Cloud Storage Forensic Analysis

Darren Quick
Graduate Diploma in Science (Information Assurance)
University of South Australia

Thesis submitted to the University of South Australia
School of Computer & Information Science
in partial fulfilment of the
requirements for the degree of
Master of Science (Cyber Security and Forensic Computing)

Supervisor: Dr Kim-Kwang Raymond Choo

Adelaide, South Australia
28 October 2012
Chapter Guide

Chapter 1 – Introduction ........................................................................................................ 1
Chapter 2 – Literature Review .............................................................................................. 9
Chapter 3 – Research Methodology ..................................................................................... 26
Chapter 4 – Digital Forensic Analysis Cycle ....................................................................... 39
Chapter 5 - Dropbox Analysis ............................................................................................... 48
Chapter 6 - Microsoft SkyDrive Analysis ............................................................................ 76
Chapter 7 - Google Drive Analysis ...................................................................................... 116
Chapter 8 - Forensic Collection of Cloud Storage Data ....................................................... 150
Chapter 9 – Summary ........................................................................................................... 167
References ............................................................................................................................ 181
Appendix A – Legislation ...................................................................................................... 187
Appendix B – Example of Examination Report .................................................................... 194
Appendix C – ACPO and NIJ Guidelines ............................................................................ 196
Appendix D – Publications submitted for consideration ..................................................... 198
# Table of Contents

Chapter Guide ........................................................................................................... ii  
Table of Contents ..................................................................................................... iii  
List of Figures .......................................................................................................... v  
List of Tables ........................................................................................................... vi  
List of Abbreviations ............................................................................................... xi  
Abstract ................................................................................................................... xiv  
Declaration ............................................................................................................... xvi  
Acknowledgements ................................................................................................. xvii  

## Chapter 1 – Introduction ......................................................................................... 1  
1.1 Background ........................................................................................................... 1  
1.2 Motivation ............................................................................................................ 3  
1.3 Research Objectives .............................................................................................. 5  
1.4 Thesis Structure ................................................................................................... 6  

## Chapter 2 – Literature Review ................................................................................ 9  
2.1 Cloud computing and cloud storage .................................................................... 9  
2.2 Digital investigations ........................................................................................... 12  
2.3 Cloud storage and digital investigations ............................................................... 13  
2.3.1 Mobile device analysis ..................................................................................... 16  
2.3.2 The need for common procedures ..................................................................... 16  
2.3.3 Current research ............................................................................................... 18  
2.4 Issues at each stage of a forensic investigation ...................................................... 19  
2.4.1 Identification of data ......................................................................................... 19  
2.4.2 Preservation of data ........................................................................................ 20  
2.4.3 Analysis of data ............................................................................................... 20  
2.4.4 Presentation of data ......................................................................................... 21  
2.5 Additional considerations ..................................................................................... 22  
2.6 Summary .............................................................................................................. 23  

## Chapter 3 – Research Methodology ...................................................................... 26  
3.1 Research problem ................................................................................................. 26  
3.2 Research purpose ................................................................................................. 27  
3.3 Research questions ............................................................................................... 27  
3.3.1 Research Question 1 ....................................................................................... 27  
3.3.2 Research Question 2 ....................................................................................... 28
3.4 Research method .................................................................................................................................. 29
  3.4.1 Research Question 1 Experiment Process ......................................................................................... 30
  3.4.2 Research Question 2 Experiment Process ......................................................................................... 35
3.5 Research equipment ............................................................................................................................... 36
3.6 Research limitations ............................................................................................................................... 37
3.7 Summary ............................................................................................................................................. 38

Chapter 4 – Digital Forensic Analysis Cycle ............................................................................................... 39
  4.1 Introduction ......................................................................................................................................... 39
  4.2 Proposed Digital Forensic Analysis Cycle ........................................................................................... 40
    4.2.1 Commence (Scope) ......................................................................................................................... 42
    4.2.2 Prepare and Respond ..................................................................................................................... 42
    4.2.3 Identify and Collect ......................................................................................................................... 43
    4.2.4 Preserve (Forensic Copy) ................................................................................................................ 44
    4.2.5 Analysis ........................................................................................................................................ 45
    4.2.6 Presentation .................................................................................................................................. 45
    4.2.7 Feedback ...................................................................................................................................... 46
    4.2.8 Complete ...................................................................................................................................... 46
  4.3 Applying the cycle ............................................................................................................................... 46
  4.4 Summary ............................................................................................................................................. 47

Chapter 5 - Dropbox Analysis .................................................................................................................... 48
  5.1 Introduction ......................................................................................................................................... 48
  5.2 Dropbox Analysis: Windows 7 Computer Environment .................................................................... 49
    5.2.1 Commence (Scope) ......................................................................................................................... 49
    5.2.2 Prepare and Respond ..................................................................................................................... 50
    5.2.3 Identify and Collect ......................................................................................................................... 51
    5.2.4 Preserve (Forensic Copy) ................................................................................................................ 51
    5.2.5 Analysis ........................................................................................................................................ 51
    5.2.6 Presentation .................................................................................................................................. 61
    5.2.7 Feedback ...................................................................................................................................... 64
    5.2.8 Complete ...................................................................................................................................... 64
  5.3 Dropbox Analysis: Apple iPhone 3G ................................................................................................. 65
    5.3.1 Commence (Scope) ......................................................................................................................... 65
    5.3.2 Prepare .......................................................................................................................................... 65
    5.3.3 Identify and Collect ......................................................................................................................... 66
    5.3.4 Preserve (Forensic Copy) ................................................................................................................ 66
    5.3.5 Analysis ........................................................................................................................................ 66
    5.3.6 Presentation .................................................................................................................................. 67
List of Figures

Figure 1 - NIST Cloud Service Model (adapted from Lumley (2010)).......................... 10

Figure 2 - Block diagram of Research scope .................................................................. 31

Figure 3 - Proposed Digital Forensic Analysis Cycle..................................................... 41

Figure 4 - Folder list of the AppData\local\SkyDrive folder (X-Ways 16.5)............... 81

Figure 5 - Thumbcache picture for Enron jpg file (X-Ways 16.5)................................. 92

Figure 6 - Dropbox web page in Mozilla Firefox Browser .............................................. 154

Figure 7 - Google Drive web page in Mozilla Firefox Browser ....................................... 157

Figure 8 - Microsoft SkyDrive web page in Mozilla Firefox Browser ......................... 159
## List of Tables

Table 1 – FBI and SAPOL data extrapolated to forecast growth in digital evidence................................................. 14
Table 2 - Configurations of Virtual PCs.................................................................................................................... 32
Table 3 - Data preservation steps for Cloud Storage.................................................................................................. 36
Table 4 - Configuration of Host PC.......................................................................................................................... 37
Table 5 - Dropbox Windows software files with MD5 values..................................................................................... 53
Table 6 - Output from RegRipper for Dropbox........................................................................................................ 56
Table 7 - Dropbox IP’s and registered organisation details........................................................................................ 59
Table 8 - Data observed in client software memory (VMEM)..................................................................................... 61
Table 9 - Summary of Dropbox analysis findings....................................................................................................... 63
Table 10 - Output from .XRY Web-History.txt file #2 ................................................................................................. 67
Table 11 - Output from .XRY Web-History.txt file #14 .............................................................................................. 67
Table 12 - Example of Timeline (Dropbox Case Study)................................................................................................. 72
Table 13 - SkyDriveSetup information (using X-Ways 16.5)...................................................................................... 80
Table 14 - SkyDrive executable information (using X-Ways 16.5)............................................................................. 81
Table 15 - example of SyncDiagnostics.log file contents.......................................................................................... 82
Table 16 - OwnerID INI file contents.......................................................................................................................... 83
Table 17 - SkyDrive.EXE Prefetch (X-Ways 16.5).......................................................................................................... 88
Table 18 - Wordpad.EXE Prefetch (X-Ways 16.5).......................................................................................................... 88
Table 19 - Notepad.EXE Prefetch (X-Ways 16.5)............................................................................................................ 89
Table 20 - SkyDrive.lnk Link file contents (X-Ways 16.5)............................................................................................. 90
Table 21 - Enron3111.lnk Link file contents (X Ways 16.5).......................................................................................... 90
Table 22 - Thumbcache information for Enron jpg file (X-Ways 16.5)................................................................. 91
Table 23 - Output from RegRipper for SkyDrive........................................................................................................ 93
Table 24 - $Recycle.Bin $I information for deleted SkyDrive folder (X-Ways 16.5) ...
94
Table 25 - SkyDrive IP addresses observed in network traffic .................................... 96
Table 26 - Unencrypted data observed in network traffic ............................................. 98
Table 27 - Summary of SkyDrive analysis findings ..................................................... 104
Table 28 - Output from .XRY Web-History.txt file #6 ............................................... 107
Table 29 - Output from .XRY Web-History.txt file #5 ............................................... 107
Table 30 - Example of Timeline (SkyDrive Case Study) ............................................. 113
Table 31 - example of snapshot.db SQLite file contents for enron.jpg file ................. 121
Table 32 - Output from RegRipper for Google Drive .................................................. 128
Table 33 - $Recycle.Bin $I data for deleted Google Drive folder (X-Ways 16.5)....... 129
Table 34 - example of snapshot.db SQLite file contents and IEF output ................... 130
Table 35 - Google Drive IP addresses observed in network traffic ............................ 132
Table 36 - Summary of Google Drive analysis findings .............................................. 139
Table 37 - Google Drive ‘snapshot.db’ file contents (abbreviated) ........................... 159
Table 38 - Microsoft SkyDrive ‘syncdiagnostics.log’ file contents (abbreviated) ...... 161
Table 39 - File Dates and Times for Dropbox, Google Drive, and Microsoft SkyDrive
downloaded files using a Browser compared with using Client Software
(Encase 6.19.4) ........................................................................................................ 163
Table 40 - File Dates and Times for Dropbox, Google Drive, and Microsoft SkyDrive
downloaded files using a Browser compared with using Client Software (X-
Ways 16.5 and FTK 1.81.6) ......................................................................................... 163
Table 41 - File Date and Time nomenclature: MFT Parser, FTK, X Ways, and Encase.
.................................................................................................................................. 164
Table 42 - Client software log files and databases ...................................................... 173
Table 43 - Keyword search terms ............................................................................ 173

x
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$MFT$</td>
<td>Windows Master File Table</td>
</tr>
<tr>
<td>ACPO</td>
<td>Association of Chief of Police Officers</td>
</tr>
<tr>
<td>AD1</td>
<td>AccessData Logical Evidence File</td>
</tr>
<tr>
<td>AGIMO</td>
<td>Australian Government Information Management Office</td>
</tr>
<tr>
<td>AS</td>
<td>Apple Safari</td>
</tr>
<tr>
<td>CIO</td>
<td>Chief Information Officer</td>
</tr>
<tr>
<td>CTR</td>
<td>X-Ways Evidence File Container</td>
</tr>
<tr>
<td>DBAN</td>
<td>Dariks’ Boot and Nuke</td>
</tr>
<tr>
<td>DFRWS</td>
<td>Digital Forensics Research Workshop</td>
</tr>
<tr>
<td>DOCX</td>
<td>Windows Document Format</td>
</tr>
<tr>
<td>DoD</td>
<td>Department of Defence</td>
</tr>
<tr>
<td>E01</td>
<td>Encase Physical Evidence Format</td>
</tr>
<tr>
<td>EXIF</td>
<td>Exchangeable image file format</td>
</tr>
<tr>
<td>EXT3</td>
<td>Linux Extended File System</td>
</tr>
<tr>
<td>FCS</td>
<td>Forensic Computer Section</td>
</tr>
<tr>
<td>FF</td>
<td>Mozilla Firefox</td>
</tr>
<tr>
<td>FTK</td>
<td>Forensic Tool Kit</td>
</tr>
<tr>
<td>GC</td>
<td>Google Chrome</td>
</tr>
<tr>
<td>HFS+</td>
<td>Apple Hierarchical File System</td>
</tr>
<tr>
<td>HR</td>
<td>Human Resources</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HTML</td>
<td>Hypertext Markup Language</td>
</tr>
<tr>
<td>IaaS</td>
<td>Infrastructure as a Service</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>IE</td>
<td>Microsoft Internet Explorer</td>
</tr>
<tr>
<td>IEF</td>
<td>Internet Evidence Finder</td>
</tr>
<tr>
<td>IO</td>
<td>Investigating Officer</td>
</tr>
<tr>
<td>iOS</td>
<td>Apple iPhone Operating System</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>ISP</td>
<td>Internet Service Provider</td>
</tr>
<tr>
<td>IT</td>
<td>Information Technology</td>
</tr>
<tr>
<td>JPG</td>
<td>Joint Picture Group</td>
</tr>
<tr>
<td>L01</td>
<td>Encase Logical Evidence Format</td>
</tr>
<tr>
<td>LAN</td>
<td>Local Area Network</td>
</tr>
<tr>
<td>MD5</td>
<td>Message Digest</td>
</tr>
<tr>
<td>NIJ</td>
<td>National Institute of Justice</td>
</tr>
<tr>
<td>NIST</td>
<td>National Institute of Standards and Technology</td>
</tr>
<tr>
<td>OS</td>
<td>Operating System</td>
</tr>
<tr>
<td>OSX</td>
<td>Apple Operating System</td>
</tr>
<tr>
<td>OWADE</td>
<td>Offline Windows Analysis and Data Extraction</td>
</tr>
<tr>
<td>PaaS</td>
<td>Platform as a Service</td>
</tr>
<tr>
<td>PC</td>
<td>Personal Computer</td>
</tr>
<tr>
<td>PCAP</td>
<td>Network traffic capture file</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>PLIST</td>
<td>Property List</td>
</tr>
<tr>
<td>RAM</td>
<td>Random Access Memory</td>
</tr>
<tr>
<td>RTF</td>
<td>Rich Text Format</td>
</tr>
<tr>
<td>SaaS</td>
<td>Software as a Service</td>
</tr>
<tr>
<td>SAPOL</td>
<td>South Australia Police</td>
</tr>
<tr>
<td>SHA</td>
<td>Secure Hash Algorithms</td>
</tr>
<tr>
<td>SOP</td>
<td>Standard Operating Procedure</td>
</tr>
<tr>
<td>SSH</td>
<td>Secure Shell</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>TOR</td>
<td>The Onion Router</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>URL</td>
<td>Uniform Resource Locator</td>
</tr>
<tr>
<td>USB</td>
<td>Universal Serial Bus</td>
</tr>
<tr>
<td>USGAO</td>
<td>US Government Accountability Office</td>
</tr>
<tr>
<td>VM</td>
<td>Virtual Machine</td>
</tr>
<tr>
<td>VMDK</td>
<td>Virtual Machine Disk</td>
</tr>
<tr>
<td>VMEM</td>
<td>Virtual Memory File</td>
</tr>
<tr>
<td>XRY</td>
<td>Forensic software from Microsystemation for mobile device analysis</td>
</tr>
</tbody>
</table>
Abstract

Cloud storage has been identified as an emerging challenge to digital forensic researchers and practitioners in a range of literature. There are various types of cloud services with each type having a potentially different use in criminal activity. One area of difficulty is the identification and acquisition (also known as “preservation” in the digital forensic community) of potential data when disparate services can be utilized by criminals. Without knowing where data may reside can impede an investigation, as it could take considerable time to contact all potential service providers to determine if data is being stored within their cloud service. There is a need for a sound digital forensic framework relating to the forensic analysis of client devices to identify potential data holdings. In this thesis a digital forensic framework is proposed which expands on current methodologies.

Using popular cloud storage services; Dropbox, Microsoft SkyDrive, and Google Drive, the proposed framework was applied when undertaking research into the data remnants on a Windows 7 computer and an Apple iPhone 3G. A variety of circumstances were examined, including a variety of methods to store, upload, and access data in the cloud. By determining the data remnants on client devices, this research contributes to a better understanding of the types of artifacts that are likely to remain for digital forensics practitioners and examiners at the ‘Identification’ stage of an investigation. Potential information sources identified include client software log files, prefetch files, link files, network traffic capture and memory captures. Of note is that it was possible to locate the username and password in cleartext for the three services examined.

Once it is determined that a cloud storage account has potential evidence of relevance to an investigation, an examiner can communicate this to legal liaison points within service providers to enable them to respond and secure evidence in a timely manner. In addition, a jurisdiction which has legal provisions to collect data available to a computer or device, the process may involve accessing an account to collect the data. This research explores the process of collecting data from a cloud storage account using a browser, downloading files using client software, and comparing these with the original files.
Analysis of the resulting data determined that there were no changes to the contents of files during the process of upload, storage, and download to three popular cloud storage providers; Dropbox, Google Drive, and Microsoft SkyDrive. However, the timestamps of the files were changed, and this needs to be considered when forming assumptions in relation to the created, modified, or accessed times attached to files downloaded via a browser and via client software. Timestamp information may be a crucial aspect of an investigation, prosecution, or civil action, and therefore it is important to record the information available, and to understand the circumstances relating to a timestamp on a file.
Declaration

I declare that:

This thesis presents work carried out by myself and does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; to the best of my knowledge it does not contain any materials previously published or written by another person except where due reference is made in the text; and all substantive contributions by others to the work presented, including jointly authored publications, is clearly acknowledged.

Signed: Darren Quick Date: 28 October 2012
Acknowledgements

This research and thesis was undertaken through the Information Assurance Research Group, School of Computer & Information Science, University of South Australia. Support was received from a range of people, and I take this opportunity to thank you all. If I have missed anyone, be comforted by the fact you have helped me.

The understanding of my wife, Ruth, and daughter, Sami, through all those times I was in my study working on this research is very much appreciated. The holiday to Point Turton was a welcome break, and the walks around the block with Bubbles all served to keep me sane. I also thank Brendon for perfect timing with pizza and movies.

Without the tireless assistance of my UniSA supervisor Dr Kim-Kwang Raymond Choo this research and thesis would not have been possible. His availability in the evenings, on weekends, public holidays, and when he was overseas was unsurpassed, and his vast knowledge and experience helped to guide my research and achieve more than I set out to at the beginning of this year.

In previous years, the assistance of Dr Elena Sitnikova in refining my research abilities, methods, and writing, set the base for which I was able to undertake this research. In addition, all the UniSA Information Assurance lecturers and guest supervisors provided invaluable training and guidance. In particular, Professor Jill Slay and Detective Senior Sergeant Barry Blundell were instrumental in establishing this program, and seeing it grow to where it is today.

I am eternally grateful for the tireless efforts of Detective Chief Superintendent Jim Jeffery in getting me back to SAPOL as an Electronic Evidence Specialist, thank you! Matt Simon provided much needed feedback in relation to the Dropbox paper. To all my colleagues at SAPOL and ECS, thank you for your support.

I must also thank; Iron Maiden, Star Trek, Community, MythBusters, and the other DVDs I had playing whilst undertaking this research. Finally, I would like to thank Bill Phillips, what more can I say; thank you!

The views and opinions expressed in this thesis are those of the author alone and not the organisations with whom the author has been associated or supported. This thesis was hosted, and some parts written, using cloud storage.
Chapter 1 – Introduction

1.1 Background

Cloud computing is a relatively recent term to describe computer resources available as a service accessible over a network, such as internally to a corporation, or externally available over the Internet. The definition introduced by the National Institute of Standards and Technology is that: ‘Cloud computing is 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’ (Mell & Grance 2011). There are a range of different cloud services, which are grouped as; infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) (Taylor et al. 2011). Cloud storage, or file hosting, is the...
storage of electronic data on remote infrastructure, rather than local storage which is attached to a computer or electronic device.

In recent years, there has been a marked increase in the adoption of cloud computing (AGIMO 2011; Gantz 2012). A study sponsored by Microsoft, for example, estimated that ‘spending on public cloud IT services in 2011 was [US]$28 billion, compared with more than [US]$1.7 trillion in spending on total IT products and services — it already has begun changing how IT delivers economic value to countries, cities, industries, and small businesses’ (Gantz 2012). A recent audit by the US Government Accountability Office found that ‘[e]ach of the seven agencies [selected for audit] has also identified opportunities for future cloud implementations, such as moving storage and help desk services to a cloud environment’ (USGAO 2012). A report by Gartner predicts that ‘[e]merging cloud services will become the glue that connects the web of devices that users choose to access during the different aspects of their daily life’ (Kleynhans 2012). The availability of cloud storage services is becoming a popular option for consumers, business, and government to store data that is accessible via a range of devices, such as personal computers, tablets, and mobile phones.

Cloud storage is a popular option for users to store their data and be able to access it via a range of internet connected devices. A report by Gartner, for example, has highlighted the trend in client computing is shifting from a focus on personal computers (PCs) to include portable devices, and that personal cloud storage will replace PCs as the main method of users’ storage by 2014 (Kleynhans 2012).

There are a range of cloud storage hosting providers, and many offer free cloud storage services; such as Dropbox™, Microsoft® SkyDrive® and Google Drive™. Accessing the various cloud storage services can be undertaken in a variety of ways; a user can install client software on a personal computer (PC), use a web browser to access the cloud storage service, or use a browser on a mobile portable device, such as a web-enabled mobile phone or tablet.

Cloud storage is used by criminals to store and distribute data, such as child exploitation material, terrorism-related material, and illicit drug material. Cloud storage offers criminals an ability to avoid the scrutiny of law enforcement and national security agencies, and to provide for difficulty in attributing ownership or an association with illicit data (Biggs & Vidalis 2009). Cloud storage can also be subject to attacks by cyber
criminals, who may be able to hijack and use resources for criminal purposes, thus adding to the challenge of growing volumes of digital evidence in cases under investigation.

1.2 Motivation

The use of cloud computing by criminals or their victims means that data of interest may be virtualised, geographically distributed, and transient. This presents technical and jurisdictional challenges for identification and seizure by law enforcement and national security agencies, which can impede digital forensic investigators and potentially prevent agencies from acquiring digital evidence and forensically analysing digital content in a timely fashion (Taylor et al. 2010). Data fragmentation and distribution across the globe within numerous datacenters raises corresponding legal and jurisdiction issues.

Digital forensic analysis is the process of examining electronic evidence for legal purposes, such as criminal investigations or civil disputes. Taylor et al. (2011) explained that ‘in legal terms, cloud computing systems will make it potentially more difficult for the computer forensic analyst to acquire and analyse digital evidence to the same standards as that currently expected for traditional server-based systems, due to the difficulty in establishing what data was stored or processed by what software on what specific computing device’. One issue facing forensic examiners is the identification of service providers and accounts, including usernames and passwords. The analysis of user devices, such as computer hard drives, network traffic, or mobile devices such as an Apple iPhone or Android mobile phone, may provide this information.

The timely identification of the use of cloud storage is a concern for law enforcement and investigation agencies. Current legislation, such as in Australia, allows for remote data to be seized at the time of executing a warrant. An on-scene triage process may not identify cloud stored data, which may come to light with subsequent forensic analysis, potentially months later. By that time the data may have been moved, or become unavailable for seizure, therefore there is a need to identify cloud stored data in a timely manner.
In addition, another challenge ‘is to have the data translated into a form that is acceptable as evidence to the courts …. Assuming that the fragile and elusive evidence can be gathered together, the prosecutor must keep in mind that he or she will one day need to be able to prove the chain of evidence. All processes will need to appropriately documented in a way that can be understood by the layman and the prosecutor must be prepared if necessary to demonstrate that the ‘original’ digital material has not been changed or tampered with in any way’ (Pallaras 2011). Therefore to reduce the risk of digital (forensic) evidence being called into question in judicial proceedings, it is important to have a rigorous methodology and set of procedures for conducting digital forensic investigations and examinations that are flexible enough to be able to work with future providers offering new cloud storage services.

For an examiner, it is also important to have a contemporary understanding of the location and type of data remnants left behind by cloud storage users on the devices they use to access their data (Guo, Shang & Jin 2012). The timely identification of potential data sources is an area that can impede an investigation. An investigation will stall if there is a requirement to contact every cloud service provider to determine whether data of relevance is held with the provider. This further demonstrates a need to be able to quickly identify if suspected data is stored remotely, and identify any associated user accounts and service providers.

After cloud stored data is identified, investigators need to undertake legal processes to secure data via service providers. In a digital forensic investigation, the steps of preservation and analysis follow after identification. Preservation can potentially be problematic due to delays or jurisdictional issues. As a hypothetical example, if an investigator in Australia discovers a suspect has data stored with a cloud storage provider based in the United States of America (USA), and the data is spread across multiple countries in the companies storage facilities, the process of securing and gaining access to the data relies upon communication and legal processes across the various countries, with different legal system requirements having to be met along the way.

Some jurisdictions have legal provisions which allow for data to be secured at the time of serving a warrant, i.e. at the time of search and seizure. In Australia, Section 3LA of the Crimes Act 1914 (Cth), has a provision for the executing officer of a search and
seizure warrant to access data which includes data not held at the premises, i.e. accessible from a computer or data storage device. This may not be the case in a number of jurisdictions, including the USA.

Current methods available to collect data from cloud storage accounts include via a browser or using client software to download data. This can have cause for concern, as interaction with identified data needs to be done in accordance with common guidelines to ensure acceptance in court, such as the Association of Chief of Police Officers (ACPO) ‘Good Practice Guidelines for Computer Based Evidence V4.0’(ACPO 2006), the National Institute of Justice (NIJ) ‘Electronic Crime Scene Investigation: A Guide for First Responders, Second Edition’(NIJ 2008), and Standards Australia ‘Handbook 171-2003 Guidelines for the Management of IT Evidence’(SAI 2003). (A summary of the guidelines is included in Appendix C).

The motivation for conducting research into cloud storage forensic identification, preservation, analysis, and presentation can be summarised in the following three points; Cloud storage is increasingly being used by consumers, businesses, and government users to store growing amounts of data, which can be accessed with computers or mobile electronic devices. Criminals are embracing the opportunity to store illicit data in cloud file hosting services, which contributes to difficulties in proving ownership and interaction. Investigations can stall if identification and preservation of potential evidence is not able to be undertaken, or this is not able to be done in a timely manner.

1.3 Research Objectives

The focus of this research is to discover whether there are any cloud storage data remnants on popular client devices, for the purposes of this research; a computer running Windows 7, and an Apple iPhone running iOS 4.2.1. When commencing this research it was determined that there was a need to have a framework to guide the investigation process, and hence a framework was developed and assessed using the practical research to examine the benefits of using a framework when undertaking forensic analysis of cloud remnants. This was built upon the guidelines for digital forensic analysis to provide for a process which follows a common digital forensic
examination, enabling forensic examiners to apply the framework to real-world investigations, and also serve to research the data remnants using a real-world process of analysis.

To achieve the aims of the research, the following research objectives are defined;

**Objective 1:** To examine current research published in literature relating to cloud storage and identified cloud storage analysis methodologies.

**Objective 2:** To develop a digital forensic analysis framework that will assist practitioners, examiners, and researchers follow a standard process when undertaking forensic analysis of cloud storage services.

**Objective 3:** To conduct research using popular cloud storage services; Dropbox, Microsoft SkyDrive, and Google Drive, and determine whether there are any data remnants which assist digital forensic analysis and investigations.

**Objective 4:** To examine the forensic implications of accessing and downloading cloud stored data from popular cloud storage services; Dropbox, Microsoft SkyDrive, and Google Drive.

By determining the data remnants the aim is to provide a better understanding of the type of artifacts that are likely to remain, and the access point(s) for digital forensics examiners to assist in the ‘identification’, ‘preservation’, ‘analysis’, and ‘presentation’ stages of an investigation.

### 1.4 Thesis Structure

This thesis consists of an introduction and follows with the chapters that describe the research in detail, and summarise the findings. The thesis also includes an overall summary, acknowledgements, table of contents, list of figures, tables, and a glossary of common terms.

Chapter one introduces the overall topic, including background information regarding cloud storage, forensic analysis, and the issues for investigators. The motivation for conducting the research is discussed, and the objectives are listed. The structure of the thesis is then outlined.
Chapter two examines current literature focusing on cloud forensics. The first section outlines cloud computing, digital forensic analysis, and cloud storage implications. Issues relating to identification, preservation, analysis, and presentation are outlined. Additional issues are outlined, and a summary concludes the chapter.

Chapter three outlines the research methodology used in this thesis. The research problem, purpose, questions and hypotheses are outlined. The two main research questions and associated hypotheses are discussed, and the research methodology for each question is detailed. Finally, the limitations of the research are explained.

Chapter four outlines the proposed Digital Forensic Analysis Cycle and how this can be applied to forensic analysis of cloud storage. Each step of the cycle is explained; Commence (Scope), Prepare and Respond, Identify and Collect, Preserve (Forensic Copy), Analyse, Present, Feedback, and Complete.

Chapters five, six, and seven describes the application of the framework to the analysis of three popular cloud storage services; Dropbox, Microsoft SkyDrive, and Google Drive. In each case, the data remnants on a Windows 7 computer are first examined using the proposed framework. Next, an Apple iPhone is examined to further assess the framework and to determine the data remnants on a popular portable device. A Case Study is presented to outline the application of the research findings. A summary for each chapter highlights the relevant information for each provider.

Material presented in the chapters four to seven were submitted for publication and are currently under peer review:


Chapter eight addresses issues and challenges associated with the preservation of data identified in the previous chapters. Firstly the forensic implications of accessing evidence are discussed, and then the process is outlined. The research is described for each of the three popular cloud storage providers; Dropbox, Microsoft SkyDrive, and
Google Drive. This is undertaken by accessing live data via a browser, and then using the client software of each provider. Both methods are then compared for forensic suitability.

Material presented in chapter eight was submitted for publication and is currently under peer review:


Chapter nine summarises the overall thesis. The benefits of the proposed framework are outlined, and then the findings in relation to analysis of Dropbox, Microsoft SkyDrive, and Google Drive are discussed. The research findings in relation to preservation of cloud stored data are then summarised. The original objectives are listed, and the method used to answer each question is explained. Areas for future research are then summarised.
This chapter examines current literature focusing on cloud storage and digital forensic investigation. The first section outlines cloud computing and cloud storage. The second and third sections discuss digital forensic analysis and cloud storage implications. Issues relating to identification, preservation, analysis, and presentation are outlined. Additional issues are discussed, and a summary concludes the chapter.

Chapter 2 – Literature Review

2.1 Cloud computing and cloud storage
The definition proposed by the National Institute of Standards and Technology (NIST) for cloud computing encompasses convenient, on-demand network access to shared configurable computing resources that can be rapidly provisioned and released with minimal management (Mell & Grance 2011). There are a range of different cloud services, which are grouped as; infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) (Taylor et al. 2011). Infrastructure as a Service
consists of the rental and use of hardware, with the user deploying an operating system and applications. Platform as a Service consists of the use of a supplied operating system, and the user deploying the software to run on virtually provisioned systems. Software as a service consists of the use of provided software over a network connection on a subscription basis, with limited control by the user (Lumley 2010). Figure 1 from Lumley (2010) shows the different levels of user control in relation to the various cloud computing services. NIST also describe four deployment models for cloud services; private (within an organisation), community (shared by organisations with common concerns), public (available to the general public), and a hybrid (composed of a mix of private, community, or public) (Mell & Grance 2011).

<table>
<thead>
<tr>
<th>Service</th>
<th>Cloud provider</th>
<th>Layer</th>
<th>Cloud customer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure as a Service (IaaS)</td>
<td>No control</td>
<td>Application</td>
<td>Total control</td>
</tr>
<tr>
<td></td>
<td>Admin control</td>
<td>Middleware</td>
<td>No control</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Operating System</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hypervisor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hardware</td>
<td></td>
</tr>
<tr>
<td>Platform as a Service (PaaS)</td>
<td>Admin control</td>
<td>Application</td>
<td>Limited programmability</td>
</tr>
<tr>
<td></td>
<td>Total control</td>
<td>Middleware</td>
<td>No control</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Operating System</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hardware</td>
<td></td>
</tr>
<tr>
<td>Software as a Service (SaaS)</td>
<td>Admin control</td>
<td>Application eg Docs</td>
<td>Limited admin control</td>
</tr>
<tr>
<td></td>
<td>Total control</td>
<td>Middleware</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Operating System</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hardware</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 - NIST Cloud Service Model (adapted from Lumley (2010))

Watson (2009) states that the origins of cloud computing are from the Advanced Research Projects Agency (ARPA) ‘Intergalactic Network’ concept from the 1960’s. During the 1960’s and 1970’s there was a growth and reliance on mainframe computers with access via dumb terminals. In the 1980’s, personal computers introduced by IBM put local storage and processing onto users desks. In the 1990’s, HTML combined with the Internet moved towards a potential implementation of the Intergalactic Network concept. According to Desisto, Plummer and Smith (2008), services which could be considered as SaaS offerings first appeared in the late 1990’s, although these services
were not discussed in ‘cloud’ terms. A cloud icon is commonly used in network diagrams to represent general network infrastructure, and from the mid 1990’s was often used to represent the Internet. Ultimately, cloud computing is a new model that offers an economic solution to the increasing need for IT infrastructure (Reilly, Wren & Berry 2011).

Cloud storage provides users with virtual storage space to host documents, pictures, music, and other files (Tadjer 2010). Some services also offer the ability to work with the stored data, such as; editing documents, viewing pictures, or playing music files (i.e. Google Docs or Microsoft SkyDrive). According to Chung et al. (2012) of the various cloud services, consumers mostly use storage as a service, which is available to client computers and portable devices.

Cloud storage can be used by criminals to store illicit data, and provide a distribution point which distances the owner or users from the illicit data. The latter can include a wide range of material, such as child exploitation material, terrorism-related material, or illicit drug material. Cloud storage offers criminals an ability to distance themselves from the service, and hence avoid the scrutiny of law enforcement and national security agencies. This also serves to provide a difficulty in attributing ownership or an association with illicit data (Biggs & Vidalis 2009). Cloud storage can also be subject to attacks by cyber criminals, who may be able to gain access to a victim’s account for the data contained therein, or hijack an account to use the resource for criminal purposes, such as distributing illicit data, thus increasing the challenge of investigation cyber-crimes or traditional crimes carried out in the cyber domain.

Reilly, Wren and Berry (2011) conclude that security issues in relation to cloud computing are well addressed, and were so from the outset, as security is a well understood aspect of IT applications. On the other hand, the issues of forensic readiness, forensic response, and forensic analysis, are not as well understood, and hence were not initially considered in relation to cloud computing. Birk (2011) observes that ‘current discussion centers mostly around security, privacy and data protection issues’ and that forensic considerations have so far been neglected. Taylor et al. (2011) also raise the issue that each of the cloud service models, PaaS, IaaS, and SaaS, pose different challenges to a forensic examiner.
2.2 Digital investigations

Forensic computer analysis is a recent discipline which has resulted from an increase in the use of computers and storage devices by criminals (Reilly, Wren & Berry 2011). The application of forensic computer practices has been around for several decades, and now faces challenges on a variety of fronts; the increase in data required to be analysed, encryption of data, the increase in the use of mobile devices with different operating and storage systems, an increase in users storing data via Internet enabled applications (cloud), and the increase in the use of social media to store and disseminate information.

A widely accepted definition of forensic computer analysis was outlined by McKemmish (1999) and consists of the process of ‘identifying, preserving, analyzing and presenting digital evidence in a manner that is legally acceptable’. The United Kingdom Association of Chief Police Officers (ACPO) Guidelines outline four principles for forensic computer analysis; no action should change data, when it is necessary to access original data the persons accessing data should be competent to do so, a record of processes should be made, and the investigator in charge is responsible to ensure the principles are adhered to (ACPO 2006). The US National Institute of Justice has two publications addressing forensic computer analysis and first response (NIJ 2004, 2008). They outline principles to adhere to which include; any actions should not affect the integrity of data, persons undertaking analysis should be trained, activity should be documented (NIJ 2004). RFC 3227, Guidelines for Evidence Collection and Archiving, also outlines steps to undertake (Brezinski & Killalea 2002). Australian Standard HB 171:2003, Guidelines for the Management of IT Evidence also addresses forensic response and analysis (SAI 2003).

Digital evidence can include a wide range of potential data. Files can be stored on a computer hard drive (at rest), evidential data may be briefly located in memory (in execution), which disappears when power is disconnected, and data can be in-transit across a network (in motion) (Birk 2011; Reilly, Wren & Berry 2011).

Digital evidence must also satisfy the same requirements as other evidence, in that it must be; authentic, reliable, complete, believable, and admissible (Reilly, Wren & Berry 2011). In a cloud environment, the basic forensic principles and processes still apply (Guo, Shang & Jin 2012).
2.3 Cloud storage and digital investigations

It is recognized that law enforcement agencies have a need to access data stored in the cloud (Mason & George 2011). Service agreements of some providers may outline the process for forensic investigations, including the jurisdiction the provider is operating under, and the process relating to data seizure (Reilly, Wren & Berry 2011). The difficulty in accessing the physical hardware to locate evidential data is a challenge for investigators (Birk 2011). Key evidence may also be spread across multiple data centers in different countries. To preserve data, there may be a need to examine several thousands of physical data items for data chunks of files, and then reconstruct the chunks into the relevant files. As the data may be spread across many countries, there may be legal and jurisdictional issues which need to be complied with (AGIMO 2011).

