Cloud Storage Forensic: hubiC as a Case-Study

Ben Blakeley, Chris Cooney, Ali Dehghantanha, Rob Aspin
School of Computing, Science, and Engineering, University of Salford
Manchester, United Kingdom
ben@benblakeley.com, christopherpiercecooney@gmail.com, A.Dehghantanha@Salford.ac.uk, R.Aspin@salford.ac.uk

Abstract— In today's society where we live in a world of constant connectivity, many people are now looking to cloud services in order to store their files so they can have access to them wherever they are. By using cloud services, users can access files anywhere with an internet connection. However, while cloud storage is convenient, it also presents security risks. From a forensics perspective, the increasing popularity of cloud storage platforms, makes investigation into such exploits much more difficult, especially since many platforms such as mobile devices as well as computers are able to use these services. This paper presents investigation of hubiC as one of popular cloud platforms running on Microsoft Windows 8.1. Remaining artefacts pertaining different usage of hubiC namely upload, download, installation and uninstallation on Microsoft Windows 8.1 are presented.

Keywords— Cyber Investigation; Cloud Forensics; hubiC

I. INTRODUCTION

Cloud Computing may be defined as "a model for enabling ubiquitous, 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 or service provider interaction." [1] and has provided a powerful platform for criminal use [2]. It has been used to store, share and transfer malicious content. Ergo, investigations of cloud platforms rise as a priority for forensic experts. Investigation of cloud storage services poses new challenges arising from the federated nature of information which presents multiple legal issues [3].

Utilizing a forensics investigation to analyze Amazon S3, Dropbox, Evernote and Google Docs on Motorola Droid with Android v2.2.2, iPhone 4 with iOS v4.3.5, Mac and Windows Machines [4] identified a variety of digital remnants including; username, downloaded and uploaded filenames. Moreover, investigations of Dropbox, Google Docs, PicasaWeb and Flicker, carried out on Windows 7 PCs, have recovered artefacts of web browser and desktop applications [5]. Dykstra and Sherman analysed export features of Amazon EC2 to illustrate a cloud evidence collection method using different forensics tools such as EnCase and AccessData FTK [6]. Hale also recovered a variety of artifacts, such as the installation path and registry modifications [7], from an Amazon cloud drive on Windows XP when analyzing both server and client sides of a private cloud storage service [9], authentications credentials, file contents and timestamps were found, allowing for the full recovery of server files. Oother investigations, highlighted that, despite the challenge of obtaining logs from cloud services, these services held substantial worth to a forensic investigator [11] and a tool was provided for accessing the logs on a cloud service.

Martini and Choo proposed a collection process, consisting of six steps, for investigating VMware vCloud and showed that by using vCloud's REST API, a variety of evidential data could be collected [14]. Similarly, a cloud data imager was developed to collect evidences from Google Drive, Dropbox and Skydrive [15], providing interaction (browsing and imaging) with folder trees. Google Drive was employed as a case study to test and demonstrate a forensic analysis framework [16]. This was subsequently tested against SkyDrive [17], Dropbox [18], SugarSync [19] and UbuntuOne [13] cloud services.

A summary of cloud storage forensic research is outlined in Table I.

| TABLE I | A SNAPSHOt OF EXISTING CLOUD FORENSICS RESEARCH |
|---------|-----------------------------------------------|
| General Cloud Research | Public cloud | Private cloud |
| Mell & Grance [1] | Amazon Cloud Drive [7] | Dropbox, Box and SugarSync [8] |
| Daryabar et al. [2] | XtreemFS [12] | Dropbox [18] |
| Damshenas et al [3] | Ubuntu One [13] | SugarSync [19] |
| Chung et al [4] | Dropbox, Google Drive and Microsoft SkyDrive [15] | |
| Marturana et al. [5] | Dykstra & Sherman [6] | Google Drive [16] |
| Martini & Choo [9] | Skydrive [17] | |
| Wen et al. [10] | Martini & Choo [14] | |
| Zawoad et al [11] | | |
| Martini & Choo [14] | | |
| Dezfoli et al. [20] | | |
| Daryabar et al. [21] | | |
| Daneshenas [22] | | |

hubiC is a rapidly growing cloud storage service, created by OVH (On Vous Héberge) at the end of 2011. This cloud service celebrated 500,000 unique users on 27th March 2015, with ambitions to achieve 1,000,000 users by 2020. hubiC was originally solely available to users in France until 2013, after which the ability to register accounts for international users was offered. However, literature suggests that investigation of hubiC as a cloud storage platform has been left untouched. Therefore, the research reported here aims to provide insight into the remnants of data left behind by the hubiC cloud service on Windows 8.1 (the current active, stable mass...
consumer operating system). This research paper aims to answer following questions:

