Digital Forensic Evidence Collection of Cloud Storage Data for Investigation

Sathishkumar Easwaramoorthy[1], Sankar Thamburasa[2], Guru Samy[3], S.Bharath Bhushan[1], Karrothu Aravind[1]

[1]Research Scholar, SITE, VIT University, Vellore, India
[2]Information security Analyst, Synophic System Private limited, Bangalore, India
[3]PG Scholar, Department of Information Technology, PSG College of Technology, Coimbatore, India

***Abstract*— In recent days Cloud services such as storage is more familiar to business and Individuals. This storage services are found as a problem to examiners and researchers in the field of forensics. There are many kind of storage services available in cloud and every service face a diverse issues in illegitimate action. The evidence identification, preservation, and collection are hard when dissimilar services are utilized by offenders. Lack of knowledge regarding location of evidence data can also affect investigation and it take more time to meet every cloud storage providers to decide where the evidence is saved within their infrastructure. In this study two popular public cloud service providers (Microsoft One Drive and Amazon cloud drive) are used to perform forensics evidence collection procedure through browser and service providers software on a Windows 7 computer. By identifying the evidence data on a client device, provide a clear idea about type of evidences are exist in machine for forensics practitioners. Possible evidence determined throughout this study include file timestamps, file hashes, client software log files, memory captures, link files and other evidences are also obtainable to different cloud service providers.**

Keywords—Digital forensics, web browser, client software, Microsoft One Drive, Amazon cloud Drive, Windows 7.

I. INTRODUCTION

Cloud is a heterogeneous Information technology capacity (platform or infrastructure or software) distributed as a service through internet in a pay based on usage policy. There are a variety of storage services available in the industry and each is classified as a service based on software, platform and infrastructure [2]. Cloud storage is a data storage technique which save data in a remote computer, in place of user hard drive integrated with Personal computer . There are a variety of storage providers and every one offer few amount of data space as free such as Icloud, Mega, One Drive, Box and Amazon cloud drive. Users can enter cloud services through different system such as install software on a user computer or employ a browser to enter the storage space.

Cloud storage Environment is hard for investigator to acquire and examine evidence data compare to traditional forensics evidence examination and acquisition process. The reason for this difficulty is because of lack of details about where the evidence data is saved and in which file format it is saved and by whom it is saved this question make the examiner to feel hard to collect evidence. In order to eliminate risk of evidence data integrity is questionable in courtroom, it is essential to possess an accurate procedure and a group of technique for performing evidence acquisition and examination in cloud service and that will also helpful to acquire evidence in future storage service provider. Furthermore, it is essential to have a contemporary interpretation of the type of evidence that remain on user computer and location of the data in the user machine while user accessing cloud data.

II. DIGITAL FORENSICS

The forensic analysis is the procedure of acquisition, preservation and examination of evidence data from computer systems, networks, and storage devices for judicial purposes, such as criminal investigations or civil cases. The digital forensic process can be divided into four different parts such as user artifacts identification and acquisition, preservation, examination, reporting and presentation as shown in the following Fig. 1[2].

A. Evidence Data Identification and Preservation

This stage is regarding the investigator identifies the location of electronic evidence data on the user machine and then preserves the evidence. In the initial part of searching, origin of evidence data determined on the user machine (e.g. Laptops, mobile equipments and desktops). In the second part of searching, it mainly focuses on the identification of storage service company details important for investigation, potential evidence stored with the storage service and process for protection of this user data.

B. Evidence Collection

During this stage, the examiner collects user evidence data from dissimilar kinds of media e.g., computer main memory, electronic mail, portable mobile devices, tablet and hard disk, etc. Additionally, forensics examiner wants to protect the integrity of the user data without being modified.

C. Examination and Analysis of Evidence

In this stage, the forensic examiner acquires and examines the evidence and its associated properties. In the next stage, the examiner understands, examine and coordinate the collected evidence in order to derive a final decision, which can be shown as evidence in civil or illegal disputes.

