Client Forensics: An Assessment of Existing Research And Future Directions

Rose Shumba *shumba@usna.edu*
Cyber Science Department
United States Naval Academy
Annapolis, MD

Abstract

This paper provides an assessment of processes for identifying artifacts, left on client devices after a cloud storage interaction. It focuses on those artifacts that may be used to prove usage of a cloud service, as proposed by the current research. Besides providing the current state of knowledge in client forensics, this paper (1) provides a summary of current research in the area of client forensics, (2) presents similarities and differences among proposed processes and identified artifacts, and (3) presents some possible future work. Investigators need to understand how devices and cloud storage services interact, the types of evidential artifacts that are likely to remain on the devices after cloud storage interaction, and how they may be used to prove usage. Not knowing if a cloud service was accessed, or which cloud service or the location of digital evidence can potentially impede an investigation.

Keywords: Cloud Forensics, Cloud Storage Services, Client Forensics.

1. INTRODUCTION

A cloud can be defined as a scalable, virtualized, distributed computing platform, whose shared resources are accessed remotely by users through a network. There are three primary cloud service models; Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). The fourth delivery service model, which has emerged because of the ever-increasing cloud storage options, is Storage as a Service (StaaS). StaaS has been widely adopted by governments, organizations, and individuals, with Gartner forecasting that a third of user data will be stored in the cloud[1]. Each service model mentioned above is deployable as a public, private, hybrid or community. A description of the service and deployment models is beyond the scope of this paper.

As the use of cloud services continues to transform IT Services by providing benefits such as increased flexibility, efficiency, and costs, the security of corporate data is becoming a concern. Cyber-crime, such as theft of intellectual property, espionage, acceptable use policy violations, and data breaches is on the rise. When such incidents occur, it may be necessary to conduct investigations. The National Institute of Standards and Technology [2] and several researchers identified over 65 challenges associated with cloud forensics.[3]; [4,5,6]; [8]; [9]; [10],

Depending on the deployment architecture and service model, it is possible that investigators may not have access to the physical servers to conduct server analysis making it hard to determine the legal framework to follow to obtain the evidence. It leaves the investigator with three options:

1) To attempt and recover evidence from seized local devices known to have interacted with the cloud,
2) To attempt to eavesdrop network traffic between local devices and the cloud network, and

3) To request a court in a foreign jurisdiction to seize evidence directly from a cloud server[3].

The last option brings additional legal challenges, such as the problem of identifying and addressing issues of jurisdiction for legal access to data and the lack of adequate channels for international communication and cooperation during cyber forensic investigations. Section 2 of this paper provides a summary of the research work, including the experiments carried out by the different researchers, similarities, and differences between processes used and identified artifacts, section 3 presents further questions for future research.

2. LANDSCAPE OF EXISTING RESEARCH WORK

Unique to client forensics analysis, research has focused on identifying remnants that can be used to prove usage. A cloud storage service is accessible through installed client software or a browser. Researchers carried out some experiments that involved accessing a storage service from one or multiple devices:

1) Google Drive and Dropbox from a Windows 7 PC and an iPhone 3G [4,5,6]
2) Amazon S3, Dropbox, Evernote, and Google docs from Windows XP/Vista/7, a Mac PC, and an iPhone 3G [7].
3) Dropbox, Google Drive, and SkyDrive from a Windows 7PC and iPhone 3G [9].
4) Copy and ownCloud from Windows 8.1 PC[8].
5) Google Documents, Flickr, PicasaWeb, Dropbox from a Windows 7 PC [3]
6) 360 and Baidu from a Windows 7 PC[10]

Researchers used a combination of static and dynamic processes to identify the artifacts. The static approach assumes that the investigator has a forensic image and can use forensic tools and prior acquired knowledge and skills to locate remnants [8]. The dynamic method uses software tools, such as Disk Pulse and RegShot to find artifacts[9], while the experiment activity is underway and the PC being used to access the service is on.

Table 1 shows for each researcher, the cloud storage service, and the devices used to access the service. Also, the PC platform, in order of popularity and associated browsers is presented. Mobile access used an iPhone 3G with inbuilt Safari browser.

The experiments core activities involved installing, accessing cloud service through a web browser and the client software, involved uploading/copying and deleting the user data files.