It is also recognized that the volume of data stored in the cloud is increasing, and it is anticipated that this will increase in the future (Bursztein, Cassidian & Martin 2011; Mason & George 2011; Kleynhans 2012). As reported by Ruan et al. (2011) the Federal Bureau of Investigation (FBI) reported that the average size of a digital forensic case rose from 83 GB in 2003 to 277 GB in 2007, an increase of 35% per year. In addition, the number of cases is also increasing. Turnbull, Taylor and Blundell (2009) conducted analysis of the number of jobs and devices presented for analysis at the South Australia Police (SAPOL) Electronic Crime Section from 1999 to 2008, which showed an overall increase of 20% per annum for requests, and an overall increase in the number of exhibits each year. This trend has continued, and the number of jobs for the 2011-2012 financial year (i.e. 1 July 2010 to 30 June 2011) was 1,014, which maintains a 20% increase per annum. The number of items presented for analysis has also increased, from 226 items in 2000-2001 to 3,231 items in 2011-2012, representing a 28% increase per annum (Blundell B, personal communication, October 16, 2012).

When the figures from SAPOL and the FBI are combined, it demonstrates that an increase in the number of requests, the number of items, and the size of data. Extrapolating the data to indicate the expected figures for 2016, shows that in four years the average size of a case will be over four (4) terabytes and SAPOL will have over 2,000 requests per year to examine over 8,500 computers and other electronic devices (Table 1 displays the extrapolated figures in italics, and known figures in bold). Ruan et al. (2011) conclude that the increase in cloud computing exacerbates the problem of scale of digital forensic activities, and also creates new challenges. During their
research, Biggs and Vidalis (2009) identified that few high tech crime units were ready or prepared to undertake investigations of cloud based crimes.

<table>
<thead>
<tr>
<th>year</th>
<th>size of data (GB) - FBI</th>
<th>number of SAPOL requests</th>
<th>number of SAPOL exhibits</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td></td>
<td>140</td>
<td>226</td>
</tr>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td>277</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td>374</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td></td>
<td>505</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td>682</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td>920</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td>1,242</td>
<td>3,231</td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td>1,677</td>
<td>4,120</td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td>2,264</td>
<td>5,252</td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td>3,056</td>
<td>6,697</td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td>4,126</td>
<td>8,538</td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td>5,570</td>
<td>10,886</td>
</tr>
<tr>
<td>2018</td>
<td></td>
<td>7,519</td>
<td>13,880</td>
</tr>
<tr>
<td>2019</td>
<td></td>
<td>10,150</td>
<td>17,697</td>
</tr>
<tr>
<td>2020</td>
<td></td>
<td>13,703</td>
<td>22,564</td>
</tr>
</tbody>
</table>

Table 1 – FBI and SAPOL data extrapolated to forecast growth in digital evidence

There is also the issue of co-mingling data, and the difficulty in seizing data related to a specific suspect from a data center (Reilly, Wren & Berry 2011). The MegaUpload investigation, as reported in the media, is an example of the consequences that an investigation involving a file storage service can have. Innocent users of a popular hosting service are affected, and also the volume of data involved, including the associated cost of maintaining the billions of digital files involved, all become costly issues (Sandoval 2012). It is alleged that MegaUpload had rented 1,103 servers from Carpathia Hosting, which hold upwards of 28 petabytes of data, and costs US$9,000 a day to maintain. As MegaUpload’s assets are frozen, Carpathia have requested the investigators pay for the preservation costs in relation to the data, which may be required for court proceedings (Kirk 2012).
The implications of cloud computing in the context of civil disputes and criminal prosecution, and preserving and storing electronic evidence, is also an issue highlighted by Mason and George (2011) who state that the ‘admission of evidence obtained from the cloud will be of greater significance in the future’. Wolthusen (2009) outlines a number of challenges in relation to distributed environments, including; computer environments are now more challenging in scope and diversity, and the standards of evidence and investigation practices are becoming more stringent. There is difficulty in seizing physical data, and examiners must rely on snapshots. The attribution of data and determining ownership is another issue raised, along with the stability of the evidence, and the difficulties in the presentation of complex data. As the original evidence resides in the cloud, there will be a need to rely on the ‘next best evidence rule’ (Mason & George 2011). There is also a need to show that the original data remains the same, and has not been altered (Pallaras 2011).

The identification of the actual suspects within the cloud environment is also an issue (Taylor et al. 2011). When a physical computer or device is seized, there may be some provenance to demonstrate ownership or association, but with cloud data this provenance may not be possible. The use of anonymous networks, such as ‘The Onion Router’ (TOR), to access to a cloud storage account may hinder investigations.

Crimes involving cloud storage can be categorized using common electronic crime definitions. Electronic crime can involve offences where a computer is used as a tool, target, or storage device (Etter 2000). Cloud stored data can be a target of criminals, and cloud storage also used to store illicit data, or data related to a crime. Cloud services can also be used as a tool in the commission of a crime. It has been reported that a virtual server from the Amazon EC2 cloud service was used to launch the attack that led to the Sony PlayStation Network outage (Alpeyev, Galante & Yasu 2011).

Grabosky (2000) explains that computer related crime is explained by three factors; motivation, opportunity, and the absence of capable guardianship. The growth in the internet and technology has resulted in a vast increase in opportunity for offenders. Just as banks and financial institutions have moved their business online, so do criminals chasing the money. Karyda (2009) also focuses on electronic crime as a crime of opportunity. As online services increase, criminals seek to exploit these opportunities.
2.3.1 Mobile device analysis

Traditional forensic analysis is undertaken on computer hard drives and common storage, such as USB devices and optical media. There is a growing shift towards the greater use of mobile devices which increasingly use cloud storage, as internal storage on each device is limited. A report from Gartner, for example, highlights the trend in client computing is shifting from a focus on personal computers to include portable devices, and also predicts that personal cloud storage will replace the PC as the main method of users’ storage by 2014 (Kleynhans 2012). This presents challenges to forensic examiners and methodologies for a number of reasons, such as the technical differences between the devices, and the increase in the use of cloud storage.

The variety and difference between operating systems and file systems amongst portable devices presents technical difficulties for examiners (Taylor et al. 2011). The numerous operating systems have little technical information published about the inner workings, requiring forensic tool developers to reverse engineer the operating and file system to access data. In addition, even where there is published information regarding a file system, there is still limited support. For example, YAFFS2 is not supported by the major forensic tools (Encase, FTK, X-Ways) and is used on a large number of mobile phones and tablets using the Google Android operating system, which is increasing in popularity (Quick & Alzaabi 2011). Although current mobile phone forensic tools, such as Microsystemation XRY and Oxygen Forensics, are able to extract increasing amounts of data from smartphones, evidence in relation to cloud storage may not be able to be extracted (Zhu 2011).

2.3.2 The need for common procedures

As concluded by Reilly, Wren and Berry (2011), cloud providers have yet to address forensic readiness, and forensic practitioners have yet to outline common procedures for the cloud investigations. Standard processes are needed to assist in the investigation of cloud computing services, such as cloud storage of data. According to Birk (2011), traditional forensic approaches are no longer practical due to the decentralized nature of the cloud environment. Ruan et al. (2011) identify cloud forensic analysis as a subset of network forensic analysis. The first part of the Digital Forensics Research Workshop (DFRWS) definition of network forensic analysis is as follows; ‘The use of
scientifically proven techniques to collect, fuse, identify, examine, correlate, analyze, and document digital evidence from multiple, actively processing and transmitting digital sources’ (DFRWS 2001). Ruan et al. (2011) conclude that as cloud computing is based on network accessible data, forensic investigations of cloud data should be based on a network forensic process, with consideration for unique cloud techniques in each phase. Hence, a common forensic framework is appropriate to guide an investigation involving cloud storage, with considerations to include network collection of cloud data.

Taylor et al. (2011) explained that ‘in legal terms, cloud computing systems will make it potentially more difficult for the computer forensic analyst to acquire and analyse digital evidence to the same standards as that currently expected for traditional server-based systems, due to the difficulty in establishing what data was stored or processed by what software on what specific computing device’. A common digital investigation process leads to seizure of a physical computer or device which is submitted to an examiner for analysis. A bit-for-bit image of the hard drive is made via a write-blocker, which is verified with a hash algorithm. A working copy of the evidential hard drive or media is made, and the original evidence and chain of custody is preserved. Analysis is conducted on the forensic copy of the data to examine a variety of stored information, such as; internet cache, cookies, registry files, swap files, recycle bin information, and unallocated space from which deleted data may be recovered. In a cloud computing environment the files and artifacts present and able to be examined on a seized physical machine may not be present.

There are guidelines and procedures in relation to conducting digital investigations which are used by practitioners to reduce the risk of evidence being excluded from a legal process, such as ACPO (2006), NIJ (2008) and SAI (2003). When an investigation involves a cloud storage environment, the application of the guidelines may be difficult (Birk & Wegener 2011; NIST 2011; Taylor et al. 2011; Zatyko & Bay 2012). It is important to note that there are several types of cloud services and each type with its potential different use in criminal activity, there will be variation in the way criminal investigation is carried out in each type of cloud service (Martini & Choo 2012).
2.3.3 Current research

This research examines cloud storage, for which there are very few publications examining or outlining a process to identify, preserve, analyse, and present stored data.

McClain (2011) discusses Dropbox client software from a forensic perspective, highlighting the information available from the client software system files installed at the time (May 2011). However, Dropbox version 1.2.48 introduced on October 31, 2011, resulted in the configuration files being encrypted to prevent the theft of account credentials by copying these files to another computer (Dropbox 2011). No published research was located examining Dropbox client software since the update at the time of writing this thesis. Chung et al. (2012) examine Dropbox client software version 1.1.35, which is prior to the version examined by McClain (2011).

Chung et al. (2012) also examine other cloud storage services in addition to the Dropbox client software, including Amazon S3, Google Docs, and Evernote. Microsoft Internet Explorer is examined on a Windows system and Mozilla Firefox on an Apple OSX system, but Google Chrome and Apple Safari browsers are not included in the research. Although Dropbox client software access and remnants were researched, access via a browser was not examined. Google Docs access through a browser and iOS application were researched, and in addition, on an Android installation. However, other popular cloud storage services, such as Microsoft SkyDrive or Google Drive were not examined. Memory remnants were also not examined. Chung et al. (2012) include a case study to demonstrate an application of their research.

Zhu (2011) examines Skype, Viber, Mail, and Dropbox mobile cloud applications on Android and iPhone iOS mobile devices in depth as three scenarios. In relation to Dropbox, Zhu reported that filename and username information was able to be extracted in relation to Dropbox use, but the file contents were not able to be retrieved.

Reese (2010) describes a process of using snapshots using EBS Boot volumes in Amazon cloud services where there is the ability to snapshot a system, however, this is not applicable to a cloud storage service as this process is outside a user’s control.

Clark (2011) examines picture Exif metadata remnants in Microsoft SkyDrive, Windows Azure, and Flickr, focusing on the information within pictures that are publicly shared, discovering that personally identifiable information was available in
many publicly shared pictures, which could be of relevance in an investigation, such as global positioning information.

Whilst the research of McClain, Chung et al., and Zhu provides information to address some of the research objectives, there are information gaps that need to be answered prior to forming any conclusions.

### 2.4 Issues at each stage of a forensic investigation

The process outlined by McKemmish (1999) lends itself as a framework for forensic investigation, and is used in the following sections to explore the issues at each stage of the process in relation to forensic analysis in a cloud environment.

#### 2.4.1 Identification of data

When conducting analysis with data stored with cloud services, one of the main issues is to identify where potential data resides. The previous Director of US Department of Defence Computer Forensics Laboratory and the previous Chief Scientist at US Air Force Research Laboratory Information Directorate, posited that ‘more research is required in the cyber domain, especially in cloud computing, to identify and categorize the unique aspects of where and how digital evidence can be found. End points such as mobile devices add complexity to this domain.’ (Zatyko & Bay 2012).

According to Birk (2011), in cloud computing environments the source of evidence on a client system in most scenarios will be the user agent, such as a web browser, and hence the information gathered from a browser should be included in forensic examinations. In addition, Taylor et al. (2011) state that digital evidence may be located in multiple client browser history caches, which also presents difficulty in collection, collation and verification.

Delays in accessing data due to legal or logistical issues could be detrimental to an investigation. Taylor et al. (2010) conclude that delays could result in data being deleted before access is granted. Therefore timely identification and preservation are crucial.
2.4.2 Preservation of data

As observed by Mason and George (2011), data stored within a cloud storage environment can be as transitory as a real cloud. A cloud in the sky can appear and disappear quite rapidly, and is very hard to collect. Data stored in ‘the cloud’ can move from location to location rapidly, disappear just as quickly, and is difficult to collect in a forensically sound manner. This needs to be considered in line with the volatility of data when undertaking collection from cloud sources (Ruan et al. 2011).

In jurisdictions with the legal provision to collect data by accessing an account (such as Australia and Norway) there are two potential methods available; via browser login and via client software. Software which collects of data stored in cloud storage account in a forensically sound manner is not currently available. Cumulus Data have stated that a future upgrade to their software will include the ability to collect data from ‘Google Docs, cloud storage such as Dropbox and iCloud, and social media such as LinkedIn, Facebook, and Twitter’ (Roach 2012). Dykstra and Sherman (2012) examine the process of securing evidence from a remote infrastructure-as-a-service (IaaS) computer running via Amazon EC2, but this relies upon the examiner being able to install an Encase Servlet or an FTK Agent. An alternative they raise is to request the data from the service provider, but this relies on trusting the service provider to collect and provide accurate data. In addition, requesting data from the service provider may not be timely due to legal and jurisdiction issues. When identified data is stored within a cloud storage service (file-hosting), it is not possible to make use of an Encase Servlet or an FTK Agent, as a user is only able to access data stored in the account, not to run software in the storage environment.

Ruan, Carthy and Kechadi (2011) conducted a survey of digital forensic practitioners in relation to cloud forensic issues, and found that there is a need for a procedure and set of toolkits to proactively collect data in the cloud.

2.4.3 Analysis of data

When dealing with computers and electronic data, the volume of data can easily overwhelm an analyst, and also become difficult and costly to manage. There is a need to refine the data to that which is relevant to the investigation, and ignore extraneous data. This can be difficult to undertake, as at the beginning of an investigation, it is
difficult to determine what is and what is not relevant. This may not become apparent until later in an investigation, resulting in a need to be able to include and exclude data as required, without affecting the admissibility of the data as evidence. This also highlights a need to identify the purpose of an investigation early in the process, as per Taylor et al. (2011), and is included in the process listed by Guo, Shang and Jin (2012) at ‘Step 1: Determine the purpose of the forensics requirement’.

Analysis of data will also require more time and resources due to the increase in devices that may need to be examined (Taylor et al. 2011). In addition, there is a need to aggregate disparate data into a single unified format, such as that proposed by Bursztein, Cassidian and Martin (2011) with the Offline Windows Analysis and Data Extraction tool (OWADE). In addition, the amount of data associated with an investigation is outgrowing the ability to conduct analysis in a timely manner (Ruan et al. 2011). One security method utilised by corporations is to encrypt data prior to uploading it to a cloud service (Taylor et al. 2010). This practice may inhibit analysis of data by a forensic examiner as the decryption key may not be available.

Hash analysis is a common technique to reduce data and identify known illicit files or files of interest. Hegarty et al. (2011) identify that the volume and distribution of data in cloud platforms presents a barrier to the use of signature detection methodologies, and propose the development and application of a distributed process when hash analysis is required across a cloud platform. Other challenges include the lack of operating system metadata, such as registry files, unallocated space, and temporary files which are important in a traditional forensic analysis process, but absent when examining data stored in the cloud (Grispos, Glisson & Storer 2011).

### 2.4.4 Presentation of data

Digital evidence exists in a logical context and a physical context. Data is stored on physical media, and can also be interpreted by software and processes for users to understand the information (Reilly, Wren & Berry 2011). For an investigator, prosecutor, and those involve in the court process, there is a need to understand the importance of the information being presented to them, and the implications of the information. Although electronic evidence is becoming more commonplace in today’s digital age, when presenting digital information in court, ‘the challenge is to have the
data translated into a form that is acceptable as evidence to the courts …. Assuming that the fragile and elusive evidence can be gathered together, the prosecutor must keep in mind that he or she will one day need to be able to prove the chain of evidence. All processes will need to be appropriately documented in a way that can be understood by the layman and the prosecutor must be prepared if necessary to demonstrate that the ‘original’ digital material has not been changed or tampered with in any way’ (Pallaras 2011).

Timeline presentation of data can assist to enable an understanding of the data in context (Reilly, Wren & Berry 2011). The timeline process is one aspect of intelligence analysis, and has been refined over the many decades that it has been practiced. This method of analysis and presentation is used in the case studies in this research to demonstrate a method to present information from various sources in context.

Taylor et al. (2011) conclude that prosecution cases may ultimately rely upon the information on a client device, as opposed to information acquired from the cloud. However, Bursztein, Cassidian and Martin (2011) state that most user activity is moving to the cloud, and that analysis of a user’s computer may not be sufficient. Hence, there is a need to understand the data remnants left behind on a client device when cloud storage is used, and to also understand the forensic implications of collecting data from a cloud storage service to enable evidence to be presented and accepted. Guo, Shang and Jin (2012) conclude that examiners will need to develop a broad technical knowledge across various hardware platforms and operating systems as well as a deep understanding of the various technologies, applications and networks.

2.5 Additional considerations

There are some positive implications for digital investigations and the growth in cloud services. As highlighted by Reilly, Wren and Berry (2010), cloud offerings have the potential to address many of the issues facing forensic examiners, such as; encryption, central storage, logging, and in-built hash authentication. Ruan et al. (2011) also highlight some opportunities, including; the low cost of cloud infrastructure enables companies that cannot afford forensic services to access these services at a lower cost, and with data being spread across data centers, full data deletion may not be possible,
and hence crucial data may remain elsewhere. In addition, cloud services may provide a solution to the analysis of large volumes of data, by being able to distribute the analysis across forensic examiners on a global scale as a service.

Cloud computing services can in some instances be of assistance in an investigation. IaaS services where virtual computers are used for criminal purposes, may be paused rather than shut down, and potentially able to be preserved for investigation (Taylor et al. 2010). Also, Dropbox splits each file into 4 megabyte (MB) chunks which have a hash value calculated using the SHA-256 algorithm which may be of assistance in an investigation to identify data (Mulazzani et al. 2011).

Roussev et al. (2009) examine the implementation of scalable development platforms to leverage distributed resources, comparing MPI MapReduce and Hadoop. They conclude that MapReduce demonstrates a substantial improvement for common indexing tasks, and aim to further research the potential to develop distributed tools to process large data collections. However, the security implications of using cloud resources for processing evidence need to be a consideration in deploying this process outside of a forensic laboratory environment. Collaborative forensic analysis, such as that outlined by Mabey and Ahn (2011) and Lee and Hong (2011) are also concepts that provide benefits of deployment and collaboration across a cloud-style network.

2.6 Summary

Cloud computing is a rapidly growing service which seeks to address economic concerns and provide large scale computing power for minimal cost. Cloud storage is an increasingly popular service used by consumers, business and government to store vast amounts of data. Cloud storage is popular because it provides virtual storage for users to store a range of files, such as pictures, videos, documents, emails, and other data. Some services also allow a user to edit files, and collaborate on documents. Cloud services are used by consumers, accessing their hosted data via computers, mobile phones and other portable devices. Cloud storage is used by criminals, and is a target of cyber-criminals. Security of cloud services are well addressed, however forensic investigations are rarely considered, posing challenges to forensic readiness, forensic response and forensic
analysis. As cloud storage can be accessed by a variety of devices, and is increasingly being accessed by users with mobile devices, there are additional challenges as the forensic analysis of portable devices is not as easy or well documented as the analysis of data remnants on computers and hard drive storage media.

Law enforcement agencies and investigators have a need to access data stored in cloud storage accounts. Difficulties arise when attempting to apply traditional computer forensic methods of investigation to a cloud environment. In a traditional computer investigation, the physical hardware is seized, and/or a bit-for-bit forensic copy is made, and analysis is undertaken on the copy, with the original media preserved. In a cloud storage environment, the physical hardware will be hosted in a large data center, which may be hosted off-shore in another country, and the data may be spread across many such data centers across the globe, hence physical analysis would be a challenge to say the least.

In addition, the initial identification of the existence of evidential data being hosted within a cloud storage service can be difficult, if the service provider and user details are not known. Hence, there is a need for a forensic examiner to know where to look for signs of cloud storage being used by a suspect. In undertaking secondary research in relation to published material, it was apparent there is minimal information available to guide investigators where to look for data remnants or indicators of cloud storage use. Information is known in relation to older versions of the Dropbox client software, however in relation to other popular cloud storage services, such as Microsoft SkyDrive, and Google Drive, there was no information located. This highlights a need for research to be undertaken in relation to the data remnants and identification points for an examiner to quickly look for signs that cloud storage is being used, service providers, and account details. This is applicable for computers, and mobile devices, highlighting a need to research methods of analysis for computers and mobile devices.

Once identification of data is made, there are no tools available, or methods published, to guide an examiner in how to collect identified data, when they have legal access to an account. Current methods rely on communication with the identified service provider, and then a reliance on technicians to preserve the identified data. If this is not undertaken in a forensically sound manner, the entire investigation in relation to the data may be put at risk, and the evidence not accepted in court or legal proceedings.
Research into the forensic implications of an examiner accessing and downloading data from a cloud service account is required to enable examiners to understand the implications of different methods of collection.

In addition the secondary research highlighted a need to develop a framework, and procedures and guidelines for examiners to undertake in relation to cloud storage, which has additional demands when compared to traditional analysis. Cloud storage is deemed to be an aspect of network forensic analysis, which is an aspect of forensic analysis. This highlights the need to build upon current methodologies, rather than create new frameworks to apply to investigations. The common digital forensic process of; identify, preserve, analyse, and present, is commonly used by examiners, however, some additional steps have been highlighted for inclusion, such as defining the scope of an investigation, and reviewing the results of an investigation.

In the following chapter, the research methodology for this thesis is discussed, and then in the subsequent chapter a proposed digital forensic analysis cycle is outlined, which is subsequently used to guide the research process in relation to cloud storage analysis.
This chapter describes the research method undertaken to address the objectives listed in chapter one. Information from the literature review in chapter two is used to design the main research questions, and partially answer some of the sub questions. Limitations of the research are then listed.

Chapter 3 – Research Methodology

3.1 Research problem

As outlined in the previous chapter, cloud storage is being used to store large amounts of data, including illicit data and evidential data. Computers and mobile devices are increasingly used to access cloud storage. There is a lack of information relating to forensic analysis and data remnants to identify associated service provider information, username details, and other potential evidence. There is also a lack of information relating to methods to preserve data, and what changes, if any, may occur when collecting data using software or a web browser. As discussed in Chapter two, published
literature does not answer the research objective to determine whether there is a method to identify and preserve data and remnants relating to popular cloud services in a timely manner by investigators.

3.2 Research purpose

The purpose of this research is to locate answers to the research problem outlined in section one. This is to identify whether data remnants are available after cloud storage is accessed or used, and whether there is a method available to identify and preserve data relating to cloud storage services. This is undertaken to assist forensic examiners, such as in law enforcement, security, and private industry, to identify cloud storage use in a timely manner, and to provide a methodology to undertake identification and preservation in a forensically sound manner.

3.3 Research questions

Analysis of the results of the literature review highlighted the issues in relation to forensic analysis of cloud storage as; whether there are any data remnants from cloud storage use, and what methods are available to preserve data stored in a cloud storage account.

3.3.1 Research Question 1

The first primary research question is thus defined as:

\[ Q_1 \]

What data remnants result from the use of cloud storage to identify its use?

\[ Q_1 \]

leads to the following hypotheses;

\[ H_0 \] - There are no data remnants from cloud storage use to identify the service provider, username, or file details.

\[ H_1 \] – There are remnants from cloud storage use which enable the identification of the service, a username, or file details.
These lead to the following sub questions;

Q1a) What data remains on a Windows 7 computer hard drive after cloud storage client software is installed and used to upload and store data with each hosting provider.

Q1b) What data remains on a Windows 7 computer hard drive after cloud storage services are accessed via a web browser with each hosting provider?

Q1c) What data is observed in network traffic when client software or browser access is undertaken?

Q1d) What data remains in memory when client software or browser access is undertaken?

Q1e) What data remains on an Apple iPhone 3G after cloud storage services are accessed via a web browser with each hosting provider?

Q1f) What data remains on an Apple iPhone 3G after cloud storage services are accessed via an installed application from each hosting provider?

3.3.2 Research Question 2

The second primary research question is defined as:

\[ Q_2 \quad \text{What forensically sound methods are available to preserve data stored in a cloud storage account?} \]

\[ Q_2 \] leads to the following hypotheses;

\[ H_0 \quad \text{the process of downloading files from cloud storage does not alter the internal data or the associated file metadata.} \]

\[ H_1 \quad \text{the process of downloading files from cloud storage alters the internal file data and the associated file metadata.} \]
H$_2$ the process of downloading files from cloud storage does not alter the internal data, but does alter the file metadata.

H$_3$ the process of downloading files from cloud storage alters the internal data, but not the associated file metadata.

This leads to the following sub question, Q2a;

*What data can be acquired and preserved from a cloud storage account using existing forensic tools, methodologies, and procedures when applied to cloud storage investigations?*

Some of the above questions are able to be partially answered from the literature review;

The research by Zhu (2011) concluded that Dropbox files were not able to be retrieved from an Apple iPhone, however, file and user information was able to be located in a plist file, ‘com.getdropbox.Dropbox.plist’. This partially answers research Question 1(f) in relation to data remnants on an iPhone after using Dropbox client software. No information was able to be located in relation to other cloud storage providers, Microsoft SkyDrive and Google Drive.

The research by McClain (2011) and Chung et al. (2012) concluded that Dropbox file history was available in client software versions prior to October 2011, however, as these files are potentially now encrypted, this issue needs to be examined with updated client software. The information from McClain and Chung et al. partially answers research question 1(a) in relation to Dropbox cloud storage.

### 3.4 Research method

The first step was to undertake literature research to identify if methods relating to cloud storage forensic analysis have been published. This was undertaken and whilst some
information assisted in answering research sub questions, the overall aim of the research was not answered in published literature.

### 3.4.1 Research Question 1 Experiment Process

To answer **Research Question 1**, a quasi-experimental nonrandomized control pretest-posttest quantitative research method is used. It is a quasi-experimental process as the nature of cloud storage means the devices are connected to the Internet to gather data and establish the circumstances for review, which can lead to changes to the devices outside the scope of the experiment. Hence the results may not be repeatable within the exact same parameters in future, depending on changes to the websites visited, operating system updates, and other processes not controlled by a user. The experiment is nonrandomized to enable analysis to be undertaken to determine changes to various control circumstances, and the use of randomized samples would reduce the scope of the research, therefore known content control media is used within virtual computer infrastructure and then copied and used to create various circumstances of access.

Control base image devices are established at the first stage of the experiment process, and are used to provide a pretest set of data for comparison. The experiment is undertaken, and then analysis compares the results with the pretest media, to determine any changes posttest. Each instance is preserved to capture the changes made when undertaking each process. Various circumstances are created using different browsers to access different cloud storage services.

Anti-forensic methods are also undertaken, and compared with the control and different instances to determine if these methods serve to hide identified data. The block diagram in Figure 2 outlines the process to be undertaken; starting from a control base system, each cloud service is examined individually with multiple circumstances involving the client software and different browsers, and an iPhone is also included in the process. The Windows test systems memory, hard drive, and network data is captured and preserved for analysis. The iPhone is examined with Microsystemation XRY.
This process is applied to answer the research questions in relation to the use of cloud storage services; Dropbox, Microsoft SkyDrive, and Google Drive. A variety of virtual machines were created. It was decided to examine a variety of circumstances of a user accessing cloud storage, and also to examine any differences when using different browsers. Multiple scenarios were explored; each made use of cloud storage with a different browser; Internet Explorer (IE), Mozilla Firefox (FF), Google Chrome (GC), and Apple Safari (AS). Multiple Virtual Machines (VMs) were created for each browser to replicate different circumstance of usage, as outlined in Table 2 - Configurations of Virtual PCs.

Had this been attempted using physical hardware, the expense and time to undertake setup, erasing, copying and re-installing would have been onerous.
### Virtual PCs Details

<table>
<thead>
<tr>
<th>Virtual PCs</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>IE, FF, GC or AS Base-VM</td>
<td>Windows 7 Home Basic SP1, 1 GB RAM, 20 GB Hard Disk Drive. Browser for each test installed; Microsoft Internet Explorer (IE), Mozilla Firefox (FF), Google Chrome (GC) or Apple Safari (AS).</td>
</tr>
<tr>
<td>IE, FF, GC or AS Upload-VM</td>
<td>Cloud Storage Windows Client software downloaded and installed. Test account accessed. Enron sample data uploaded to the cloud storage account.</td>
</tr>
<tr>
<td>IE, FF, GC or AS Access-VM</td>
<td>Browser used to access the cloud storage website and sign in to the user test account. Each file stored in the cloud storage account opened but not purposely downloaded.</td>
</tr>
<tr>
<td>IE, FF, GC or AS Download-VM</td>
<td>Browser used to access the cloud storage website at <a href="http://www.dropbox.com">www.dropbox.com</a> and sign in to the user test account. Each file downloaded to the VM Hard Drive Desktop and opened.</td>
</tr>
<tr>
<td>IE, FF, GC or AS Eraser-VM</td>
<td>Using the three copies of the Download-VM’s, the cloud storage and Enron files were erased using Eraser software.</td>
</tr>
<tr>
<td>IE, FF, GC or AS CCleaner-VM</td>
<td>Using the three copies of the Eraser-VM’s, CCleaner was downloaded, installed and run.</td>
</tr>
<tr>
<td>IE, FF, GC or AS DBAN-VM</td>
<td>Using the three copies of the CCleaner-VM’s, DBAN was used to perform a complete wipe.</td>
</tr>
</tbody>
</table>

| Table 2 - Configurations of Virtual PCs |

The UC Berkeley Enron Email subset data file was used as the sample data, and was downloaded from the project website (http://bailando.sims.berkeley.edu/enron_email.html) on the 9th February 2012. The text of email 3111 was copied and used to create a text file saved in the ‘rtf’ format. The text was also copied and used to create a Word 2010 document saved in the ‘docx’ format. A picture file in ‘jpg’ format was created using a screen snapshot (ALT+PrintScreen) of the text file open in Notepad++. MD5 values were calculated for these files and key terms were selected to enable searching and location of the data and files in subsequent analysis.

Virtualised PCs were created using VMware® Player 4.0.1. For each scenario, a base image was created, and Windows 7 Home Basic was installed on a 20 GB virtual hard drive with 1 GB RAM. The Base-VM files were used as control media to determine the files created when user activity was undertaken in each scenario. The different actions undertaken were as follows;

1. The first step was to install the browser software into separate Base-VM’s for each browser; Mozilla Firefox 13.0.1, Internet Explorer 9.0.8112.16421IC, Google Chrome 19.0.1084.56m, and Apple Safari 5.1.7 for Windows 7.
2. Next was to make a copy of the Base-VM for each browser. These four VMs were labeled IE, FF, GC or AS Upload-VM, and were used to access the cloud storage website and download and install the client software. An account was created for this research, and sign-in undertaken using the client software. The Enron sample files were uploaded to the account using the client software.

3. Additional copies of the Base-VM for each browser were made, labeled Access-VM. These were used to access the cloud storage website using each installed browser. The sign-in option was used to log in to the user account created in step 2. Each of the stored files were opened within the browser, but not deliberately downloaded.

4. Copies were made of the four Base-VMs, labeled Download-VM. Each installed browser was used to access the cloud storage website. The sign-in option was used to log in to the user account created in step 2. All of the files in the cloud storage account were downloaded to the Virtual Machine hard drive as a zip file. The contents of the zip file were extracted to the Desktop and then each file was opened and closed.

5. Copies were made of the four Download-VMs, labeled Eraser-VM. Eraser v6.0.10.2620 was installed and used to erase each of the Enron data files and the zip files.

6. Copies were made of the four Eraser-VMs, labeled CCleaner-VM. CCleaner v3.19.1721 was installed and run across the virtual machine hard drives to remove the browsing history and file references.

7. Copies were made of the Four CCleaner-VMs, labeled DBAN-VM. Darik’s Boot and Nuke (DBAN) version 2.2.6 Beta was used to boot each DBAN-VM. DBAN was run with the option to erase the entire 20 GB hard drive with the ‘US DoD level 3’ setting.

Whilst preparing each VM, Wireshark 1.6.5 was run on the host computer to capture the network traffic from the Virtual Machine network interface. Memory capture was facilitated by copying the Virtual Memory (VMEM) files created by VMWare. The VMEM files were copied whilst the VM was running, just prior to shut down. The
network capture files were saved at various points whilst the VMs were running, and after shut down.

A forensic copy was then made of each virtual hard drive (VMDK file) using AccessData FTK Imager CLI 2.9.0 in the Encase Evidence format (E01). In regard to the each memory file (VMEM) and network capture (PCAP) file, a forensic copy was made using Encase version 6.19.4, in the Logical Evidence format (L01), and the X-Ways Evidence File Container format (ctr). Hash values (MD5) were used to ensure the forensic integrity of the data.

Each of the forensic copies of the virtual computer hard drives, memory and network captures were examined using a range of forensic analysis tools, including; X-Ways Forensic version 16.5, Guidance Software EnCase version 6.19.4, AccessData FTK version 1.81.6 and version 4.01, Network Miner 1.0, Wireshark 1.6.5, Magnet Software (formerly JadSoftware) Internet Evidence Finder 5.52, and RegRipper version 20080909. Many of these tools are widely used for digital forensic analysis by law enforcement agencies and the private sector. In addition, Encase 6.5 and FTK Imager 2.5.3.14 have been tested by the Office of Law Enforcement Standards of the National Institute of Standards and Technology (NIST 2012).

When assessing the use of cloud service on an iPhone, the following process is undertaken; an iPhone 3G was selected which had not been used to access each cloud service previously as the device history and usage was known. MicroSystemation .XRY 6.2.1 was used to extract a logical image of the contents, excluding audio and video files, prior to undertaking the research. This extract was analysed to confirm there was no service provider related data on the device. Next, the inbuilt Apple Safari iOS browser was used to access the cloud service user account created for this research and view the sample files and Enron files stored remotely. A logical extract was then conducted using .XRY. The cloud service iOS Application was then downloaded and installed to the iPhone 3G. The test user account was then accessed using the Application, and the files stored in the account were viewed. A third logical extract with .XRY was then conducted.
3.4.2 Research Question 2 Experiment Process

To answer Research Question 2, a different experimental process was undertaken. Again, a quasi-experimental nonrandomized control pretest-posttest quantitative research method is used. Control data is identified, and then used to load into cloud storage services; Dropbox, Microsoft SkyDrive, and Google Drive. A different computer system is used to access the cloud storage accounts via browser and client software, and the files previously uploaded are identified and downloaded to the hard drive or iPhone memory. Analysis is then conducted to compare the original control files to the files downloaded using the browser and the ones downloaded using the client software on the computer system.

The process is outlined in Table 3, and was used for this research to preserve data from cloud storage accounts. This was undertaken by accessing an account (created for the purposes of this research) using a browser running in a virtual computer (VM) and recording the VM window using Microsoft Expression Encoder 4. Network traffic was captured using Wireshark running on the host computer. The data in each account was downloaded using a browser, and synchronised and downloaded using the client software. The VM computer was then paused and stopped. The VM files, video, and network captures were then preserved within logical evidence files (Encase L01, AccessData AD1, or X-Ways CTR files).

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Use a host computer with connectivity to the Internet (forensic analysis computers are generally not connected to the Internet, and hence a PC setup for the purpose of accessing the Internet was used, rather than a standard analysis computer).</td>
</tr>
<tr>
<td>2.</td>
<td>Create a clean base image virtual machine (VM) to access the cloud storage account, with a new installation of Windows 7 Home Basic operating system.</td>
</tr>
<tr>
<td>3.</td>
<td>Run Wireshark on the Host computer to capture the network traffic between the VM and the Internet.</td>
</tr>
<tr>
<td>4.</td>
<td>Run Microsoft Windows Expression Encoder 4 SP2 to video record the VM window to retain a record of the process undertaken.</td>
</tr>
<tr>
<td>5.</td>
<td>Using the VM, connect to the Service Provider URL using a browser, in this case Mozilla Firefox.</td>
</tr>
<tr>
<td>6.</td>
<td>Login to the cloud storage account using the username and password created for this research.</td>
</tr>
<tr>
<td>7.</td>
<td>Browse the folders and files, noting the dates and times associated with the files and folders, creators and modifiers, and computers linked to an account.</td>
</tr>
</tbody>
</table>
8. Select the files in the account and choose to download them to the VM Downloads folder.
9. Select the option to download the client software.
10. Install the client software to the VM hard drive and synchronise the account. Observe the process of the contents of the cloud stored account files and folders being downloaded to the VM.
11. Pause and close the VM.
12. Preserve the VM files (including the network PCAP, VMDK, VMEM, Expression Video) from the Host computer into a Logical evidence file (L01, AD1 or CTR).

| Table 3 - Data preservation steps for Cloud Storage |

VM’s were created using VMware Player 4.0.1. For each cloud storage service, a VM was created and Windows 7 Home Basic was installed on a 20 GB virtual hard drive with NTFS as the file system. Selected files from the UC Berkeley Enron Email subset data file was used for sample data, downloaded from the project website (http://bailando.sims.berkeley.edu/enron_email.html) on the 9th of February 2012. Hash values (MD5) were calculated for the selected files. The files were previously uploaded into cloud storage accounts created for the purposes of this research.