978-1-4673-9802-2/16/$31.00

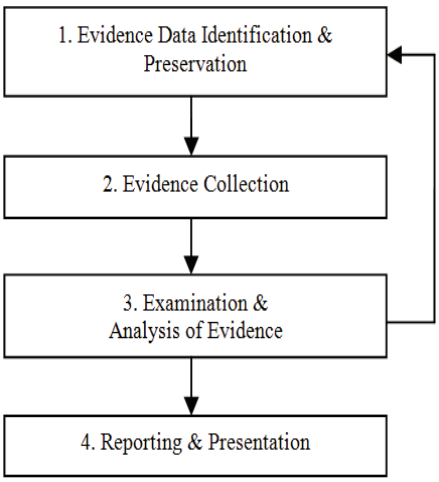

Fig.1. Digital Forensic Analysis Cycle

D. Evidence Report

In this section it covers regarding the judicial appearance of the user data collected by the examiner. The final investigation report must covers detail such as all techniques, software, and technologies utilized for digital forensics investigation.

III. CLOUD FORENSICS CHALLENGES

Cloud forensic have several challenges in the cloud computing environment. Some problems undergone by forensics investigators are determining the company details which provide of service and user information such as user credentials. In cloud storage, it is difficult to locate evidence data, the acquisition is impossible without knowing data location and there would be no analysis of evidence without acquisition. In a cloud infrastructure, data is partitioned and stored in distributed computing systems usually spanning different jurisdictions [3]. Hence, it would be very difficult to locate evidence. End user save data in the remote computer in an encrypted data format such that even if evidence data is identified and acquisitions are performed successfully, it does not be useful to the digital investigator. In cloud forensics, there is also an issue of identity, where it is more complex to associate a particular person to the data stored in the remote machine. In a cloud, user files are stored in remote locations and accessed using either client software or browsers. It is a challenge in a cloud environment to single out a cloud user from a large number of users distributed globally and assume them to be the owner of the file. Absence of good relationship with other countries may affect evidence collection process and it causes a delay in forensics investigation process.

The forensic practitioners are no more well known about the various kinds of storage technologies available in the industry and also about variety of evidence data every company might cause to remain after the user operation activity was finished in the cloud. The objective of this research is to announce investigators regarding various evidence exist in the user machine after the user accessed the One Drive, and amazon drive. The study was handled to reveal a procedure for retrieving and downloading client files through a browser, as well as windows cloud software. We then evaluate these two methods and summarize our result.

IV. RELATED WORKS

Darren Quick et al. [4] have developed an approach in the paper. In this research, one major area of difficulty in cloud storage is the collection of user artifacts when various cloud facilities are utilized by end user. Lack of knowledge about various user data exist in the user machine after end user accessed a particular cloud. It causes a delay in the examination process and consumes extra time to meet all company to verify whether evidence is saved inside its company cloud. By identifying the data evidence remains on the user machine, we present an improved interpretation of the kind of evidences that will exist in user machine for examination done by practitioners.

Darren Quick et al. [5] have developed an approach in the paper. In this research, recognition of user artifacts related to cloud in user machine are still remains a problem in forensics examination. By treating One Drive as a company providing services, examiner determined the various kinds of evidence data that exist on a end user machine and which are the access points for forensics investigators to examine and collect evidence data on time period.

V. PROPOSED TECHNIQUES

During this part, we propose and present a technique for forensic investigation in cloud storage. The proposed technique addresses issues such as:

- What Evidence data exist on user machine storage device following a One Drive and Amazon cloud Drive are utilized by end users through web browser or cloud software and the place of the user evidence stored in a user machine?
- What end user evidence exists in main RAM memory of the computer system used for connecting with cloud storage service?
- What end user evidence can be observed in Internet packets passing through the network when connected with cloud storage service provider?
- Whether the end user data is change throughout the procedure of transfer, saving data in cloud and saving data in user machine from the user account.
- Whether any timestamp detail is modified or remains constant without any change during the process of upload, download, and storage from cloud storage accounts.

A. Cloud Storage Service Providers

1) Microsoft One Drive: Microsoft One Drive was formerly called as Microsoft SkyDrive and Microsoft Live SkyDrive, and offers a 15 GB of free cloud storage space [5]. One Drive is connected through a software application or through a browser. Microsoft is an American company with headquarters in Washington.

