic, Cybercrimes

1 Introduction

Due to the rapid development of information technology, information community has become an important application on the internet. And information technology brings people convenience in their daily lives and work. On the other hand, the cyberspace also becomes a hotbed for the growth of crime while cybercrimes are formed quickly in cyberspace. As a result, cybercrimes investigate not only become very complex but also very difficult. Digital forensics including obtain, preserve, analyze, and document digital evidence in a court of law. Investigating officers should be retained or detained digital equipment in the crime scene when a cybercrime occurred, then they collect digital evidences and analyze them. Besides, we also establish a forensic report that can be treated as evidence in court. DFL provides us a large storage media as a centralized database of digital evidence, while giving us an efficient digital evidence analysis platform. It provides a thematic and individual digital evidence collection, which digital evidence can be given forensic officers to assist in cybercrime occurred after investigation section.

We implement a systematic approach to collecting, preservation, analysis and presentation of digital evidence. DFL answer the basic questions of a cybercrime reconstruction, such as what happened, where, when, how, who was involved, and why. We focus on the establishment of a cloud computing platform-based digital forensics laboratory. Forensic tools for windows, Linux, and Android platform are collected, and then they are saved to a centralized digital forensics lab. Forensic officers can download these latest forensic tools to collect digital evidence from digital forensics laboratory. And then analyze evidence and generate forensic report automatically.

This research focuses on implement the proposed DFL based on cloud computing platform. Besides, it provides a centralize storage for digital evidences and performs forensic tools for analyzing evidence. It can also provide tools for android device, and it should be helpful for the rapid growth of smart phones.

* Kaohsiung Normal University, No.116, Heping 1st Rd., Lingya Dist., Kaohsiung City 802, Taiwan (R.O.C.). jimmytine@gmail.com
† Kaohsiung Normal University, No.116, Heping 1st Rd., Lingya Dist., Kaohsiung City 802, Taiwan (R.O.C.)
2 Related Works

2.1 Digital forensics

The digital evidence is a series of binary digit numbers on transmission, or stored information files on the electronic device [6]. The file formats of digital evidence include audio, video, images, digital data and so on. Although digital evidence is not physical presence, there are still some features, such as unlimited copy, modify. And it is difficult to identify the original resource. Fig. 1 is Four-way linkage theory [10], it explains the relationship between the victims, suspects and digital evidence. Element relationships with each other be established, the greater the probability if successfully solving the cybercrime case.

![Fig. 1: Four-way linkage theory](image)

During the cybercrime investigation, the procedures must be performed according to a scientific procedure in order to have legal effect of digital evidence [13]. Digital forensics are obtaining, preserving, analyzing, and documenting digital evidence from digital devices, such as tablet PC, server, digital camera, PDA, fax machine, iPod, smart phone, and various memory storage devices [8][19]. Generally speaking, Computer forensics divide into six phases that are identify, collect, collection, acquisition, analysis and presentation [11][1], illustrated in Fig. 2.

The following is a brief description of the six phases.

1. Evidence identification: This phase is looking for possible digital evidence.
2. Evidence collection: This phase gathers physical items those contain potential digital evidences.
3. Evidence acquisition: This phase establishes a copy of the information according to the defined sets. For example, we usually use the computer forensics tool to create a disk image.
4. Evidence preservation: This phase is focus on using the methods that are reliable and verifiable.
5. Evidence analysis: This phase extracts digital information that is significant to criminal investigation.
6. Evidence presentation: The final phase is documenting the analyzing results. And it could be present the digital evidence.

Digital evidence stored on digital devices play an important role in a wide range of types of crime, including murder, extortion, computer intrusion, espionage, and child pornography in proof of a fact about what did or did not happen [8][14]. However, digital information is fragile because it can be easily modified, copied, stored or destroyed. All digital evidence will be analyzed to determine the type of information stored in digital devices. In this point, special tools are used because they could display useful information format to investigators. Most popular forensic tools, such as FTK and EnCase, are all commercial software. And the price is high for the small enterprises or individual. During the investigation, the investigating officers must ensure the digital evidence cannot be modified without proper authorization.