- What data can be recovered on the hard drive of a Windows 8.1 machine after the use of the hubiC cloud storage service?
- What data can be recovered from the physical memory (RAM) of a Windows 8.1 machine after the use of the hubiC cloud storage service?
- What data is transmitted in the network traffic during communication with the hubiC cloud storage service during the upload and download of files?

The following sections present the research experiment setup, followed by detailed discussion of detected remnants. The paper concludes providing a summative review of findings and proposals of future works.

II. EXPERIMENT SETUP

The experiment was conducted over various VMs. A base VM was created, containing a clean installation of Windows 8.1.9.3.9600. This VM was then cloned to the following VMs:

**The Access VM (1.1)** – To detect remnants of accessing and logging into hubiC via the web browser.

**The Upload VM (1.2)** – To investigate evidences of uploading files to hubiC.

**Download/Open VM (1.3)** – To detect evidences during the downloading and opening of a file from the web browser.

**Delete VM (1.4)** – Investigating file remnants left behind when a file is deleted via the web browser.

**Install Desktop VM (1.5)** – Searching for evidences after installation of the hubiC desktop application.

**Upload Desktop VM (1.5.1)** – Searching for file remnants after upload of a test file via the hubiC desktop application.

**Uninstall Desktop VM (1.5.1.1)** – Searching for evidences of hubiC desktop application existence and test files.

**Download Desktop VM (1.5.2)** – Investigating evidences available after using hubiC desktop application to download a test file.

**Delete Desktop VM (1.5.2.1)** – Searching for file remnants after deleting test file using hubiC desktop application.

Internet Explorer v11.0.9600.16384, Google Chrome v40.0.2214.115 and Mozilla Firefox v35.0.1 were used during web-browser based experiments. When using Chrome, ChromeCacheView v1.66 was used to view data remnants left behind when accessing the hubiC website. Additionally, the HxD hex editor V1.7.7.0 was used for the memory forensics portion of the investigation. Email 22 from the /maildir/delanney-d/inbox directory of the Enron email dataset (http://www.cs.cmu.edu/~enron/) (10/04/2015), was used to upload to the hubiC server. The network traffic was investigated using Wireshark 1.12.4.

III. RESULTS AND ANALYSIS

In this section, results of investigating the Access VM (1.1), Upload VM (1.2), Download/Open VM (1.3), Delete VM (1.4), Install Desktop VM (1.5), Upload Desktop VM (1.5.1), Uninstall Desktop VM (1.5.1.1), Download Desktop VM (1.5.2) and Delete Desktop VM (1.5.2.1), including memory forensics and analysis of temporary and log files, which provide insight into the internal operations of hubiC are presented. Network monitoring was also conducted during these investigations.

A. Network Traffic Analysis

During analysis of the network traffic, it was observed that when the initial request was made to www.hubic.com, a redirect to https on port 8080 was returned. This indicates that there is no plaintext traffic accepted by the hubiC website, a subsequent redirection response is then sent, with the status code 302 (moved temporarily), which finally redirects to https://www.hubic.com/en.

All communication between the hubiC client and server applications was encrypted using TLSv1, preventing successful extraction of data of any sniffed traffic. Therefore, no clear indication of the e-mail address or the password could be detected within the login stream. Once login was completed, a session id was created, although this was not saved into the cookies. This value was a combination of “hubiC.com” and a timestamp. All involved domain names and their respective IP address were analyzed. Most of the IP addresses did not have an available DNS mapping. During analysis, we managed to infer the certificate type, by following the TCP packets in transit using Wireshark as shown in Figure 1.