2) Amazon Cloud Drive: Amazon cloud drive is a new offer from Amazon.com Incorporation. Amazon cloud drive is able to store any file types as long as it does not exceed the size of 2 GB per file in cloud storage space. It offers a 5 GB of free cloud storage space while opening an account [7]. Although, the storage space can also be expand in size by procuring more quantity of user account space from the company. This cloud drive provides software application for Linux distributions, Windows operating system, Mac ios and android. Amazon.com Incorporation is an American company with headquarters in Seattle.

B. Method for Digital Forensics Investigation of Cloud storage services

The digital forensics practitioners collect and examine evidence from all computing equipments that are possessed by the user to connect a cloud account. Such devices include Personal Computers, Smart phones and laptops, but this study analyzes only Personal computers which are the commonly used devices. In Windows operating system, the investigator first analyze whether it is feasible to acquire main RAM memory. Next, the forensics examiner acquires and preserves the evidences stored in main RAM space. Consequently, the hard disk evidence data is also obtained; examiner can collect evidence from the system files, history files, file timestamps and root folders.

An investigator examines user data obtained from Personal computer and then analyze even if any traces of a user account present in the collected client data. If so, the forensic examiner validates whether any evidence regarding username and password is exists. If user credential and any further details which allow connecting with a user storage account are available then provide a search warrant to collect data [10].

The forensic examiner has a warrant then examiner can examine and acquire user details stored in the account. Next, it is feasible to examine evidence acquired from user account and the evidence available in user machine. If the examiner not have warrant given by the concerned legal entity, it is only feasible to examine only the artifacts that are available in Personal computer [10]. The scope of this research is to tell examiners about the evidence left after when a One Drive and Amazon drive is used from a Windows 7 machine. Fig.2 explains the flow of the forensic examination system.

A various number of virtual machines were newly created for this research purpose. It was decided to analyze and acquire a various diverse environment of an end user accessing cloud account storage space and also to analyze

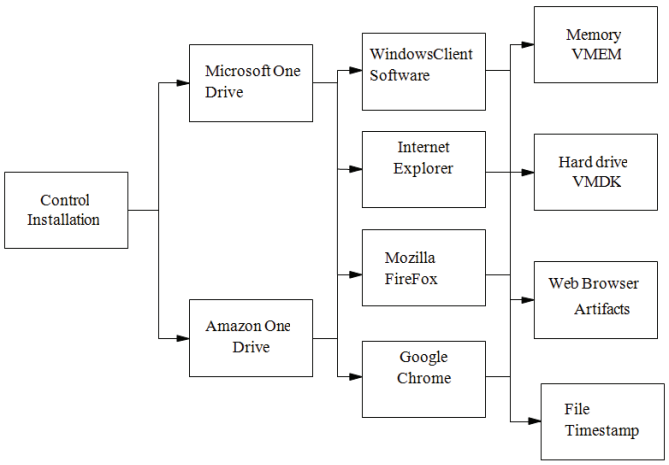

Fig.2. Block Diagram of System Design

whether any deviations found when using variety of browsers such as chrome, Internet Explorer, and firefox. A variety of special Virtual Machines (VMs) were created for each browser to replicate various background of usage.

For this study purpose, particularly selected files from the Enron Email dataset files were used as a testing input files and were save in the user machine from the concerned webpage on the 15th October 2015. Hash values such as SHA and MD5 are computed for the selected sample input files. The selected input sample data were initially uploaded into storage account opened for this study purpose [6].

Windows 7 virtual computers were produced using vmware edition 12.0.1. For every end-user usage scenario, a bottom stand virtual machine files were created and Windows 7 was deployed on a 25 GB virtual storage with 1 GB main memory. The bottom stand virtual machine was used as control medium to decide when user activity such as cloud account creation, input file uploading and downloading was performed under the different scenario. The variety of actions performed for this research purpose were as follows. The Proposed methodology for evidence data acquisition, collection and preservation are shown in Fig. 3 [8]. A virtual computer image copy are subsequently created for each virtual storage using the tool OSF image and OSF clone to make a forensic image in the Advanced Forensics Format (AFF) filetype. Hash values such as SHA and MD5 were obtained to confirm the integrity of the user account data using Hash compare tool.