The typical goal is using the generally accepted method for gathering evidence to make evidence in court is accepted and recognized. And the final forensic report must include the following three items [14]:

1. Where the evidence was stored?
2. Who had obtained to the evidence?
3. What had been done to the evidence?

Digital forensics can be classified into live forensics and dead forensics [9]. A live forensics means that the suspicious system is analyzed while it is running while a dead forensics means the suspicious system is analyzed while it is shutdown. Many digital forensics research use dead-analysis, but this approach may lose the data because the machine is turned off or...
remove the connector. Volatile data collection is very important for forensic analysis. Because volatile data may include hardware information, installed software packages, as well as the state of all programs [4].

Due to the acquisition of a piece of evidence on the destination system will also affect the other evidence in the destination. In order to produce the best quality of the evidence, we will perform a binary executable file known useful hash of all the evidence, and data collected according to the order of volatile.

2.2 Cloud computing

Though cloud computing is not the novel technology, nowadays, there are still many applications. Our work platform transfer from computer to cloud. Since computer softwares, such as word, excel and so on, could be operated through the internet. So the same information can be stored in cyberspace. Besides, using services provided by computer group that is called cloud computing [15].

The NIST definition of cloud computing is enabling a ubiquitous, convenient and on-demand network access is configured computer resource sharing library model [17]. This mode consists of five basic characteristics, three service models, and four deployment models. Table 1 is NIST cloud structure.

<table>
<thead>
<tr>
<th>Table 1: NIST cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essential characteristics</td>
</tr>
<tr>
<td>Service models</td>
</tr>
<tr>
<td>Deployment models</td>
</tr>
</tbody>
</table>

2.3 Digital forensic tools

Consider the commercial version of the forensic tools are expensive, this study integrate forensic tools of open source for cloud forensic.

2.3.1 AIR (Automated Image Restore)

This study use the command "dd" and "dcffild" made graphical interface function to facilitate forensic staff to create an image, but that only support linux. Its advantage is that users may use dd or dcffild command to produce image, using MD5 or SHA-1 to authenticate image, and we can use gzip/bzip2k for compression, etc.

2.3.2 TSK (The Sleuth Kit)

TSK is open source forensic software and it can detect UNIX or Microsoft operating system files and partitions. Besides, it helps forensic officers restore files, produce image files and rootkit hidden files. The TSK's architecture can be divided into four parts: File system layer, Data layer, Meta data layer and User interface layer.

2.3.3 AFB (Autopsy Forensic Browser)

It is a frontend for the sleuth kit. Owing to it is graphical interface as well as it also provides digital investigation tools in TSK. We can easily use it and these tools allow us to investigate the file system and volumes of a computer. AFB includes File and directory browser, Sector browsing, Inode browse and Sector search.

3 The Proposed Digital Forensic Lab

Cloud computing according to the user's need to provide services automatically, without having to let users know the location of his/her data stored in the cloud platform. Cloud computing digital forensics provides an excellent and convenient platform, especially during the analysis of digital evidence. Digital forensics laboratory crawl forensic software by such a cloud host and provide a centralized location to perform forensic analysis and storage of forensic reports. The proposed DFL (Digital Forensic Laboratory) support the demand for computing and storage. Fig. 3 is DFL system structure.

When crime occurs, investigating officers will connect remotely the victim machine to the DFL site portal. DFL is based on responsive web design to automatically detect the type of machine and guide the investigating officers to download the appropriate evidence capture tool. After installing the forensic tool, digital evidence will be uploading to the DFL and then that will be using for forensic analysis and evidence present. In the DFL, multiple forensic analysis tools are simultaneously executed, and the analysis results could be detailed compare. The final forensic report is generated and the investigating officers can check
report through a web.