![Figure 1 – Certificate Data](image1)

During uninstallation, packets were found to be exchanged with the hubiC server. The TCP conversation revealed that there was no communication between the hubiC client and server during the uninstallation process, indicating that all changes were performed locally.

During the upload of the file, hubiC communicated with 37.59.76.98, which, again, is an OVH server without a domain. Initiation of upload traffic streams are tagged with “hubiC.com” and a timestamp. All involved domain names and their respective IP address were analyzed. During analysis of the network traffic, it was observed that there was no communication between the hubiC client and server during the uninstallation process, indicating that all changes were performed locally.

During the upload of the file, hubiC communicated with 37.59.76.98, which, again, is an OVH server without a domain. Initiation of upload traffic streams are tagged with “hubiC.ovh.net” as shown in Figure 2. Investigation of uploaded traffic shows utilization of compression algorithms before files are sent. For example, corresponding stream of a 1318 bytes uploaded file was just 1307 bytes.

![Figure 2 – hubiC string in TCP stream](image2)

During the upload, a GIF file was found that was named “prepareUpload.gif”. This provided information about what was happening during the usage of the machine, but did not present any information about the file itself.
B. Analysis of Access VM (1.1)

During analysis of the Access VM (1.1), the browser cache was used to attempt to obtain knowledge of when the access occurred, along with temporary internet files (such as browsing history and cookies). Using the Chrome cache viewer, Javascript, HTML and Image files were obtained that were downloaded during transactions with the hubiC browser interface (Figure 3). Analysis also revealed the e-mail address and password of the user that was used to log in to hubiC (Figure 4). This was found within the process memory of Internet Explorer using “password” as a keyword, by searching directly through the memory using HxD hex editor. Additionally, a remnant of the page title was found in memory (Figure 5).

C. Analysis of Upload VM (1.2)

For VM 1.2 the remnants that were left behind during the uploading of a file via the web browser, to the hubiC server, were analysed.

SHA1 checksums of the file before and after the upload were exactly the same.

| Properties |
|------------------|------------------|
| File name:       | 1.3.mezarella.css |
| URL:             | https://hubic.com/home/privacy/styles/browser.min.css |
| Content Type:    | text/css          |
| File size:       | 61,368            |
| Last Accessed:   | 02/09/2015 16:22:13 |
| Server Time:     | 02/09/2015 16:22:13 |
| Expires Time:    | 30/09/2015 15:52:52 |
| Name:            | _l_1_1_7_1_._b_.%e |
| Source Name:     | nuxbox            |
| Server Response: | HTTP/1.1 200 OK  |
| Content Encoding:| f_0_0_0_0_0_0     |
| E Tag:           | 513163344710      |
| URL Length:      | 60               |

Figure 3 – Example of content in CacheMemory

Figure 4 – E-Mail address and password of the user in memory

D. Analysis of Download/Open VM (1.3)

The first assessment performed was the hash comparison. The hash comparison matches indicated that the contents of the file had not been changed during the download process. The file metadata was then compared with the original. As expected, both the date created and date accessed were the time that the file was downloaded to the disk, however the date modified field was set to two seconds after.

During the memory analysis, the downloaded file was found in multiple locations. The first location represented the file that was loaded into memory by Internet Explorer during the download (the same file was not found on all other browsers), the second copy located on the hard disk. These file locations were as expected.

E. Analysis of Delete VM (1.4)

The bulk of the analysis came from investigation into the cache. While most browsers did not cache any files pertinent to communication with the hubiC web client, beyond HTML fragments, the Chrome Cache contained references to a gif file. In this instance, the file was named “delete”. This file was uploaded when the delete command was sent to the server.