Every individual virtual computer forensic image version consist of the virtual machine storage, RAM, and packet traffic in the network captures was acquired and analyzed independently with the help of a variety of forensic tools like OS Forensics version 3.3.1, Mandiant Redline edition 2.5, Digital Forensics Framework (DFF) version 1.3.0 and Magnet Forensics IEF Version 6.6. The storage companies are usually a web page oriented storage hosting mechanism, it is important to acquire and examine files about browser cookies, history and log files. These files are saved in user root directory. The log files exist in the paths are shown in the Table I.

Digital Forensics Analysis Methodologies for Cloud Storage

Step 1: Evidence Identification

Identify cloud storage account (Microsoft one drive and Amazon Cloud drive) details such as Username and Password.

Step 2: Evidence Preservation

1. Employ a computer with internet connection to connect with cloud account for this research and arrange one more computer to perform forensic analysis without any internet connection because forensic analysis computers are not connected to the internet.
2. Build a base virtual machine (VM) to access the cloud storage data, with a new installation of Windows 7 operating system.
3. Execute network traffic capture tool wireshark on the Host computer to capture the network traffic between the Virtual Machine and the Internet.
4. Execute Microsoft Windows Expression Encoder edition 4 to record the Virtual Machine window screen as a video and retain a record of the procedure undertaken.
5. Using the Virtual Machine, connect to the cloud storage service provider website using a browser, such as Google chrome, Internet Explorer and Mozilla Firefox.
6. Login to the cloud account using the username and password created for this study.
7. Browse the files and folders, note the times and dates associated with the folders and files, such as file created, accessed and modified time before uploading to the cloud account.
8. Choose files in the cloud account and then download it to the local VM Downloads folder.
9. Choose the option to download the client software from the service provider website.
10. Install the client software to the Virtual Machine and synchronize the cloud account files with virtual machine.
11. Preserve the VM files (including the network PCAP, VMDK, VMEM, Expression Video) from the Host computer into a Logical evidence file (AFF, AD1 or CTR).

Step 3: Examination and Analysis of Evidence

Execute Examine and analysis of the collected evidence as per standard forensics methodology for user files and data.

Step 4: Evidence reporting and Presentation

Submit digital forensics report to the court for legal proceedings.

Fig. 3: Proposed Data Preservation process For Cloud storage

TABLE I. Significant Files and Paths (Mozilla Firefox, Internet Explorer, Google Chrome)

Web Browser	OS Version	Data	Path
Mozilla Firefox	Windows 7	Cache Cookie Session History	C:\Users\<username>\AppData\Local\Mozilla\Firefox\Profiles\tjqw24ha.default_CACHE_CLEAN C:\Users\<username>\AppData\Roaming\Mozilla\Firefox\Profiles\tjqw24ha.default\cookiesqlite C:\Users\<username>\AppData\Roaming\Mozilla\Firefox\Profiles\tjqw24ha.default\sessionstore.js C:\Users\<username>\AppData\Roaming\Mozilla\Firefox\Profiles\tjqw24ha.default\places.sqlite
Internet Explorer	Windows 7	Cache Cookie History Download	C:\Users\<username>\AppData\Local\Microsoft\Windows\Caches C:\Users\<username>\AppData\Local\Microsoft\Windows\Cookies\index.dat C:\Users\<username>\AppData\Local\Microsoft\Windows\Histroy\Histroy.IE9\index.dat C:\Users\<username>\AppData\Local\Microsoft\Windows\IEDownloadHistory\index.dat
Google Chrome	Windows 7	Cache Cookie Histroy	C:\Users\<username>\AppData\Local\Google\Chrome\Userdata\Default\cache\index C:\Users\<username>\AppData\Local\Google\Chrome\Userdata\Default\cookies C:\Users\<username>\AppData\Local\Google\Chrome\Userdata\Default\Histroy

C. Evidences related to cloud storage Providers

1) Microsoft One Drive Forensics Analysis

Our research then focused on Microsoft One Drive and the process outlined was undertaken using a new virtual machine. We accessed a One Drive account for this research using each web browser and client software.

a) Microsoft One Drive Client Software: When Microsoft One Drive client software is installed on a Windows 7 operating system. The OneDrive client software was synchronized and downloads the data from storage webpage to the root file folder, in this case, "D:\Account user\Account ID\ One Drive\". Examination of the Syncdiagnostic file placed variety of details of evidence. The metadata and network related details enclosed within the evidence file such as one drive owner name.