The process outlined in Table 1 is to be applied to Dropbox, Google Drive and Microsoft SkyDrive. Analysis will then be conducted on the logical evidence files to determine any changes made during the process of downloading them from cloud storage using this methodology.

3.5 Research equipment

The details of the host computer system are outlined in Table 4. In relation to the VM’s, each of the virtual PCs consisted of a fresh installation of Windows 7 Home Basic on a 20 GB virtual hard drive with 1 GB RAM allocated to the VM. A wide range of software was used, and is detailed in each section where used. Handwritten notes were made, as per common forensic practice, at each stage of the experiments. These proved to be invaluable, and were referred to numerous times whilst undertaking the research.
<table>
<thead>
<tr>
<th><strong>Host PC</strong></th>
<th>Generic Personal Computer Tower</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OS</strong></td>
<td>Windows® 7 64-bit Professional 6.1.7601 SP1</td>
</tr>
<tr>
<td><strong>Processor</strong></td>
<td>Intel® Core i7 960 @ 3.20GHz</td>
</tr>
<tr>
<td><strong>Memory</strong></td>
<td>24.0 GB Triple-Channel DDR3 @ 532MHz (7-7-7-20)</td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>64GB Kingston SV100S264G ATA Device 2TB Seagate Hard Drive</td>
</tr>
</tbody>
</table>

**Table 4 - Configuration of Host PC**

3.6 **Research limitations**

**Version dependent:** Due to the quasi-experimental nature of the research, results apply to the versions of software available at the time of undertaking the research. Any prior versions of software may have different outcomes, and subsequent changes to software may result in different findings. In addition, as access to the account via a browser relies upon the information returned from the provider, this may change as each provider updates the HTML or other code presenting the information to a user.

**Operating system dependent:** Research was undertaken assessing changes made to files running under the Windows 7 operating system and the NTFS file system. Alternative operating systems, such as Microsoft Vista, XP, Apple Mac OSX, or Linux, may all have different data remnants. Alternative file systems, such as HFS+ or EXT3, may also result in different data remnants. Hence, there may be different findings in relation to the research questions when NTFS or Windows 7 are not involved.

**Hardware dependent:** Research was limited to a non-jailbroken Apple iPhone running iOS 4.2.1, hence other OS versions, such as iOS 5 may provide different information, and a jailbroken device may provide additional information. In addition, Google Android and other mobile device operating systems and hardware may provide different data remnants and outcomes.
3.7 **Summary**

This chapter has described the research purpose which resulted from the literature review and research problem. The research questions were listed, and the methodology outlined the process to be undertaken in subsequent chapters. The equipment used to undertake the research was described and limitations of the research outlined.

In the next chapter, a proposed forensic analysis cycle is explained, which addresses the aspects of an overall investigation highlighted in the literature review. This also serves to address the nature of investigations in a cloud environment which result in information identification occurring during the stage of analysis, resulting in a need to return to previous steps to preserve data.
As highlighted in the literature review, there is a need for a methodology and framework to guide forensic investigations where cloud storage is involved. This chapter outlines a proposed framework, which endeavors to enlarge upon the process used for traditional forensic computer analysis. This serves to expand the common framework, to be applicable when dealing with cloud storage environments.

Material presented in this and the next chapters was submitted for publication and is currently under peer review:


Chapter 4 – Digital Forensic Analysis Cycle

4.1 Introduction

Over the years, a number of digital forensics models have been proposed (see Slay et al. (2009) for an overview of digital forensics models). However, these existing digital forensics methods may not be fit-for-purpose in the cloud environment (Birk 2011). There is a need for a forensic analysis framework to guide investigations, which is flexible enough to be able to work with future providers offering new services. Any proposed framework needs to be generic enough to apply to a range of circumstances and services, but be able to step an investigation through a formalized process to ensure
information sources are identified and preserved. There are generally accepted standards, rules, and procedures that digital forensic practitioners and examiners follow (NIJ 2004; ACPO 2006; NIJ 2008). McKemmish (1999) identifies and specifies four stages of a forensic computing investigation: ‘identification of digital evidence, preservation of digital evidence, analysis of digital evidence, and presentation of digital evidence’. When investigating cloud services there will be a cycle of identification and preservation which may branch off from the analysis of evidence already seized. For example, an examiner should not stop analysis of already seized evidence to wait for identified data stored in the cloud to be preserved and provided for investigation. The examiner should continue analysing the evidence at hand, and when the cloud data is provided, should include it for analysis. It is possible that analysis of the cloud data may identify further cloud stored data which needs to be preserved and analysed. Hence the digital forensic process should be viewed as a cyclic process.

4.2 Proposed Digital Forensic Analysis Cycle

The practice of intelligence analysis concerns itself with data analysis, and has been refined over the years. As outlined by Ratcliffe (2003), the intelligence process is a continuous cycle of tasking, collection, collation, analysis, dissemination and feedback. Whilst the definition of forensic analysis by McKemmish (1999) lends itself to form a model consisting of the following steps; identify, preserve, analyse, and present; when investigating cloud services there will be a cycle of identification and preservation which may branch off from the analysis of evidence already seized. For example, an examiner should not stop analysis of already seized evidence to wait for identified data stored in the cloud to be preserved and provided for investigation. The examiner should continue analyzing the evidence at hand, and when the cloud data is provided, should include that for analysis. It is possible that analysis of the cloud data may identify further cloud stored data which needs to be preserved and analysed. Hence the digital forensic process should be viewed as a cyclic process, akin to the intelligence analysis process, rather than a formal step-by-step process where the investigation follows from one step the next. Furthermore, there are some additional steps which can be adapted from the intelligence cycle; namely, tasking and feedback. The proposed framework (Figure 3) is based on the intelligence analysis cycle and the process of forensic analysis, and consists of the steps of; Commence (Scope), Prepare and Respond,
Identify and Collect, Preserve (Forensic Copy), Analyse, Present, Feedback, and Complete.

In the proposed framework, the process is cyclic and iterative as it is common that during an investigation a forensic examiner may need to return to a previous step. For example during the analysis phase, an examiner may uncover information relating to data stored with a particular cloud storage provider. The examiner may return to a previous step, such as ‘Prepare and Respond’ or ‘Identify and Collect’ and undertake enquiries to locate, identify, and collect the newly identified data using legal processes. At the same time, the forensic analysis of other data already collected would continue. Once the (new) data has been acquired from the identified provider, the examiner will preserve the collected data by creating a forensic copy, and include the data under the scope of the investigation for analysis.

Figure 3 - Proposed Digital Forensic Analysis Cycle
4.2.1 Commence (Scope)
At the beginning of an investigation it is important to outline the scope, nature, and background of the analysis. Taylor et al. (2011) state that ‘it is important that the purpose of a computer forensic investigation is clearly defined so that the full scope of the investigatory process can be decided.’ Investigators and forensic examiners need to understand the ‘What, Where, When, Who, Why and How’ of an investigation, and to determine the boundaries of an investigation. This should all be documented, as examiners and investigators will refer to this during the examination process. The scope would include the persons involved, any data or evidence already seized, keyword terms, any urgent timeframes, and other relevant information. The initial scope may be quite generic, and as the investigation proceeds, will refine to focus on issues as they arise.

4.2.2 Prepare and Respond
Once the scope is determined, the next step of any investigation, criminal or civil, is to understand the requirements and ensure that the correct equipment and information is available. Preparation can include training and equipment acquisition. There is a need for an examiner to have the correct skills, as per ACPO Principle 2 (ACPO 2006) ‘competency’ and this can be addressed by undertaking the appropriate training prior to undertaking an examination. Advances in ICT such as the cloud computing environments will require ongoing training and research in digital forensics, and this will include for general ICT professionals. The latter is becoming more actively involved in the investigation and prosecution of cybercrime. For example, ICT professionals may be called upon to help facilitate compliance with legal obligations, developing and operating secure computer and cloud computing systems to ensure the privacy of protected information is not compromised. Training would equip ICT professionals with a working knowledge of key legal challenges and issues they are likely to encounter in the course of professional activities.

Preparation can also include research and development, undertaken to gain an understanding of a particular issue or aspect of an investigation. For example, if the scope of an investigation relates to a particular cloud storage service, an examiner can conduct research using virtual computers or available equipment to gain an
understanding prior to an investigation commencing. This can also occur during an investigation if cloud storage becomes an aspect of an investigation. Inferences can be outlined and tested in a controlled environment to determine outcomes, which can then be applied to an investigation to answer particular questions, or gain an understanding of the presence of data or information, and form hypotheses.

Preparation also includes other aspects of an investigation, such as timely response, timeframe, personnel, duties, and locations of interest. An investigation plan can outline the various issues that need to be considered and addressed. The investigation scope would be outlined to the relevant people, and the appropriate equipment organized and available. Additional expertise can be sought to ensure the process can flow in a timely manner, or an examiner can undertake research to ensure they have the requisite knowledge to undertake an examination. Furthermore, general forensic issues which include note-taking, logging, auditing, documentation, the integrity of evidence and methods, chain of custody and other issues can be outlined and addressed (Haggerty & Taylor 2006; Peisert, Bishop & Marzullo 2008).

4.2.3 Identify and Collect

The next stage of the proposed framework (and in a typical forensic investigation) is to identify and collect the relevant items which may contain information or evidence. In some cases this may involve the compilation of a formal collection plan. The compilation of a collection plan will ensure that all the relevant data and information relating to the investigation is identified and preserved. A collection plan can outline and address personnel requirements, anticipated locations, timeframes, contingencies, documentation to be used such as evidence logs and chain of custody for each evidence item, equipment required, and any legal permissions or orders are obtained, such as court warrants. The collection plan also details the process to undertake when confronted with running (live) computer systems, the secure storage of seized data, maintaining a chain of evidence and a log of access.

During the process of analyzing a PC, an examiner may identify that cloud storage has been used to store information relevant to an investigation. The examiner will liaise with the investigating officer, and contact will need to be made with the identified service provider, such as Dropbox, with relevant information, such as a username, file
information, and dates and times of access. The cloud service provider will then assist according to relevant legal process to identify and preserve the identified data, and enable the investigator access to the data. In the case of Dropbox the first point of contact would be through the email address listed on the Dropbox ‘Contact’ webpage; legal@dropbox.com (Dropbox 2012). The investigator may need to assist in advising the service provider on the methods to preserve the data in a manner suitable for a legal process or Court, such as those outlined in ACPO (2006) and NIJ (2004, 2008). This should include MD5 and/or SHA hash values for the preserved data, and storing the data on write-once media, such as CD or DVD. If the data is too large to copy to optical media, hash values are crucial to be able to show the data has not been changed. Whilst the examiner is waiting for the identified data to be provided, analysis of the other data would continue.

4.2.4 Preserve (Forensic Copy)

One of the basic tenets of a digital forensic investigation is the ability to conduct analysis on a forensic copy, rather than interacting with or altering the original source (NIJ 2004; ACPO 2006; NIJ 2008). At the preservation stage, the physical items are examined and details documented, such as hard drive manufacturer, serial numbers, model numbers, and other information. Data is copied in a forensic manner, using write-protection and creating a bit-for-bit copy, verified with an MD5 and/or SHA hash algorithm.

In the case of cloud stored data, liaison with the service provider legal contacts may be the first step to preserve data. An investigator may need to assist the service provider with advice on how to preserve the data. This could include providing advice on using forensic software, such as AccessData FTK Imager, to create a logical container of the identified files. Write-protection may not be possible, and hence the person collecting the data will need to make notes (at the time) of the steps they undertook to preserve the data, including the reasons for the steps they are undertaking. Notes are crucial as legal examination may not occur for many years, and relying on memory may result in intricate details being forgotten. ICT professionals employed by the service provider may be responsible for preserving the identified data, and hence there is a need for training in forensic response and preservation methods. Consultation with the ACPO
guidelines (2006), NIJ publications (2004, 2008), and Standards such as HB 171-2003 (SAI 2003) are also recommended for those tasked with preserving data.

4.2.5 Analysis
In relation to the proposed framework, analysis of the forensic data is undertaken to locate information and test hypotheses in relation to the investigation parameters. This can draw on a range of different processes, using various software and hardware to address and answer questions in relation to the scope of the investigation (Willassen 2008). During the analysis there may be discovery of other sources of data, in which case the process returns to the ‘Prepare and Respond’ stage for the newly identified data. Analysis would continue with the already copied data, and the new data would be analysed when available. At any stage of the framework, a range of people can undertake different steps, including having multiple examiners. Practitioners with specific expertise can be brought in to undertake analysis of specific aspects, drawn together by a manager of an investigation.

4.2.6 Presentation
This next stage of the proposed framework is to explain the information in a manner which is understandable to investigators, judiciary (including juries) and other decision makers. For the findings of the analysis to be meaningful, they must be reported in a legally acceptable manner. The information must be explained in a thorough manner to ensure false conclusions are not reached, but still be understood. This communication can be verbal or written, and should include information about the analysis process, the systems used, and any limitations (McKemmish 1999; Yeager 2006).

A spreadsheet timeline of events can be created using data and information from the various sources, which can assist to explain the course of events. Merging the information from the hard drive, network captures, memory and the cloud storage account into one timeline may also be of assistance during the analysis process to gain an understanding of the sequence of events (examples are in the Case Study sections).
4.2.7 Feedback

Feedback is the next step of the proposed framework, and is the process of gathering information from those involved in the investigation, relating to the validity of the analysis and reported findings. The digital forensic model of Baryamureeba and Tushabe (2006) includes the final step of ‘Review’. Feedback is an additional component to the ‘Review’ process and is an important consideration to ensure that;

- The questions of an investigator have been answered,
- The process undertaken was correct for the circumstances,
- All avenues for further analysis are identified, and
- The processes and practices for future investigations can be recommended.

4.2.8 Complete

The final step is the completion of an investigation. A decision is made whether further analysis is required, based on the feedback from the investigator or legal counsel. If further enquires are required, the process returns to the ‘Prepare and Respond’ step. If there are no further avenues for investigation, the case can be completed, data archived and backed-up as required. As prosecution or legal processes can continue for many years, or data may be required sometime in the future, archiving must be done in a manner that information can be accessed at a later date, and not destroyed.

4.3 Applying the cycle

In the proposed framework, the processes can be cyclic and iterative as it is common that during an investigation a forensic examiner may need to return to a previous step. For example during the analysis phase, an examiner may uncover information relating to data stored with a particular cloud storage provider. The examiner may return to a previous step, such as ‘Prepare and Respond’ or ‘Identify and Collect’ and undertake enquiries to locate, identify, and collect the newly identified data using available legal processes. At the same time, the forensic analysis of other data already collected would continue. Once the (new) data has been acquired from the identified provider, the
examiner will preserve the collected data by creating a forensic copy, and include the new data under the scope of the investigation for analysis.

4.4 Summary

This chapter outlined the proposed framework, which serves to enlarge upon the process used for traditional forensic computer analysis with the inclusion of the initial steps of Commence (Scope) and Prepare. In addition, final steps of Feedback and Complete are also proposed. The framework is cyclic, and an examiner can return to previous steps, whilst the overall investigation progresses. This serves to expand common digital forensic analysis frameworks to be applicable when dealing with cloud computing environments, such as when data identification can cause an investigation to branch in different directions.

This framework is applied in the following chapters discussing the research of cloud storage providers to (a) provide a guiding framework to step the process through, as would be the case in a digital forensic investigation, and (b) test the application of the proposed framework in a variety of circumstances. This follows the process of a common digital forensic examination, enabling forensic examiners to apply the proposed framework to real-world investigations, and also serve to research the data remnants using a real-world applicable process of analysis.

In the following chapter, the framework is used in the analysis of the Dropbox cloud storage service in relation to using a Windows 7 computer to access the service via a browser and via client software, using an Apple iPhone to access the service through the in-build browser and Dropbox software, and a case study is used to demonstrate the application of the research findings.
5 Dropbox Analysis: Data Remnants on User Machines

In the previous chapters, the focus of the research was outlined, current literature was discussed, and the proposed Digital Forensic Analysis Cycle was explained. This chapter focuses on Dropbox cloud storage and explores the process of determining data remnants using the proposed framework to guide the research.

Material presented in this and the previous chapters was submitted for publication and is currently under peer review:


Chapter 5 - Dropbox Analysis

5.1 Introduction

Dropbox is a file hosting service that enables users to store and share files and folders. There is a free service which allows 2 GB of data storage, and additional storage space can be acquired by referring or signing up new users or subscribe to a paid service of up to 100GB for USD 200 per year. The service can be accessed using a web browser (e.g. Internet Explorer) or client software. Dropbox client software is available for Microsoft
Windows operating system (OS), Apple Mac OSX, Linux, Apple iOS, Android, Blackberry and Windows Phone devices.

This research focuses on identifying data remnants of the use of cloud storage, in this chapter the storage examined is Dropbox, on a Windows 7 PC, and an Apple iPhone 3G. This was undertaken to determine the remnants an examiner should search for when cloud storage, and in particular Dropbox, is suspected. This research also included the circumstances of a user employing anti-forensic methodology to hide evidence of cloud storage usage, and whether remnants remain to identify cloud storage use. The aim is to identify the cloud storage used, any usernames or passwords, files or data associated with a cloud account, and any other remnants that may assist an investigation.

In this first case study, a standard personal computer (PC) environment with Windows 7 OS installed is used to locate data remnants. It was determined that the use of a virtual computer with a standard installation of the Windows 7 OS would enable different configurations to be quickly setup and analysed, without having to re-configure and copy physical hard drives, memory or undertake external network capture. This allowed the examination of a variety of test PCs in several configurations to facilitate forensic analysis of the Dropbox client software and different browsers, namely Microsoft Internet Explorer®, Google Chrome™, and Mozilla Firefox®.

In the following chapter, the process of locating data remnants of a user accessing Dropbox in a variety of ways, and also undertaking anti-forensics to hide the use of cloud storage, on a Windows 7 PC is undertaken using the proposed framework. The framework is then applied in the research and analysis of an Apple iPhone 3G to determine the data remnants when using the inbuilt browser and when installing the Dropbox iOS Application. A case study is used to demonstrate application of the research findings. The chapter concludes with a summary of the findings and recommends areas for future research.

5.2 Dropbox Analysis: Windows 7 Computer Environment

5.2.1 Commence (Scope)

The scope of this research is to determine the data remnants on a Windows 7 PC for the use of Dropbox, such as; username, password, browser or software access, files stored
within the account, data remnants from the files stored, and the associated dates and times. As the research aims to locate data remnants, control (base) files are needed to compare subsequent processes to. There is a need to undertake known interactions with Dropbox, such as creating and accessing accounts in a range of ways including different browsers and client software use. There is also a need to undertake anti-forensic processes, and be able to compare this with the control (base) data. There will be a large amount of data created to be analysed, and this needs to be undertaken in a timely manner. This all needs to be documented and reported.

5.2.2 Prepare and Respond

To gather the data required to answer the research questions in relation to the use of Dropbox, a variety of virtual machines were created. The process undertaken was outlined in Chapter 3. It was decided to examine a variety of circumstances of a user accessing Dropbox, and also to examine any differences when using different browsers. Multiple scenarios were explored; each made use of Dropbox with a different browser; Internet Explorer (IE), Mozilla Firefox (FF), and Google Chrome (GC). Multiple Virtual Machines (VMs) were created for each browser to replicate different circumstance of usage, as outlined in Chapter 3.

Specific software used in the preparation of the Dropbox research, in addition to the software outlined in Chapter Three, includes;

- Dropbox client software version 1.2.52
- Mozilla Firefox (FF) 10.0.2
- Internet Explorer (IE) 9.0.8112.16421IC
- Google Chrome (GC) 18.0.1025.162m
- Eraser 5.8.8
- CCleaner v3.17.1689
- Darik’s Boot and Nuke (DBAN) version 2.2.6 Beta
5.2.3 Identify and Collect

In the context of this research, files were identified which would contain the information needed to conduct the analysis. In this instance files included the VMDK files in each VM folder for the hard drive, each VMEM file for the memory capture, and each saved network capture file (PCAP) for the network captures. These were identified for each VM.

5.2.4 Preserve (Forensic Copy)

For this research, a forensic copy was made of each virtual hard drive (VMDK file) using AccessData FTK Imager CLI 2.9.0 in the E01 container format. A forensic copy of each memory file (VMEM) and network capture (PCAP) file was made in the AccessData Logical Image (AD1) format, and an MD5 hash value for each original file was calculated and verified with each forensic copy.

5.2.5 Analysis

Each of the forensic copies of the virtual computer hard drives, memory and network captures were examined using a range of tools, such as Guidance Software EnCase version 6.19.4, AccessData FTK version 1.81.6 and version 4.01, Network Miner 1.0, Wireshark 1.6.5, Magnet Software (formerly JadSoftware) Internet Evidence Finder 5.52, and RegRipper version 20080909. Many of these tools are widely used for digital forensic analysis by law enforcement agencies and the private sector. In addition, Encase 6.5 and FTK Imager 2.5.3.14 have been tested by the Office of Law Enforcement Standards of the National Institute of Standards and Technology (NIST 2012).

5.2.5.1 Keyword search terms

Keyword search terms were determined from the filenames observed, and the text from within the Enron data files. These included the following:


• the username and password of the Dropbox account created for this research,

• ‘filecache.dbx’, ‘dataset.zip’,


### 5.2.5.2 Control (Base) images

Analysis of the three control Base-VM hard drives confirmed there was no data originally present relating to the Enron sample test data and Dropbox files. References were found for the term ‘dropbox’ in ‘index.dat’ files, ‘msjnt40.dll.mui’, ‘pagefile.sys’, and unallocated clusters. This should be borne in mind, as this indicates the presence of the keyword term ‘dropbox’ on a hard drive does not necessarily indicate that Dropbox has been used. As is usual for a digital forensic examination, the context of a search result needs to be analysed to determine the reason for a keyword match, rather than drawing a conclusion at face value of the presence of data. The control VM’s in this case have shown that data matches will occur, even when user activity in relation to Dropbox has not been undertaken.

### 5.2.5.3 Dropbox Client Software

In the three Upload-VM’s, it was observed that the Dropbox client software installed under Windows 7 does not install into the ‘C:\Program Files\’ folder, and instead installs into the Users ‘C:\Users\[username]\AppData\Roaming\’ folder. Dropbox sample files and folders were observed on the hard drive, at the default Dropbox folder location (‘C:\Users\[username]\Dropbox’). These included files; “Getting Started.pdf”, two “How to use” text files, and three sample pictures. The files, locations, and MD5 values for the software and sample files are listed in Table 5. When a new release of the Dropbox Windows client software is released, the hash values of the standard files and sample files may need to be recalculated.
Previously, the Dropbox (Windows) client software included a filecache.db file, which contained a history of filenames synchronized with Dropbox (McClain 2011). In the Dropbox client software used for this research, (version 1.2.52), this file is not present. However, there is a Filecache.dbx file, which appears to be encrypted. There is a ‘host.db’ file which includes the path for Dropbox file storage which is in Base64 string encoded text. For example, in a host.db file the following string was located; ‘QzpcVXNlcnNcY2xvdWRcRHJvcGJveA==’. Converting this as a Base64 string results in the text; ‘C:\Users\cloud\Dropbox’.

<table>
<thead>
<tr>
<th>File</th>
<th>MD5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropbox Windows Client Software: Dropbox 1.2.52.exe</td>
<td>c05a03f72386b7c9d5cc7dcefa8555da</td>
</tr>
<tr>
<td>File: Dropbox\Getting Started.pdf.</td>
<td>d0a31650c916d07316458ad11e4093c3</td>
</tr>
<tr>
<td>File: Dropbox\Photos\How to use the Photos folder.txt,</td>
<td>e6655fb94380d9a943c5d13976253</td>
</tr>
<tr>
<td>File: Dropbox\Photos\Sample Album\Boston City Flow.jpg,</td>
<td>5fcd8abc87de2629a3e75598999944da</td>
</tr>
<tr>
<td>File: Dropbox\Photos\Sample Album\Costa Rican Frog.jpg,</td>
<td>30bf1fadcfc52c796d143563a9d4484f</td>
</tr>
<tr>
<td>File: Dropbox\Photos\Sample Album\Pensive Parakeet.jpg,</td>
<td>b44d238643412e46d5ec7a6fe95d5e8f</td>
</tr>
<tr>
<td>File: Dropbox\Public\How to use the Public folder.txt,</td>
<td>12313c2e338ee010a8ddec97ed33d9158</td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\bin\Dropbox.exe,</td>
<td>9ed3cfe54cd2e797dc001e89</td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\bin\DropboxExt.14.dll,</td>
<td>6d74290856347cf8682277a54b433d4b</td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\bin\msvcp71.dll,</td>
<td>561fa2abb31d94a8fab762145f81667c2</td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\bin\msvcr71.dll,</td>
<td>86f1895ac85e8b17d99e7e68a70732</td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\bin\Uninstall.exe,</td>
<td>6420486a64af043b5b0e0a11b15c3e23</td>
</tr>
<tr>
<td>Hash values are not included for these as they will change according to usage;</td>
<td></td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\bin\Dropbox.exe.log</td>
<td></td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\bin\itag</td>
<td></td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\config.db</td>
<td></td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\config.dbx</td>
<td></td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\filecache.dbx</td>
<td></td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\host.db</td>
<td></td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\host.dbx</td>
<td></td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\sigstore.dbx</td>
<td></td>
</tr>
<tr>
<td>File: \AppData\Roaming\Dropbox\unlink.db</td>
<td></td>
</tr>
</tbody>
</table>

Table 5 - Dropbox Windows software files with MD5 values
5.2.5.4 Dropbox account information

It was observed that the Dropbox web account retains a record of computers used to access and synchronise with an account. The information is shown when using a web browser to access the Dropbox account, and is available under a tab labeled “My Computers” and also displays the timestamp of the last activity. There is also the ability to rename and unlink a linked computer. With the client software installed, it was possible to gain full access to an account from the Dropbox icon in the system tray, which has a link to ‘launch Dropbox website’ without requiring the username or password.

Accessing a Dropbox account via a web browser also has an option to show deleted files, which are available for 30 days for free accounts, and unlimited time for paid accounts (McClain 2011). When a file is deleted it is not visible unless the option to ‘Show Deleted Files’ is selected, however, the last modified date is not visible. The deleted files can be restored or permanently deleted from the browser. Dropbox retains a snapshot of a file each time it is modified and the history of a file can be viewed. A timeline of previous events can also be viewed, such as when files were uploaded, deleted, restored, or other computers were linked to the account, including the date and time for the event.

5.2.5.5 Directory Listing

For each of the VM hard drives, a directory listing was produced using AccessData FTK Imager 3.1.0. Analysis was also conducted using Encase to view the filenames. Analysis of the directory and file listings of the three control Base-VM hard drives revealed no references to the Enron sample data filenames, Dropbox files, or Dropbox sample files. It was concluded that references to Dropbox and the Enron sample files were not present in the control media directory listings prior to installing the software or accessing the files.

There were many references in the other VMs. In the three Upload-VM’s, there was a substantial amount of filenames seen, including on the User’s Desktop, in the Downloads folder, and in the Dropbox folder; ‘C:\User\[username]\Dropbox’. Dropbox URL references were also observed in the three Upload-VMs file listings, but not seen.
in the three Access-VMs or the Download-VMs, indicating this URL reference was created as part of the Dropbox client software installation.

When a browser was used to access a Dropbox account there were less filename references, in comparison with when the Dropbox client software was used to access cloud storage. However, with the browser being used to access the Dropbox account stored files, there were sufficient file name references remaining on the hard drives to identify the filenames that were accessed, including the Enron sample data files and the Dropbox sample files. When the files were downloaded from the Dropbox cloud storage, there were additional filename references observed on the hard drive, including link files. This indicates when a user uses Dropbox there will be references in the $MFT and directory listing to indicate this, and also to determine the filenames associated with Dropbox use.

Analysis of the three Eraser-VMs and three CCleaner-VM hard drives identified references to link files for the Enron sample files and Dropbox sample files. When all options were used for Eraser and CCleaner, the references to the filenames were removed. The use of DBAN completely erased the hard drives, and there was no data to analyse in these VMs.

5.2.5.6 Prefetch

Prefetch files are used by Windows to store information in relation to software activity, including the number of times software has run, and associated files used by the software. Analysis of the VM hard drives identified that prefetch files contained information relating to the filenames of the Dropbox executable, Dropbox sample files, and Enron test data filenames in all the VM hard drives except the control Base-VM hard drives. It was observed that even after erasing with both Eraser and CCleaner, there was enough information in prefetch files, such as notepad.exe.pf, wordpad.exe.pf, explorer.exe.pf and dllhost.exe.pf to indicate the presence and path of the Enron sample data files and the sample Dropbox files.
5.2.5.7 Link files

There were no Dropbox associated link files found within the three control Base-VMs. Link files relating to Dropbox and the Enron sample data files were observed for all Upload-VM, Download-VM, Eraser-VM and CCleaner-VM hard drives. The link files observed related to the filenames and folder names for the Dropbox executable, Dropbox sample files, such as ‘Boston City Flow.lnk’, ‘Photos.lnk’, and ‘Dropbox.lnk’, and the Enron test data. These were located in the ‘Windows\Recent’ and ‘Windows\Start’ folders in the Users ‘AppData’ directory. There were no Dropbox or Enron link files within the three Access-VM hard drives, indicating that if files were not downloaded from Dropbox to the computer, link files were not created.

5.2.5.8 Registry

Registry files were analysed using AccessData Registry Viewer 1.6.3.34 and RegRipper version 20080909. Analysis revealed there were no references to Dropbox or the Enron sample files in the three control Base-VM hard drives. In the Upload-VM, Access-VM and Download-VM registry files there were references to the Dropbox URL, Dropbox software files and folders, the Dropbox sample files, and the Enron test files. For example, the ‘RecentDocs’ key in the NTUSER.dat registry file provided a list of Dropbox and Enron related files and folders, a sample RegRipper output is listed in Table 6.

```
RecentDocs - recentdocs
**All values printed in MRUList\MRUListEx order.
Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs
LastWrite Time Fri Apr 20 12:13:24 2012 (UTC)
  14 = Public
  13 = How to use the Public folder.txt
  12 = Photos
  11 = How to use the Photos folder.txt
   8 = Sample Album
   10 = Pensive Parakeet.jpg
    9 = Costa Rican Frog.jpg
     7 = Boston City Flow.jpg
      3 = Dataset
      5 = Enron3111.rtf
      2 = Enron3111.jpg
      6 = Enron3111.docx
      4 = 311.txt

Table 6 - Output from RegRipper for Dropbox
```
There were no references to the Dropbox username within the registry files in any of the VM’s. Deleted Dropbox information was located within the NTUSER.dat registry files in the Chrome and IE CCleaner-VM’s, such as filenames and URL references, but not in the Firefox CCleaner-VM. Dropbox software file references were located in the SOFTWARE and SYSTEM registry hives when the Dropbox client software was installed, but not when a browser was used to access or download data from Dropbox. References were also found within the ‘UsrClass.dat’ registry files when the Dropbox client software was used, but not when a browser was used to access Dropbox.

5.2.5.9 Thumbcache

Analysis of the thumbnails stored within the thumbcache files within the three control Base VMs determined that there were no Dropbox Sample pictures or Enron sample picture thumbnails present prior to installing or accessing Dropbox. Thumbnails for the Dropbox sample pictures and the Enron sample picture were located in the Upload-VM, Download-VM and Eraser-VMs, however none were found in the Access-VMs or the CCleaner-VMs. This indicates that the thumbnail cache is a source for possible data relating to Dropbox use, but results may not be definitive. Dropbox can be used in certain circumstances without leaving any trace in thumbnails, such as accessing files using a browser but not downloading the files to a computer.

5.2.5.10 Event logs

The keyword term ‘dropbox’ was found within Windows ‘Windows Firewall.evtx’ event log files when client software was installed and run. For example; the text; ‘a rule has been added to the Windows Firewall exception list’ was located in the log file. There were also references within ‘GroupPolicy_Operational.evtx’ files relating to Dropbox and Enron sample data files when Internet Explorer was used to download data from a Dropbox web account. However the details of these entries could not be interpreted with Event Log Explorer, EvtParser or the inbuilt Windows Event Viewer. Similar entries were not seen when the Google Chrome or Mozilla Firefox browser were used.
5.2.5.11 Browser Analysis

Internet browsing information was analysed using Encase 6.19.4, Magnet Software (formerly JADsoftware) Internet Evidence Finder (IEF) v5 Standard Edition, Digital Detective NetAnalysis v1.53, and SQLite Database Browser v1.3. It was confirmed there was no Dropbox or Enron sample data in the internet history of the three control Base-VM files. The Dropbox account username could be determined when Mozilla Firefox and Google Chrome were used, but there was no record of this when Internet Explorer was used. However, if the username and password were stored when using IE version 9, the details were able to be retrieved by using Nirsoft IE PassView v1.26. The username was stored in Firefox in the ‘formhistory.sqlite’ database and in Chrome in the Autofill ‘Web Data’ file.

Dropbox website information was located in Cookie files, web history, FavIcons, and in the FileSlack of other files, when Internet Explorer, Mozilla Firefox, or Google Chrome was used to access Dropbox. Filenames for files downloaded, including the Dropbox sample files, were also located in the web history for the browsers. When using the Dropbox client software, references to the Enron sample data filenames and file path was located in the Internet Explorer History when Mozilla Firefox was used. Using all options with CCleaner removed all references in the Mozilla Firefox and Internet Explorer VMs, but information remained in the Google Chrome ‘FavIcon’ history to identify Dropbox usage, which included dates and times for last access.

5.2.5.12 Network Analysis

Analysis of the network traffic capture files was undertaken using Network Miner v1.0 and Wireshark Portable 1.6.5. The network traffic was seen on TCP Port 80 (HTTP) and 443 (HTTPS) only. When accessing Dropbox accounts using the Client software or a web browser, a session with an IP in the range 199.47.216.0 – 199.47.219.255, (registered to Dropbox), was first established on Port 80, and then a session with an IP in the range 199.7.48.0 – 199.7.63.255 or 199.16.80.0-199.16.95.255 on Port 80 then Port 443, which is registered to VeriSign/Thawte. Certificates appeared to be approved using VeriSign/Thawte services. The next IP’s accessed appeared to be in the range of 74.125.0.0 – 74.125.255.255, (which are registered to Google), and appear to be part of the Google Analytics service. IP numbers allocated to AmazonAWS, the Amazon Web
Services (EC2) service, were then observed. See Table 7 for a list of IP number ranges and the registered owner for the range.

<table>
<thead>
<tr>
<th>Organization</th>
<th>IP Address Ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>DROPBOX</td>
<td>199.47.216.0 - 199.47.219.255</td>
</tr>
<tr>
<td>VERISIGN</td>
<td>199.7.48.0 - 199.7.63.255</td>
</tr>
<tr>
<td></td>
<td>199.16.80.0 - 199.16.95.255</td>
</tr>
<tr>
<td>GOOGLE</td>
<td>74.125.0.0 - 74.125.255.255</td>
</tr>
<tr>
<td>AMAZON-AWS</td>
<td>23.20.0.0 - 23.23.255.255</td>
</tr>
<tr>
<td></td>
<td>50.16.0.0 - 50.19.255.255</td>
</tr>
<tr>
<td></td>
<td>107.20.0.0 - 107.23.255.255</td>
</tr>
<tr>
<td></td>
<td>184.72.0.0 - 184.73.255.255</td>
</tr>
<tr>
<td></td>
<td>174.129.0.0 - 174.129.255.255</td>
</tr>
<tr>
<td></td>
<td>204.246.160.0 - 204.246.191.255</td>
</tr>
<tr>
<td>SoftLayer Technologies Inc.</td>
<td>75.126.0.0 - 75.126.255.255</td>
</tr>
<tr>
<td>WestHost, Inc.</td>
<td>69.36.160.0 - 69.36.191.255</td>
</tr>
</tbody>
</table>

**Table 7 - Dropbox IP’s and registered organisation details**

When the Dropbox Windows Client software was downloaded, the URL name made reference to Amazon Cloudfront. Amazon Web Services EC2 was also listed for many URL’s, with additional information referencing Dropbox, such as ‘photos1.dropbox’, ‘photos2.dropbox’ and ‘photos3.dropbox’. IP numbers also observed in the packet captures are registered to SoftLayer Technologies and WestHost Inc. and according to the URL were related to client-stats.Dropbox.com and ocsp.digicert. OCSP is the Online Certificate Status Protocol, used by VeriSign/Thawte.

Viewing the network data, it appeared there were different URL names when using Dropbox client software in comparison with using a web browser, even if the same IP number was used. When the client software was used, the URL name seen in the network capture commenced with ‘v-client’ or ‘client’. When a browser was used, the URL commenced with ‘v-www’ or ‘dl-web’. This appears to be an indicator that when a Dropbox cloud account is accessed with client software the URL displayed ‘client’, and when a browser was used, ‘www’ or ‘web’ was displayed.

Network Miner was used to rebuild files from the captured packets, Dropbox webpage icon files, such as ‘landingboxbig.png’ were observed. However, no Enron sample data files or the Dropbox sample files were seen in the network traffic. There was no evidence of the filenames for the Enron or Dropbox sample files, nor any Dropbox user
credentials seen in the network traffic. On examining the captured network data, it appeared to be encrypted, with no readable text observed after a Dropbox user account logged in.