From Table 1, it is clear that most the Windows 7 using IE browser has been the commonly used platform and the 3G iOS device. This shows the areas of future experiments.

2.1 Conclusions From The Experiments Carried Out – Current Research

A closer look at the processes and identified artifacts, from the experiments described in the previous section, shows the following:

1) Accessing a cloud storage service through a Web browser or client software creates a substantial amount of artifacts that can be used to prove usage of the service. Examples of remnants include the cloud storage service used, installation location, installed version, usernames, passwords, URLs of downloaded client software, prefetch files, link files, and file references related to the account. These artifacts play a significant role in an investigation as they may lead an investigator to the potential location of other remnants promptly.
2) The identified residues are dependent on the browser used to access the storage. From the experiments, we can see that Ephani experimented with Mozilla Firefox and Internet Explorer[9], Chung used Internet Explorer[7], and Quick, and Choo used Mozilla Firefox, Google Chrome, Safari and Internet Explorer with Dropbox and Google Drive[4,5,6].

3) Quick and Choo's found a Google Drive account username through browser analysis with Mozilla and Google Chrome, but none with Apple Safari. For Dropbox, the account username was identified from browser history with Mozilla Firefox and Google Chrome browsers.
4) The principal sources of artifacts depend on the device used to access the storage. Where a PC is used the three sources of remnants are the hard drive, the RAM and the eavesdropped network traffic between the device and the cloud network. When an iPhone 3G the specific locations for artifacts were database files, XML files, and plist files, as identified from the logical extract.
5) The different experiments, using the same cloud service and device, identified similar artifacts. Tables 2 and 3 compare identified remnants when Dropbox and Google Drive are accessed from a PC as presented by Ephani[9] and Quick [4]). The experiment activities involved downloading the client software for installation, performing various operations including uploading and deleting data files, accessing storage service through a browser and client software and uninstalling the service application. For example, if a suspect installed the Google Drive client software, the investigator can look in the Program Files folder for proof of installation.
6) With Google Drive and Dropbox, Choo [4,5] showed that additional artifacts to prove usage are obtainable from the network traffic analysis. These include the IP address, registered owners, and the digital certificates. When Internet Explorer was used to access Google Drive, the username was observed in the network traffic. RAM analysis after accessing drive through a browser also showed the username.
7) Except for Choo [4,5] and Long [10], the presented research did not follow a process. Choo and Long [10] proposed methods for client forensics analysis on a PC client. Choo [4,5] developed a standard analysis process for practitioners, examiners, and researchers to follow for client forensics. The proposed method can be used by an investigator to find out Google Drive data remnants when accessed from a local device.

The process is cyclic and consists of the following phases:

a. setting the scope of the investigation,
b. setting up the virtual machines, installing the tools and browsers,
c. identifying and collecting the virtual hard drive image, memory image, and network capture file,
d. making forensic copies of each image,
e. using a range of forensic tools to analyze each image,
f. stopping the investigation, or further analysis, if required.

Long developed a process that involved analyzing the registry on the user device, collecting evidence from the browser and client and then determining the possible event sequence by examining the correlations and rules of user activities among different time, different targets, and behavior intention. The process was tested with 360 cloud and Baidu cloud services.

2.2 Further Questions And Future Work

The main conclusion from the current research work is that some evidentiary artifacts that can be used to prove usage are obtainable through an exhaustive analysis of the client devices, without accessing the cloud server.

Recovering evidence from client devices known to have interacted with the cloud has the advantages that the devices can easily be accessed, and the cost of the forensic analysis is relatively low.

Depending on the deployment model in use, much of the evidence may be on the client. An exhaustive analysis of the client devices may help us obtain details, such as account credentials that can be used to access the server side. Location of evidence and having precise knowledge of what information is on the device can expedite current casework.

There are further questions that need to be addressed as far as client forensic analysis research is concerned.

1) Presented study limits the scope of an investigation to a single device. How do we merge the artifacts/evidence from several devices belonging to a single user into a timeline of events to gain an accurate depiction of the user activities?
2) How can the identified artifacts be used to speed up an investigation?
3) How can this work be extended to include more commonly used mobile device platforms; Android, Windows and latest iPhone?
4) How can artifacts gather at the client side be used to access account data on the server side?
5) What is the relationship between the remnants collected on the client side and those on the server side?
6) Does the version of the client software matter as to the types of identified artifacts?