b) Microsoft One Drive Client software dates and times: When analyzing the data saved from webpage through One Drive application software, below specified variations were noticed, the Last Accessed time, Modified, created timestamp of the saved data were similar to the properties file was downloaded in the one drive. Next, like one drive, the file written time of the saved files and the original are similar.

c) Microsoft One Drive Browser dates and times: Examine the file timestamps of the data saved through the browser, there are many variations found. When utilizing a browser to save a data, the data created and accessed time of

the original and saved files are same to the timestamp of file transferred to the One Drive account. No entries in saved file timestamps are matched with the original. No deviations were observed between various browsers tested and all gave the similar results.

d) *Microsoft One Drive Analysis findings:* In this research, a different data leftover was identified once a end user accessed OneDrive to save data. OneDrive related file saved locations were shown in Table II. The concentration is to determine what data leftovers were left to identify if OneDrive accessed in user machine. Performing a search of the word 'Onedrive' was done and it gives a better result found the evidence present in the system when OneDrive used. The OneDrive user credentials were identified from a variety of locations, such as history files, pagefile.sys and IEF output. When a browser was used to access OneDrive there were many pieces of evidence found in the outcome of forensic evidence tools such as Internet Evidence Finder.

Filenames, timestamps, and dates was noticed in the SyncDiagnostic file it gives Owner name . Some files give details was also noticed in the Registry key Recent Docs file, Prefetch files and $MFT entries such as Wordpad, DLLhost, and Notepad. A forensics report should include information about Analysis findings of all processes, the tools, and applications used for forensics investigation.

TABLE II. Microsoft OneDrive Setup information

Microsoft One Drive File Path and Registry Path
C:\Users\cloud\AppData\Local\Microsoft\OneDrive C:\Users\<username>\AppData\Local\ Microsoft\OneDrive C:\Users\<username>\OneDrive
HKEY_NTUSER\Software\Microsoft\CurrentVersion \Uninstall HKEY_NTUSER\Software\Classes\OneDriveClient.

2) Amazon Cloud drives Forensics Analysis

Our research then focused on Amazon cloud Drive and the process outlined was undertaken using a new virtual machine. We accessed an Amazon cloud Drive account for this research using each web browser and client software.

a) Amazon Cloud Drive Client Software: When Amazon Cloud Drive client software version 1.15.6464.0228 was installed on a Windows system. The software application was noticed the synchronize and saved the data in Amazon cloud Drive to the root folder in user machine, in this case, "D:\Account name\Account ID\ Amazon Cloud Drive\". Analysis of Amazon Cloud drives setup file that was saved from the Cloud account to personal computer. To execute 'Amazon Cloud drive setup' was moved to 'C:\Program Files\Amazon\Cloud Drive\' folder, and 'Amazon Cloud drive setup.exe.dll' was created. The installation executable files downloaded from Amazon in the file repository as follows; 'D:\Account\AccountID\AppData\Local\Amazon\Cloud Drive \'. This directory contains two files of attention, '1319b5c6-2672-49b4-b623-bf5a33fd4c40.db'and a file created in local PC called'ADrivenativeclientservice.log'.

b) Amazon Cloud Client software dates and times: Next we examined the data saved from webpage through amazon Drive application software, below specified variations were noticed, the Last Accessed time, Modified, created timestamp of the saved data were similar to the properties file was downloaded in the one drive. Next, like one drive, the file written time of the saved files and the original are similar.

c) Amazon Cloud Drive Browser dates and times: Examine the file timestamps of the data saved through the browser, there are many variations found. When utilizing a browser to save a data, the data created and accessed time of the original and saved files are same to the timestamp of file transferred to the Amazon drive account. No entries in saved file timestamps are matched with the original. No deviations were observed between various browsers tested and all gave the similar results.

d) Amazon cloud drives Analysis findings: In this research, different of files leftovers were noticed at the time of user accessing Amazon Cloud Drive to save. Amazon Cloud Drive application software and its locations are elaborated in Table III. The focus is identify type of data exist Amazon Drive are accessed by the user machine with Windows. AdriveNativeClientService log file resides in the root directory folder in the client users root drive; it contains helpful details for investigation concerning utilization of the cloud application software. AdriveNativeClientService log contains details regarding each successful download or upload. The report should include information about Analysis findings of all processes, the tools, and applications used for forensics investigation.