### 5.2.5.13 RAM Analysis

Analysis of the memory captures (VMEM files) was undertaken using Encase 6.19.4 and AccessData FTK 4.01. The term ‘dropbox’ was located in the memory captures of the Internet Explorer and Google Chrome control Base-VM files. None were located in the Mozilla Firefox control Base-VM file. The entries in the Chrome and Explorer memory appeared with lists of other websites, such as ‘drugstore.com’.

Observed within the memory captures when client software was used was the Dropbox username near to text; ‘u’email’; as listed in Table 8. This text can be used to search and locate potential usernames for Dropbox accounts. However, when a browser was used to access a Dropbox account the text; “login_email” was observed preceding the username. Surprisingly, the **password in cleartext** for the user account was able to be located in memory captures when the client software was installed (Upload-VM memory captures) near to the following extracted text;

- “... $0.00  free   name   periods   $0.00...”, or
- “... name  2 GB  periods $99.00  free   ...

Observed within memory captures were Dropbox and Enron filename references in all VM’s except the control Base-VM and the DBAN-VM. Website details such as ‘www.dropbox.com’ were recovered in all Upload-VM’s, and the majority of the Access-VM’s, Download-VM’s and CCleaner-VM’s. The full text of the Enron data files and the Dropbox sample files was recovered in the Upload-VM, Access-VM and Download-VM memory captures. Data carving was undertaken and resulted in the recovery of thumbnail pictures and partial picture files of the Dropbox sample files and the Enron sample pictures from the memory captures for Access-VM, Upload-VM, and Download-VM files.
Analysis of Pagefile.sys files was not as conclusive as the memory capture files. Undertaking data carving of the pagefile.sys files produced no matches to Dropbox sample files or Enron data pictures or thumbnails. Some thumbnail icons were located within the Mozilla Firefox Upload-VM, and a partial screen capture of the Dropbox website was found in the Google Chrome Access-VM pagefile, indicative of a thumbnail icon used in the Google Chrome navigation tab. The username or password was not located in the pagefile.sys files.

5.2.5.14 Eraser, CCleaner and DBAN (anti-forensics)

Remnants indicating the use of Dropbox were located after Eraser and CCleaner were used. This indicated that Eraser or CCleaner did not necessarily remove all data remnants in relation to the use of Dropbox or completely remove the information relating to the previous presence of the Enron sample files or the Dropbox sample files. The use of a full erasing tool, in this case Darik’s Boot and Nuke (DBAN), was found to remove all traces, but this also erased the operating system and user files. The operating system and user files would need to be re-installed, which could be a lengthy process and may dissuade the average user from undertaking this process to remove evidence of Dropbox usage.

5.2.6 Presentation

5.2.6.1 Analysis findings

A variety of data remnants were located when a user used Dropbox to store or access data. Dropbox client software, sample data files, and common locations were outlined in Table 5. This information enables an examiner to conduct hash analysis and keyword searches for the filenames listed, and will assist to determine if Dropbox client software has been used.
From the analysis undertaken, the task of identifying whether a Dropbox account has been accessed either using client software or a web browser can be determined in several ways. Information such as URL’s in the network traffic, filename data in prefetch and link files, filename references in directory listings, RecentDocs and typed URL’s in registry files, browser history, and information recovered from memory files all proved conclusive to identify Dropbox use, filenames and actual data from the files of interest. Once forensic analysis has determined that a Dropbox account has potential evidence of relevance to an investigation, the examiner can communicate this to relevant persons to enable them to respond to secure evidence in a timely manner. The findings are summarized in Table 9.

Network analysis determined that most of the network traffic was encrypted, and the file information was not recovered. Some of the network traffic was not encrypted, and hence useful to identify Dropbox access, and also to identify if Dropbox client software or a browser was used to access cloud storage. Analysis of the RAM captures recovered full files, filenames, the Dropbox username, and other data. This indicates that memory capture is important, and when possible should be undertaken.

It was also observed that Dropbox retains a record of computers used to access and synchronise with an account. This information would be beneficial to determine whether a particular computer was synchronised to a Dropbox account. There is a timeline history of files and computer synchronisation, and previous versions of files are also available through the Dropbox account when accessed with a browser.
<table>
<thead>
<tr>
<th>Control (Base-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username, Password, Software, URL, Enron KWS terms</td>
<td>Nil</td>
</tr>
<tr>
<td></td>
<td>Matches to ‘dropbox’ in ‘index.dat’ files, ‘msjint40.dll.mui’, ‘pagefile.sys’, unallocated clusters, and memory captures.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Client Software (Upload-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>Memory capture files near; ‘u’email’;</td>
</tr>
<tr>
<td>Password</td>
<td>Located in RAM – search for ‘free name periods’</td>
</tr>
<tr>
<td>Software</td>
<td>Dropbox 1.2.52.exe file located when downloaded.</td>
</tr>
<tr>
<td>URL</td>
<td>Dropbox Software installation under ‘[User]\AppData\Roaming\Dropbox’</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Dropbox sample files and folders at location ‘C:\Users[username]\Dropbox’</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>When software downloaded, URL’s included <a href="http://www.dropbox.com">www.dropbox.com</a></td>
</tr>
<tr>
<td>KWS terms</td>
<td>Multiple locations, including Prefetch, Link files, $MFT, Registry.</td>
</tr>
<tr>
<td></td>
<td>Located in Sync folder under User\Dropbox. Thumbnail pictures in Thumbcache</td>
</tr>
<tr>
<td></td>
<td>‘dropbox’ in Event Logs</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Browser Access (Access-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>FF and GC History; ‘formhistory.sqlite’ and ‘Autofill’. Memory capture files ‘login_email’</td>
</tr>
<tr>
<td>Password</td>
<td>Nil</td>
</tr>
<tr>
<td>Software</td>
<td>Nil</td>
</tr>
<tr>
<td>URL</td>
<td>Multiple locations; cookie, history, icons, pagefile.sys and unallocated</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Sufficient to identify files accessed with references to the filenames in Registry and Browsing History</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>Full text in RAM</td>
</tr>
<tr>
<td>KWS terms</td>
<td>Multiple matches to KWS terms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Browser Download (Download-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>FF and GC History; ‘formhistory.sqlite’ and ‘Autofill’.</td>
</tr>
<tr>
<td>Password</td>
<td>Nil</td>
</tr>
<tr>
<td>Software</td>
<td>Nil</td>
</tr>
<tr>
<td>URL</td>
<td>Multiple locations; cookie, history, icons, pagefile.sys and unallocated</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Sufficient to identify files accessed with references in $MFT, Link, Registry and Prefetch files</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>Via uncompressed zip or folder name; ‘Documents.zip’</td>
</tr>
<tr>
<td>KWS terms</td>
<td>References in Event Logs (IE)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eraser (Eraser-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>FF and GC History; ‘formhistory.sqlite’ and ‘Autofill’.</td>
</tr>
<tr>
<td>Password</td>
<td>Nil</td>
</tr>
<tr>
<td>Software</td>
<td>Nil</td>
</tr>
<tr>
<td>URL</td>
<td>Multiple locations; cookie, history, icons, pagefile.sys and unallocated</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Sufficient to identify files accessed with references to the filenames in Prefetch and Link files</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>筝Multiple matches to KWS terms</td>
</tr>
<tr>
<td>KWS terms</td>
<td>Thumbnail pictures in Thumbcache</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CCleaner (CCleaner-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>Nil</td>
</tr>
<tr>
<td>Password</td>
<td>Nil</td>
</tr>
<tr>
<td>Software</td>
<td>Nil</td>
</tr>
<tr>
<td>URL</td>
<td>Google Chrome FavIcon history</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Sufficient to identify files accessed with references to the filenames in Prefetch and Link files</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>Nil</td>
</tr>
<tr>
<td>KWS terms</td>
<td>Multiple matches to KWS terms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DBAN (DBAN-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username, Password</td>
<td>All Data erased, no information located</td>
</tr>
<tr>
<td>Software</td>
<td></td>
</tr>
<tr>
<td>URL</td>
<td></td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td></td>
</tr>
<tr>
<td>Enron sample files</td>
<td></td>
</tr>
<tr>
<td>KWS terms</td>
<td></td>
</tr>
</tbody>
</table>

Table 9 - Summary of Dropbox analysis findings
5.2.7 Feedback

During this research, one finding was that conducting research using virtual computers was beneficial to compare a range of different configurations of software and data access. The addition of the steps of ‘prepare’ and ‘feedback’ to a forensic analysis framework was beneficial to adequately prepare for analysis and also to incorporate lessons learnt into subsequent investigations. It is important to consider the positive and negative experiences during an examination and incorporate these into future examinations. It is also important to report any issues or improvements to other examiners, and the digital forensic community in general.

5.2.8 Complete

The final step is the completion of an investigation. For the purposes of this research, it was determined that analysis of other cloud storage services; Microsoft SkyDrive and Google Drive, will serve to widen the scope of the research and further assess the proposed framework. The files created for this research were archived, and working copies were available to compare when undertaking the next stages of the research.
5.3 Dropbox Analysis: Apple iPhone 3G

As discussed in Chapter Two, the previous Director of US Department of Defence Computer Forensics Laboratory stated that mobile devices add complexity to the identification of digital evidence associated with cloud computing (Zatyko & Bay 2012). To explore this in the context of this research, an Apple iPhone 3G was selected to determine the data remnants on a portable device, and to further assess the proposed Digital Forensic Analysis Cycle in different circumstances. An iPhone can be used to access a Dropbox account via the inbuilt browser, or client software (Dropbox App) can be used to access an account. Both processes will be tested for data remnants in this research.

5.3.1 Commence (Scope)

The scope of this part of the research is to examine the data remnants on an iPhone 3G running iOS 4.2.1 when used to access Dropbox via a browser, and when using the Dropbox iOS application.

5.3.2 Prepare

An iPhone 3G was selected which had not been used to access Dropbox previously as the device history and usage was known. MicroSystemation .XRY 6.2.1 was used to extract a logical image of the contents, excluding audio and video files, prior to undertaking the research. This extract was analysed to confirm there was no Dropbox related data on the device. Next, the inbuilt Apple Safari iOS browser was used to access the Dropbox user account created for this research and view the Dropbox sample files and Enron files stored remotely. A logical extract was then conducted using .XRY. The Dropbox iOS Application was then downloaded and installed to the iPhone 3G. The Dropbox account was then accessed using the Application, and the files stored in the account were viewed. A third logical extract with .XRY was then conducted.
5.3.3 Identify and Collect

Files were identified which would contain the information needed to conduct the analysis, in this instance being the .XRY extract files and the output of the .XRY software, including PDF reports and the files exported using .XRY. These were identified for each of the extracts; Base, Browser, and Application.

5.3.4 Preserve (Forensic Copy)

To observe the principles of forensic computer analysis (NIJ 2004; ACPO 2006) a forensic copy was made of the .XRY extract files, the file output, and the reports. As these were logical files, this was done in the X-Ways Evidence File Container format (ctr). MD5 hash values were used to ensure the forensic integrity of the data.

5.3.5 Analysis

Each of the forensic files was examined using forensic tools including; X-Ways Forensic version 16.5 and Guidance Software EnCase version 6.19.4. PList Explorer v1.0 was used to examine the contents of the Apple plist files extracted by .XRY. SQLite Database Browser 2.0 was used to examine the sqlite files.

5.3.5.1 Control - Base-XRY

Analysis of the control Base-XRY extract confirmed there was no data originally present relating to the Dropbox sample files, the Enron sample test data, or the Dropbox Application. In addition, no references were located for the term ‘dropbox’ or the website URL (www.dropbox.com).

5.3.5.2 Dropbox accessed via the iOS Safari Browser

The research username or account password was not located in the Browser-XRY extracts. Filenames for the Dropbox sample files and Enron test files were located in History.plist. This information was also extracted by XRY in the Web-History.txt file, as displayed in Table 10.
The dropbox website URL was also located in the History.plist and Web-History.txt files. An example is listed in Table 11. The URL was also located in the Cookies.binarycookies file.

<table>
<thead>
<tr>
<th>Web-History #</th>
<th>Application:</th>
<th>Web Address:</th>
<th>Access Count</th>
<th>Accessed:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web-History #</td>
<td>Application:</td>
<td>Web Address:</td>
<td>Access Count</td>
<td>Accessed:</td>
</tr>
</tbody>
</table>

### 5.3.5.3 Dropbox Application (App) used to access the research account

Analysis of the third XRY extract was able to determine the username used to access Dropbox account. This was located in the ‘com.getdropbox.Dropbox.plist’ file near the text; ‘DBAccountInfo’, and in the ‘keychain-backup.plist’. This information does not appear to be parsed by XRY. To determine if the application is present, information in the ‘iTunesstored2.sqlitedb’ file also referenced Dropbox. The text from the Enron files was not located in any of the extracts. No password was located in any of the extracts or files.

### 5.3.6 Presentation

#### 5.3.6.1 Analysis findings

Several data remnants were located when searching for evidence of Dropbox use on the test item (iPhone 3G). Initial analysis of the control extract identified no matches to the term ‘dropbox’, but matches in the subsequent extracts indicated that conducting a keyword search for the word ‘dropbox’ could indicate if Dropbox had been used.
When the in-built browser was used to access the Dropbox account, entries were left in the History.plist file, which were also listed in the .XRY Web-History.txt file. The username and password were not located in the extract. The Enron filenames were listed in the History.plist file and in the .XRY Web-History.txt file.

When the Dropbox application was installed and used to access the research account, the username was able to be determined from the ‘com.getdropbox.Dropbox.plist’ file. Information in the ‘com.apple.AppStore.plist’ file lists ‘http://itunes.apple.com/au/app/dropbox/id327630330’. However, the text from the Dropbox sample files or the Enron files was not located in any of the extracts, nor was it possible to locate the password for the account.

This research was limited by not being able to jailbreak the iPhone. .XRY and other iPhone extract solutions require an iPhone to be jailbroken to undertake a Physical extract. A future research opportunity is to undertake a physical extract of an iPhone and compare this to a logical extract to determine what information is available in comparison with the logical extract, and also to compare other iPhone forensic tools.

### 5.3.7 Feedback

This research was limited by not being able to jailbreak the iPhone. .XRY and other iPhone extract solutions require the iPhone to be jailbroken to undertake a Physical extract. A future research opportunity is to undertake similar research using an iOS 5 device, and also widen the devices examined to include other popular mobile device operating systems, such as Google Android and Microsoft Windows. In addition, the comparison of a physical extract of an iPhone to a logical extract to determine what information is available in comparison with the logical extract, and also to compare the process with other iPhone forensic software and hardware processes.

### 5.3.8 Complete

In the context of this research, the data created and used may be required in future research opportunities, and hence has been stored on multiple hard drives to enable future use.
5.4 **Dropbox Analysis: Case Study**

To illustrate an application of the research findings, the following hypothetical case study outlines where the information identified in this chapter can assist in an investigation. The circumstances relate to a hypothetical case of a person suspected of managing a botnet. The circumstance of the case study are greatly simplified for the purposes of this paper, and explore a possible response using currently available software and hardware forensic solutions;

“Police receive confidential information that Bob Smith has been seen managing a botnet. A close acquaintance was visiting Bob, and observed files relating to a ‘Poseidon’ botnet on the screen of a computer in Bob’s study.

An investigation case file was raised, and issued to IO Jones. IO Jones undertook background enquiries and determined that there had been prior allegations about Bob being involved in botnet offences, and that he had previous convictions for these offences. A Warrant was issued by the Court for IO Jones to attend the premises, and search and seize evidence to investigate these allegations.

IO Jones attended Bob’s home address and seized a PC, laptop, mobile phone, and some USB devices. These items were submitted to the Police Forensic Computing Section (FCS) for analysis”

5.4.1 **Commence (Scope)**

IO Jones completed a Request for Analysis form, outlining the scope of the investigation, namely to search for evidence of material related to the ‘Poseidon’ botnet.

5.4.2 **Prepare and Respond**

As this type of analysis is a process undertaken by the FCS on a regular basis, the examiners are trained and experienced in the analysis of the type of devices seized by IO Jones. A Standard Operating Procedure (SOP) relating to the analysis of computers and mobile devices is documented and referred to by examiners, and outlines the
process. Notes are an important aspect of the SOP, and all examiners maintain accurate notes of the process they undertake.

5.4.3 Identify
The devices submitted for analysis included a custom tower PC with 2 hard disk drives, a generic laptop with one hard disk drive, and an Apple iPhone 3G. The USB devices were identified as wireless adapters with no data stored within them. The USB devices were photographed and details noted for future reference.

5.4.4 Preserve (Forensic Copy)
The details of the hard drives were documented, including make, model and serial number. A forensic copy of the hard drives was then made using FTK Imager 3.1.0 and a Tableau T35es-R2 eSATA Forensic Bridge (hardware write blocking), and verified with MD5 and SHA1 hashes. Microsystemation .XRY was used to extract the logical data from the iPhone. Appropriate entries were made in the exhibit log relating to access to the evidence by the examiner.

5.4.5 Analysis
Analysis was then undertaken on the forensic copies of the data using X-Ways Forensic 16.5. ‘Poseidon’ botnet related files were located on the second hard drive in the tower PC, and were extracted and reports created relating to the information located.

A search was conducted using the filenames identified, and references were located indicating the files were also associated with a Dropbox account. Further analysis was undertaken and URL’s relating to Dropbox access were located, as was a username; bob@badstuff.com. Dropbox client software was located under ‘C:\Users\bob\AppData\Roaming\Dropbox\’. Additional names of further files associated with the Dropbox account were located in link files and prefetch files, such as dllhost.pf, indicating additional evidence was stored within the Dropbox account.
Analysis of the logical extract of the iPhone located the Dropbox application and the username in the ‘com.getdropbox.Dropbox.plist’ file and in the ‘keychain-backup.plist’ file. Filenames associated with Dropbox were located in the ‘History.plist’ file, and the information was also reported in the .XRY ‘Web-History.txt’ file.

IO Jones communicated with the legal section at Dropbox to ensure the data in the account was preserved and made available in relation to the investigation.

5.4.6 Presentation

A report was prepared outlining the process and the findings of the analysis. The selected files were copied to DVD, along with the associated reports for each file, outlining creation and access dates, file details, and hash values. A timeline was used to demonstrate interaction with the files from the tower PC and the iPhone, as listed in Table 12. In the timeline it is possible to see; when the documents were created via the metadata, when the documents were copied to the tower PC, link files related to the documents are created, Dropbox is accessed in relation to Poseidon, Dropbox software is searched and installed onto the PC, and files read on the PC and the iPhone.
### Table 12 - Example of Timeline (Dropbox Case Study)

<table>
<thead>
<tr>
<th>Source</th>
<th>Key Date</th>
<th>Event type</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-Ways Docx Metadata</td>
<td>9/02/2012 15:24:00</td>
<td>Doc created metadata</td>
<td>Poseidon.docx</td>
</tr>
<tr>
<td>X-Ways Docx Metadata</td>
<td>25/02/2012 12:02</td>
<td>Doc created metadata</td>
<td>Poseidon.rtf</td>
</tr>
<tr>
<td>Event Log Explorer</td>
<td>9/06/2012 12:27:09</td>
<td>Logon</td>
<td>cloud user</td>
</tr>
<tr>
<td>IEF Firefox</td>
<td>9/06/2012 12:34:00</td>
<td>Date Visited</td>
<td><a href="http://www.mozilla.com/en-US/firefox/13.0/firstrun/">http://www.mozilla.com/en-US/firefox/13.0/firstrun/</a></td>
</tr>
<tr>
<td>IEF Firefox</td>
<td>9/06/2012 12:34:00</td>
<td>Date Visited</td>
<td><a href="http://www.mozilla.org/en-US/firefox/13.0/firstrun/">http://www.mozilla.org/en-US/firefox/13.0/firstrun/</a></td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:41:49</td>
<td>Created</td>
<td>Poseidon.txt</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:41:49</td>
<td>Created</td>
<td>Poseidon.tar.gz</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:41:49</td>
<td>Created</td>
<td>Poseidon.docx</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:41:49</td>
<td>Created</td>
<td>Poseidon.jpg</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:41:49</td>
<td>Created</td>
<td>Poseidon.rtf</td>
</tr>
<tr>
<td>X-Ways Link Data</td>
<td>9/06/2012 12:41:49</td>
<td>Target Created</td>
<td>Poseidon.Lnk</td>
</tr>
<tr>
<td>X-Ways Link Data</td>
<td>9/06/2012 12:41:49</td>
<td>Target Created</td>
<td>Poseidon.Lnk</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:41:55</td>
<td>Created</td>
<td>Poseidon.Lnk</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:42:03</td>
<td>Created</td>
<td>NOTEPAD.EXE-EB1B961A.pf</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:42:08</td>
<td>Created</td>
<td>Poseidon.Lnk</td>
</tr>
<tr>
<td>iPhone XRY extract</td>
<td>9/06/2012 12:43:30</td>
<td>accessed</td>
<td>Poseidon - Dropbox</td>
</tr>
<tr>
<td>Internet History</td>
<td>9/06/2012 12:46:00</td>
<td>Start Date</td>
<td>DropboxSetup.exe</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:46:34</td>
<td>Created</td>
<td>Dropbox.exe</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:46:34</td>
<td>Created</td>
<td>Dropbox.lnk</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:46:34</td>
<td>Created</td>
<td>Dropbox.lnk</td>
</tr>
<tr>
<td>X-Ways Link Data</td>
<td>9/06/2012 12:46:34</td>
<td>Target Created</td>
<td>Dropbox.Lnk</td>
</tr>
<tr>
<td>X-Ways Link Data</td>
<td>9/06/2012 12:46:34</td>
<td>Target Created</td>
<td>Dropbox.Lnk</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:47:01</td>
<td>Created</td>
<td><a href="mailto:cloud@linux.live">cloud@linux.live</a>[1].txt</td>
</tr>
<tr>
<td>iPhone XRY extract</td>
<td>9/06/2012 12:48:30</td>
<td>accessed</td>
<td>Poseidon.rtf download</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:48:30</td>
<td>Created</td>
<td>Poseidon.lnk</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:50:16</td>
<td>Prefetch last run</td>
<td>NOTEPAD.EXE-EB18961A.pf</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:50:28</td>
<td>Created</td>
<td>Poseidon (2).lnk</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:50:35</td>
<td>Created</td>
<td>WORDPAD.EXE-1BCC3DB7.pf</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:50:36</td>
<td>Created</td>
<td>Poseidon (3).lnk</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:50:36</td>
<td>Prefetch last run</td>
<td>WORDPAD.EXE-1BCC3DB7.pf</td>
</tr>
<tr>
<td>X-Ways file list</td>
<td>9/06/2012 12:51:30</td>
<td>Accessed</td>
<td>Dropbox</td>
</tr>
</tbody>
</table>

### 5.4.7 Feedback

The examiner was concerned that Bob may have used other cloud storage services to store copies of the ‘Poseidon’ files, and communicated this with IO Jones. IO Jones interviewed Bob, but Bob would not disclose if other accounts existed. The examiner undertook additional searches for common cloud storage providers, such as Microsoft SkyDrive and Google Drive, identifying additional accounts held by Bob. Further enquiries were undertaken with the legal sections of the identified service providers to determine if these accounts held any relevant data.

### 5.4.8 Complete

At the completion of the original examination, the extracted data, notes and other material related to the investigation was appropriately stored in preparation for future legal action.
5.5 Dropbox Analysis: Summary

With the increasing use of cloud computing, particularly cloud storage services, by both public and private sectors, the continued development of the digital forensic discipline is more important than ever. Digital forensic disciplines need to have sound methodologies to facilitate rigorous standards of evidence. Methods and techniques developed for use in forensic analysis must be done so in a scientific manner. Outputs and conclusions from any forensic science can be brought into question where rigorous and scientific processes (inference, testing and hypotheses) have not been followed.

Risk areas associated with the use of cloud computing in Government identified by AGIMO (2011) include:

- The dynamic nature of the cloud can result in confusion as to where information physically resides (or is transitioning through) at any given point in time, and impede data/information retrieval by forensic investigators.
- Digital forensic examiners need to be aware of the different legislative and regulatory requirements in other jurisdictions as data/information is likely to be hosted off-shore.
- Skill requirements: There will be an increasing demand for forensic investigators to be trained in data/information retrieval in the cloud computing environment.

In addition, there is a need for training of prosecutors and other persons involved in the legal process, as well as ICT professionals to enable them to understand the evidence and information relating to cloud storage and cloud computing.

Research is ongoing and formal methodologies are still required to form a solid basis for the field in general. One issue for digital forensic practitioners, examiners and researchers is the constant and rapid change and output of new technologies. In order to develop tools that can be used by digital forensic practitioners and examiners, researchers must first develop techniques or frameworks to encompass the process of examination. The proposed framework is suitable for forensic analysis in both cloud and non-cloud-based investigations.
In relation to the storage of data using cloud service providers, the initial focus of an investigation will be to determine the provider and particular user details. This will enable examiners to identify the potential location of data, and act to secure this data. In the case study, an examiner can identify Dropbox use by undertaking hash analysis and examine common file locations to locate software and files. A Dropbox username can be determined from browser history when web access has been undertaken with Mozilla Firefox and Google Chrome, and when a user has stored their username with Microsoft Internet Explorer. By searching for a range of keyword terms, an examiner can determine whether Dropbox has been used on seized hardware. As outlined, there are a wide range of investigation points for an examiner to determine the use of Dropbox, such as; directory listings, prefetch files, link files, thumbnails, registry, browser history, and memory captures.

There are many Cloud Storage Service Providers, including Dropbox, Microsoft SkyDrive and Google Drive. The next chapter will apply the research methodology and proposed framework to Microsoft SkyDrive. This serves two purposes;

- To determine data remnants from other cloud service providers and encompass a methodology that can be used for current and future cloud storage providers or cloud computing providers, and;

- To test and validate the proposed framework to determine suitability and application.

The outcomes of this research proved to be beneficial for a current Police investigation when information was located which identified the use of Dropbox cloud storage to store illicit material. The details of the investigation cannot be expanded as it is ongoing, but it is important to highlight the real-world application of this research is already achieving the goals of expanding the information and methodology available in forensic examinations to identify when cloud storage services are being used.
This was the first experiment conducted for the thesis and the Apple Safari browser was not originally included in the scope. Monitoring the popularity of browsers on W3Counter (2012), it became apparent that the use of Safari is increasing, and hence was determined to include Safari in the overall experiment. Safari was included in the subsequent experiments, and was revisited in relation to Dropbox. The findings in relation to accessing Dropbox with the Apple Safari browser were the same as for the other cloud storage services, SkyDrive and Google Drive.
6 Digital Droplets: Microsoft SkyDrive forensic data remnants

In the previous chapter, the forensic analysis of cloud storage using Dropbox was explored when a user uploads, accesses, or downloads data using a Windows 7 PC or an Apple iPhone 3G. This was achieved using the proposed Digital Forensic Analysis Cycle, and data remnants were located in numerous locations. In this chapter, the focus is on Microsoft SkyDrive. The proposed framework is again used to step through the process in a logical manner.

Material presented in this chapter was submitted for publication and is currently under peer review:


Chapter 6 - Microsoft SkyDrive Analysis

6.1 Introduction

Microsoft® SkyDrive® offers a potential 25GB of free storage to users, and was previously known as Windows Live Folders and Windows Live SkyDrive. SkyDrive can be accessed using a web browser or client software. SkyDrive client software is available for installation on a computer or portable devices such as an Apple iPhone.
The focus of this chapter is to discover the remnants left on a client device, a computer running the Windows 7 operating system, and an Apple iPhone 3G, after a user accesses Microsoft SkyDrive, and to examine the benefits of using a framework when undertaking forensic analysis of a cloud computing environment.

The research will examine what data remnants remain after a user has uploaded, accessed, or downloaded data from SkyDrive using either the SkyDrive client software or a browser. Popular browsers include Microsoft Internet Explorer, Mozilla Firefox, Google Chrome, and Apple Safari (W3Counter 2012). These four browsers will be examined as part of this research to determine any differences in the ability to retrieve data remnants.

In the following discussion, again using the proposed digital forensic analysis cycle, the preparation and analysis of SkyDrive access using Windows 7 is detailed. Following this the analysis of SkyDrive access using an iPhone 3G running iOS version 4.2.1 is discussed, and then a hypothetical case study is used to explore the relevance of the research. The last section summarises the research findings and potential future research opportunities.

### 6.2 SkyDrive Analysis: Windows 7 PC

The framework in Chapter Four is again used to guide the analysis process. This is undertaken with a second cloud storage provider to validate the framework; to ensure that it is forensically sound, and is flexible enough to work with different cloud storage services, as there will most certainly be variation in the way criminal investigation is carried out for each type of cloud service.

#### 6.2.1 Commence (Scope)

The scope of this research is to determine the remnants such as a username, password, filenames, dates and times, or the presence of client software, to indicate which cloud service, if any, had been used on a Windows 7 PC. It is also assumed criminals would use anti-forensic tools such as Eraser or CCleaner to remove evidence of using SkyDrive, and hence, simulate such a scenario in the research. Memory analysis is increasingly becoming a source of evidence which should be captured and examined.
when possible (ACPO 2006; NIJ 2008) and is included in the scope. Network data capture is another potential source of information which will form part of the scope to determine what data is present within the network traffic from a host PC.

6.2.2 Prepare

To gather the data required to answer the research questions in relation to the use of SkyDrive, a variety of virtual computers (VMs) were created. It was decided to examine a variety of circumstances of a user accessing SkyDrive, and also to examine any differences when using different browsers. Multiple scenarios were explored, each making use of SkyDrive with the following browsers; Internet Explorer (IE), Mozilla Firefox (FF), Google Chrome (GC), and Apple Safari (AS). Multiple Virtual Machines (VMs) were created for each browser to replicate different circumstance of usage, as outlined in Chapter 3.

Specific software used in this research, in addition to the software outlined in Chapter Three, includes;

- SkyDriveSetup.exe 2012 16.4.3347.0416 Beta and 16.4.4111.0525
- Mozilla Firefox 13.0
- Internet Explorer 9.0.8112.16421IC
- Google Chrome 19.0.1084.56m
- Apple Safari 5.1.7 for Windows 7
- Eraser v6.09.2343 or v6.0.10.2620
- CCleaner v3.18.1707 or v3.19.1721
- Darik’s Boot and Nuke (DBAN) version 2.2.6 Beta
6.2.3 Identify and Collect

In the context of this research, files were identified that would contain the information needed to conduct the analysis; the virtual hard drives (VMDK files) in each VM folder, each memory instance (VMEM files), and each saved network capture file (PCAP). These were identified for each of the VMs.

6.2.4 Preserve (Forensic Copy)

To observe the principles of digital forensic acquisition and analysis (ACPO 2006; NIJ 2008), a forensic copy was made of each virtual hard drive (VMDK file) using AccessData FTK Imager CLI 2.9.0 in the Encase Evidence format (E01). In regard to the each memory file (VMEM) and network capture (PCAP) file, a forensic copy was made using Encase version 6.19.4, in the Logical Evidence format (L01), and the X-Ways Evidence File Container format (ctr). Hash values (MD5) were used to ensure the forensic integrity of the data.

6.2.5 Analysis

For this research, each of the forensic copies of the virtual computer hard drives, memory and network captures were examined using a range of forensic analysis tools, including; X-Ways Forensic version 16.5, Guidance Software EnCase version 6.19.4, AccessData FTK version 1.81.6 and version 4.01, Network Miner 1.0, Wireshark 1.6.5, Magnet Software (formerly JadSoftware) Internet Evidence Finder 5.52, and RegRipper version 20080909. Many of these tools are widely used for digital forensic analysis by law enforcement agencies and the private sector. In addition, Encase 6.5 and FTK Imager 2.5.3.14 have been tested by the Office of Law Enforcement Standards of the National Institute of Standards and Technology (NIST 2012).

6.2.5.1 Control - Base-VMs

Analysis of the four control Base-VM hard drive images confirmed there was no data originally present relating to the Enron sample test data and SkyDrive client software files. References were located for the term ‘skydrive’ in locations including; Temporary
Internet Files such as ‘welcome[1].htm’, and ‘ie[1].htm’; and in system files such as ‘$MFT’, ‘$UsnJrnl.$J’, ‘$LogFile’ and ‘AgRobust.db’. References were also located in unallocated clusters. This should be borne in mind, as this indicates the presence of the keyword term ‘skydrive’ does not necessarily indicate that SkyDrive has been used. This highlights that the context of a search result needs to be analysed to determine the reason for a keyword match, rather than drawing a conclusion at face value of the presence of data.

6.2.5.2 SkyDrive Client Software

Analysis of the four Upload-VMs located the SkyDriveSetup.exe file that was downloaded from the Microsoft SkyDrive website. When run, the SkyDrive software was installed into the Users ‘C:\Users\[username]\AppData\Local\Microsoft\’ directory in a folder called ‘SkyDrive’. SkyDrive synchronised files and folders were observed at the default SkyDrive folder location (‘C:\Users\[username]\SkyDrive’), and the following folders were created; ‘Documents’, ‘Pictures’, and ‘Public’. Unlike Dropbox™ (which was used as the case study in the first experiment), SkyDrive had no sample files, and the folders were empty when first installed. Information regarding the setup file for the client software is listed in Table 13. When a new version of the SkyDrive Windows client software is released, the hash values of the standard files will need to be recalculated. In comparing the files across two different client software releases, hash matches were observed for the following files; ‘CollectSkyDriveLogs.bat’, msvcp.dll, sqmapi.dll and msvcr.dll. All other installed files had different hash values in the subsequent release.

<table>
<thead>
<tr>
<th>Name</th>
<th>SkyDriveSetup.exe Path</th>
<th>SkyDriveSetup.exe Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Path</td>
<td>\Users\cloud\AppData\Local\Microsoft\SkyDrive\16.4.3347.0416</td>
<td>\Users\cloud\AppData\Local\Microsoft\SkyDrive\16.4.4111.0525</td>
</tr>
<tr>
<td>Size</td>
<td>4.7 MB</td>
<td>4.9 MB</td>
</tr>
<tr>
<td>Created</td>
<td>27/05/2012 17:07</td>
<td>10/06/2012 11:43</td>
</tr>
<tr>
<td>Int. creation</td>
<td>16/04/2012 9:18</td>
<td>25/05/2012 16:39</td>
</tr>
<tr>
<td>Hash</td>
<td>C95FBC79481F3B705ADF289B94DA9D4B</td>
<td>6961F3EC7F861C65091B8FB35086561</td>
</tr>
<tr>
<td>Metadata signed</td>
<td>‘true’</td>
<td>‘true’</td>
</tr>
<tr>
<td>CA</td>
<td>Microsoft Timestamping PCA</td>
<td>Microsoft Timestamping PCA</td>
</tr>
<tr>
<td>Signing date</td>
<td>16/04/2012 19:02:46</td>
<td>26/05/2012 02:25:52</td>
</tr>
</tbody>
</table>

Table 13 - SkyDriveSetup information (using X-Ways 16.5)
Approximately 152 files were created when the installation software was run. The observed folder structure is listed in Figure 4, and version number is displayed as a folder. Within this folder was the executable file. The Logs folder is of particular interest as there is a \texttt{SyncDiagnostics.log} file which has a list of the files synchronised, sizes, dates and times, an example is listed in Table 15. The information in this file is particularly important as it displays the history of files associated with a SkyDrive account, and is discussed in the next section.

![Folder list of the AppData\local\SkyDrive folder (X-Ways 16.5)](image)

Details for the SkyDrive executable are listed in Table 14, including the MD5 hash value, which can be used to quickly locate the presence of SkyDrive client software; even in the instance it has been renamed. As the software is updated, a forensic examiner will need to calculate new hash values. This can be done by installing the software to a clean PC, and calculate hash values for the installed files.

<table>
<thead>
<tr>
<th>Name</th>
<th>SkyDrive.exe</th>
<th>SkyDrive.exe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path</td>
<td>\Users\cloud\AppData\Local\Microsoft\SkyDrive</td>
<td>\Users\cloud\AppData\Local\Microsoft\SkyDrive</td>
</tr>
<tr>
<td>Size</td>
<td>290 KB (296,672)</td>
<td>290 KB (296,672)</td>
</tr>
<tr>
<td>Created</td>
<td>27/05/2012 17:07:59</td>
<td>09/06/2012 12:46:34</td>
</tr>
<tr>
<td>Int. creation</td>
<td>16/04/2012 09:18:43</td>
<td>25/05/2012 16:40:09</td>
</tr>
<tr>
<td>Hash</td>
<td>D943C6ADF16045E7FE00233FF9C15376</td>
<td>4F7C80E5A420E47B584055EB1AC61562</td>
</tr>
<tr>
<td>Metadata</td>
<td>signed true CA Microsoft Timestamping PCA Signing date 16/04/2012 18:50:38</td>
<td>signed true CA Microsoft Timestamping PCA Signing date 26/05/2012 02:12:21</td>
</tr>
</tbody>
</table>

\textbf{Table 14 - SkyDrive executable information (using X-Ways 16.5)}
Analysis of the SyncDiagnostics.log file located a range of information of interest. There is file system and metadata information contained within the log file, an example is listed in Table 15. The SkyDrive OwnerID is listed for each file, and the associated file or folder number. The filename and path is listed, with the creation and modified times. The creation and modTime information appears to be in Unix Numeric Value format, which can be decoded using conversion software such as Digital Detective’s DCode v4.02a; for example, for the Enron3111.jpg file listing in the Cloud Metadata section, the listed information is modTime=1328766794. This is converted with DCode to ‘Thu, 09 February 2012 15:23:14 +0930’, which is the original modified time for the Enron jpg file. The file size is listed, as are file and folder totals.