A fundamental observation is that software developers are continually providing additional functionality, closing security holes; this affects how much remnants are collectible from the client side. For example in 2012, Chung [7] research identified a database file (file cache.db) in Dropbox, as being of high value to the investigator, since it contained historical synchronization logs. In 2013, Choo [4] and Ephani[9] showed that the file was encrypted and a unique tool was now needed to decrypt it. Also, several different storage services apps are now available, making it a challenge to collect remnants for each. How can a more unified digital investigation process be developed to cater for different client software? Is there a better way of investigating storage services?

3. CONCLUSION

The growing popularity of cloud storage services means that this media will be used for cybercrime, resulting in more investigation cases. One challenge is maintaining a chain of custody in the cloud. There is a need or more research in this area.

Future work should include accessing popularly used cloud storage services from commonly used mobile platforms; Android, Windows, and the latest iPhone. A series of experiments that involve installing, accessing, uploading and downloading some documents; uninstall the client software, and then using anti-forensics techniques (deletion, uninstalling and clearing the browser history) have been performed for the popularly used cloud services, accessed from Android. For each mobile device, a physical extract will be used. We have started work on rooting an Android device, in order to collect the main system folders.

4. REFERENCES

[1] Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results. *Digital Investigation*, *10*(1), pp 34-43 [2016].

[2] Stamford. (n.d.). Gartner Says That Consumers Will Store More Than a Third of Their Digital Content in the Cloud by 2016. Available: http://www.gartner.com/newsroom/id/2060215

[3] NIST. (2014). *NIST Cloud Computing Forensic Science Challenges* (NISTIR 8006).

[4] Marturana, F., Me, G., & Tacconi, S.. A Case Study on Digital Forensics in the Cloud. In Proceedings of the *2012 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery [2012]*.

[5] Quick, D., & Choo, K. R.. Dropbox analysis: Data remnants on user machines. *Digital Investigation*, *10*(1), pp 3-18 [2013].

[6] Quick, D., & Choo, K. R. Google Drive: Forensic analysis of data remnants. *Journal of Network and Computer Applications*, *40*, pp 179-193 [2014].

[7] Quick, D., Martini, B., & Choo, K. R. Forensic Collection of Cloud Storage Data. *Cloud Storage Forensics*, pp 153-174 [2014].

[8] Chung, H., Park, J., Lee, S., & Kang, C.. Digital forensic investigation of cloud storage services. *Digital Investigation*, *9*(2), pp 81-95 [2012].

[9] Malik, R., Shashidhar, N., & Chen, L.. *Analysis of Evidence in Cloud Storage Client Applications on the Windows Platform*. In the Proceedings of the Int'l Conf. Security and Management [2015].

[10] Epifani, M.. *Cloud Storage Forensics*. Paper presented at SANS European Digital Forensics Summit, Prague [2013].

[11] Long, C., & Qing, Z.. Forensic Analysis to China’s Cloud Storage Services. *International Journal of Machine Learning and Computing*, *5*(6), pp 467-470 [2015].

TABLE 1: Summary of Experiments

Researcher	**Storage Service**	**Device used for access**											
		PC									**Mobile device**		
		Platform					**Browser**						
		Win 7	**Win 10**	**Win XP**	**Win 8**	**Mac OS X**	**Firefox**	**IE**	**Chrome**	**Safari**	**Android**	**iOS**	**Windows**
Chung	Amazon s3	✓		✓	✓			✓				✓	
	Dropbox	✓		✓	✓	✓		✓				✓	
	Evernote	✓		✓	✓	✓		✓				✓	
	Google docs					✓		✓				✓	
Quick	Dropbox	✓					✓	✓	✓			✓	
	SkyDrive	✓										✓	
	Google Drive	✓					✓	✓	✓			✓	
Epifani	Dropbox	✓					✓	✓				✓	
	Google Drive	✓					✓	✓				✓	
	SkyDrive	✓					✓	✓				✓	
	iCloud						✓	✓				✓	
Long	Baidu							✓					
	360							✓					
Malik	Copy				✓				✓				
	ownCloud				✓				✓				
	Dropbox				✓				✓				