F. Analysis of Install Desktop VM (1.5)

Using regedit.exe, the registry of the Install VM was scanned to find references to hubiC post installation. An assessment into the registry found entries at HKEY_CURRENT_USER/Software/OVH/hubiC-sync. After install, logs were present in AppData/Local/Temp. The names of these log files contained the time when installation was performed. The only configuration file of hubiC (hubiC.exe.config) was located at %programfiles%\OVH\hubiC. This file indicated some basic setup for how the application should parse the URL.
G. Analysis of Upload Desktop VM (1.5.1)

The first stage was to obtain the process ID of the hubiC process. This was discovered using Resource Monitor. Once the application was started up, the log files were assessed to find relevant changes. The logs written to during installation were not changed, but new logs were created in the AppData/Roaming/hubiC folder. This folder contained the application.txt file, the log file, and hubiC.db file. hubiC.db was an SQLite file. This file was subsequently opened using the Firefox plugin SQLite Manager 0.8.3. An assessment of this database before an upload had occurred showed a collection of empty tables. These tables were named backupDirs, backupFileVersions, backupFiles, files, localHashCache, remoteHashCache and synchronizedDirs. This file was then uploaded and the state of the SQLite DB was monitored. After logging in to the hubiC client application, a new database was made named “hubiC.db-journal”. Initially, hubiC.db was analysed. Data was now found in the files. Data was also found in the synchronizedDirs table as shown in Figure 6.

![Figure 6 – New entry in table](image)

Once a hubiC directory was chosen, three separate locations were placed into the files table. These folders corresponded to folders on the hubiC cloud server, indicating synchronization. When attempting to access the hubiC.db file, an error message was presented, indicating that the database file was either encrypted or corrupted. To combat this, the file was carved using a hex editor, however the encryption remained. As a way to glean information from this, the bytes were analyzed in an attempt to calculate the structure of the database table and identify any data contained. DDL syntax was found, indicating the structure of the database. This DDL syntax seemed to create the tables file, in the same structure as found in the hubiC.db file. An analysis of the file metadata revealed that there was no change to the original file during the upload process. Registry files were also observed during the upload process. The login details used are clearly stored at HKEY_CURRENT_USER/Software/OVH/hubiC-sync. The key “AccountEmail” contains the reference to the email address of the logged in user, along with SecureAuthToken, which is an oAuth 2.0 secure token, granting the hubiC software access to the API.

Memory analysis was again carried out on the virtual machine using a hex editor. Within this, the e-mail address used to log into the application was discovered as shown in Figure 7. In addition to this, the name of the file uploaded to the hubiC server was found in memory.

![Figure 7 – Email account found in memory](image)

H. Analysis of Uninstall Desktop VM (1.5.1.1)

The uninstallation was performed using the Windows uninstaller: hubiC does not have its own uninstaller application.

Analysis of the log files stored in AppData/Local/Temp revealed they were not removed during the uninstallation process; however they contained no information within them about the application being uninstalled. All Registry entries and hubiC.db and hubiC.db-journal files were removed during the uninstallation process, with the exception of the application.txt file, which was located in AppData/Local/Temp.

The sync folder (C\Users\Alice\hubiC), was left untouched and all files remained intact, without change to their metadata. Program installation was deleted from the C\Program Files directory.

During the analysis of the memory, the e-mail address of the user was found in the memory of the current machine state.

I. Analysis of Download Desktop VM (1.5.2)

The previously discovered log files and database files were analysed for notable changes that could pertain to the current activity. Additionally, the file was hashed before upload to the server, on a different machine, and hashed after download, to compare content. Analysis of the metadata, of the test file, revealed the date created field was the time when the file was uploaded to the server originally, but the access date matched the time of file synchronisation during the download.

Assessing the content of the hubiC.db file, after download completion, resulted in a new record contained within the files table. This new file identifier, at ID 4, represented the test file used. This indicates that the hubiC.db file is a constant database representation of the contents of the sync folder. Comparing the file hashes, showed no change to the original file contents during the download process.

During the analysis of the logs, in the application.txt file in AppData/Roaming/hubiC, there were entries reporting on directories created as shown in Figure 8.

![Figure 8 – Logs indicating folder creation](image)

During analysis of the memory of the download VM, the downloaded file was found contained in memory, with some stray bits added.

J. Analysis of Delete Desktop VM (1.5.2.1)

Assessing contents of the log files showed records of the deletion were available in the Application.txt file (AppData/Roaming/hubiC). When the test file was deleted from the sync folder, it was moved to the recycle bin, by effect of hubiC utilising Windows Explorer to enact this deletion. Additionally, when the deletion occurs, hubiC updates the hubiC.db file located in AppData/Roaming/hubiC. During the
memory analysis of the delete VM, the deleted file was found four times in the memory.