<table>
<thead>
<tr>
<th>Cloud Metadata:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- folder XXXXAB89F707D01106 'C:\Users\cloud\SkyDrive\Dataset', creationTime=1338104478, modTime=1338104478</td>
</tr>
<tr>
<td>- file XXXXAB89F707D01107 'C:\Users\cloud\SkyDrive\Dataset\3111.txt', size=2734, creationTime=1338104500, modTime=1100097650</td>
</tr>
<tr>
<td>- file XXXXAB89F707D01109 'C:\Users\cloud\SkyDrive\Dataset\Enron3111.docx', creationTime=1338104500, modTime=1328766856</td>
</tr>
<tr>
<td>- file XXXXAB89F707D01110 'C:\Users\cloud\SkyDrive\Dataset\Enron3111.jpg', size=315868, creationTime=1338104500, modTime=1328766794</td>
</tr>
<tr>
<td>- file XXXXAB89F707D01111 'C:\Users\cloud\SkyDrive\Dataset\Enron3111.rtf', size=34215, creationTime=1338104500, modTime=1303133578</td>
</tr>
<tr>
<td>- file XXXXAB89F707D01108 'C:\Users\cloud\SkyDrive\Dataset\enron_with_categories.tar.gz', size=4523350, creationTime=1338104500, modTime=1308027150</td>
</tr>
<tr>
<td>- folder XXXXAB89F707D01104 'C:\Users\cloud\SkyDrive\Documents', creationTime=1335323578, modTime=1335323578</td>
</tr>
<tr>
<td>- folder XXXXAB89F707D01102 'C:\Users\cloud\SkyDrive\Pictures', creationTime=1335323577, modTime=1335323577</td>
</tr>
<tr>
<td>- folder XXXXAB89F707D01103 'C:\Users\cloud\SkyDrive\Public', creationTime=1335323577, modTime=1335323577</td>
</tr>
</tbody>
</table>

Cloud Total: 4 folders, 6 files, 5607286 bytes

<table>
<thead>
<tr>
<th>File System:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanning 'C:\Users\cloud\SkyDrive'</td>
</tr>
<tr>
<td>- file 'C:\Users\cloud\SkyDrive\Desktop.ini' ignored, size=96, creationTime=1339299083, modTime=1339299083</td>
</tr>
<tr>
<td>- folder 'C:\Users\cloud\SkyDrive\Desktop'</td>
</tr>
<tr>
<td>- file 'C:\Users\cloud\SkyDrive\Documents'</td>
</tr>
<tr>
<td>- folder 'C:\Users\cloud\SkyDrive\Pictures'</td>
</tr>
<tr>
<td>Folder 'C:\Users\cloud\SkyDrive' Total: 4 folders, 1 files</td>
</tr>
</tbody>
</table>

Scanning 'C:\Users\cloud\SkyDrive\Dataset' |
| - file 'C:\Users\cloud\SkyDrive\Dataset\3111.txt', size=2734, creationTime=1339299089, modTime=1100097650 |
| - file 'C:\Users\cloud\SkyDrive\Dataset\Enron3111.docx', size=14072, creationTime=1339299107, modTime=1328766856 |
| - file 'C:\Users\cloud\SkyDrive\Dataset\Enron3111.jpg', size=315868, creationTime=1339299108, modTime=1328766794 |
| - file 'C:\Users\cloud\SkyDrive\Dataset\Enron3111.rtf', size=34215, creationTime=1339299109, modTime=1303133578 |
| - file 'C:\Users\cloud\SkyDrive\Dataset\enron_with_categories.tar.gz', size=4523350, creationTime=1339299090, modTime=1308027150 |

Folder 'C:\Users\cloud\SkyDrive\Dataset' Total: 0 folders, 5 files

Filesystem Total: 4 folders, 6 full files, 0 stub files, 5607286 bytes, quotaState = 1

Table 15 - example of SyncDiagnostics.log file contents
Located in the ‘settings’ folder was a data file with the name matching the Owner ID number of the SkyDrive account, which also listed the file and folder names of the files and folders synchronised with the SkyDrive account, such as:

```
SkyDrive\settings\XXXXA4AB89F707D0.dat.
```

Also located in the settings folder was an INI file with the Owner ID as the name, `XXXXA4AB89F707D0.ini`. Contained within this file were the computer name, OwnerID, and SkyDrive sync folder information, as listed in Table 16;

```
nickname =  
installID = 1  
installName = WIN-TH8PP2HRCAG  
maxNumRequests = 10  
originatorID = a083ec80-17ac-4537-89c9-c723a05a2c64  
library = 1 4 XXXXA4AB89F707D0I101 1339211794 "SkyDrive" Me personal  
"C:\Users\[username]\AppData\Local\Microsoft\SkyDrive\16.4.4111.0525\folder, called ‘CollectSkyDriveLogs.bat’. Examining the code in this file indicated it will package the SkyDrive log files into a CAB file with the date and time of preparation. The CAB file will be placed on the User’s Desktop by default. The hash value of the .bat file remained the same across two different client software releases, and was; 40A379C64F2A1B473D8A5F8B760FC7C8.`
Located at the root of the hard drive was a temporary folder for SkyDrive, which had a sub-folder named with the Base Machine SID and the Users Windows ID number, for example:

```
C:\SkyDriveTemp\S-1-5-21-XXXX090636-1120983347-324155541-1000
```

The SkyDrive username was located in a cookie file at the following location; `\Users\[username]\AppData\Roaming\Microsoft\Windows\Cookies\cloud@login.live[1].txt`. This was also located in the $MFT, and $LogFile. However, in the Internet Explorer Upload-VM, the username was in `\Users\[username]\AppData\Roaming\Microsoft\Windows\Cookies\WGYG41GV.txt` and also in the $MFT.

When installed, the SkyDrive client software ran automatically when the Windows operating system started, and logged in to the user account without prompting for a password. This can be of assistance in an investigation, as a forensic copy of a seized computer hard drive can be used with software that allows the forensic copy of a hard drive to be run within a virtual environment. For example, Virtual Forensic Computing (VFC) or LiveView will scan the image of a hard drive and prepare the requisite files to run an operating system on a hard drive within VMWare Player. If the forensic copy of the hard drive contains the SkyDrive client software with a user account and password already stored, when started, the PC will automatically sign in to the SkyDrive account and provide an examiner access to the files stored within the SkyDrive user account.

Care would also need to be taken when connecting a forensic image to the Internet. Legal authority would be required to ensure an examiner has the appropriate authority within their jurisdiction to examine the data stored within the cloud storage account, which could potentially be stored overseas, or in another jurisdiction. In Australia, Section 3LA of the *Crimes Act 1914* (Cth), has a provision for the executing officer of a warrant, or a constable assisting, to access data which includes data not held at the premises, i.e. accessible from a computer or data storage device.\(^1\)

For this research the process of copying all the files contained within the installed SkyDrive folders to another PC was undertaken, however, this did not allow access to the SkyDrive account without knowing the username and password. The process of

\(^1\) Chapter 8 further discusses legislation regarding the process of accessing remote data.
copying the software files from a PC previously worked with Dropbox client software (see McClain (2011) and Chung et al. (2012)).

6.2.5.3 SkyDrive Account when accessed via a browser

It was observed that the SkyDrive web account (accessed via the website https://skydrive.live.com) retains a record of computers used to access and synchronise with an account. The information is shown when using a browser to access the SkyDrive account, and is available on the left of the displayed page with the heading; ‘Computers’. When selecting a computer name there was an attempt to connect to the selected computer. When unable to connect, a message was displayed. There is also the ability to unlink a linked computer. The username was displayed at the top right of the browser, and the OwnerID number was observed in the URL, for example; https://skydrive.live.com/?cid=XXXXXXXXXXXXXXXX#

When selecting a file in a SkyDrive account using a browser, there was an option to display the metadata within the file by clicking on an ‘Information’ link on the right hand of the page. Folder information was also displayed under Information, and included the time and date a folder was added, and modified. Version history was displayed for files when viewing the file, and previous versions could be viewed or downloaded.

When downloading folders from SkyDrive via the browser, the folder and files were packed into an uncompressed ZIP file. When done from the browser, choosing the root SkyDrive folder, the ZIP was named ‘skydrive’ with the date, for example; ‘skydrive-2012-06-09.zip’. When a download was undertaken from a folder, the ZIP was named with the folders name but not the date, for example; ‘Dataset.zip’. As these are uncompressed ZIP files, the contents would be located when data carving a hard drive, ignoring sector or cluster boundaries.

6.2.5.4 Keyword Search Terms

Keyword search terms were determined from the filenames observed, and the text from within the Enron data files. These included the following:
the username and password of the SkyDrive account created for this research,

- ‘skydrive’,
- ‘skydrivesetup.exe’,
- ‘dataset.zip’, ‘3111.txt’, and ‘enron3111’,
- ‘Enron Wholesale Services’,
- ‘enabledogfood’,
- SkyDrive OwnerID

6.2.5.5 Directory Listings

Analysis was conducted using X-Ways Forensic 16.5 and Encase 6.19.4 to view the filenames stored within the VMs, determined from the directory listing ($MFT^2 files). Analysis of the directory and file listings of the four control Base-VM hard drives revealed no references to the Enron sample data filenames, or SkyDrive client software filenames. It was concluded that references to SkyDrive and the Enron sample files were not present in the control media directory listings prior to installing the software or accessing the files. However, references were located for deleted SkyDrive icon files in the Internet Explorer Base-VM. This highlights that whilst there were no references to the SkyDrive software or sample data, it is possible to locate references to filenames with SkyDrive in the name. The presence of the search term does not indicate the use of SkyDrive, as in this instance it was related to the installation and update of Internet Explorer, and not any use of the client software or intentional browsing to the SkyDrive website. The reason for the presence of a search term or filename needs to be determined prior to a conclusion to be reached.

^2 ‘The NTFS file system contains at its core, a file called the master file table (MFT). There is at least one entry in the MFT for every file on an NTFS volume, including the MFT itself’ Microsoft (2008), How Ntfs Reserves Space for Its Master File Table (Mft), <http://support.microsoft.com/kb/174619>..
Not surprisingly, there were references to the SkyDrive client software in the Upload-VMs, which correlated to the VMs where the client software was used. There were no references to the client software in the other VMs. There were references to the filenames for the Enron sample files in all the other VMs (excluding the Base VM’s), and when the files were downloaded via the browser from the root folder. In the Upload-VM’s, there was a substantial amount of filenames seen, including on the User’s Desktop, in the Downloads folder, and in the SkyDrive folder; ‘C:\User\[username]\SkyDrive’.

When a browser was used to access a SkyDrive account there were fewer filename references for the Enron files, (in comparison with when the client software was used to access cloud storage). However, with a browser being used to access the SkyDrive account stored files, there were sufficient file name references remaining on the hard drives to identify the filenames that were accessed, such as the Enron sample data filenames. In the circumstances when the files were downloaded and opened from the SkyDrive cloud storage, there were additional filename references observed on the hard drive, including link files. This indicates when a user accesses SkyDrive there will potentially be references in the $MFT directory listing to indicate this, and also to potentially determine the filenames associated with SkyDrive use by searching for SkyDrive filename entries in the $MFT.

Link file references remained after Eraser and CCleaner had been used to remove the files and ‘clean’ the hard drive. Using CCleaner with all options selected removed the link files (however sufficient information remained in other locations such as prefetch files, relating to SkyDrive and the Enron sample files, discussed in the next section). The use of Darik’s Boot and Nuke (DBAN) completely erased the hard drives, and there was no data to analyse in these VMs.

6.2.5.6 Prefetch Files

Analysis of the VM hard drives identified that prefetch\(^3\) files stored information relating to the filenames of the SkyDrive executable and Enron test data filenames. This was

\(^3\) ‘Windows keeps track of the way a computer starts and which programs are commonly opened. Windows saves this information as a number of small files in the prefetch folder.’ Microsoft (2012b),
located in all the VM hard drives except the control Base-VM hard drives. It was observed that even after running CCleaner there was enough information in prefetch files, such as notepad.exe.pf, wordpad.exe.pf, explorer.exe.pf and dllhost.exe.pf, to indicate the presence and path of the Enron sample data files. Also located were skydrive.exe.pf, skydriveconfig.exe.pf and skydrivesetup.exe.pf prefetch files when the client software was installed. Information located within the prefetch files included the number of times run and last run time and date (Table 17).

![Table 17 - SkyDrive.EXE Prefetch (X-Ways 16.5)](image)

Analysis was conducted using X-Ways Forensic 16.5, (which parses information out of Windows Prefetch files). For example, an extract from the c:\windows\prefetch\wordpad.exe.pf prefetch file from the Internet Explorer CCleaner-VM displays the Run Count, Last run time, and the path of the files opened, including an entry for the Enron3111.docx file used as a sample, highlighted in bold in Table 18;

![Table 18 - Wordpad.EXE Prefetch (X-Ways 16.5)](image)

An example from the Notepad.exe pf prefetch file from the Google Chrome Upload-VM displays the path which includes the SkyDrive default storage location (highlighted in bold in Table 19):

<table>
<thead>
<tr>
<th>*** Prefetch ***</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOTEPAD.EXE-EB1B961A</td>
</tr>
<tr>
<td>Run Count: 1</td>
</tr>
<tr>
<td>Last Run: 10/06/2012 11:45:23</td>
</tr>
<tr>
<td>\DEVICE\HARDDISKVOLUME1\USERS[username]\SKYDRIVE\DATASET\3111.TXT</td>
</tr>
</tbody>
</table>

Table 19 - Notepad.EXE Prefetch (X-Ways 16.5)

6.2.5.7 Link Files

Analysis of Link files was undertaken using X-Ways 16.5, which parses the information within link files and displays this in a Preview pane. There were no SkyDrive associated link files found within the four control Base-VMs, or the Access-VM’s. Link files relating to SkyDrive and the Enron sample data files were observed for all Upload-VM, Download-VM, Eraser-VM and CCleaner-VM hard drives. When CCleaner was used with all options selected, there were no logical link files containing SkyDrive or the Enron sample data filenames (however the information from the deleted Link files was still present and recoverable from the hard drives).

The link files observed related to the filenames and folder names for the SkyDrive executable, such as:

- `\Users\[username]\Links\SkyDrive.lnk` (see Table 20)
- `\Users\[username]\AppData\Roaming\Microsoft\Windows\Recent\SkyDrive.lnk`
- `\Users\[username]\AppData\Roaming\Microsoft\Windows\Start Menu\Programs\Microsoft SkyDrive.lnk`

---

Table 20 - SkyDrive.lnk Link file contents (X-Ways 16.5)

Link Files relating to the Enron test data were also located. These were located in the ‘Windows\Recent’ and ‘Windows\Start’ folders in the Users ‘AppData’ directory, such as Table 21;

Table 21 - Enron3111.lnk Link file contents (X-Ways 16.5)
There were no SkyDrive or Enron link files within the four Access-VM hard drives, indicating that if files were not downloaded from SkyDrive to the computer and opened, link files were not created.

**Thumbcache Files**

Analysis of the thumbnail pictures stored within the thumbcache⁵ files within the four control Base VMs determined that there were no Enron sample picture thumbnails present prior to installing or accessing SkyDrive. Thumbnails for the Enron sample picture were located in the Upload-VM, Download-VM and Eraser-VMs, however none were found in the Access-VMs or the CCleaner VMs. This indicates that the thumbnail cache is a source for possible data relating to SkyDrive use, but results may not be definitive. It is possible for SkyDrive to be used in certain circumstances without leaving thumbnails, for example; accessing files using a browser but not downloading the files to a computer. An example output for thumbcache information is included in Table 22 and Figure 5;

<table>
<thead>
<tr>
<th>Name</th>
<th>3DE592C2E9065148</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>existing file</td>
</tr>
<tr>
<td>Type</td>
<td>jpg</td>
</tr>
<tr>
<td>Type status</td>
<td>newly identified</td>
</tr>
<tr>
<td>Type descr.</td>
<td>JPEG</td>
</tr>
<tr>
<td>Category</td>
<td>Pictures</td>
</tr>
<tr>
<td>Evidence object</td>
<td>Safari 5 erase, Partition 1</td>
</tr>
<tr>
<td>Path</td>
<td>\Users[\user]\AppData\Local\Microsoft\Windows\Explorer\thumbcache_256.db</td>
</tr>
<tr>
<td>Size</td>
<td>11.7 KB (11,932)</td>
</tr>
<tr>
<td>Modified</td>
<td>10/06/2012 13:24:38</td>
</tr>
<tr>
<td>Int. ID</td>
<td>57845</td>
</tr>
<tr>
<td>Int. parent</td>
<td>55325</td>
</tr>
<tr>
<td>Report table</td>
<td>thumbcache carved</td>
</tr>
<tr>
<td>Comment</td>
<td>file:C:/Users/[user]/Desktop/skydrive-2012-06-09/Dataset/Enron3111.jpg</td>
</tr>
</tbody>
</table>

**Table 22 - Thumbcache information for Enron jpg file (X-Ways 16.5)**

⁵ ‘Windows 7 creates small thumbnail images of graphic files. There are files named Thumbcache_32.db, Thumbcache_96.db, Thumbcache_256.db & Thumbcache_1024.db which correspond to the thumbnails stored for that specific user account and size’ Mueller, L (2010), Windows 7 Forensics - Part Iv - Thumbcache*.Db', [http://www.forensickb.com/2010/01/windows-7-forensics-part-iv.html]..
6.2.5.8 Event Log Files

Event Log Files can be good sources of information relating to system, software, and other events recorded by Windows operating systems. In this research, there were no records relating to SkyDrive or the Enron sample files within the event log files in any of the VM’s created for this research. Research into the other cloud service client software suggested that event log files can contain information of interest, and hence, event log files should not be precluded when conducting analysis.

6.2.5.9 Registry Files

Windows Registry files were parsed using RegRipper version 20080909 and parsed and analysed with X-Ways Forensic 16.5. Analysis revealed there were no references to SkyDrive or the Enron sample files in the four control Base-VM hard drives. There were no references to the SkyDrive username or password within the registry files in any of the VM’s. When the SkyDrive client software was used, there were references found in the Upload-VM’s for all browsers in NTUser.dat, UsrClass.dat, SOFTWARE and SYSTEM Registry Files. This information was located using keyword searches for the term ‘skydrive’ across the Registry files and also the parsed text from the Registry files (output from RegRipper and the X-Ways HTML Reports). Specific keys were created when the client software was installed, such as SkyDrive.SyncFileInformationProvider in UsrClass.dat, and Software\Microsoft\SkyDrive in NTUSER.dat files.
When a browser was used to access SkyDrive accounts, there were references to the filenames of the files in the Download-VM registry files for all browsers. When Internet Explorer was used for accessing but not downloading the files (IE Access-VM) there were also references, but not for the other browsers Access-VM’s. There were references to SkyDrive URLs, SkyDrive software files and folders, and the Enron test files located in the browser VMs. For example, the ‘RecentDocs’ key in the NTUSER.dat registry file provided a list of SkyDrive and Enron related files and folders, a sample RegRipper output is listed in Table 23;

```
RecentDocs - recentdocs  
**All values printed in MRUList\MRUListEx order.  
Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs  
LastWrite Time Sat Jun  9 03:42:27 2012 (UTC)  
3 = Dataset  
6 = Enron3111.rtf  
5 = Enron3111.jpg  
4 = Enron3111.docx  
2 = 3111.txt  
1 = Downloads  
0 = skydrive-2012-06-08.zip  
```

Table 23 - Output from RegRipper for SkyDrive

Deleted information regarding SkyDrive was located within the NTUSER.dat registry files after CCleaner was run, such as the previously observed filenames and URL references. References were also found within the UsrClass.dat registry files when the SkyDrive client software was used, such as the download of the file ‘skydrive-2012-xx-xx.zip’.

### 6.2.5.10 $Recycle Bin

Files that were deleted were easily located in the $Recycle.Bin folder in a folder with the SID of the user. The information files (beginning with $I) included the data relating to the original file and when it was deleted, such as the following example relating to the SkyDrive folder extracted from the zip file (Table 24);
Table 24 - $Recycle.Bin $I information for deleted SkyDrive folder (X-Ways 16.5)

Also located in the Recycle Bin were the file contents files (beginning with $R), which contained the original file or folder contents. In the scope of this research this included the unzipped contents of the zip file downloaded from SkyDrive account using the Firefox browser.

6.2.5.11  Data Carve

Data carving is the process of searching through allocated or unallocated data to locate files based on known headers and footers, such as 0xFFD8 and 0xFFD9 for a jpg file. This was undertaken using X-Ways 16.5 File Header Signature Search across all VM’s created for this research. Thumbnail icons and large size pictures were carved from all VM’s, except the Base-VM’s.

6.2.5.12  Browser Analysis

Internet browsing information was analysed using Magnet Software (formerly JADsoftware) Internet Evidence Finder (IEF) v5.52 Standard Edition, X-Ways Forensic 16.5 64-bit, and SQLite Database Browser v1.3. It was confirmed there were no references to SkyDrive or Enron sample data in the internet history of the four control Base-VMs. The SkyDrive account username could be determined when Mozilla Firefox and Google Chrome were used to access the account using the browser, but the username was not located when Internet Explorer or Safari was used. The username was stored in Firefox in the ‘formhistory.sqlite’ database and in Chrome in the Autofill ‘Web Data’ file. If the username and password were stored by the user when using IE version 9, the details were able to be retrieved by using Nirsoft IE PassView v1.26.

SkyDrive website URL’s (skydrive.live.com or login.live.com) were located in a range of areas; such as; Cookie files, web history, FavIcons, in the FileSlack of other files, unallocated space, and Pagefile.sys. This occurred for all browsers. The previously
mentioned term ‘enabledogfood’ found in the SkyDrive client software; was not located in the extracted internet history.

Filenames for files downloaded were located in the web history for all browsers (except for the Base-VMs). Data was located in Browser history files, such as index.dat files and downloads.sqlite. When using the Google Chrome browser to access SkyDrive, references were located in a range of browser system files; such as Top Sites, History, Shortcuts, Preferences, History Provider Cache, FavIcons, Current Tabs and Current Sessions. When using the SkyDrive client software, references to the Enron sample data filenames and file-paths was located in the Internet Explorer History even when other browsers were used; such as Mozilla Firefox, Google Chrome and Apple Safari.

Using CCleaner with standard options did not remove the references previously located in relation to SkyDrive use. Using all options on CCleaner removed all references in the Internet Explorer VMs, but information remained in the other browsers; Firefox, Chrome and Safari, to identify SkyDrive usage, which included filenames and times and dates of access. Interestingly, this included filename and path history stored in Internet Explorer index.dat, but not web browsing history.

IEF 5.5 introduced the ability to report information relating to cloud storage, including SkyDrive. Using this option, data was located in the Upload-VM files. This data included filenames and folder names, the owner name and the owner ID number. Information was also located in the Internet Explorer and Safari Access-VM’s. It appears information is read from the ‘AppData\local\Microsoft\SkyDrive\setting\[OwnerID].dat’ file. Information was recovered from the Internet Explorer Access, Download, Eraser and CCleaner-VMs, and also the Safari Access-VM. Information was also recovered from unallocated space, pagefile.sys and fileslack. No additional information was located in the Google Chrome or Firefox VM’s (apart from when the client software was used, as previously mentioned). Running IEF 5.52 over the captured RAM files resulted in files and folder names relating to SkyDrive when the client software was used for all browsers (Upload-VM’s), and the Internet Explorer Access-VM and Download-VM’s, highlighting the potential information available from RAM captures. No information was reported when IEF was run over the Wireshark network PCAP capture files.
6.2.5.13 **Metadata**

Using X-Ways 16.5, it is possible to extract metadata from files, such as the EXIF data from JPG files, or information from documents. This can then be used to conduct analysis. It was possible to determine the picture files which related to the Enron test data by searching on the metadata contents. It was also possible to locate the rtf document files based on the Author information stored within the file.

6.2.5.14 **Network Analysis**

Analysis of the network traffic capture files was undertaken using Network Miner v1.0, Wireshark Portable 1.6.5 and X-Ways 16.5. The network traffic used Port 80 (HTTP) and 443 (HTTPS) only. When accessing SkyDrive accounts using the client software or a browser it appears login sessions are first established with Microsoft Live using IP numbers registered to Microsoft, such as in the range 64.4.0.0 to 64.4.63.255 and 65.52.0.0 to 65.55.255.255 (login.live.com, skydrive.live.com). When accessing files or data in an account, IP’s in the range 70.37.0.0 to 70.37.181.255 were observed. See Table 25 for a list of IP number ranges observed in the captured network traffic, and the registered owner for the range.

<table>
<thead>
<tr>
<th>Registered Owner</th>
<th>IP Start</th>
<th>IP Finish</th>
<th>URL’s observed in network traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS Hotmail</td>
<td>64.4.0.0</td>
<td>64.4.63.255</td>
<td>skydrive.live.com, msnmessenger</td>
</tr>
<tr>
<td>Microsoft</td>
<td>65.52.0.0</td>
<td>65.55.255.255</td>
<td><a href="http://www.skydrive.com">www.skydrive.com</a>, login.live.com, skdrivesync.policies.live.net</td>
</tr>
<tr>
<td>Microsoft global online</td>
<td>70.37.0.0</td>
<td>70.37.191.255</td>
<td>s1-word-view, S1-word-edit, S1-excel.vo, S1-powerpoint.vo, S1-onenote.vo</td>
</tr>
<tr>
<td>services</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microsoft Singapore</td>
<td>111.221.16.0</td>
<td>111.221.31.255</td>
<td>word-view.officeapps.live.com</td>
</tr>
<tr>
<td>Microsoft</td>
<td>131.253.12.0</td>
<td>131.253.18.255</td>
<td>us.c01.cb3.glbdns.microsoft.com</td>
</tr>
<tr>
<td>Microsoft-global-net</td>
<td>207.46.0.0</td>
<td>207.46.255.255</td>
<td>sn2.storage.msn.com</td>
</tr>
<tr>
<td>Microsoft internet data center</td>
<td>213.199.160.0</td>
<td>213.199.191.255</td>
<td>ncsi.glbdns.microsoft.com</td>
</tr>
<tr>
<td>Akamai Technologies</td>
<td>118.214.0.0</td>
<td>118.215.255.255</td>
<td>e2044.c.akamaiedge.net</td>
</tr>
<tr>
<td>Akamai Technologies</td>
<td>122.252.32.0</td>
<td>122.252.63.255</td>
<td>secure.shared.live.com.edgekey.net</td>
</tr>
<tr>
<td>Akamai Technologies</td>
<td>173.222.0.0</td>
<td>173.223.255.255</td>
<td>e4344.g.akamaiedge.net</td>
</tr>
<tr>
<td>Verisign Global Registry</td>
<td>199.7.48.0</td>
<td>199.7.63.255</td>
<td>ocsp.verisign.net</td>
</tr>
</tbody>
</table>

Table 25 - SkyDrive IP addresses observed in network traffic
Security certificates appeared to be approved using VeriSign/Thawte services, in the 199.7.48.0 to 199.7.63.255 range. The URL; ocsp.verisign.net was also observed in the network traffic, OCSP is the Online Certificate Status Protocol, used by VeriSign/Thawte.

When the SkyDrive Windows Client software was downloaded, the URL name made reference to Microsoft Live Service, and the following path was observed in the network traffic;