TABLE III. Amazon Cloud Drive File Path

Amazon Cloud Drive File Path and Registry Path
C:\Users\cloud\AppData\Local\Amazon\Cloud Drive C:\Users\<username>\AppData\Roaming\Microsoft\windows\start menu\Programs\Amazon\Cloud Drive\: shortcut to Amazon Cloud drive
HKEY_CURRENT_USER\SOFTWARE\Amazon\ AmazonCloudDrive HKEY_CURRENT_USER\Software\javasoft\Prefs\com\amazon

VI. CONCLUSION

After performing this forensics examination and analysis of the One Drive and Amazon cloud drive service providers, the investigators found that during initial stages of an analysis will be to identify the service provider's information and its corresponding user account details. This user account information will help the investigator to predict the location of user data and take suitable action to collect, analyze and protect this evidence in a timely manner. It was found that an investigator can predict cloud storage account usage details by

performing various activities such as hash comparison, Registry analysis, RAM analysis, keyword searches and network packet traffic analysis.

By analyzing One Drive and Amazon cloud drive, the examiner found that the various evidence data is stored in user personal computer. Most important artifacts are stored in log files, database files, system configuration and setup files, which give information about every action performed within Mega cloud drive and One Drive.

It was also found that an examiner can identify the One Drive username and the computer name from the files generated by One Drive. The investigator was also able to collect the username and password of the One Drive account by collecting and examining a RAM dump of the computer.

By analyzing configuration and log files generated by Amazon cloud drive, it is possible for an examiner to identify service provider details and it was found that an examiner can collect username and password of the mega account by a analyzing the network packet traffic.

As described, there is a lot of investigation evidence available for an examiner to identify the use of One Drive and Amazon cloud drive accounts. The proposed framework will helpful to guide the examiners in a digital forensic evidence collection process from starting to completion. In addition, it was able to find account details such as username information and passwords in plain text.

REFERENCES

[1] Darren Quick, Kim-Kwang Raymond Choo, "Forensic collection of cloud storage data: Does the act of collection result in changes to the data or its metadata?", Science direct, Digital Investigation, vol. 10, Issue 3, pp. 266-277, October 2013.

[2] Hyunji Chung, Jungheum Park, Sangjin Lee, Cheulhoon Kang, "Digital forensic investigation of cloud storage services", Science direct, Digital Investigation, vol. 9, Issue 2, pp. 81-95, November 2012.

[3] Ben Martini, Kim-Kwang Raymond Choo, "An integrated conceptual digital forensic framework for cloud computing", Science direct, Digital Investigation, vol. 9, Issue 2, pp. 71-80, November 2012.

[4] Darren Quick, Kim-Kwang Raymond Choo, "Dropbox Analysis: Data Remnants on User Machines", Science direct, Digital Investigation, vol. 13, pp. 63-93, June 2013.

[5] Darren Quick, Kim-Kwang Raymond Choo, " Digital droplets: Microsoft SkyDrive forensic data remnants", Science direct, Future Generation Computer Systems, vol. 29, Issue 6, pp. 1378-1394, August 2013.

[6] Darren Quick, Ben Martini, Kim-Kwang Raymond Choo, "Google Drive: Forensic Analysis of Cloud Storage Data Remnants", Science direct, Cloud Storage Forensics, vol. 40, pp. 179-193, April 2014.

[7] Jason S hale, "Amazon cloud drive Forensic analysis", Science direct, Digital Investigation, vol. 10, Issue 3, pp. 259-265, October 2013.

[8] Keyun Ruan, Joe Carthy, Tahar Kechadi, Ibrahim Baggili, "Cloud forensics definitions and critical for cloud forensic capability: An overview of survey results", Science direct, Digital Investigation, vol. 10, Issue 1, pp. 34-43, June 2013.

[9] Ben Martini, Kim-Kwang, Raymond Choo, "Cloud storage forensics: ownCloud as a case study", Science direct, Digital Investigation, vol. 10, Issue 4, pp. 287-299, December 2013.

[10] Deoyani shirkhekar, Sulabha Patil, "Design of Digital Forensic technique for Cloud Computing" , International Journal of Advanced Research in Computer science and Management studies, Vol. 2, pp. 192-194, June 2014.