In our research, the victim machine is divided into two categories. One class is that the victim machine is still running while another class is the victim machine already shut down. Victim machine may be having installed Windows, Linux or Android operating system, for instance, we will discuss the proposed DFL computer forensics procedure. Fig. 4 is DFL system flow chart.

If the destination host is still operation when the investigating officer is arrival, he / she should be quick to collect volatile information, including which TCP and UDP ports are opened, user login history, services that are activated, cryptographic key on RAM and so on [12]. This volatile information will be disappearing from the victim machine while the machine is shutdown. Based on the proposed DFL, we write a script to collect volatile information. If the victim machine is still in operation, we execute the script to perform real-time digital evidence capture and collect volatile information. The investigating officers upload the evidence file to the DFL and then shutdown victim machine.

If the victim machine was shut down, then we will use self-developed bootable DVD / USB to restart the machine and make a disk image [3]. Our DVD / USB include AIR disk image tool (http://air-imager.sourceforge.net/), defldd (http://defldd.sourceforge.net/) and other forensic tools [2]. All digital evidence will be uploaded to DFL, and then we can perform analysis and generate a forensic report.

4 System Implementation

4.1 Development tools and Test environment

This study use PHP as a development language. PHP's main purpose is deal with dynamic pages and we can modify the PHP code to control page rendering while forensic officers browses. Table 2 is our development tools and test environment.

<table>
<thead>
<tr>
<th>Table 2: development environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
</tr>
<tr>
<td>Web server</td>
</tr>
<tr>
<td>Program</td>
</tr>
<tr>
<td>Database</td>
</tr>
</tbody>
</table>

4.2 System functions
We use PHP as the developing language and we implement a cloud-based digital forensics system. First of all, forensic officers must login the portal which could identify user using the list of forensic officers. Fig. 5 is the login checking procedures which can assure the identity. Before logging into the portal, we verify logger's username and password. Then we will start the session to identify operator is the same person who has been defined. By checking username and password, we can confirm their identity. If their identity is forensic officers, the system will automatically directed to the forensic page. Otherwise, it will show a login failed webpage. Fig. 6 is DFL login page.

After logging into the proposed system, web page will be redirected to the main page of the forensic system. In order to provide immediate assistance, this study has several individual page links, such as: evidence upload, tools download, forensic analysis, and report. Fig. 7 is the proposed DFL main page.

Forensic officers can select local digital evidence from web page, and then they can upload them to the cloud platform. Fig 8 is the local digital evidence uploading image.

However, In order to ensure that the investigating officer can instantly get the updated forensic tools, DFL provides the investigating officer with the latest forensic tools. Figure 9 is the image of forensic tool
After uploading the digital evidence, we can continue to analyze the digital evidence. For example, we use AFB as our analyzing interface to execute digital evidence forensic analysis. First of all, we enable the autopsy program from terminal as well as we specify evidence storage directory and connection address. Then we click the autopsy option button, using it as forensic analysis tool to analyze the digital evidence. Fig. 10 is executed autopsy program.

Due to each case will have a folder to store related files, forensic officers can easily choose file to conduct a forensic analysis. Fig. 11 is AFB web page. The image analysis can be carried out in accordance with the actual needs. Its common features include: file format, single sector analysis and Meta Data analysis. We can use AFB to display the name of the file that has been deleted, the filing time, file size etc. After completing the forensic analysis, the forensic report will be stored on this cloud digital forensics system where we can design a friendly user interface to query related forensic information.

5 Conclusion

In recent years, the numbers of cybercrime case are growing. And investigating officers must collect digital evidence of suspicious digital device after cybercrime occurred. However, most existing digital forensic software is commercial version. And they could not provide the latest forensic software to cope with the rapidly changing digital devices, such as: iPad or android tablet.

The proposed DFL is a cloud-based platform. It would fast response the cybercrime investigate. Forensic investigators could collect and store digital evidence in the DFL. And DFL is having a friendly user interface for investigators.

Acknowledgment

This study was supported in part by research grant NSC 101-2221-E-017-013 from the National Science Council of Taiwan.

References