```
```

The same URL names were observed when using either the SkyDrive client software or accessing via a browser. Previous analysis of Dropbox network traffic determined Dropbox uses different URL’s when using a browser compared to when using the client software.

Network Miner 1.0 was used to rebuild files from the captured packets. These included SkyDrive certificate files, such as ‘skydrive.live.com.cer’. The contents of the Enron sample data files were not seen in the network traffic, suggesting the data was encrypted. Analysis of the files downloaded from SkyDrive revealed the download was contained in an uncompressed ZIP, so if this was not encrypted it would be possible to recover the contents of files from the network traffic. Of note was the presence of the filenames and some information in cleartext (not encrypted) for the Enron files. Information extracted from the Firefox Upload-VM is included in Table 26. The filename (enron3111.docx) and SkyDrive OwnerID are highlighted. The details of the file are easily discerned from the text, including the file size, created and modified times (in UTC), and the parent resource (folder) information is also listed. These highlight that relevant information may be present in network traffic, even if the traffic is expected to be encrypted and unreadable, and should be a major consideration when determining whether to capture this information if possible.
Table 26 - Unencrypted data observed in network traffic

No username or password information was observed in the cleartext network traffic. The term ‘skydrive’ was observed in all network traffic, including the control Base-VM’s. However, the website URL’s ‘www.skydrive.com’ and ‘https://skydrive.live.com’ were not observed in the Base-VM network traffic. The website URL’s were observed in all Access-VM’s and Download-VM’s, and also in the Google Chrome Upload-VM and Eraser-VM network traffic. SkyDrive icon pictures were recovered from Upload-VM’s, associated with the setup.exe files.

6.2.5.15 System Volume Information

When either Eraser or CCleaner was used, information was still located in various System Volume Information files (also known as Restore Points or Volume Shadow Copies). The information included the username for the SkyDrive account, the URL’s for the skydrive websites (https://skydrive.live.com or www.skydrive.com), the filenames and full text of the Enron files was located, and the SkyDrive OwnerID.
When limited information was located in during analysis, such as only filenames, the System Volume Information files contained a wider range of information, such as file contents, or deleted URL’s. System Volume Information was highlighted as an important source of information when anti-forensic methods had been used, in this research; Eraser and CCleaner.

6.2.5.16 Memory (RAM) Analysis

Analysis of the memory captures (VMEM files) was undertaken using X-Ways 16.5, Encase 6.19.4, and AccessData FTK 4.01. The term ‘skydrive’ was located in all the memory captures, including the control Base-VM memory files. However, the URL’s (‘www.skydrive.com’ and ‘https://skydrive.live.com’) were not located in the control Base-VM memory files, but were located in all other memory capture files. The SkyDrive account username was located in all Upload-VMs, and in the Firefox, Google Chrome, and Internet Explorer Access and Download-VM’s. The SkyDrive username was located in memory capture files near the text; ‘skydrive’ or ‘live.com’ in the format of; ‘&login=[username] &passwd=[password]’. The text ‘&login’ and ‘&passwd’ can be used to search and locate potential usernames for SkyDrive accounts (also Hotmail and Windows Live accounts). The password was also located in freetext in the Upload-VM memory files for Firefox, Internet Explorer and Safari browsers.

The username was also located in the Eraser-VM memory captures, in the Firefox and Google Chrome CCleaner-VM memory captures, and also in the Google Chrome memory capture file when CCleaner was used with all options. However when the Safari browser was used, the username was only recovered in the Upload-VM memory capture, near the text; ‘&login’. Also located in the memory files was the SkyDrive Owner ID, near the text ‘?cid=’ or ‘?id=’, such as; https://skydrive.live.com/?cid=FD79E4BB87F9XXXX&id=FD79E4BB87F9XXXX.

The full text of the Enron data files was recovered in the Upload-VM, Access-VM and Download-VM memory captures. In addition, the text was located in the Google Chrome and Internet Explorer Erase-VM memory files. No text was located in the control Base-VM memory files. Observed within memory captures were Enron filename references in all VM memory files, except the control Base-VM memory files.
The previously mentioned term, ‘enabledogfood’ was located in the Upload VM memory files, associated with the website. ‘http://g.live.com/8SESkyDrive/EnableDogFood’.

Data carving was undertaken across the memory capture files and resulted in the recovery of thumbnail pictures, partial and full picture files of the Enron sample pictures and SkyDrive logo’s from the memory captures for all VM memory files except the Base-VM memory files.

Analysis of Pagefile.sys files was also undertaken. The username was only located in the Safari Access-VM pagefile.sys file. The website URL’s (‘www.skydrive.com’ and ‘https://skydrive.live.com’) were located in the Google Chrome, Internet Explorer and Safari browser pagefile.sys files, but not the Firefox browser pagefile.sys. The filenames for the Enron test files were located in browser Upload, Access and Download-VM pagefile.sys files when Firefox, Chrome and Internet Explorer were used, but not when Safari was used to download files. Filenames were also located after Eraser and CCleaner were used, again highlighting pagefile.sys as an area for examination if anti-forensic methods are suspected of being used.

The full text of the Enron email was found in the pagefile.sys file when Firefox and Google Chrome browsers were used, but not when Internet Explorer or Safari were used. The SkyDrive OwnerID was located when Firefox, Internet Explorer and Safari were used, but not when Google Chrome was used.

6.2.5.17 Eraser, CCleaner and DBAN

As discussed, many remnants indicating the use of SkyDrive and the presence of the Enron sample files were located after Eraser and CCleaner were used. This indicated that Eraser or CCleaner did not necessarily remove all data remnants in relation to the use of SkyDrive or completely remove the information relating to the previous presence of the Enron sample files. The use of a full erasing tool, in this case Darik’s Boot and Nuke (DBAN), was found to remove all traces, but this also erased the operating system and user files. The operating system and user files would need to be re-installed, which could be a lengthy process and may dissuade the average user from undertaking this process to remove evidence of SkyDrive usage.
6.2.6 Presentation

6.2.6.1 Analysis findings

In this case study, a variety of data remnants were located when a user used SkyDrive to store or access data. The focus (or scope) was to determine what data remnants were left to identify if SkyDrive had been used on a Windows 7 PC. Conducting a search for the term ‘skydrive’ was shown to be inconclusive to demonstrate use, as it was shown that the term is present even when SkyDrive has not been accessed, as outlined in the analysis of the control Base-VM’s. The SkyDrive username was able to be discerned from a variety of locations, such as cookie files, memory captures (searching for ‘slogin=’), in pagefile.sys, from SQLite database files, and IEF output.

Surprisingly, the password for the SkyDrive account was able to be located unencrypted in the memory captures from the Upload-VM’s. Passwords are commonly used in a variety of locations, and when conducting analysis to determine a password to view encrypted files, one practice is to build an index of data to use with password analysis tools, such as AccessData Password Recovery Toolkit, Passware, or Elcomsoft password analysis tools. With the actual password in cleartext this would vastly speed up the process of password discovery. Hence, memory capture and analysis should be a key consideration when determining the possible sources of data in step 3 of the framework (Identify and Collect) for indexing purposes. However, if the password is not located in memory or if memory capture is not possible, it is still possible to access a SkyDrive account using the forensic image of a hard drive where the SkyDrive client software has been used to access an account. This can be achieved by running the hard drive forensic image as a Virtual Machine. When the VM starts up it will automatically synchronise with the SkyDrive cloud service and update files in the local SkyDrive folder. Appropriate legal authority would be required to ensure the account and information can be accessed and used for analysis.

Analysis to determine the method of access; whether the client software was used, a browser used to upload, access or download, or a combination of both, is possible. When the client software was downloaded, there was a SkyDriveSetup.exe file downloaded to the local hard drive. Hash values for the client software can be calculated and searched across forensic image files, memory and network captures. In addition, registry entries may indicate the use of the setup software or the client software, as will link files. Prefetch files were observed for the client software, and also
include the number of times run, and associated dates and times. The file ‘CollectLogFiles.bat’ appears to have a consistent hash value across SkyDrive client software releases and the hash value for this file can be used to conduct searches. Also, searching for the term ‘enabledogfood’ may determine if the client software has been installed on a PC. A folder was observed with the client software version as the folder name, in the AppData\local\ folder.

When a browser was used to access SkyDrive there were many references in the output from Internet Evidence Finder (IEF), but when access was undertaken with no downloading there were no references to the filenames in the IEF output. When bulk files were downloaded using the web account from the root folder, an uncompressed ZIP file was observed with the name ‘skydrive’ and the date.

Filenames, dates, and times were observed in the ‘SyncDiagnostics.log’ file and the ‘.dat’ file with the OwnerID as the filename. Filenames were also observed in the Registry ‘RecentDocs’ keys, $MFT$ entries, and Prefetch files such as DLLHost.pf, Wordpad.pf and Notepad.pf. Link files and the IEF output also listed the filenames. The contents of the files were also recovered from temporary internet files, thumbcache, memory captures, pagefile.sys, system volume restore points, and unallocated space. Eraser and CCleaner were not effective in removing all data remnants, and information was able to be determined from these VM’s relating to the SkyDrive accounts, filenames, dates and times and file contents.

In summary, it has been demonstrated that it is possible to determine the username, password, method of access, filenames, contents, and dates and times of access when SkyDrive is used to store, access or download data from cloud storage. Memory captures and Network files were an important source of information, especially locating the username and password in memory files when the client software was setup and used. Table 27 summarises the information located.

Once forensic analysis has determined a SkyDrive account has potential evidence of relevance to an investigation, the examiner can communicate this to relevant persons to enable them to respond to secure evidence in a timely manner. Knowledge of the username and OwnerID details would enable identification of SkyDrive accounts, to preserve and secure potential evidence. It was observed that SkyDrive retains a record of computers used to access and synchronise with an account. This information would
be beneficial to determine whether a particular computer was synchronised to a SkyDrive account.

In addition, a (spreadsheet) timeline of events can be created using the identified data and information from the various sources, which can assist to explain the course of events. Merging the information from the hard drive, network captures, memory and the SkyDrive account into one timeline may also be of assistance during the analysis process to gain an understanding of the sequence of events. This will be demonstrated in the Case Study section of this Chapter.

<table>
<thead>
<tr>
<th>Control (Base-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username, Password, Software, URL, Enron sample filenames or files KWS terms</td>
<td>Nil</td>
</tr>
<tr>
<td>Nil</td>
<td></td>
</tr>
<tr>
<td>Nil</td>
<td></td>
</tr>
<tr>
<td>Nil</td>
<td></td>
</tr>
<tr>
<td>Matches to ‘skydrive’ relating to icon files (Internet Explorer Base-VM)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Client Software (Upload-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username OwnerID Password Software URL Enron sample filenames Enron sample files KWS terms</td>
<td>Cookie files ‘<a href="mailto:username@login.live">username@login.live</a>’, $MFT and $LogFile. Memory capture files &amp;login=’ OwnerID found in SkyDrive software installation .dat and .ini files Located in RAM – search for &amp;login= and &amp;passwd= SkyDriveSetup.exe file located when downloaded. SkyDrive Software installation under User\AppData\Local\Microsoft SyncDiagnoses.log file includes OwnerID and File information, dates and times When software downloaded, URL’s included <a href="https://skydrive.live.com">https://skydrive.live.com</a> and <a href="http://www.skydrive.com">http://www.skydrive.com</a> Multiple locations, including prefetch, Link files, $MFT, Registry. Filenames in Network PCAP files Located in Sync folder under User\SkyDrive ‘EnableDogFood’ found in SkyDrive client software files</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Browser Access (Access-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username OwnerID Password Software URL Enron sample filenames Enron sample files KWS terms</td>
<td>FF and GC History; ‘formhistory.sqlite’ and ‘Web Data’. Memory capture files &amp;login=’ Observed in URL; <a href="https://skydrive.live.com/?cid=XXXXXXXXXXXXXXXXX">https://skydrive.live.com/?cid=XXXXXXXXXXXXXXXXX</a> (FF and GC) +RAM Nil Nil Multiple locations; cookie, history, icons, pagefile.sys and unallocated Sufficient to identify files accessed with references to the filenames in Registry and Browsing History Full text in RAM and System Volume Information Multiple matches to KWS terms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Browser Download (Download-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username OwnerID Password Software URL Enron sample filenames Enron sample files KWS terms</td>
<td>FF and GC History; ‘formhistory.sqlite’ and ‘Web Data’. Memory capture files &amp;login=’ Observed in URL; <a href="https://skydrive.live.com/?cid=XXXXXXXXXXXXXXXXX">https://skydrive.live.com/?cid=XXXXXXXXXXXXXXXXX</a> (FF and GC) +RAM Nil Nil Multiple locations; cookie, history, icons, pagefile.sys and unallocated Sufficient to identify files accessed with references in $MFT, Link, Registry and Prefetch files Via uncompressed zip; ‘skydrive-YYYY-MM-DD.zip’ or folder name; ‘Documents.zip’ Full text in RAM and System Volume Information</td>
</tr>
</tbody>
</table>
### Eraser (Eraser-VM)

<table>
<thead>
<tr>
<th>Username</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF and GC History; 'formhistory.sqlite' and 'Web Data' Memory capture files' &amp; login=</td>
<td>Observed in URL; <a href="https://skydrive.live.com/?cid=XXXXXXXXXXXXXXXXXX">https://skydrive.live.com/?cid=XXXXXXXXXXXXXXXXXX</a> (FF and GC) + RAM</td>
</tr>
<tr>
<td>Nil</td>
<td>Multiple locations; cookie, history, icons, pagefile.sys and unallocated</td>
</tr>
<tr>
<td>Nil</td>
<td>Sufficient to identify files accessed with references to the filenames in $MFT, Link and Prefetch files</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Full text in RAM and System Volume Information</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>Multiple matches to KWS terms</td>
</tr>
<tr>
<td>KWS terms</td>
<td></td>
</tr>
</tbody>
</table>

### CCleaner (CCleaner-VM)

<table>
<thead>
<tr>
<th>Username</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF and GC History; 'formhistory.sqlite' and 'Web Data' Memory capture files' &amp; login=</td>
<td>Observed in URL; <a href="https://skydrive.live.com/?cid=XXXXXXXXXXXXXXXXXX">https://skydrive.live.com/?cid=XXXXXXXXXXXXXXXXXX</a> (FF and GC) + RAM</td>
</tr>
<tr>
<td>Nil</td>
<td>Multiple locations; cookie, history, icons, pagefile.sys and unallocated</td>
</tr>
<tr>
<td>Nil</td>
<td>Sufficient to identify files accessed with references to the filenames in $MFT, Link and Prefetch files</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Full text in RAM and System Volume Information</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>Multiple matches to KWS terms</td>
</tr>
<tr>
<td>KWS terms</td>
<td></td>
</tr>
</tbody>
</table>

### DBAN (DBAN-VM)

<table>
<thead>
<tr>
<th>Username</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Data erased, no information located</td>
<td></td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td></td>
</tr>
<tr>
<td>Enron sample files</td>
<td></td>
</tr>
<tr>
<td>KWS terms</td>
<td></td>
</tr>
</tbody>
</table>

Table 27 - Summary of SkyDrive analysis findings

### 6.2.7 Feedback

During this research, it confirmed that conducting research using virtual computers was beneficial to compare a range of different configurations of software and data access. The overall findings are summarised in the final section, and highlight the information that should be promulgated to the wider digital forensic community.

### 6.2.8 Complete

The final step is the completion of an investigation. For the purposes of this research, it was determined that analysis of another popular cloud storage services; Google Drive, will serve to widen the scope of the research and further assess the proposed framework. The files created for this research were archived, and working copies were available to compare when undertaking the next stages of the research.
6.3 SkyDrive Analysis: Apple iPhone 3G

The use of portable devices is rapidly increasing. Gartner have highlighted the trend in client computing is shifting from a focus on personal computers to include portable devices, and that personal cloud storage will replace the PC as the main method of users’ storage by 2014 (Kleynhans 2012). The Apple iPhone is one such (popular) device that can be used to access SkyDrive cloud storage either using the built in Safari browser, or installing the Microsoft SkyDrive application (App). Analysis of portable devices is a growing area for forensic practitioners, and there are both hardware and software solutions to assist this process, such as; MicroSystemation .XRY, Cellebrite UFED, and Radio Tactics Aceso. In the scope of this research, it is relevant to consider what information can be determined from a portable device in relation to the use of SkyDrive. This also serves to further assess the suitability of the framework being applied in a variety of circumstances.

6.3.1 Commence (Scope)

The scope of this part of the research is to examine the data remnants on an iPhone 3G running iOS 4.2.1 when used to access SkyDrive via a browser, and when using the SkyDrive iOS application (app). Analysis also needs to determine if there are any remnants prior to accessing SkyDrive on a standard device.

6.3.2 Prepare

An iPhone 3G was selected which had not been used to access SkyDrive previously. The device history and usage was known, and MicroSystemation .XRY 6.2.1 was used to extract a logical image of the contents, excluding audio and video files, prior to accessing SkyDrive or installing the SkyDrive Application. This first extract was analysed to confirm there was no SkyDrive related data on the device prior to undertaking the research. The inbuilt Apple Safari iOS browser was used to access the SkyDrive user account created for this research and view the Enron files stored remotely. A logical extract was then conducted using .XRY. The Microsoft SkyDrive iOS app was then downloaded and installed to the iPhone 3G. The research account was
then accessed using the Application, and the files stored in the SkyDrive account were viewed. A third logical extract with .XRY was then conducted.

6.3.3 Identify and Collect

In the context of this research, files were identified which would contain the information needed to conduct the analysis, in this instance being the .XRY extract files and the output of the .XRY software, including PDF reports and the files exported using .XRY. These were identified for each of the extracts; Base, Browser, and Application.

6.3.4 Preserve (Forensic Copy)

To observe the principles of digital forensic analysis (ACPO 2006; NIJ 2008) a forensic copy was made of the .XRY extract files, the file output, and the reports. As these were logical files, this was done in the Encase Logical format (L01), and the X-Ways Evidence File Container format (ctr). MD5 hash values were used to ensure the forensic integrity of the data.

6.3.5 Analysis

For this research, each of the forensic logical files was examined using forensic tools including; X-Ways Forensic 16.5 and EnCase 6.19.4. PList Explorer v1.0 was used to examine the contents of the Apple plist files extracted by .XRY.

6.3.5.1 Control - Base-XRY

Analysis of the control Base-XRY extract confirmed there was no data originally present relating to the Enron sample test data or the SkyDrive app. In addition, no references were located for the term ‘skydrive’ or the website URL’s (skydrive.live or skydrive.com).
6.3.5.2 SkyDrive accessed via the iOS Safari Browser

The username was not located in the Browser-XRY extracts; however the Client ID number was located in a URL for skydrive.live.com (https://skydrive.live.com) in the History.plist file. This information was also extracted by .XRY in the Web-History.txt file (output listed in Table 28). PList Explorer 1.0 was also used to view the data in the plist files to verify the information in the XRY report. The SkyDrive OwnerID was also located in Cookies.binarycookies file.

<table>
<thead>
<tr>
<th>Web-History #</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application:</td>
<td>Safari (Apple)</td>
</tr>
<tr>
<td>Web Address:</td>
<td><a href="https://skydrive.live.com/browse.aspx?cid=XXXXA4AB89F707D0">https://skydrive.live.com/browse.aspx?cid=XXXXA4AB89F707D0</a></td>
</tr>
<tr>
<td>Page Title:</td>
<td>Dataset - SkyDrive</td>
</tr>
<tr>
<td>Access Count:</td>
<td>1</td>
</tr>
<tr>
<td>Accessed:</td>
<td>4/07/2012 3:24:05 AM UTC (Device)</td>
</tr>
</tbody>
</table>

Table 28 - Output from .XRY Web-History.txt file #6

Filenames for the Enron test files were located in the History.plist, for the txt, rtf and docx files. This information was also extracted by XRY in the Web-History.txt file. A full picture file was located in /private/var/mobile/Media/DCIM/102APPLE/ within the XRY extract (Table 29).

<table>
<thead>
<tr>
<th>Web-History #</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application:</td>
<td>Safari (Apple)</td>
</tr>
<tr>
<td>Web Address:</td>
<td><a href="https://goq9bw.sn2.livefilestore.com/y1pnjQfg9cr138aplsHG0P8vBe4CqfSiUO-2INcwF2eGQ8oVOO9LCPWbf5CeeHl4aGyyhoaxPRmYG2OXTMzPVowvoGr5G0/Enron3111.rtf?download&amp;psid=1">https://goq9bw.sn2.livefilestore.com/y1pnjQfg9cr138aplsHG0P8vBe4CqfSiUO-2INcwF2eGQ8oVOO9LCPWbf5CeeHl4aGyyhoaxPRmYG2OXTMzPVowvoGr5G0/Enron3111.rtf?download&amp;psid=1</a></td>
</tr>
<tr>
<td>Access Count:</td>
<td>1</td>
</tr>
<tr>
<td>Accessed:</td>
<td>4/07/2012 3:25:25 AM UTC (Device)</td>
</tr>
</tbody>
</table>

Table 29 - Output from .XRY Web-History.txt file #5

6.3.5.3 SkyDrive Application used to access the research account

Analysis of the third XRY extract was able to determine the username used to access SkyDrive account. This was located in the ‘keychain-backup.plist’ near the text; useridentity, windows live, and ‘.com.microsoft.skydrive’. This information
appears to be parsed by XRY in the file ‘Device-Accounts.txt’. However, the information is listed by .XRY as a Windows Live ID without mention of SkyDrive for the entry. To determine if the client application is present, information in the ‘com.apple.AppStore.plist’ file lists ‘http://itunes.apple.com/au/app/skydrive’. Information is also stored in the ‘IconState.plist’ file as ‘com.microsoft.skydrive’. The text from the Enron files was not located in any of the extracts. No password was located in any of the extracts or files.

6.3.6 Presentation

6.3.6.1 Analysis findings

In this research, several data remnants were located when searching for evidence of SkyDrive use on the test item (iPhone 3G). Recall that the aim was to determine what data remnants were left to identify if SkyDrive was accessed, and whether this was via the browser or an installed application. Initial analysis of the control extract identified no matches to the term ‘skydrive’, but matches in the subsequent extracts indicating that conducting a keyword search for the word ‘skydrive’ could indicate if skydrive has been used.

When the browser was used to access the skydrive account, entries were left in the History.plist file, which is also listed in the .XRY Web-History.txt file. The SkyDrive OwnerID was listed in the Cookies.binarycookies file, but the username was not located in the extract. The filenames accessed were listed in the History.plist file and in the .XRY Web-History.txt file. A full picture file was located within the .XRY extract.

When the SkyDrive application was installed and used to access the SkyDrive research account, the username was able to be determined from the keychain-backup.plist file and in the XRY ‘Device-Accounts.txt’ file. Information in the ‘com.apple.AppStore.plist’ file lists ‘http://itunes.apple.com/au/app/skydrive’, and in the ‘IconState.plist’ file was ‘com.microsoft.skydrive’. However, the text from the Enron files was not located in any of the extracts, nor was it possible to locate the password for the SkyDrive account.
This research was limited by not being able to jailbreak the iPhone. XRY, and other iPhone extract solutions, require the iPhone to be jailbroken to undertake a Physical extract. A future research opportunity is to undertake a physical extract of an iPhone and compare this to a logical extract to determine what information is available in comparison with the logical extract, and also to compare other iPhone forensic tools.

6.3.7 Feedback

Again, this research was limited by not being able to jailbreak the iPhone. The comparison of a physical extract of an iPhone to a logical extract to determine what information is available in comparison with the logical extract is a potential research opportunity. Also a comparison of the process with other iPhone forensic software and hardware processes would be beneficial. Another future research opportunity is to undertake similar research using an iOS 5 device, and also widen the devices examined to include other popular mobile device operating systems, such as Google Android and Microsoft Windows.

6.3.8 Complete

In the context of this research, the data created and used may be required in future research opportunities, and hence has been stored on multiple hard drives to enable future use.
6.4 SkyDrive Analysis: Case Study

To illustrate the relevance of the research, the following case study outlines where the information previously identified can assist in an investigation. The circumstances relate to a hypothetical case of an employee suspected of copying intellectual property from Acme Corp (a fictional organisation), the type of situation a business could experience. The circumstances of the case study are greatly simplified for the purposes of this case study;

*IT Security in Acme Corp observed unapproved IP addresses, including 70.37.10.100 and 70.37.179.232, in network logs originating from the Marketing Department. A ‘whois’ lookup determined these IP addresses to be associated with Microsoft Global Online Services, and are suspected of being associated with cloud storage. This network traffic was flagged as unusual, as the Marketing Department do not have a need to use online cloud services, nor are they permitted to in the user agreement. Discreet enquiries were made with the manager of the Marketing Department, who confirmed that no-one had asked for permission to use cloud storage. It was determined that an investigation was necessary to confirm whether company data had been released without authority.*

6.4.1 Commence (Scope)

Authority and approval to undertake an investigation was received from the chief information officer (CIO), with the scope and limitation being to determine and identify any unauthorised use of cloud storage within the Marketing Department, and any workstations and persons involved.

6.4.2 Prepare and Respond

Analysis of the network log files identified the computer associated with the IP addresses identified. Encase Enterprise had been previously deployed throughout the organisation as a security measure, and was used to conduct analysis of the computer in the marketing department.
6.4.3 Identify


6.4.4 Preserve (Forensic Copy)

Memory and network traffic for the identified computer was forensically copied using Encase and Wireshark. This data is the most volatile and should be preserved as soon as reasonably possible (ACPO 2006, NIJ 2008). IT Security liaised with the senior HR Manager to attend the Marketing Department to identify who was using the computer. The workstation was located and was unattended. The power plug was disconnected from the back of the PC (ACPO 2006, NIJ 2008), with the understanding that volatile data had already been preserved. The PC was immediately sealed in an exhibit bag and signed by IT security with HR witnessing this.

A forensic expert was contracted to undertake analysis to determine if any breaches of the company policy had occurred. The forensic expert reviewed the process undertaken thus far to ensure the process was legally and forensically sound. The expert then commenced Exhibit logs and chain of custody sheets to document the process. A forensic copy of the PC hard drive was then made using a Tableau T35es-R2 eSATA Forensic Bridge (hardware write blocking), and verified with MD5 and SHA1 hashes. Appropriate entries were made in the exhibit log relating to access to the PC and hard drive by the expert. Forensic copies were also made of the memory capture and network traffic in Encase Logical format (L01) and verified with hash values.

6.4.5 Analysis

Analysis was then undertaken on the forensic copies of the data. A username and password were located in the memory capture, after searching for the terms ‘$login=’ and the term ‘$passwd=’. SkyDrive client software was also located under ‘C:\Users\SmithT\AppData\Local\Microsoft\SkyDrive\’. Filenames of files associated with the SkyDrive account were located in link files and prefetch files such as dllhost.pf. The contents of the files was able to be discerned from data carving of unallocated space, and the
metadata within the files further identified the computer and user account used to create the documents and pictures. Dates and times were able to be discerned from the SkyDrive syncediagnostics.log file and the $MFT to indicate when the SkyDrive software had been installed, and when the files had been created and accessed.

A company-issued iPhone was located on the desk near the workstation, and placed in a faraday bag\textsuperscript{6} to prevent network interaction. It was conveyed to the forensic expert who undertook a logical extract using Microsystemation .XRY. A physical extract was considered but this requires the phone to be jailbroken, which is in breach of company policy and hence was not undertaken. Analysis of the logical extract located the SkyDrive application and the username in the keychain-backup.plist file and in the .XRY ‘Device-Accounts.txt’ file. A picture file was located in /private/var/mobile/Media/DCIM/102APPLE/ within the .XRY extract. This picture was a jpg file which appeared to be a screenshot of a confidential company document.

The employee was interviewed by HR and when shown the information located so far, agreed to cooperate. The employee allowed the forensic expert to access the SkyDrive account, providing the password and written permission to access the account for the purposes of the investigation. The expert created a virtual machine and recorded the VMWare Player window as a video file. The expert logged in to the SkyDrive account and examined the account information. The computers linked to the account were listed on the browser screen and included the company computer as being linked to the account. This however did not rule out other persons accessing the account in addition to the linked computers.

\textbf{6.4.6 Presentation}

A report was prepared for the HR director and the CIO outlining the process and the findings of the analysis. A meeting was held with the forensic expert where the findings were outlined in a manner that both the CIO and the HR Director were able to understand. A timeline of the data from the various sources was used to demonstrate

\textsuperscript{6} A ‘Faraday bag’ is intended to shield a mobile phone or similar small device to prevent unwanted applications being invoked remotely, such as wiping the memory, or to prevent possible problems with veracity of evidence’ Duffy, A (2010), Faraday Bag Testing in a Reverberation Chamber – a Preliminary Study, <http://www.faradaybag.com/faraday-bag-testing/demontfort-university-faraday-bag-testing.html>.
interaction with the files from the tower PC and the iPhone (Table 30). In the timeline it is possible to see; when the documents were initially created via the metadata, when the documents were copied to the work PC, link files related to the documents are created, SkyDrive is accessed, SkyDrive software is installed onto the PC, and files read on the PC and the iPhone.

<table>
<thead>
<tr>
<th>Source</th>
<th>Key Date</th>
<th>Event type</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEF RAM</td>
<td>14/06/2011 14:22:00</td>
<td>Last Modified</td>
<td>enron_with_categories.tar.gz</td>
</tr>
<tr>
<td>IEF RAM</td>
<td>9/02/2012 15:23:00</td>
<td>Last Modified</td>
<td>Enron3111.txt</td>
</tr>
<tr>
<td>X-Ways Docx Metadata</td>
<td>9/02/2012 15:24:00</td>
<td>Last Modified</td>
<td>Enron3111.docx</td>
</tr>
<tr>
<td>IEF RAM</td>
<td>25/02/2012 11:02:00</td>
<td>Last Modified</td>
<td>Enron3111.rtf</td>
</tr>
<tr>
<td>X-Ways Docx Metadata</td>
<td>25/02/2012 12:02:00</td>
<td>Doc created metadata</td>
<td>Enron3111.docx</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Event Log Explorer</th>
<th>9/06/2012 12:37:09</th>
<th>Logon</th>
<th>cloud user</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEF Firefox</td>
<td>9/06/2012 12:34:00</td>
<td>Date Visited</td>
<td><a href="http://www.mozilla.org/en-US/firefox/13.0/firstrun/">http://www.mozilla.org/en-US/firefox/13.0/firstrun/</a></td>
</tr>
<tr>
<td>X-Ways Docx Metadata</td>
<td>9/06/2012 12:41:49</td>
<td>Created</td>
<td>3111.txt</td>
</tr>
<tr>
<td>X-Ways Docx Metadata</td>
<td>9/06/2012 12:41:49</td>
<td>Created</td>
<td>enron_with_categories.tar.gz</td>
</tr>
<tr>
<td>X-Ways Docx Metadata</td>
<td>9/06/2012 12:41:49</td>
<td>Created</td>
<td>Enron3111.docx</td>
</tr>
<tr>
<td>X-Ways Docx Metadata</td>
<td>9/06/2012 12:41:49</td>
<td>Created</td>
<td>Enron3111.rtf</td>
</tr>
<tr>
<td>X-Ways Link Data</td>
<td>9/06/2012 12:41:49</td>
<td>Target Created</td>
<td>3111.txt</td>
</tr>
<tr>
<td>X-Ways Link Data</td>
<td>9/06/2012 12:41:49</td>
<td>Target Created</td>
<td>Enron3111.txt</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:42:03</td>
<td>Created</td>
<td>NOTEPAD.EXE-EB18E61A.pdf</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:42:08</td>
<td>Created</td>
<td>Enron3111.txt</td>
</tr>
<tr>
<td>IEF Firefox</td>
<td>9/06/2012 12:45:00</td>
<td>Date Visited</td>
<td><a href="http://www.google.com/search?q=skydrive%20preview%20for%20windows&amp;ie=utf-8&amp;oe=utf-8&amp;aq=t&amp;rls=org.mozilla:en-US:official&amp;client=firefox-a&amp;source=hp&amp;channel=np">http://www.google.com/search?q=skydrive%20preview%20for%20windows&amp;ie=utf-8&amp;oe=utf-8&amp;aq=t&amp;rls=org.mozilla:en-US:official&amp;client=firefox-a&amp;source=hp&amp;channel=np</a></td>
</tr>
<tr>
<td>IEF Firefox</td>
<td>9/06/2012 12:45:00</td>
<td>Date Visited</td>
<td><a href="https://apps.live.com/skydrive/go/ba6e47d1-606a-4816-a246-90f548f7a3b9">https://apps.live.com/skydrive/go/ba6e47d1-606a-4816-a246-90f548f7a3b9</a></td>
</tr>
<tr>
<td>Encase Internet</td>
<td>9/06/2012 12:45:00</td>
<td>Date Visited</td>
<td><a href="http://g.live.com/8SESkydrive/WinDownload">http://g.live.com/8SESkydrive/WinDownload</a></td>
</tr>
<tr>
<td>Encase Internet</td>
<td>9/06/2012 12:46:00</td>
<td>Start Date</td>
<td>SkyDriveSetup.exe</td>
</tr>
<tr>
<td>IEF Firefox</td>
<td>9/06/2012 12:46:00</td>
<td>Date Visited</td>
<td><a href="https://live.com/?rewlive=skydrive/skydrive_setup">https://live.com/?rewlive=skydrive/skydrive_setup</a></td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:46:20</td>
<td>Created</td>
<td>CollectSkyDriveLogs.bat</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:46:26</td>
<td>Created</td>
<td>SKYDRIVESETUP.EXE-FC6CAE9C.pdf</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:46:34</td>
<td>Created</td>
<td>Microsoft SkyDrive.link</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:46:34</td>
<td>Created</td>
<td>SkyDrive</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:46:34</td>
<td>Created</td>
<td>SkyDrive.exe</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:46:34</td>
<td>Created</td>
<td>SkyDrive.exe</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:46:34</td>
<td>Created</td>
<td>SKYDRIVECONFIG.EXE-1F1120F.pdf</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:46:34</td>
<td>Created</td>
<td>SkyDriveSetup.exe</td>
</tr>
<tr>
<td>X-Ways Link Data</td>
<td>9/06/2012 12:46:34</td>
<td>Target Created</td>
<td>Microsoft SkyDrive.link</td>
</tr>
<tr>
<td>X-Ways Prefetch Data</td>
<td>9/06/2012 12:46:34</td>
<td>Prefetch last run</td>
<td>SKYDRIVE.EXE-76A83CF5.pdf</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:46:44</td>
<td>Created</td>
<td>SKYDRIVE.EXE-76A83CF5.pdf</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:47:01</td>
<td>Created</td>
<td><a href="mailto:cloud@login.live">cloud@login.live</a>[1].txt</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:47:55</td>
<td>Created</td>
<td>fd79e46d7f905000.dat</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:47:55</td>
<td>Created</td>
<td>fd79e46d7f905000.ini</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:48:34</td>
<td>Created</td>
<td>SyncDiagnostics.log</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:50:16</td>
<td>Accessed</td>
<td>3111.txt</td>
</tr>
<tr>
<td>X-Ways Prefetch Data</td>
<td>9/06/2012 12:50:16</td>
<td>Prefetch last run</td>
<td>NOTEPAD.EXE-EB18E61A.pdf</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:50:25</td>
<td>Accessed</td>
<td>Enron3111.txt</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:50:28</td>
<td>Created</td>
<td>Enron3111 (x)</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:50:35</td>
<td>Created</td>
<td>WORDPAD.EXE-1BC83D07.pdf</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:50:36</td>
<td>Created</td>
<td>Enron3111 (x)</td>
</tr>
<tr>
<td>X-Ways Prefetch Data</td>
<td>9/06/2012 12:50:36</td>
<td>Prefetch last run</td>
<td>WORDPAD.EXE-1BC83D07.pdf</td>
</tr>
<tr>
<td>Encase file list</td>
<td>9/06/2012 12:51:30</td>
<td>Accessed</td>
<td>SkyDrive</td>
</tr>
<tr>
<td>Phone XRY extract</td>
<td>4/07/2012 12:54:05</td>
<td>accessed</td>
<td>Dataset - SkyDrive</td>
</tr>
<tr>
<td>Phone XRY extract</td>
<td>4/07/2012 12:55:20</td>
<td>accessed</td>
<td>Enron3111.rf download</td>
</tr>
</tbody>
</table>

Table 30 - Example of Timeline (SkyDrive Case Study)
6.4.7 Feedback

The forensic expert reviewed the entire process and was concerned about the lack of documentation from Acme Corp in relation to the seizure of the PC. The expert contacted the ACME CIO and forwarded a recommendation relating to the use of notes, exhibit logs and chain of custody forms by Acme IT Security in future, in accordance with the NIJ First Responders Guidelines (NIJ 2008) and Standards Australia HB171 recommended practices (SAI 2003). A training package was prepared and forwarded to the company to ensure IT staff could be trained in appropriate forensic response procedures to ensure potential evidence would be secured in accordance with best practice for future incidents. Not adhering to these recommended practices could jeopardise the chance of having information admitted as evidence in a court or tribunal.

6.4.8 Complete

The forensic expert was notified that the matter was finalised and ensured the forensic copies, notes and other material related to the investigation was appropriately stored in case of future unfair dismissal claims or legal action was undertaken. As this could potentially be many years later, the data and information was stored in a manner that could be retrieved in future if required, such as storing the forensic copies on both optical and magnetic media, and relevant reports were also printed out in hardcopy.
6.5 SkyDrive Analysis: Summary

When investigating the storage of data using cloud service providers, the initial stages of an investigation include the identification of a cloud service and user account details. This will enable examiners to identify the potential location of data, and act to secure this data in a timely manner. In the case study, it was determined that an examiner can identify SkyDrive account use by undertaking keyword searches, hash comparison, and examine common file locations to locate relevant information (as detailed in Table 27).

A SkyDrive username was able to be located within the forensic images. Of great interest is that a SkyDrive account password can, in some instances, be located as plain text in memory captures. As outlined, there are a wide range of investigation points for an examiner to determine the use of SkyDrive, such as; directory listings, prefetch files, link files, thumbnails, registry, browser history, and memory captures. It has been identified that conducting research into physical extracts of data from an iPhone would be of relevance, when compared to the information obtained from a logical extract.

The next chapter consists of research into the remnants of the cloud storage service Google Drive. This will assist to determine data remnants from other cloud service providers and encompass a methodology that can be used to identify cloud storage providers and cloud computing providers. The aim is to encompass future developments in the field of remote storage, and developing a consistent digital forensic framework, such as the one proposed in this research.
In the previous two chapters, the forensic analysis of cloud storage using Dropbox and SkyDrive was explored when a user uploads, accesses, or downloads data using a Windows 7 PC or an Apple iPhone 3G. This was achieved using the proposed Digital Forensic Analysis Cycle, and data remnants were located in numerous locations. In this chapter, the focus is on Google Drive. The proposed framework is again used to step through the process in a logical manner.

Material presented in this chapter was submitted for publication and is currently under peer review:


Chapter 7 - Google Drive Analysis

7.1 Introduction

The focus of this chapter is to discover the remnants left on a client device, in this case a Windows 7 computer and an Apple iPhone, after a user accesses Google Drive. In addition, the benefits of using a framework to guide an investigation and follow a
standard process when undertaking forensic analysis of a cloud computing environment is also examined.

In the following discussion, the analysis of Google Drive access using Windows 7 is detailed. Following this, the analysis of Google Drive access using an iPhone 3G running iOS version 4.2.1 is explored. A hypothetical case study is used to demonstrate the relevancy of the research findings. The findings are then summarised, and future research opportunities are outlined.

7.2 Google Drive Analysis: Windows 7 PC

In Chapter Four, the steps of; Commence (Scope), Prepare and Respond, Identify and Collect, Preserve (Forensic Copy), Analyse, Present, Feedback, and Complete were outlined. This is the third experiment to validate the framework to ensure that it is forensically sound and is flexible enough to work with different cloud storage services, as there will most certainly be variation in the way criminal investigation is carried out for each type of cloud service. The framework is used to guide the process of researching the forensic analysis of client devices relating to Google Drive use.

7.2.1 Commence (Scope)

The focus of this research is to determine what data remnants are locatable after a user has up-loaded, accessed, or downloaded data from Google Drive using either the Google Drive client software or a browser. Popular browsers include Microsoft Internet Explorer, Mozilla Firefox, Google Chrome, and Apple Safari (W3Counter 2012). These four browsers will be used in this research to determine any differences in the ability to retrieve data remnants.

The scope of this research is to determine the remnants such as; username, password, filenames, dates and times, or the presence of client software; to indicate which cloud service, if any, had been used on a Windows 7 PC. In addition, circumstances were created to simulate a user running Eraser or CCleaner to remove evidence of using Google Drive (i.e. anti-forensics). Memory capture and analysis is included as this is an
important source of evidence which should be captured and examined when possible (NIJ 2004; ACPO 2006; NIJ 2008). Network data capture is another potential source of information which will form part of the scope to determine what data is present within the network traffic from a host PC.

7.2.2 Prepare

To gather the data required to answer the research questions in relation to the use of Google Drive, a variety of virtual computers (VMs) were created. It was decided to examine a variety of circumstances of a user accessing Google Drive, and also to examine any differences when using different browsers. Multiple scenarios are explored, each making use of Google Drive with the various browsers; Internet Explorer (IE), Mozilla Firefox (FF), Google Chrome (GC), and Apple Safari (AS). Multiple Virtual Machines (VMs) were created for each browser to replicate different circumstance of usage, as outlined in Chapter Three.

Specific software used in this research, in addition to the software outlined in Chapter Three, includes;

- Googledrivesync.exe 1.2.3101.4994 and 1.3.3209.2688
- Mozilla Firefox 13.0.1
- Internet Explorer 9.0.8112.16421IC
- Google Chrome 19.0.1084.56m
- Apple Safari 5.1.7 for Windows 7
- Eraser v6.0.10.2620
- CCleaner v3.19.1721
- Darik’s Boot and Nuke (DBAN) version 2.2.6 Beta
7.2.3 Identify and Collect

In the context of this research, files were identified which would contain the information needed to conduct the analysis; the virtual hard drives (VMDK files) in each VM folder, each memory instance (VMEM files), and each saved network capture file (PCAP). These were identified for each of the VMs.

7.2.4 Preserve (Forensic Copy)