TABLE 2: Summary of Activities for Google Drive

Activity	Google drive (Choo)	Google Drive (Ephani)
Installation	• Client software prefetch file and link files • Created registry keys. • Program files folder for installtion • Sync_config.db - user email used to access the Google Drive account;	• Installs in Program File folder • Client configuration in user profile • New registry keys created • Registry –installed version • Link file and prefetch files for the client software • sync_config.db –client version, sync root path, user email • snapshop.db –file names, created, modified, URL, size, shared
Uploading *(sample documents)*	• Thumbcache.db -thumbnails for any uploaded pictures • Snapshot.db- URL and Resource-ID identifier.	
Downloading *(client software and test data file)*	• Prefetch files and link files for filenames of client executable and test data files. • RecentDocs key in NTUSER,dat for recently used documents. • Network Traffic analysis; IP addresses registered to Google and URL were observed • URL, from which client software from cookies, history, icons, pagefioe.sys, unallocated, Temp Internet files.	
Deletion	• Link file references remaining after deletion • Prefetch files for client and uploaded files remained. • Deleted information on Google Drive and test files in NTUSER.dat registry files. • Deleted files in $Recycle.Bin folder with the SID of user	• snapshot.db remained -filename
Accessing: storage through a browser,	• Username –Web browser, history, cookies • URL for client software –Cookie files, Web history, • Downloaded files –web history • IP addresses and registered owners (Network	• Sync_log.log –sync sessions, files created, file saved, file deleted, deleted files, version history, recent activities, • RAM Analysis (username and password in clear text)

	traffic). • RAM analysis -full text of sample files, username in pagefile.sys file, and filenames for the test files	
Access through the client	• Login sessions • URL for security certificate	• RAM analysis – session information, user email, version number, snapshot.db, sync_config.db path, local sync folder path
Uninstall		• Client config folder is removed • Prefetch and link files not deleted,

TABLE 3: Summary of Activities for Dropbox

Activity	Dropbox (Quick, Choo)	Dropbox (ephani)
Installation	• Installs in the user profile • Filecache.dbx - server path, local filename, local creation time, local modified time, local size • Host.db –path for Dropbbox file storage	• Installed user profile • New registry keys created • Registry –installed version, install location • Prefetch and link file for the client software • Host.db – local folder used to sync the account • Filecache.dbx –server path, local filename, local creation time, local modified time, local size
Upload	• Dropbox URL-from file listing. • References to Dropbbox URL, software files and folders, sample files, and test data files. • NTUSER.dat provided a list of Dropbox and sample files. • Prefetch files for Dropbox executable, Dropbox sample files, and test data files. • Link files for filenames and folder names for the Dropbox executable, Dropbox sample files and sample files. • Thumbnails for the Dropbox sample pictures, • RAM Analysis –password in clear text, sample filename references,	
Download	• Access through browser- References to filenames. • Website information located in the Cookie files, Web history, FavIcons • Filenames for downloaded files from Web history of the browsers. • Link files and prefetch files for Dropbox executable, Dropbox sample files • Thumbnails for the Dropbox sample pictures,	
Deletion/erase	• Deleted Dropbox information –NTUSER.dat	• snapshot.db remained -filename

	registry (filesnames, URL references) • Cleaning removes all file references in the browsers, but information remained in Google Chrome FavIcon for Dropbox use	
Accessing: drive through a browser,	• synched devices, timestamp of the last activity, the IP address for the last connection, and the version, deleted files, timeline of previous events, • Browser access to account: filename references for accessed files and Dropbox sample files, prefetch files for Dropbox executable, Dropbox sample files, Dropbox website information, network traffic analysis - a session with an IP in the Range, digital certificates,	• deleted files, • devices connected to the account • for every file version history • last browser sessions • RAM Analysis- client access (user email, display name, filecache.dbx path, server time, file list, deleted file) • RAM- browser access (login email, login password)
Access through the client	• prefetch files for Dropbox executable, Dropbox sample files, • References to sample files in the browser history.	• RAM analysis Session information (User email, version number, snapshot.db, sync_config.db path, local sync folder path
Uninstall	• Only Dropbox.exe marked deleted, others remained. • Synch folder and file contents remained on hard drive. • All remnants were unaffected.	• Client config folder is removed • Prefetch not deleted • Local copy of the file not deleted • Can recover registry keys about recent files, link files, browser history and cach, thumbnails,