IV. CONCLUSION AND FUTURE WORKS

In reference to the initial questions posed in the introduction, the following answers were gained.

The data found on the hard drive provided insight into the inner workings of the hubiC application. The Chrome Cache Viewer enabled file remnants to be obtained during the use of the hubiC web client. Username and password of the active user were found in the hubiC process memory. Additionally, the test file, deleted from the hubiC cloud service, was found, after deletion. An analysis of the network traffic provided insight into the communication occurring between the hubiC client and server, giving IP addresses of the hubiC servers and the security system used, TLSV1. These present significant risks to the secure use of the hubiC system both while the client is installed and after it has been uninstalled.

Direct future applications of this research could apply similar methodology to the investigation of other cloud platforms, (e.g. ADrive, eCloud, etc) on Windows 8.1. This would increase the scope of the investigation and provide greater insight to an investigator, potentially revealing common flaws across several cloud platforms. Extending this common approach to applications of other categories of forensic techniques such as cloud malware forensics [23-27], mobile device cloud applications forensics [28,29], social network platforms investigation on cloud systems [30, 31], and cyber-crime and cyberwar investigation techniques in cloud environments [32,33] would explain this to a comprehensive body of knowledge. Finally, as forensics investigation process may reveal many private data [36,37], it is important to consider privacy effects of forensics investigation in cloud environments in future works.

V. REFERENCES

[1] P. Mell and T. Grance, “The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology” Natl. Inst. Stand. Technol. Inf. Technol. Lab., vol. 145, p. 7, 2011.

[2] Farid Daryabar, Ali Dehghantanha, Nur Izura Udzir, Nor Fazlida binti Mohd Sani, Solahuddin bin Shamsuddin, “A Review on Impacts of Cloud Computing on Digital Forensics,” Ijcsdf, vol. 2, no. 2, p. 18, 2013.

[3] M. Damshenas, A. Dehghantanha, R. Mahmoud, and S. bin Shamsuddin, “Forensics investigation challenges in cloud computing environments,” in Proceedings Title: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), 2012, pp. 190–194.

[4] H. Chung, J. Park, S. Lee, and C. Kang, “Digital forensic investigation of cloud storage services,” Digit. Investig., vol. 9, no. 2, pp. 81–95, 2012.

[5] F. Marturana, G. Me, and S. Tacconi, “A Case Study on Digital Forensics in the Cloud,” in 2012 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 2012, pp. 111–116.

[6] J. Dykstra and A. T. Sherman, “Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques,” in Digital Investigation, 2012, vol. 9, no. SUPPL.

[7] J. S. Hale, “Amazon Cloud Drive forensic analysis,” Digit. Investig., vol. 10, no. 3, pp. 259–265, 2013.

[8] G. Grispos, W. B. Glisson, and T. Storer, “Using smartphones as a proxy for forensic evidence contained in cloud storage services,” in Proceedings of the Annual Hawaii International Conference on System Sciences, 2013, pp. 4910–4919.

[9] B. Martini and K. K. R. Choo, “An integrated conceptual digital forensic framework for cloud computing,” Digit. Investig., vol. 9, no. 2, pp. 71–80, 2012.

[10] Y. Wen, X. Man, K. Le, and W. Shi, “Forensics-as-a-Service (FaaS): Computer Forensic Workflow Management and Processing Using Cloud,” CLOUD Comput. 2013. Fourth ..., no. c, pp. 208–214, 2013.

[11] S. Zawood, A. K. Dutta, and R. Hasan, “SecLaaS,” in Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security - ASIA CCS ’13, 2013, p. 219.

[12] B. Martini and K.-K. R. Choo, “Distributed filesystem forensics: XtreemFS as a case study,” Digit. Investig., vol. 11, no. 4, pp. 295–313, Dec. 2014.

[13] M. Shariati, A. Dehghantanha, B. Martini, and K.-K. R. Choo, “Ubuntu One Investigation: Detecting Evidences on Client Machines,” Cloud Secur. Ecosyst. Syngress Impr.Elsevier, 2015.