To observe the principles of forensic computer acquisition (NIJ 2004; ACPO 2006; NIJ 2008), a forensic copy was made of each virtual hard drive (VMDK file) using AccessData FTK Imager CLI 2.9.0 in the Encase Evidence format (E01). In regard to the each memory file (VMEM) and network capture (PCAP) file, a forensic copy was made using Encase version 6.19.4, in the Logical Evidence format (L01), and the X-Ways Evidence File Container format (ctr). Hash values (MD5) were used to ensure the forensic integrity of the data.

7.2.5 Analysis

For this research, each of the forensic copies of the virtual computer hard drives, memory and network captures were examined using a range of forensic analysis tools, including; X-Ways Forensic version 16.5, Guidance Software EnCase version 6.19.4, AccessData FTK version 1.81.6 and version 4.01, Network Miner 1.0, Wireshark 1.6.5, Magnet Software (formerly JADSoftware) Internet Evidence Finder 5.52, and RegRipper version 20080909. Many of these tools are widely used for digital forensic analysis by law enforcement agencies and the private sector. In addition, Encase 6.5 and FTK Imager 2.5.3.14 have been tested by the Office of Law Enforcement Standards of the National Institute of Standards and Technology (NIST 2012).

7.2.5.1 Control - Base-VMs

Analysis of the four control Base-VM hard drive images confirmed there was no data originally present relating to the Enron sample test data and Google Drive client software files. References were only located for the website address details ‘drive
google.com in the Google Chrome Base-VM, in the ‘chrome.dll’ file and ‘chrome.7z’. No references were located in the other VM’s.

7.2.5.2 Google Drive Client Software

Analysis of the four Upload-VMs located the ‘googledrivesync.exe’ file that was downloaded from the Google Drive website to the Users ‘Downloads’ folder. When run, the executable file ‘googledrivesync.exe’ was copied to ‘C:\Program Files\Google\Drive\’ folder, and ‘googledrivesync32.dll’ was created. Different hash values were observed for the executable files, the only consistent hash value was ‘1de85d9907bb78b381d1433b8330880d’ for the executable installed to the Program Files folder, created on 13 June 2012 at 4:30 pm (ACST). The installation executable files downloaded from Google at different times had varying hash values, suggesting changes are made regularly, and hence the hash values were different. For an examiner to use hash values to locate client software files, there would be a need to maintain a database of hash values, including historical software hash values for the different client software releases.

Created in the Users folder is a directory as follows; ‘C:\Users\[username]\AppData\Local\Google\Drive\’. Within this folder were two files of interest; ‘sync_config.db’ and ‘snapshot.db’. Both files are SQLite format 3 database files, able to be viewed in SQLite reader software such as the SQLite Manager Application for Mozilla Firefox (https://code.google.com/p/sqlite-manager/). Data stored in the sync_config.db file included the local path for the sync folder where the files in the account are downloaded and synchronised (\??\C:\Users\[username]\Google Drive) and the user email used to access the Google Drive account. Data stored in the snapshot.db file included the file details for files stored in the Google Drive account. Information such as filenames, modified and created times, URL, Size, Resource ID, and a checksum value which matched the MD5 value for the associated file. Example data from one of the snapshot.db files is included in Table 31 with the data for the Enron.jpg test file. The modified and created times appear to be Unix Numeric time format, which can be decoded using DCode (www.digital-detective.co.uk) and were found to match the times of the files stored in the sync folder. The information in the snapshot.db file is of particular importance to a forensic investigator.
Google Drive synchronised files and folders were observed at the default Google Drive folder location (‘C:\Users[username]\Google Drive’). Within this folder was a ‘desktop.ini’ file with the following contents; ‘InfoTip=Your Google Drive folder contains files you’re syncing with Google.’ and ‘IconFile=C:\Program Files\Google\Drive\googledrivesync.exe’. A Prefetch file was also created in the Windows Prefetch folder with the name ‘GOOGEDRIVESYNC.EXE-XXXXXXXXXX.pf’. Client Software filename references were also located in a variety of locations, such as; $LogFile, $MFT, $UsnJrnl, hiberfil.sys, and pagefile.sys. Link files were created on the Windows Desktop and in the Windows Start Menu.

Also of note is that the password for the Google Drive user account was located in cleartext within the file; ‘C:\Users[username]\AppData\Local\Microsoft\Internet Explorer\Recovery\Last Active\{XXXX6D05-B2C2-11E1-B53F-000C29985EDE}.dat’ near the text ‘&Passwd=’, such as;

‘&ktl=&ktf=&Email=username@mail.com&Passwd=XXXXXXXX&PasswdAgain=XXXXXXXX’

This information was located within the IE Upload-VM, and also in System Volume Information Restore Points. The password in the file was not observed when the other browsers were used. Whilst this may be beneficial to a forensic investigation, this also presents a security risk to users. Subashini and Kavitha (2011) explained that ‘[m]alicious users can exploit weaknesses in the data security model to gain unauthorized access to data’. A password and username stored within files on a hard drive could easily be discerned by a criminal user who has gained remote access to a victim’s computer using a common exploit, and could be used to extract data from an account, or to store illicit data in the victim’s account.
When installed, the Google Drive client software ran automatically when the Windows operating system started, and logged in to the user account without prompting for a password. This can be of assistance in an investigation, as a forensic copy of a seized computer hard drive can be used with software which will allow the forensic copy of a hard drive to be run within a virtual environment. Software such as Virtual Forensic Computing (VFC) or LiveView will scan the forensic image of a hard drive and prepare the requisite files to run an operating system on a hard drive within VMWare Player. In tests conducted, when the forensic copy of a hard drive contained the Google Drive client software with a user account and password already stored, the PC when started in a VM automatically signed in to the Google Drive account. This process will provide an examiner access to the files stored within the Google Drive user account (once synchronised).

In addition there was an option with the Google Drive client software from the Google Drive icon at the bottom right of the desktop on Windows 7 labeled ‘visit Google Drive on the web’ which when selected resulted in the opening of the default browser and provided full access to the Google Drive account, including the ability to view user activity, all items, and view the modified, edited and last opened dates for files. In a forensic environment, care would need to be taken when connecting a forensic image to the Internet. Legal authority would be required to ensure an examiner has the appropriate authority within their jurisdiction to examine the data stored within the cloud storage account, which could potentially be stored overseas or in another jurisdiction.

In Australia, Section 3L of the Crimes Act 1914 (Cth) has a provision for the executing officer of a warrant to access data which includes data not held at the premises, i.e. accessible from a computer or data storage device. Where assistance is required in operating the equipment and the executing officer believes on reasonable grounds that the material is liable to be destroyed, modified or tampered with they may ‘do whatever is necessary to secure the equipment, whether by locking it up, placing a guard or otherwise’ (see Section 3L(4)) for a period of up to 24 hours or until expert assistance can be obtained, whichever happens first (see subsection 3L(6)). This period may be extended under Section 3L(7), and notice of these arrangements must also be given to the occupier under Sections 3L(5) and 3L(8).
Section 3LA also makes it an offence for persons with knowledge of computers or computer networks of which computers form a part, or measures applied to protect data held in, or accessible from, computers, to fail to provide any information or assistance that is reasonable and necessary to allow access to data held in, or accessible from, a computer that is on warrant premises, to copy the data to a data storage device, or to convert the data into documentary form. Failure to comply carries a maximum penalty of two years imprisonment. Provisions such as Section 3LA of the Crimes Act 1914 (Cth) are designed to overcome the efforts of accused persons to conceal data through the use of passwords or encryption, including in cloud storage services.

For this research it was also attempted to copy the files contained within the installed Google Drive folders to another PC, however, this did not allow access to the Google Drive account without knowing the username and password. The process of copying the software files from a PC previously worked with Dropbox client software (McClain 2011; Chung et al. 2012).

7.2.5.3 Google Drive Account when accessed via a browser

It was observed that the Google Drive web account (accessed via https://drive.google.com) displays the username at the top right of the browser. When selecting a file in a Google Drive account using a browser, there is an option to display the metadata for a file by clicking on the ‘Manage Revisions’ link using the right-mouse-button (RMB). Folder information is also displayed, and includes the last modified time.

When downloading files from Google Drive via the browser, the folders and files are packed into a compressed ZIP file. There is an option to download ‘All Items’ and the ZIP file is named ‘documents-export’ with the date, for example; ‘documents-export-2012-08-29.zip’. When a download is undertaken from a folder, the ZIP is named with the folders name and the date, for example; ‘Dataset-2012-08-29.zip’.

7.2.5.4 Keyword Search Terms

Keyword search terms were determined from the filenames observed, and the text from within the Enron data files. These included the following;
the username and password of the Google Drive account created for this research,


### 7.2.5.5 Directory Listings

Analysis was conducted using X-Ways Forensic 16.5 and Encase 6.19.4 to view the filenames stored within the VMs, determined from the directory listing ($MFT^7 files) and keyword searches across the forensic image files. Analysis of the directory and file listings of the four control Base-VM hard drives revealed no references to the Enron sample data filenames, or Google Drive client software filenames. It was concluded that references to Google Drive and the Enron sample files were not present in the control media directory listings prior to installing the software or accessing the files.

Not surprisingly, there were references to the Google Drive client software in the Upload-VMs, which correlated to the VMs where the client software was used. This included the downloaded installation program; ‘googlesdrivesync.exe’, and the previously mentioned database files; ‘sync-config.db’ and ‘snapshot.db’. There were no references to the client software in the other VMs. There were references to the filenames for the Enron sample files in the VMs (excluding the Base VM’s), and when the files were downloaded via the browser from the root folder. In the Upload-VM’s, there was a substantial amount of filenames seen, including from the source folder, and in the default Google Drive folder; ‘C:\User\[username]\Google Drive’.

When a browser was used to access a Google Drive account there were less filename references for the Enron files, (in comparison with when the client software was used to access cloud storage). However, with a browser being used to access the Google Drive account stored files, there were sufficient file name references remaining on the hard drives to identify the filenames that were accessed, such as the Enron sample data filenames. In the circumstances when the files were downloaded and opened from the Google Drive cloud storage, there were additional filename references observed on the

---

^7 ‘The NTFS file system contains at its core, a file called the master file table (MFT). There is at least one entry in the MFT for every file on an NTFS volume, including the MFT itself’ Microsoft (2008), *How Ntfs Reserves Space for Its Master File Table (Mft)*, <http://support.microsoft.com/kb/174619>..
hard drive, including link files. This indicates when a user accesses Google Drive there will potentially be references in the $MFT directory listing and in a variety of locations (such as pagefile.sys and prefetch files) to indicate this, and also to potentially determine the filenames associated with Google Drive use by searching for Google Drive filename entries in the $MFT.

Link file references remained after Eraser and CCleaner had been used to remove the files and ‘clean’ the hard drive. Using CCleaner with all options selected removed the link files (however sufficient information remained in other locations such as prefetch files, relating to Google Drive and the Enron sample files, discussed in the next section). The use of Darik’s Boot and Nuke (DBAN) completely erased the hard drives, and there was no data to analyse in these VMs.

7.2.5.6 Prefetch Files

Analysis of the VM hard drives identified that prefetch\(^8\) files stored information relating to the filenames of the Google Drive executable and Enron test data filenames. This was located in all the VM hard drives except the control Base-VM hard drives and the Access-VM hard drives. It was observed that even after running CCleaner there was enough information in prefetch files, such as notepad.exe.pf, wordpad.exe.pf, and dllhost.exe.pf, to indicate the presence and path of the Enron sample data files. Also located were googledrivesync.exe.pf prefetch files when the client software was installed. Information located within the prefetch files included the file and folder path, the number of times run, and last run time and date. This information may be important in an investigation.

---

\(^8\) ‘Windows keeps track of the way a computer starts and which programs are commonly opened. Windows saves this information as a number of small files in the prefetch folder.’ Microsoft (2012b), What Is the Prefetch Folder?, <http://windows.microsoft.com/en-AU/windows-vista/What-is-the-prefetch-folder>..
7.2.5.7 Link Files

Analysis of Link files was undertaken using Encase 6.19.4. No Google Drive associated link files were found within the four control Base-VMs, or the Access-VM’s. Link files relating to Google Drive and the Enron sample data files were observed for all Upload-VM, Download-VM, Eraser-VM and CCleaner-VM hard drives. When CCleaner was used with all options selected, there were no logical link files containing Google Drive or the Enron sample data filenames (however the information from the deleted Link files was still present and recoverable from the hard drives).

The link files observed related to the filenames and folder names for the Google Drive executable, such as; C:\Program Files\Google\Drive\googledrivesync.exe with the description: Your Google Drive folder contains files you’re syncing with Google. Link Files relating to the Enron test data were also located. These were located in the ‘Windows\Recent’ and ‘Windows\Start’ folders in the Users ‘AppData’ directory, such as; ‘C\Users\[username]\AppData\Roaming\Microsoft\Windows\Recent\3111.link’.

There were no Google Drive or Enron link files within the four Access-VM hard drives, indicating that if files were not downloaded from Google Drive to the computer and opened, link files were not created. The presence of a link file may indicate the prior presence of a file, even if the logical file has been deleted, which may be of importance in an investigation.

7.2.5.8 Thumbcache Files

Analysis of the thumbnail pictures stored within the thumbcache files within the four control Base VMs determined that there were no Enron sample picture thumbnails present prior to installing or accessing Google Drive. Thumbnails for the Enron sample picture were located in ‘thumbcache_256.db’ files in the Upload-VM, Download-VM and Eraser-VMs, however none were found in the Access-VMs or the CCleaner VMs. This indicates that the thumbnail cache is a source for possible data relating to Google Drive use, but results may not be definitive. It is possible for Google Drive to be used in

---


10 ‘Windows 7 creates small thumbnail images of graphic files. There are files named Thumbcache_32.db, Thumbcache_96.db, Thumbcache_256.db & Thumbcache_1024.db which correspond to the thumbnails stored for that specific user account and size’ Mueller, L (2010), Windows 7 Forensics - Part Iv - Thumbcache*.Db’, <http://www.forensickb.com/2010/01/windows-7-forensics-part-iv.html>.
certain circumstances without leaving thumbnails, for example; accessing files using a browser but not downloading the files to a computer. Thumbcache data may be of assistance to determine whether a file was present on a drive, even if the file has been deleted or erased. It is possible that even if a file has been erased, enough information may remain within a thumbcache file to show the contents of a picture file, conclude a file was present, and associated dates and times may also be of assistance.

7.2.5.9 Event Log Files

Event Log Files are sources of information relating to system, software, and other events recorded by Windows operating systems. In this research, there were two records relating to Google Drive within the event log files. These were in the IE Upload-VM ‘Microsoft-Windows-Diagnostics-Performance Operational.evtx’ file, and related to the googledrivesync.exe file. However, the details of these entries could not be interpreted with Event Log Explorer, EvtParser or the inbuilt Windows Event Viewer. Similar entries were not seen when the Google Chrome, Mozilla Firefox, or Apple Safari browser were used. No entries were located in the Base-VM images, nor in the other VM images. Therefore data in an Event log file may indicate if the Google Drive client software was installed on a PC, even if the software has been uninstalled. Examining the entries in context may indicate dates and times when the software was in use.

7.2.5.10 Registry Files

Windows Registry files were parsed using RegRipper version 20080909 and parsed and analysed with X-Ways 16.5. Encase 6.19.4 was also used to analyse the registry files and RegRipper output. Analysis revealed there were no references to Google Drive or the Enron sample files in the four control Base-VM hard drives. There were no references to the Google Drive username or password within the registry files in any of the VM’s. When the Google Drive client software was used, there were references found in the Upload-VM’s for all browsers in NTUser.dat, UsrClass.dat, SOFTWARE and SYSTEM registry files. This information was located using keyword searches for the term ‘googledrivesync.exe’ across the Registry files and also the parsed text from the Registry files (output from RegRipper). Specific keys were created when the client software was installed, such as ‘SOFTWARE\Google\Drive’ with the key
‘InstallLocation’, listing the path to the executable; $C:\Program Files\Google\Drive\googledrive\googledrive.exe$

There were no references to the filenames when a browser was used to access the Google Drive account files. However, when a browser was used to download and open files from Google Drive accounts, there were references to the filenames of the files in the Download-VM registry files for all browsers. For example, the ‘RecentDocs’ key in the NTUSER.dat registry file provided a list of the Enron related files and folders, a sample RegRipper output is listed in Table 32;

<table>
<thead>
<tr>
<th>RecentDocs</th>
<th>recentdocs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All values printed in MRUList\MRUListEx order.</strong></td>
<td></td>
</tr>
<tr>
<td>$Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs$</td>
<td></td>
</tr>
<tr>
<td>LastWrite Time Sun Jun 17 02:31:14 2012 (UTC)</td>
<td></td>
</tr>
<tr>
<td>1 = Dataset</td>
<td></td>
</tr>
<tr>
<td>4 = Enron3111.rtf</td>
<td></td>
</tr>
<tr>
<td>3 = Enron3111.jpg</td>
<td></td>
</tr>
<tr>
<td>2 = Enron3111.docx</td>
<td></td>
</tr>
<tr>
<td>0 = 3111.txt</td>
<td></td>
</tr>
</tbody>
</table>

- -----------------------------

Table 32 - Output from RegRipper for Google Drive

Deleted information regarding Google Drive and the Enron files and folders was located within the NTUSER.dat registry files after CCleaner was run, such as the previously observed filenames and URL references. References were also found within the UsrClass.dat registry files when the Google Drive client software was used.

7.2.5.11 $Recycle Bin$

Files that were deleted were easily located in the $Recycle.Bin$ folder in a folder with the SID of the user. The information files (beginning with $I$) included the data relating to the original file and when it was deleted, such as the following example relating to the Google Drive Dataset folder, deleted from the default Google Drive sync location (Table 33);
Also located in the Recycle Bin were the file contents files (beginning with $R), which contained the original folder and file contents. In the scope of this research this included the files synchronised from Google Drive account using the client software.

### 7.2.5.12 Data Carve

Data carving is the process of searching through allocated or unallocated data to locate files based on known headers and footers, such as 0xFFD8 and 0xFFD9 for a jpg file. This was undertaken using Encase 6.19.4 File Finder Signature Search and X-Ways 16.5 File Header Signature Search across all VM’s created for this research. Thumbnail icons and partial full size pictures were carved from all VM’s, except the Base-VM’s.

### 7.2.5.13 Browser Analysis

Internet browsing information was analysed using Magnet Software (formerly JADsoftware) Internet Evidence Finder (IEF) v5.52 Standard Edition, X-Ways Forensic 16.5, and SQLite Database Browser v2. It was confirmed there were no references to Google Drive or Enron sample data in the internet history of the four control Base-VMs. References were located for the Google Drive website (‘drive.google.com’) in the Google Chrome Base-VM. The Google Drive account username could be determined when Mozilla Firefox, Google Chrome, and Internet Explorer were used to access the account using the browser, but the username was not located when Apple Safari was used. The username was stored in Firefox in the ‘formhistory.sqlite’ database and in Chrome in the Autofill ‘Web Data’ file.

The Google Drive website URL (drive.google.com) was located in a range of areas; such as; Cookie files, web history, FavIcons, in the FileSlack of other files, unallocated space, and Pagefile.sys. This occurred for all browsers. Filenames for files downloaded were located in the web history of all hard drives, except for the Base-
VMs. Data was located in Browser history files, such as index.dat files and downloads.sqlite. When using the Google Chrome browser to access Google Drive, references were located in a range of browser system files; such as Top Sites, History, Shortcuts, Preferences, History Provider Cache, FavIcons, Current Tabs and Current Sessions.

When storing files in Google Drive, a unique Resource-ID identifier is assigned to a file, and is included in a URL reference for the file. For example; for the file 3111.txt, Google Drive (and Google Docs) assigned the following URL:

https://docs.google.com/file/d/XXwUTD1c9KXixMXXNERGJKV1FTUUE/

Examining multiple URL’s, it appears the Resource-ID Identifier consists of the first part relating to the User, and the second part relates to the file, as for multiple files the initial thirteen characters are the same, and the subsequent fifteen characters are different depending on the file. This information is stored within the snapshot.db file, and IEF also outputs the information for the URL and the filename from the web browsing history in the format listed in Table 34. An examiner can conduct searches for a known filename or Resource-ID identifier to locate instances of the file or URL being used. Searches conducted across the network captures (PCAP) also located the Resource-ID Identifiers, associated with drive.google.com cookie files.

<table>
<thead>
<tr>
<th>Firefox (snapshot.db)</th>
<th>16/06/2012 20:48:46</th>
</tr>
</thead>
<tbody>
<tr>
<td>3111.txt - Google Docs</td>
<td><a href="https://docs.google.com/file/d/XXwUTD1c9KXixMXXNERGJKV1FTUUE/edit?pli=1">https://docs.google.com/file/d/XXwUTD1c9KXixMXXNERGJKV1FTUUE/edit?pli=1</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Internet Explorer (snapshot.db)</th>
<th>10/06/2012 16:01:40</th>
</tr>
</thead>
<tbody>
<tr>
<td>3111.txt - Google Docs</td>
<td><a href="https://docs.google.com/file/d/XXwUTD1c9KXixMXXNERGJKV1FTUUE/edit?pli=1">https://docs.google.com/file/d/XXwUTD1c9KXixMXXNERGJKV1FTUUE/edit?pli=1</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chrome Web History (IEF)</th>
<th>2012-06-17 02:19:47</th>
</tr>
</thead>
<tbody>
<tr>
<td>3111.txt - Google Docs</td>
<td><a href="https://docs.google.com/file/d/XXwUTD1c9KXixMXXNERGJKV1FTUUE/edit">https://docs.google.com/file/d/XXwUTD1c9KXixMXXNERGJKV1FTUUE/edit</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chrome Web History (IEF)</th>
<th>2012-06-17 02:19:56</th>
</tr>
</thead>
<tbody>
<tr>
<td>3111.txt - Google Docs</td>
<td><a href="https://docs.google.com/file/d/XXwUTD1c9KXixMXXNERGJKV1FTUUE/edit">https://docs.google.com/file/d/XXwUTD1c9KXixMXXNERGJKV1FTUUE/edit</a></td>
</tr>
</tbody>
</table>

Table 34 - example of snapshot.db SQLite file contents and IEF output
Internet Evidence Finder 5.5 introduced the ability to report information relating to cloud storage, including Google Drive related data. Using this option, some data was located in some Upload-VM files, but not for all instances when the software or access was undertaken and not all potential data was reported. Data was recovered from the Internet Explorer, Google Chrome and Apple Safari Upload-VM’s, but not the Firefox Upload-VM. Data located included filenames, the user name and the file URL Identifiers, associated file dates and times, and file sizes. Information was also located in the Internet Explorer Access-VM. Information was reported from the Google Drive ‘snapshot.db’ file, Pagefile.sys, Unallocated Clusters, and File Slack. The information recovered was of benefit, but was not all the available information from the VM’s. Additional information was viewable from the ‘snapshot.db’ files in each of the Upload-VM files, using the SQLite Manager Application for Mozilla Firefox. Using CCleaner with standard options and then all options did not remove the references previously located in relation to Google Drive use.

### 7.2.5.14 MetaData

Using X-Ways 16.5, it is possible to extract metadata from files, such as the EXIF data from JPG files, or information from documents. This can then be used to conduct analysis. It was possible to determine the picture files which related to the Enron test data by searching on the metadata contents. It was also possible to locate the rtf document files based on the Author information stored within the file.

### 7.2.5.15 Network Analysis

Analysis of the network traffic capture files was undertaken using Network Miner v1.0, Wireshark Portable 1.6.5 and Encase 6.19.4. The network traffic was observed on Port 80 (HTTP) and 443 (HTTPS). When accessing Google Drive accounts using the client software or a browser it appears login sessions are established with Google using IP numbers such as 74.125.237.133; 134; 135 and 144 (in the range registered to Google; 74.125.0.0 - 74.125.255.255), with URLs of ‘www.google.com’. Verification is then undertaken on IP 199.7.55.72, which is registered to Verisign Global Registry Services, with the URL; ‘ocsp.verisign.net’. Security certificates appeared to be approved using VeriSign/Thawte services, in the 199.7.48.0 to 199.7.63.255 range. The URL;
‘ocsp.verisign.net’ was observed in the network traffic, OCSP is the Online Certificate Status Protocol, used by VeriSign/Thawte. When downloading the client software, IP’s 173.194.72.84 and 173.194.72.190, (which are registered to Google) were observed, with the URLs ‘accounts.google.com’ and ‘dl-ssl.google.com’. The client software download of ‘gsync.msi’ then occurred via an ISP cache location (in this case 150.101.13.84 which is registered to Internode). When the client software was run and synchronised with an account, IP’s in the range 74.125.237.128-142 were observed, with URLs of ‘docs.google.com’. These are registered to Google in the range 74.125.0.0 - 74.125.255.255.

When accessing files or data in an account via a browser, the process of accessing IP’s in the ranges previously mentioned for the client software were observed; (1) ‘www.google.com’ (2) ‘verisign’ (3) ‘account.google’ (4) ‘docs.google’. An IP (173.194.72.189) with the URL of ‘drive.google.com’ was also observed. Table 35 lists the IP number ranges observed in the captured network traffic, and the registered owner for the range, in the order observed when connecting to an account and accessing data within the account. The same URL names were observed when using either the Google Drive client software or accessing via a browser. Previous analysis of Dropbox network traffic determined Dropbox uses different URL’s when using a browser compared to when using the client software.

<table>
<thead>
<tr>
<th>Step</th>
<th>Registered Owner</th>
<th>IP Start</th>
<th>IP Finish</th>
<th>URL’s observed in network traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Google Inc.</td>
<td>74.125.0.0</td>
<td>74.125.255.255</td>
<td><a href="http://www.google.com">www.google.com</a></td>
</tr>
<tr>
<td>2</td>
<td>Verisign Global Registry</td>
<td>199.7.48.0</td>
<td>199.7.63.255</td>
<td>ocsp.verisign.net</td>
</tr>
<tr>
<td>3</td>
<td>Google Inc.</td>
<td>173.194.0.0</td>
<td>173.194.255.255</td>
<td>accounts.google.com and dl-ssl.google.com</td>
</tr>
<tr>
<td>4</td>
<td>Google Inc.</td>
<td>74.125.0.0</td>
<td>74.125.255.255</td>
<td>docs.google.com and large-uploads.google.com</td>
</tr>
<tr>
<td>5</td>
<td>Google Inc.</td>
<td>173.194.0.0</td>
<td>173.194.255.255</td>
<td>drive.google.com</td>
</tr>
</tbody>
</table>

**Table 35 – Google Drive IP addresses observed in network traffic**

The username for the Google Drive account was observed in cleartext in the network traffic when Internet Explorer was used, but not when other browsers were used. The contents of the Enron sample data files were not seen in the network traffic, suggesting the data was encrypted. No password information was observed in cleartext in the
network traffic. The website URL ‘drive.google.com’ was observed in network traffic when an account was accessed, but not in the control Base-VM, the Eraser-VM or CCleaner-VM network traffic.

**7.2.5.16 System Volume Information**

When Eraser or CCleaner were used, information was still located in various System Volume Information files, also known as Restore Points or Volume Shadow Copies. The information included the username for the Google Drive account, the URL’s for the Google Drive website (https://drive.google.com), the filenames and full text of the Enron files was also located. When limited information was located during analysis, such as filenames only, the System Volume Information files contained a wider range of information, such as file contents or deleted URL’s. System Volume Information was highlighted as an important source of information when anti-forensic methods had been used, in this research; Eraser and CCleaner.

**7.2.5.17 Memory (RAM) Analysis**

Analysis of the memory captures (VMEM files) was undertaken using X-Ways 16.5 and Encase 6.19.4. The term ‘Google Drive’ was located in all the memory captures excluding the control Base-VM memory files. The URL (drive.google.com) was not located in the control Base-VM memory files, but was located in all other memory capture files. The Google Drive account username was located in all Upload-VMs, and in the Firefox, Google Chrome, and Internet Explorer Access-VM and Download-VM’s (not in the Safari Access-VM or Download-VM). The Google Drive username was located in memory capture files near the text; ‘? –Email ’, ‘?? Email’ and ‘<email>’. The password was located in freetext in the Upload-VM memory files for Firefox, Internet Explorer and Safari browsers, near the text; ‘&passwd=[password]’ and ‘&passwdagain=[password]’.

The username was also located in the Eraser-VM memory captures, in the Firefox and Google Chrome CCleaner-VM memory captures, and also in the Google Chrome memory capture file when CCleaner was used with all options. However when the
Safari browser was used, the username was only recovered in the Upload-VM memory capture.

The full text of the Enron data files was recovered in the Upload-VM, Access-VM and Download-VM memory captures. In addition, the text was located in the Google Chrome Erase-VM memory file. No text was located in the control Base-VM memory files. Observed within memory captures were Enron filename references in all VM memory files, except the control Base-VM memory files.

Data carving was undertaken across the memory capture files and resulted in the recovery of thumbnail pictures, partial and full picture files of the Enron sample pictures and Google Drive logo’s from the memory captures for all Upload-VM and Download-VM memory files. No picture files were recovered from the Base-VM memory files, or when Eraser or CCleaner had been used.

Analysis of Pagefile.sys files was also undertaken. The username was located in the Safari Access-VM pagefile.sys file and the Internet Explorer Upload-VM pagefile. The website URL (‘https://drive.google.com’) were located in the Internet Explorer and Safari browser pagefile.sys files, but not the Firefox or Google Chrome browser pagefile.sys files. The filenames for the Enron test files were located in Upload, Download, Erase and CCleaner-VM pagefile.sys files when Chrome and Safari browsers were used, but not when Firefox was used. Filenames were located in the Internet Explorer Upload-VM pagefile.sys file, but no other Internet Explorer VM files. Filenames were located after Eraser and CCleaner were used, again highlighting pagefile.sys as an area for examination if anti-forensic methods are suspected of being used. The full text of the Enron email was found in the pagefile.sys file in the Google Chrome Erase and CCleaner-VM files only.

7.2.5.18 Eraser, CCleaner and DBAN

As discussed, many remnants indicating the use of Google Drive and the presence of the Enron sample files were located after Eraser and CCleaner were used. This indicated that Eraser or CCleaner did not necessarily remove all data remnants in relation to the use of Google Drive or completely remove the information relating to the presence of the Enron sample files. The use of a full erasing tool, in this case Darik’s Boot and
Nuke (DBAN), was found to remove all traces, but this also erased the operating system and user files. The operating system and user files would need to be re-installed, which could be a lengthy process and may dissuade the average user from undertaking this process to remove evidence of Google Drive usage.

7.2.6 Presentation

7.2.6.1 Analysis findings

In this case study, a variety of data remnants were located when a user used Google Drive to store or access data. The focus (or scope) was to determine what data remnants were left to identify if Google Drive had been used on a Windows 7 PC. Conducting a search for the website ‘drive google com’ was shown to be inconclusive to demonstrate use, as it was shown that the term is present even when Google Drive has not been accessed, as outlined in the analysis of the control Base-VM’s in the Google Chrome Base-VM ‘chrome.dll’ file and ‘chrome.7z’.

The Google Drive username was able to be discerned from a variety of locations, such as cookie files, memory captures (searching for ‘&Email=’), in pagefile.sys, from SQLite database files, and IEF output. The username was also observed in cleartext the network traffic when the Internet Explorer browser was used, but not when other browsers were used.

Surprisingly, the password for the Google Drive account was able to be located unencrypted in an Internet Explorer .dat file, as follows;

'C:Users\[username]\AppData\Local\Microsoft\Internet Explorer\Recovery\Last Active\(XXXX6D05=B2C2-11E1-B53F=000C29985EDE\).dat'

near the text ‘&Passwd=’, such as;

'&ktl=&ktf=&Email=username@mail.com&Passwd=XXXXXXXX&PasswdAgain=XXXXXXXXX'

This was not observed when the other browsers were used. This information was also located within System Volume Information Restore Point data on the IE Upload-VM image file. The cleartext password was also located in the memory captures from the Upload-VM’s for Mozilla Firefox, Internet Explorer and Apple Safari browsers, pagefile.sys
from the Safari Access-VM, and the Internet Explorer Upload-VM pagefile. The password was located in freetext near the text; ‘&passwd=[password]’ and ‘&passwdagain=[password]’.

Passwords are commonly used in a variety of locations, and when conducting analysis to determine a password to view encrypted files, one practice is to build an index of words from a forensic image to use with password analysis tools such as AccessData Password Recovery Toolkit, Passware, or Elcomsoft password analysis tools. With the actual password stored in cleartext this would vastly speed up the process of password discovery. Hence, memory capture and analysis should be a key consideration when determining the possible sources of data in step 3 of the framework (Identify and Collect) for indexing purposes.

If the password is not located in memory, or if memory capture is not possible, it is still possible to access a Google Drive account using the forensic image of a hard drive where the Google Drive client software has been used to access an account. This can be achieved by running the hard drive forensic image as a Virtual Machine. When the VM starts up it automatically synchronises with the Google Drive cloud service and updates files in the local Google Drive folder. The Google Drive client software also includes the ability to connect to the account via a browser, without requiring the username or password. There is an option from the Google Drive icon at the bottom right of the desktop on Windows 7 labeled ‘visit Google Drive on the web’ which when selected results in the opening of a browser with full access to the Google Drive account, including the ability to view user activity, all items, and view the modified, edited and last opened dates for files. Appropriate legal authority would be required to ensure the account and information can be accessed and used for analysis.

Analysis to determine the method of access; whether the client software was used, a browser used to upload, access or download, or a combination of both, is possible. When the client software was downloaded, there was a GoogleDriveSync.exe file downloaded to the local hard drive. In addition, registry entries may indicate the use of the setup software or the client software, as will link files. Prefetch files were observed for the client software, and also included the number of times run, and associated dates and times.
When using a browser to access a Google Drive web account, the username was
displayed at the top right of the browser. When selecting a file in a Google Drive
account using a browser, there was an option to display the metadata for a file by
clicking on the ‘Manage Revisions’ link using the right-mouse-button (RMB). Folder
information was also displayed, and included the last modified time. This information
may be important in an investigation and should be recorded, either as hand written
notes, screen captures, or video recording the browser window using software such as
Microsoft Expression Encoder 4. There were many references in the output from
Internet Evidence Finder (IEF) for browser access to a Google Drive account. When
bulk files were downloaded using the web account from the root folder, a compressed
ZIP file was created and downloaded with the name of the folder and the date in the
filename, such as; ‘Dataset-2012-08-29.zip’.

When using the client software, files of interest are created in the Users folder at the
following location; ‘C:\Users\{username}\AppData\Local\Google\Drive\’. Within this
folder were two files of interest; ‘sync_config.db’ and ‘snapshot.db’. Both files are
SQLite format 3 database files. Data stored in the sync_config.db file included the local
sync root path (\?\I:\Users\{username\}\Google Drive) and the user email used to access
the Google Drive account. Data stored in the snapshot.db file included the file details
for files stored in the Google Drive account. Information such as filenames, modified
and created times, Resource-ID, File Size, and a checksum value which matched the
MD5 value for the associated file.

Filenames were also observed in the Registry ‘RecentDocs’ keys, $MFT entries, and
Prefetch files such as DLLHost.pf, Wordpad.pf and Notepad.pf. Link files and the IEF
output also listed the filenames. The contents of the files were also recovered from
temporary internet files, thumbcache, memory captures, pagefile.sys, system volume
restore points, and unallocated space. Eraser and CCleaner were not effective in
removing all data remnants, and information was able to be determined from VM’s
relating to the Google Drive accounts, filenames, dates and times and file contents.

In total, it has been demonstrated that it is possible to determine the username,
password, method of access, filenames, contents, and dates and times of access when
Google Drive is used to store, access or download data from cloud storage. Memory
captures and Network files were an important source of information, including locating
the username and password in memory files when the client software was setup and used. The various access points are listed in Table 36.

Once forensic analysis has determined a Google Drive account has potential evidence of relevance to an investigation, the examiner can communicate this to relevant persons to enable them to respond to secure evidence in a timely manner. Knowledge of the username details would enable identification of Google Drive accounts, to preserve and secure potential evidence.

<table>
<thead>
<tr>
<th>Control (Base-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username, Password, Software, URL, Enron sample filenames or files KWS terms</td>
<td>Nil</td>
</tr>
<tr>
<td>Nil</td>
<td></td>
</tr>
<tr>
<td>Nil</td>
<td></td>
</tr>
<tr>
<td>Nil</td>
<td></td>
</tr>
<tr>
<td>Matches to ‘drive google.com’ when Google Chrome installed (Google Chrome Base-VM)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Client Software (Upload-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>Users\AppData\Local\Google\Drive\sync_config.db; Memory capture files ‘? -Email=’ Located on HD (IE)</td>
</tr>
<tr>
<td>Password</td>
<td>Located in RAM – search for &amp;Email= and &amp;Passwd= and &amp;PasswdAgain= GoogleDriveSync.exe file located when downloaded. Google Drive Software installation under C:\Program Files\Google\Drive\Users\AppData\Local\Google\Drive\sync_config.db and snapshot.db includes file information When software downloaded, URL’s included <a href="https://drive.google.com">https://drive.google.com</a> Multiple locations, including Prefetch, Link files, $MFT, Registry. Located in Sync folder under User\Google\Drive. Multiple matches to KWS terms</td>
</tr>
<tr>
<td>Software</td>
<td>FF and GC History; ‘formhistory.sqlite’ and ‘Web Data’; Memory capture files ‘? -Email=’</td>
</tr>
<tr>
<td>URL</td>
<td>Nil</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Multiple locations; cookies, history, icons, pagefile.sys and unallocated Sufficient to identify files accessed with references to the filenames in Registry and Browsing History Full text in Temporary Internet Files, RAM and System Volume Information Multiple matches to KWS terms</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>KWS terms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Browser Access (Access-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>FF and GC History; ‘formhistory.sqlite’ and ‘Web Data’; Memory capture files ‘? -Email=’</td>
</tr>
<tr>
<td>Password</td>
<td>Nil</td>
</tr>
<tr>
<td>Software</td>
<td>Multiple locations; cookies, history, icons, pagefile.sys and unallocated Sufficient to identify files accessed with references to the filenames in Registry and Browsing History Full text in Temporary Internet Files, RAM and System Volume Information Multiple matches to KWS terms</td>
</tr>
<tr>
<td>URL</td>
<td>Enron sample filenames</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>KWS terms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Browser Download (Download-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>FF and GC History; ‘formhistory.sqlite’ and ‘Web Data’; Memory capture files ‘? -Email=’</td>
</tr>
<tr>
<td>Password</td>
<td>Nil</td>
</tr>
<tr>
<td>Software</td>
<td>Multiple locations; cookies, history, icons, pagefile.sys and unallocated Sufficient to identify files accessed with references to the filenames in Registry and Browsing History Full text in Temporary Internet Files, RAM and System Volume Information Multiple matches to KWS terms</td>
</tr>
<tr>
<td>URL</td>
<td>Enron sample filenames</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>KWS terms</td>
</tr>
<tr>
<td>Eraser (Eraser-VM)</td>
<td>Data artefacts found</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Username</td>
<td>FF and GC History; 'formhistory.sqlite' and 'Web Data'; Memory capture files '=?-Email='</td>
</tr>
<tr>
<td>Password</td>
<td>Nil</td>
</tr>
<tr>
<td>Software</td>
<td>Nil</td>
</tr>
<tr>
<td>URL</td>
<td>Multiple locations; cookies, history, icons, pagefile.sys and unallocated files</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Sufficient to identify files accessed with references to the filenames in Registry and Browsing History</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>Full text in Temporary Internet Files, RAM and System Volume Information</td>
</tr>
<tr>
<td>KWS terms</td>
<td>Multiple matches to KWS terms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CCleaner (CCleaner-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>FF and GC History; 'formhistory.sqlite' and 'Web Data'; Memory capture files '=?-Email='</td>
</tr>
<tr>
<td>Password</td>
<td>Nil</td>
</tr>
<tr>
<td>Software</td>
<td>Nil</td>
</tr>
<tr>
<td>URL</td>
<td>Multiple locations; cookies, history, icons, pagefile.sys and unallocated files</td>
</tr>
<tr>
<td>Enron sample filenames</td>
<td>Sufficient to identify files accessed with references to the filenames in Registry and Browsing History</td>
</tr>
<tr>
<td>Enron sample files</td>
<td>Full text in Temporary Internet Files.</td>
</tr>
<tr>
<td>KWS terms</td>
<td>Multiple matches to KWS terms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DBAN (DBAN-VM)</th>
<th>Data artefacts found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username, Password</td>
<td>All Data erased, no information located</td>
</tr>
<tr>
<td>Software, URL, Enron</td>
<td></td>
</tr>
<tr>
<td>files, KWS terms</td>
<td></td>
</tr>
</tbody>
</table>

Table 36 - Summary of Google Drive analysis findings

7.2.7 Feedback