[14] B. Martini and K.-K. R. Choo, “Remote Programmatic vCloud Forensics: A Six-Step Collection Process and a Proof of Concept,” in 2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications, 2014, pp. 935–942.

[15] C. Federici, “Cloud Data Imager: A unified answer to remote acquisition of cloud storage areas,” Digital Investigation, 2014.

[16] D. Quick and K. K. R. Choo, “Google drive: Forensic analysis of data remnants,” J. Netw. Comput. Appl., vol. 40, no. 1, pp. 179–193, 2014.

[17] D. Quick and K. K. R. Choo, “Digital droplets: Microsoft SkyDrive forensic data remnants,” Futur. Gener. Comput. Syst., vol. 29, no. 6, pp. 1378–1394, 2013.

[18] D. Quick and K. K. R. Choo, “Dropbox analysis: Data remnants on user machines,” Digit. Investig., vol. 10, no. 1, pp. 3–18, 2013.

[19] M. Shariati, A. Dehghantanha, and K.-K. R. Choo, “SugarSync forensic analysis,” Aust. J. Forensic Sci., pp. 1–23, Apr. 2015.

[20] F. Dezfoli, A. Dehghantanha, R. Mahmoud, N. Sani and F. Daryabar, “Digital Forensic Trends and Future” Ijcsdf, 2013.
F. Daryabar, A. Dehghantanha, N. Udzir, N. Sani, S. Shamsuddin and F. Norouzizadeh, “A Survey on Cloud Computing and Digital Forensics” J. Next Gen. Info. Tech, 2013.

M. Damshenas, A. Dehghantanha, R. Mahmoud and S. Shamsuddin, “Cloud Computing and Conflicts with Digital Forensic Investigation” Int. J of Dig. Cont. Tech. And its App. pp. 543, May, 2013.

F. Daryabar, A. Dehghantanha, H.G. Broujerdi, “Investigation of malware defence and detection techniques” International Journal of Digital Information and Wireless Communications (JDIWC), Vol 1, Issue 3, pp. 645-650, 2011.

F. Daryabar, A. Dehghantanha, N.I. Udzir, “Investigation of bypassing malware defences and malware detections”, Information Assurance and Security (IAS), 2011 7th International Conference on, pp. 173-178, 2011.

Kaveh Shaerpour, A. Dehghantanha, “Trends in android malware detection”, Journal of Digital Forensics, Security and Law, Vol 8, Issue 3, 2013

F. Norouzizadeh Dezfouli, A. Dehghantanha, R. Mahmoud, N.I. Mohd Sani, S.B. Shamsuddin, F. Daryabar, “A Survey on Malware Analysis and Detection Techniques”, International Journal of Advancements in Computing Technology, Vol. 5, Issue 14, 2013.

M. Damshenas, A. Dehghantanha, R. Mahmoud, “A survey on malware propagation, analysis, and detection”, International Journal of Cyber-Security and Digital Forensics (IJCSDF), Vol 2, Issue 4, pp. 10-29, 2013.

S. Mohtasebi, A. Dehghantanha, H. G. Broujerdi, “Smartphone Forensics: A Case Study with Nokia E5-00 Mobile Phone”, International Journal of Digital Information and Wireless Communications (JDIWC), Vol 1, Issue 3, pp. 651-655, 2011

S. Parvez, A. Dehghantanha, H. G. Broujerdi, “Framework of digital forensics for the Samsung Star Series phone”, Electronics Computer Technology (ICECT), 2011 3rd International Conference on, 2011.

S. Mohtasebi, A. Dehghantanha, “Defusing the Hazards of Social Network Services”, International Journal of Digital Information and Wireless Communications (JDIWC), Vol. 1, Issue 2, pp. 504-515, 2011.

F.N. Dezfooli, A. Dehghantanha, B. Eterovic-Soric, K.K.R Choo, “Investigating Social Networking applications on smartphones detecting Facebook, Twitter, LinkedIn and Google+ artefacts on Android and iOS platforms”, Australian Journal of Forensic Sciences [in-print], 2015.

K. Shaerpour, A. Dehghantanha, R. Mahmood, “A Survey On Cyber-Crime Prediction Techniques”, International Journal of Advancements in Computing Technology, Vol 5, Issue 14, 2013.