Important information discovered in this research was the ease of locating a Google Drive username, the discovery of the password stored in cleartext in multiple locations, the ability to gain full access an account without knowing the username or password, and the vast amount of information able to be located in relation to Google Drive use, either via a browser or client software. It is important to share this information with the wider IT community, for both forensic analysis purposes and information security purposes, to enable practitioners to locate information when required, and to ensure security issues are understood when using Cloud storage services. The addition of the steps of ‘prepare’ and ‘feedback’ to a forensic analysis framework was beneficial to adequately prepare for analysis and also to incorporate lessons learnt into subsequent investigations. It is important to consider the positive and negative experiences during an examination and incorporate these into future examinations. It is also important to report any issues or improvements to other examiners, and the digital forensic community in general.
7.2.8 Complete

The final step is the completion of an investigation. A decision is made whether further analysis is required, based on the feedback from those involved in the process. If further enquires are required, the process returns to the ‘Prepare and Respond’ step. If there are no further avenues for investigation, the case can be completed, data archived and backed-up as required. In the context of this research, the data created and used may be required in future research opportunities, and hence has been stored on multiple hard drives to enable future use.

During the process of conducting this research, it was identified that mobile devices are becoming widely used to access cloud storage, and hence it was determined to undertake research into the ability to identify cloud storage use on a mobile device. To further test the use of the Digital Forensic Analysis Cycle, it was determined to undertake research into the data remnants on an Apple iPhone available to the researchers, which is outlined in the next section.
7.3. Google Drive Analysis: Apple iPhone 3G

The use of portable devices to access cloud stored data is increasing. A report from Gartner, for example, highlights the trend in client computing is shifting from a focus on personal computers to include portable devices, and also predicts that personal cloud storage will replace the PC as the main method of users’ storage by 2014 (Kleynhans 2012). The Apple iPhone is one such device that can be used to access Google Drive cloud storage either using the built in Safari browser, or installing the Google Drive application (App). Analysis of portable devices is a growing area for forensic practitioners, and there are hardware and software solutions to assist this process, such as; MicroSystemation .XRY, Cellebrite UFED, and Radio Tactics Aceso. In the scope of this research, it is relevant to consider what information can be determined from a portable device in relation to the use of Google Drive. This also serves to further assess the suitability of the framework being applied in a variety of circumstances.

7.3.1 Commence (Scope)

The scope of this part of the research is to examine the data remnants on an iPhone 3G running iOS 4.2.1 when used to access Google Drive via a browser. The data remnants for the Google Drive iOS application were also to be tested, however the application will only work with iOS 5 or above, which prevented installation on the test device used in the research. Analysis also seeks to determine if there were any remnants prior to accessing Google Drive on the Apple iPhone 3G.

7.3.2 Prepare

An iPhone 3G was selected which had not been used to access Google Drive previously. The device usage was known, and MicroSystemation .XRY 6.2.1 was used to extract a logical image of the contents, excluding audio and video files, prior to accessing Google Drive. This first extract was analysed to confirm there was no Google Drive related data on the device prior to undertaking the research. The inbuilt Apple Safari iOS browser was used to access the Google Drive user account created for this research and view the Enron files stored remotely. A logical extract was then conducted using .XRY.
7.3.3 Identify and Collect

In the context of this research, files were identified which would contain the information needed to conduct the analysis, in this instance being the .XRY extract files and the output of the .XRY software, including PDF reports and the files exported using .XRY. These were identified for each of the extracts; Base and Browser.

7.3.4 Preserve (Forensic Copy)

To observe the principles of forensic computer analysis (NIJ 2004; ACPO 2006; NIJ 2008) a forensic copy was made of the .XRY extract files, the file output, and the reports. As these were logical files, this was done in the Encase Logical format (L01), and the X-Ways Evidence File Container format (ctr). MD5 hash values were used to ensure the forensic integrity of the data.

7.3.5 Analysis

For this research, each of the forensic logical files was examined using forensic tools including; X-Ways Forensic 16.5 and EnCase 6.19.4. PList Explorer v1.0 was used to examine the contents of the Apple plist files extracted by .XRY.

7.3.5.1 Control - Base-XRY

Analysis of the control Base-XRY extract confirmed there was no data originally present relating to the Enron sample test data. In addition, no references were located for the term ‘Google Drive’ or the website URL (drive.google.com).

7.3.5.2 Google Drive accessed via the iOS Safari Browser

The Google Drive username was located in a ‘cookies.binarycookies’ file near the text ‘m.google.com’. Filenames for the Enron test files were located in the History.plist, for the txt, rtf and docx files. This information was also extracted by XRY in the Web-History.txt file. The text from the Enron files was not located in any of the extracts, nor was it possible to locate the password for the Google Drive account.
7.3.6 Presentation

7.3.6.1 Analysis findings

In this research, several data remnants were located when searching for evidence of Google Drive use on the test item (iPhone 3G). Recall that the aim was to determine what data remnants were left to identify if Google Drive was accessed. Initial analysis of the control extract identified no matches to the keyword search terms, such as ‘Google Drive’, and ‘drive.google.com’. When the browser was used to access the Google Drive account, entries were left in the History.plist file, which is also listed in the .XRY Web-History.txt file. The Google Drive username was listed in the ‘cookies.binarycookies’ file. The filenames accessed were listed in the History.plist file and in the .XRY Web-History.txt file.

7.3.7 Feedback

This research was limited by not being able to install the Google Drive Application, or to jailbreak the iPhone. .XRY and other iPhone extract solutions require the iPhone to be jailbroken to undertake a Physical extract. A future research opportunity is to undertake similar research using an iOS 5 device, and also widen the devices examined to include other popular mobile device operating systems, such as Google Android and Microsoft Windows. In addition, the comparison of a physical extract of an iPhone to a logical extract to determine what information is available in comparison with the logical extract, and also to compare the process with other iPhone forensic software and hardware processes.

7.3.8 Complete

In the context of this research, the data created and used may be required in future research opportunities, and hence has been stored on multiple hard drives to enable future use.
7.4 Google Drive Analysis: Case Study

To illustrate the relevance of the research findings (including the framework), the following hypothetical case study outlines where the information identified in this chapter can assist in an investigation. The circumstances relate to a hypothetical case of a person suspected of importing steroids. The circumstance of the case study are greatly simplified for the purposes of this paper, and only explore one of several possible responses using currently available software and hardware forensic solutions;

“Police receive confidential information that Suspect has been importing and selling steroids. A close acquaintance was visiting Suspect, and observed computer files relating to steroid sales on the screen of a computer in Suspect’s study.

An investigation case file was raised, and issued to IO Jones. IO Jones undertook background enquiries and determined that there had been prior allegations about Suspect being involved in steroid importation offences, and that he had previous convictions for these offences. A Warrant was issued by the Court for IO Jones to attend the premises, and search and seize evidence to investigate these allegations.

IO Jones liaised with the Police Forensic Computing Section (FCS), to arrange for a Specialist to attend the search and provide assistance to IO Jones under the terms of the warrant.”

7.4.1 Commence (Scope)

IO Jones completed a FCS Request for Assistance form, outlining the scope of the investigation, namely to attend at the addresses identified and provide assistance in relation to search and seizure of evidence of material related to the importation of steroids.
7.4.2 Prepare and Respond

As this type of attendance is a process undertaken by the FCS on a regular basis, the examiners are trained and experienced in the examination of the computers and electronic devices. A Standard Operating Procedure (SOP) relating to the on-scene examination of computers and mobile devices is documented and referred to by examiners, and outlines the process. Notes are an important aspect of the SOP, and all examiners maintain accurate notes of the process they undertake.

7.4.3 Identify

On the nominated day, Examiner Perry attended the address of Suspect in the company of IO Jones and the response team. IO Jones secured the scene, and initially identified a tower PC (running) in the study, a generic laptop in the lounge room (off), and an Apple iPhone 3G (on) in a bedroom. Examiner Perry was then called in to provide assistance in searching the PC, Laptop, and iPhone for evidence.

Examiner Perry photographed all items *in situ* and made written notes about the items, including make, model and serial numbers of hard drives and the hardware. Exhibit sheets were started for each item to record information relating to access. The iPhone was placed in a Faraday bag¹¹.

7.4.4 Preserve (PC on Scene)

A video camera was setup to record the PC screen during the examination. Data capture of the running PC was undertaken with LiveDetector (http://h11dfs.com/live-detector.php) from a CD to preserve the contents of memory and running processes. Data was saved to a USB which had been erased prior to responding.

---

¹¹ ‘Faraday bag’ is intended to shield a mobile phone or similar small device to prevent unwanted applications being invoked remotely, such as wiping the memory, or to prevent possible problems with veracity of evidence’ Duffy, A (2010), Faraday Bag Testing in a Reverberation Chamber – a Preliminary Study, <http://www.faradaybag.com/faradayバッグ-testing/demontfort-university-faradayバッグ-testing.html>..
7.4.5 Analysis (PC on Scene)

Helix3 (http://www.e-fense.com/helix3pro.php) on a CD was then used to run AccessData FTK Imager and examine key files of interest. References were located indicating potential customer data files associated with a Google Drive account in the file; ‘C:\Users\Suspect\AppData\Local\Google\Drive\ snapshot.db’. A username was located in ‘C:\Users\Suspect\AppData\Local\Google\Drive\sync_config.db’ indicating the username ‘suspect@badstuff.com’. An icon for Google Drive client software was observed running in the system tray. Examiner Perry understood that any further interaction with the live PC may make changes to current or historical data, and (with the knowledge that live data had been preserved) pulled the power plug to immediately shut off the system. The hard drive was removed from the PC and connected to a Tableau T35es-R2 eSATA Forensic Bridge (hardware write blocking). This was connected to Examiner Perry’s laptop computer to enable on scene analysis to continue without altering the contents of the hard drive. Virtual Forensic Computing software (http://www.virtualforensiccomputing.com/) was then used to run the physical write blocked hard drive as a virtual computer (VM). When the VM finished the startup process it was ready to continue analysis.

The option ‘visit Google Drive on the web’ was selected from the Google Drive icon previously observed in the System Tray. A browser window opened with access to the Google Drive account of Suspect. The details of the files in the account were recorded, including the dates and times listed, filenames, and owner details where listed. The files were selected and downloaded to the USB as a compressed ZIP file. As the live data had been preserved and on scene analysis had identified evidence stored on the PC, the PC was seized by IO Jones.

7.4.6 Preserve (iPhone on Scene)

Microsystemation .XRY was used to extract the logical data from the iPhone.

7.4.7 Analysis (iPhone on Scene)

On scene analysis of the logical extract of the iPhone located filenames associated with the previously identified Google Drive account in the ‘History.plist’ file, and the
information was also reported in the .XRY ‘Web-History.txt’ file. As relevant evidence was stored on the iPhone, it was seized by IO Jones.

7.4.8 Preserve (Laptop on Scene)
As the laptop was in a powered off state, the hard drive was removed and connected to a Tableau T35es-R2 eSATA Forensic Bridge (hardware write blocking). This was connected to Examiner Perry’s laptop computer to enable on scene analysis to be undertaken without altering the contents of the hard drive.

7.4.9 Analysis (Laptop on Scene)
On scene analysis of the laptop hard drive was undertaken with X-Ways 16.5. Link file references were located for the previously identified Google Drive files in the Start menu. As relevant evidence was stored on the laptop, it was also seized by IO Jones.

7.4.10 Preserve (Forensic Copy)
All the items were transported by IO Jones, and lodged for examination with the FCS. A forensic copy of the hard drives was then made using AccessData FTK Imager 3.1.0 and a Tableau T35es-R2 eSATA Forensic Bridge (hardware write blocking), and verified with MD5 and SHA1 hashes. Appropriate entries were made in the exhibit log relating to access to the evidence by the examiner.

7.4.11 Analysis
Analysis was then undertaken on the forensic copies of the data using X-Ways Forensic 16.5.

IO Jones communicated with the legal section of Google in relation to the data identified in the Google Drive account to ensure the data in the account was preserved and made available in relation to the investigation.
7.4.12 Presentation

A report was prepared outlining the process and the findings of the analysis. The selected files were copied to CD, along with the associated reports for each file, outlining creation and access dates, file details, and hash values. An extract from a sample examination report is included in Appendix B.

7.4.13 Feedback

The examiner was concerned that Suspect may have used other cloud storage services to information about other drug importation and sales, and communicated this with IO Jones. IO Jones interviewed Suspect, but Suspect would not disclose if other accounts existed. The examiner undertook additional searches for common cloud storage providers, such as Microsoft SkyDrive and Dropbox, identifying additional accounts held by Suspect. Further enquiries were undertaken with the legal sections of the identified service providers in relation to the identified user account files and data.

7.4.14 Complete

At the completion of the examination, the extracted data, notes and other material related to the investigation was appropriately stored in preparation for future legal action.
7.5 Google Drive: Summary

When investigating the storage of data using cloud service providers, the initial stages of an investigation include the identification of a cloud service and user account details. This will enable examiners to identify the potential location of data, and act to secure this data in a timely manner. In the research, it was found that an examiner can identify Google Drive account use by undertaking keyword searches and examine common file locations to locate relevant information (as detailed in Table 36).

It is of great interest (and also a security concern) that a Google Drive account password can in some instances be located as plain text stored on a hard drive or in memory captures or pagefile.sys files. It was also possible to access a Google Drive account without knowing the username or password, by using Virtual Forensic Computing software to create a Virtual Computer from a forensic image which previously had the Google Drive client software installed and synchronised with an account. By running the forensic image as a Virtual Computer, and selecting the ‘visit Google Drive on the web’ option from the client software, it was possible to connect to the account without entering the username or password. There are a wide range of investigation points for an examiner to determine the use of Google Drive, such as; directory listings, prefetch files, link files, thumbnails, registry, browser history, and memory captures.

Information was able to be located on an Apple iPhone to identify the use of Google Drive, including the username, and the filenames for the files accessed. It has been identified that conducting research into physical extracts of data from an iPhone would be of relevance, and comparing this to the information obtained from a logical extract.

The use of the Digital Forensic Analysis Cycle was of assistance to ensure all aspects of an investigation were followed, from initially defining the scope to remain on-track, following the standard process of identifying, preserving, analysing, and reporting, with the ability to return to a previous step whilst analysis or other steps were progressing was of benefit. In addition, considering feedback was important to ensure any new processes or information was documented and reported, and the final step of completing the process ensured the data used throughout the process was stored appropriately. It is important to develop and use a consistent digital forensic framework, such as the one examined in this research, to ensure investigations are thorough, all issues are encompassed, and the process is flexible enough to apply in different situations.
In this chapter, the preservation step of the framework is explored. The task of Preservation follows after data has been identified, which was outlined in the previous chapters. This chapter examines how to preserve identified data when legislation allows for collection under terms of a warrant, and two methods an investigator could undertake for this process are outlined; including using browser access or using client software.

Material presented in this chapter was submitted for publication and is currently under peer review:


Chapter 8 - Forensic Collection of Cloud Storage Data

8.1 Introduction

Once identification of data in a cloud storage account is made, investigators have legal recourse to secure data via service providers. This process can potentially be problematic due to delays or jurisdictional issues. As a hypothetical example, if an investigator in Australia discovers a suspect has data stored with a cloud storage
provider based in the United States of America (USA), and the data is spread across multiple countries in the company’s storage facilities, the process of securing and gaining access to the data relies upon communication and legal processes across the various countries, with different legal system requirements having to be met along the way. Some jurisdictions have legal provisions which allow for data to be secured at the time of serving a warrant, such as search and seizure.

In Australia, Section 3LA of the *Crimes Act 1914* (Cth), has a provision for the executing officer of a search and seizure warrant to access data which includes data not held at the premises, i.e. accessible from a computer or data storage device (Appendix A ‘Legislation’ lists the Acts referred to in this section). This may not be the case in a number of jurisdictions, including the USA.

In South Australia, Section 176 of the *Criminal Assets Confiscation Act 2005* (SA) subsection 1(a) allows for an order with provisions for an authorized officer to access data accessible from a computer that is on the premises, and to make a copy of the data. Section 40Q of the *Road Traffic Act 1961* (SA) subsection 5(d) provides the ‘power to access or download information that is required to be kept under an Australian road law or under an approved road transport compliance scheme’ which is accessible from equipment located in the vehicle. In addition, Section 67 of the *Summary Offences Act 1953* (SA) relates to a General Search Warrant, subsections 4(c) and 4(b)(iii) provides the power for an officer to ‘seize any such goods or things to be dealt with according to law’ where ‘there is anything that may afford evidence as to the commission of an offence’

In New South Wales, the *Law Enforcement (Powers and Responsibilities) Act 2002* (NSW), Section 75B relates to the police powers in relation to ‘Access to and downloading of data from computers (including access to computers outside premises the subject of a warrant)’. The full section is listed in Appendix A – Legislation.

As stated by Hansen (2012), discussing legislation in Norway, according to Sunde (2006), ‘Norwegian law enforcement may access electronically stored data in the same way as the account owner legally could, as long as there is a valid search warrant, and as long as the user name and access codes are available.’
Current methods of collecting data from cloud storage accounts includes via a browser or using client software to download data. In the following chapter, the legal process available for three popular cloud storage providers; Dropbox, Google Drive, and Microsoft SkyDrive, is discussed. This relates to the collection of data when there is no power or authority under legislation to access the data. In this case, contact with the legal officers in a service provider would be the first step to secure and preserve data, pending legal action through the relevant court jurisdiction to access the data. Research was then undertaken exploring the process of accessing and downloading data using a browser, and then installing and using client software for each service, which is applicable for those jurisdictions with the authority to access remote data. The two access and download methods are then compared to determine the appropriate forensic preservation method and the findings are summarised.

8.2 Cloud Service Provider Legal Contact Points

8.2.1 Dropbox

Dropbox is a file hosting service that enables users to store and share files and folders. There is a free service which allows 2 GB of data storage, and additional storage space can be acquired by referring or signing up new users or subscribe to the paid service of up to 100GB for USD 200 per year. Dropbox can be accessed using a web browser (e.g. Internet Explorer) or client software. Dropbox client software is available for Microsoft Windows operating system (OS), Apple Mac OSX, Linux, Apple iOS, Android, Blackberry and Windows Phone devices.

Dropbox Inc. is an American company based in San Francisco, California. The legal requirements of the user agreement, the State of California, or the federal provisions in the United States of America potentially must be complied with to preserve and access data related to an investigation. Information on the Dropbox contact web page lists the corporate legal team which can be contacted via the email address; legal@dropbox.com (https://www.dropbox.com/contact/).
8.2.2 Google Drive

Google Drive is a recent offering from Google Inc., and has evolved from Google Docs. By default, users have 5GB of data storage when creating an account. Additional storage ranging from 25GB to 16TB is available as a paid service. Google Drive can be accessed using a web browser or client software. Google Drive client software is available for installation on a computer or portable devices such as an Apple iPhone. Google, Inc. is an American company with headquarters at 1600 Amphitheatre Parkway, Mountain View, California (http://www.google.com/contact/). Google Inc. also has offices throughout the world.

8.2.3 Microsoft SkyDrive

Microsoft® SkyDrive® offers a potential 25GB of free storage to users, and was previously known as Windows Live Folders and Windows Live SkyDrive. SkyDrive can be accessed using a web browser or client software. SkyDrive client software is available for installation on a computer or portable devices such as an Apple iPhone. Microsoft is a global company, with headquarters in Redmond, Washington (http://www.microsoft.com/worldwide/).

8.3 Data Collection via Internet Access to a User Account

For those jurisdictions with the provision to collect data via accessing an account when serving a warrant (such as Australia) there are two current options available; via browser login and via client software synchronisation. Forensic software for the collection of data stored in cloud storage accounts is currently not available. Cumulus Data have stated that a future upgrade to their software will include the ability to collect data from ‘Google Docs, cloud storage such as Dropbox and iCloud, and social media such as LinkedIn, Facebook, and Twitter’ (Roach 2012).

Dykstra and Sherman (2012) examine the process of securing evidence from a remote infrastructure-as-a-service (IAAS) computer running via Amazon EC2, but this relies upon the examiner being able to install an Encase Servlet or an FTK Agent. In the situation where identified data is stored within a cloud storage service (file-hosting), it
is not possible to make use of an Encase Servlet or an FTK Agent, as a user is only able to access data stored in the account, not to run software in the storage environment.

The methodology for the research in relation the preservation of data was outlined in Chapter Three, in the section relating to Research Question 2. This includes accessing an account (created for the purposes of this research) using a browser running in a virtual computer (VM) and recording the VM window using Microsoft Expression Encoder 4. Network traffic was captured using Wireshark running on the host computer. The data in each account was downloaded using a browser, and synchronised and downloaded using the client software. The VM computer was then paused and stopped. The VM files, video, and network captures were then preserved within logical evidence files (Encase L01, AccessData AD1, or X-Ways CTR files)

The process outlined was undertaken for Dropbox, Google Drive and Microsoft SkyDrive. Analysis was then conducted on the logical evidence files to determine any changes made during the process of downloading them from cloud storage using this methodology.

### 8.3.1 Dropbox

When accessing the Dropbox account using the Mozilla Firefox browser, the first webpage displayed a list of folders and files, with the type of file listed and a modified date, as shown in Figure 6.

![Figure 6 - Dropbox web page in Mozilla Firefox Browser](image)
Selecting a folder opened a web page with the files in the folder listed, displaying the file type and (last) modified date on the page. This information may be relevant to an investigation, and in this case was noted and also video recorded using Expression Encoder. Using the Right-Mouse-Button on a file or folder displayed an ‘options’ menu, including the ability to show previous versions of a file. Any previous versions should also be noted. Selecting a folder provided for an option to download the folder and its contents. This was undertaken and resulted in a ZIP file being downloaded to the default download location within the VM; in this case “C:\Users\[username]\Downloads\”. The ZIP file had the name of the folder, ‘Dataset.zip’.

Next the Dropbox client software was installed to the VM hard drive. The Dropbox (research) account was logged in to, and the software was observed to synchronise the contents of the Dropbox account to the default directory; “C:\Users\[username]\Dropbox\”. This process was also video recorded.

The process of downloading individual files via the browser was also tested to determine the effect on the file dates and times. It was observed for all files downloaded from the Dropbox account, the file timestamps (modified, written, accessed, and created) were in fact the date and time of the download (rather than the other timestamps outlined in Table 39).

As per the process listed in Chapter 3, Research Question 2; Data Preservation steps for Cloud Storage (Steps 11 and 12), the files were preserved in a logical evidence file using Encase 6.19.4. Analysis was conducted on the logical evidence file, with the original source files loaded into the case as ‘single files’ for comparison. It was noted that the hash values matched; hence the contents of the files were not changed during the process of uploading, storage and downloading from the Dropbox account.

Examining the file dates and times for the files downloaded via the browser, the following differences were observed; the File Created time and the Last Accessed time of the downloaded files were the same as the original files Last Written time without the time zone offset (i.e. in UTC). The Last Written time and the Entry Modified time of the downloaded files were the time it was unzipped. None of the file dates and times matched the original file date and time information (Table 39).
When examining the files downloaded via the Dropbox client software, the following differences were observed; the File Created time, Last Accessed time and the Entry Modified time of the synchronised files were the time the file was synchronised and downloaded from the account (at the time of the research). Of note is that the Last Written time of the synchronised files was the same as the Last Written time from the original file (Table 39).

Analysis of the logical evidence file relating to the VM hard drive for log files relating to the client software located no files relevant to the synchronisation process. Previous experiments by Chung et al (2012) and McClain (2011) using Dropbox (Windows) client software (version 1.1.35) discussed config.db file and filecache.db files. These were reported to contain a history of filenames synchronized with Dropbox. In the Dropbox client software used for this research (version 1.2.52), these files do not contain a history of files. There was Config.dbx and Filecache.dbx files, both of which appear to be encrypted, but could potentially contain the file history information previously stored in the db files.

8.3.2 Google Drive

The process outlined in Chapter Three was then undertaken for Google Drive. When accessing the Google Drive account using the Mozilla Firefox browser, the first webpage displayed a list of folders, with the owner and a last modified date, as shown in Figure 7.
Selecting a folder on the page opened a web page with the files in the folder listed, including the owner and last modified date for each file. Selecting the down arrow next to ‘OWNER’ displayed a drop-down list to choose the last edited or last opened dates. This information may be relevant to an investigation, and in this case was also video recorded. Using the Right-Mouse-Button on a file or folder displayed an ‘options’ menu, including the ability to show previous versions of a file (manage revisions). Previous versions may provide relevant information. Selecting a folder provided for an option to download the folder and its contents. When this was chosen a ZIP file was downloaded to the default download location within the VM; in this case “C:\Users\[username]\Downloads\”. The ZIP file had the name; ‘documents-export-YYYY-MM-DD.zip’.

Next the Google Drive client software was installed to the VM hard drive. The Google Drive (research) account was logged in to, and the software was observed to synchronise the contents of the Google Drive account to the default directory; “C:\Users\[username]\Google Drive\”.

The process of downloading individual files via the browser was tested to determine the effect on the file dates and times. It was observed for all files downloaded that the file timestamps (modified, written, accessed, and created) were the date and time the file was downloaded.
As per the process listed in Chapter 3, Research Question 2; Data Preservation steps for Cloud Storage (Steps 11 and 12), the files were preserved in a logical evidence file using Encase 6.19.4. Analysis was conducted on the logical evidence file, with the original files loaded into the case as ‘single files’ for comparison. It was noted that the hash values matched; hence the contents of the files were not changed during the process of uploading, storage and downloading.

Examining the file dates and times for the files downloaded via the browser, the following differences were observed; when using a browser to download a file in a ZIP file, the File Created time and the Last Accessed time of the downloaded files was ‘1/1/1980’. The Last Written time and the Entry Modified time of the downloaded files was the time it was unzipped. None of the file dates and times matched the original file date and time information (Table 39).

When examining the files downloaded via the Google Drive client software, the following differences were observed; the File Created time and the Entry Modified time of the synchronised files was the time the file was synchronised and downloaded from the account. The Last Written time of the synchronised files was the same as the Last Written time of the original file, and the Last Accessed time of the synchronised files was same as the Last Written time from the original file (which was different to the Last Accessed time of the original file).

Analysis of the VM hard drive for log files relating to the Google Drive client software located a ‘Snapshot.db’ file at ‘C:\Users\[username]\AppData\Local\Google\Drive\’. This file was a SQLite database and the file was viewed with SQLite Manager (Mozilla Plugin). Within the SQLite file were details of the files synchronised, including the filename, the creation time, and modified times (Table 37). The creation and modified time information appears to be in Unix Numeric Value time format, which can be decoded using conversion software such as Digital Detective’s DCode v4.02a. For example, the file named ‘3111.txt’ has a Modified Time of ‘1100097650’. This is converted using DCode to ‘Wed, 10 November 2004 14:40:50 UTC’, which is the Last Written (modified) time for the original Enron 3111.txt file. The Created time appeared to be the time the file was originally uploaded to the Google Drive account. The file size is also listed, and the MD5 Checksum value for the file (which matched the original file and the downloaded file).
Table 37 - Google Drive ‘snapshot.db’ file contents (abbreviated)

<table>
<thead>
<tr>
<th>Filename</th>
<th>Modified</th>
<th>Created</th>
<th>Size</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>root</td>
<td>1339308961</td>
<td>1339309024</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Dataset</td>
<td>1341041780</td>
<td>1341041780</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Untitled document</td>
<td>1345973985</td>
<td>1341041868</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>.3111.txt</td>
<td>1100097650</td>
<td>1339309029</td>
<td>2734</td>
<td>f3a862114dceef80cc3eb0429b4f80a62</td>
</tr>
<tr>
<td>Enron3111.rtf</td>
<td>1330133578</td>
<td>1339309056</td>
<td>34215</td>
<td>bc99a5bf44317ae105af80d8201e831</td>
</tr>
<tr>
<td>Enron3111.docx</td>
<td>1328766856</td>
<td>1339309030</td>
<td>14072</td>
<td>3e42e0d132d27db780087925f3700fb</td>
</tr>
<tr>
<td>Enron3111.jpg</td>
<td>1328766794</td>
<td>1339309046</td>
<td>315868</td>
<td>77638319ea64cc1b70d4d4f20a56295d</td>
</tr>
<tr>
<td>enron_with_categories.tar.gz</td>
<td>1308027150</td>
<td>1339309153</td>
<td>4523350</td>
<td>f880a35b54bf9f4d18d1eede8da6f179</td>
</tr>
</tbody>
</table>

8.3.3 Microsoft SkyDrive

The process outlined in Chapter Three was then undertaken for Microsoft SkyDrive.
When accessing the SkyDrive account using the Mozilla Firefox browser, the first webpage displayed folders graphically as squares, as shown in Figure 8.

![Figure 8 - Microsoft SkyDrive web page in Mozilla Firefox Browser](image)

Selecting a folder (square) opened a web page with the files in the folder listed with the file information listed on the right hand side of the browser page for the file selected. The information displayed included whether the file is being shared, the file type,
modified date and username, added date and username, the path, and the file size. This information may be relevant to an investigation. Using the Right-Mouse-Button on a folder displayed a drop down menu, including the ability to show properties of the folder and an option to download the folder and its contents. This was undertaken and a ZIP file was downloaded to the default download location within the VM; in this case “C:\Users\[username]\Downloads\”. The ZIP file had the name of the folder, in this case, Dataset.zip.

Next the Microsoft SkyDrive client software was installed to the VM hard drive. The SkyDrive (research) account was logged in to, and the software was observed to synchronise and download the contents of the SkyDrive account to the default directory; “C:\Users\[username]\SkyDrive\”.

The process of downloading individual files via the browser was tested to determine the effect on the file dates and times. It was observed for all files downloaded that the timestamps (modified, written, accessed, and created) were the date and time of the download.

As per the process listed in Chapter 3, Research Question 2; Data Preservation steps for Cloud Storage (Steps 11 and 12), the files were preserved in a logical evidence file using Encase 6.19.4. Analysis was conducted on the logical evidence file, with the original files loaded into the case as ‘single files’ for comparison. It was noted that the hash values matched; hence the contents of the files were not changed during the process of uploading, storage and downloading.

Examining the file dates and times for the files downloaded via the browser, the following differences were observed; when using a browser to download a file in a ZIP file, the File Created time and the Last Accessed time of the downloaded and unzipped files were the time the file had been originally uploaded to the SkyDrive account. The Last Written time and the Entry Modified time of the downloaded files were the time the file was unzipped. None of the file dates and timesmatched the original file date and time information (Table 39).

When examining the files downloaded via the client software, the following differences were observed; the File Created time, Last Accessed time and the Entry Modified time of the downloaded files were the time the file was downloaded from the account. Of
note was that the Last Written time of the downloaded files was the same as the Last Written time of the original file (Table 39).

Analysis of the VM hard drive for log files relating to the SkyDrive client software located a folder at ‘C:\Users\[username]\AppData\Local\Microsoft\SkyDrive\logs\’ which contained a ‘SyncDiagnostics.log’ file which has a list of the files synchronised, sizes, dates and times (an example is listed in Table 38). There were two sections located within the file, ‘Cloud Metadata’ and ‘File System’. ‘Cloud Metadata’ appeared to be information about the files stored in the SkyDrive cloud account. ‘File System’ appeared to be related to the files stored on the VM hard drive. Analysis of the SyncDiagnostics.log file located a range of information of interest. File system and metadata information was located within the log file. The SkyDrive OwnerID is listed for each file, and the associated file or folder number. The filename and path is listed, and the creation and modified times. The creationTime and modTime information appears to be in Unix Numeric Value time format, which can be decoded using conversion software such as Digital Detective’s DCode v4.02a. For example, the file named ‘3111.txt’ has a modTime of ‘1100097650’. This is converted using DCode to ‘Wed, 10 November 2004 14:40:50 UTC’, which is the Last Written (modified) time for the original Enron 3111.txt file.

---

Metadata:

folder XXXXE4BB87F707D0!106 'C:\Users\[username]\SkyDrive\Dataset',
creationTime=1338104478, modTime=1338104478
file XXXXE4BB87F707D0!107 'C:\Users\[username]\SkyDrive\Dataset\3111.txt', size=2734,
creationTime=1338104500, modTime=1100097650
file XXXXE4BB87F707D0!109 'C:\Users\[username]\SkyDrive\Dataset\Enron3111.docx',
  size=14072,
  creationTime=1338104500, modTime=1328766856
file XXXXE4BB87F707D0!110 'C:\Users\[username]\SkyDrive\Dataset\Enron3111.jpg',
  size=315868,
  creationTime=1338104500, modTime=1328766794
folder XXXXE4BB87F707D0!103 'C:\Users\[username]\SkyDrive\Public',
creationTime=1335323577, modTime=1335323577

Cloud Total: 4 folders, 4 files, 5573071 bytes

File System:

Scanning 'C:\Users\[username]\SkyDrive'
  file 'C:\Users\[username]\SkyDrive\.lock' ignored, size=0,
  creationTime=1345976968, modTime=1345976968
  - folder 'C:\Users\[username]\SkyDrive\Dataset'
    - file 'C:\Users\[username]\SkyDrive\desktop.ini' ignored, size=96,
      creationTime=1345976888, modTime=1346228278
  Folder 'C:\Users\[username]\SkyDrive\Dataset' Total: 4 folders, 1 file
Scanning 'C:\Users\[username]\SkyDrive\Dataset'
  - file 'C:\Users\[username]\SkyDrive\Dataset\3111.txt', size=2734,
    creationTime=1345976974, modTime=1100097650
  file 'C:\Users\[username]\SkyDrive\Dataset\Enron3111.docx', size=14072,
    creationTime=1345976993, modTime=1328766856
  - file 'C:\Users\[username]\SkyDrive\Dataset\Enron3111.jpg', size=315868,
    creationTime=1345976993, modTime=1328766794
Folder 'C:\Users\[username]\SkyDrive\Dataset' Total: 0 folders, 4 files
FileSystem Total: 4 folders, 5 files, 0 symLinks, 5573071 bytes

---

Table 38 - Microsoft SkyDrive ‘syncdiagnostics.log’ file contents (abbreviated)
Within the VM hard drive was a data file with the name matching the Owner ID number of the SkyDrive account, such as 'XXXXA4AB89F707D0.dat'. This file also included the file and folder names of the files and folders synchronised with the SkyDrive account. The file was located at;

   "C:\Users\[username]\AppData\Local\Microsoft\SkyDrive\settings\"

### 8.4 Data Collection: Analysis

Of note to a forensic examiner is that analysis of the hash values for the files revealed no changes were made to the data during the process of uploading, storage, and then downloading. The method of downloading did not affect the contents of the files, as the hash values were the same.

Analysis of the resulting data from the three cloud storage providers determined that different dates and times were associated with the files depending on which service was used, and which method of downloading was used, as outlined in Table 39. The only instance when the downloaded file times matched the original file information was when using the client software the Last Written time remained the same as the original file (listed in Table 39 as ‘same’). Listed in Table 39 as ‘download time’ are the instances when the downloaded file had the time of download as the timestamp, for example; accessing Dropbox or SkyDrive and downloading a file using the client software results in the times for Last Accessed, File Created and Entry Modified being the time the file was downloaded. When accessing Dropbox using a browser, the Last Accessed time and File Created time were the Last Written time from the original file in UTC (in this case the time zone was Australian Central Standard Time (+9:30)), and the created and modified times were minus 9:30 compared to the original file Last Written time.

When accessing Google Drive using a browser, the Last Accessed time and File Created time of the downloaded files were ‘1/1/1980’. The Last Written time and Entry Modified time of the downloaded files was the time the file was unzipped from the ZIP file.
When accessing Microsoft SkyDrive using a browser, the Last Accessed time and File Created time of the downloaded files appeared to be the time and date when the file was originally uploaded to the SkyDrive cloud storage account. The Last Written time and the Entry Modified time of the downloaded files was the time the file was unzipped from the ZIP file, which is consistent with Dropbox and Google Drive.

<table>
<thead>
<tr>
<th></th>
<th>Last Accessed</th>
<th>File Created</th>
<th>Last Written</th>
<th>Entry Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dropbox</strong></td>
<td>browser</td>
<td>Last Written (UTC)</td>
<td>Last Written (UTC)</td>
<td>unZIP time</td>
</tr>
<tr>
<td></td>
<td>client</td>
<td>download time</td>
<td>download time</td>
<td>same</td>
</tr>
<tr>
<td><strong>Google Drive</strong></td>
<td>browser</td>
<td>1/01/1980</td>
<td>1/01/1980</td>
<td>unZIP time</td>
</tr>
<tr>
<td></td>
<td>client</td>
<td>last written</td>
<td>download time</td>
<td>same</td>
</tr>
<tr>
<td><strong>SkyDrive</strong></td>
<td>browser</td>
<td>upload date/time</td>
<td>upload date/time</td>
<td>unZIP time</td>
</tr>
<tr>
<td></td>
<td>client</td>
<td>download time</td>
<td>download time</td>
<td>same</td>
</tr>
</tbody>
</table>

Table 39 - File Dates and Times for Dropbox, Google Drive, and Microsoft SkyDrive downloaded files using a Browser compared with using Client Software (Encase 6.19.4).

Analysis was also conducted with X-Ways 16.5 and AccessData Forensic Toolkit 1.81.6 (demo version), and differences were noticed with the terms used for the file timestamps in comparison with Encase 6.19.4. Comparing the files created for this research it appeared that where Encase listed ‘Last Accessed’, ‘File Created’, and ‘Last Written’; X-Ways and FTK listed these as ‘accessed’, ‘created’, and ‘modified’ (as displayed in Table 40 and Table 41).

<table>
<thead>
<tr>
<th></th>
<th>accessed</th>
<th>created</th>
<th>modified</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dropbox</strong></td>
<td>browser</td>
<td>File modified time (UTC)</td>
<td>File modified time (UTC)</td>
</tr>
<tr>
<td></td>
<td>client</td>
<td>download time</td>
<td>download time</td>
</tr>
<tr>
<td><strong>Google Drive</strong></td>
<td>browser</td>
<td>1/01/1980</td>
<td>1/01/1980</td>
</tr>
<tr>
<td></td>
<td>client</td>
<td>File modified time</td>
<td>download time</td>
</tr>
<tr>
<td><strong>SkyDrive</strong></td>
<td>browser</td>
<td>Date/time uploaded</td>
<td>Date/time uploaded</td>
</tr>
<tr>
<td></td>
<td>client</td>
<td>download time</td>
<td>download time</td>
</tr>
</tbody>
</table>

Table 40 - File Dates and Times for Dropbox, Google Drive, and Microsoft SkyDrive downloaded files using a Browser compared with using Client Software (X-Ways 16.5 and FTK 1.81.6)
Using Red Wolf Computer Forensics MFT Parser to examine the MFT records from the VM hard drive showed the ‘Record Change Date’ had the same data as the Encase ‘Entry Modified’ information. This was verified with FTK Imager 3.1.0, which displayed the Record Date in the Properties tab (the nomenclature observed for file dates and times for the various forensic programs are listed in Table 41).

<table>
<thead>
<tr>
<th>MFT Parser</th>
<th>Last Access Date</th>
<th>Creation Date</th>
<th>Modified Date</th>
<th>Record Change Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTK Imager</td>
<td>Date Accessed</td>
<td>Date Created</td>
<td>Date Modified</td>
<td>Record Date</td>
</tr>
<tr>
<td>FTK 1.81</td>
<td>Acc Date</td>
<td>Cr Date</td>
<td>Mod Date</td>
<td></td>
</tr>
<tr>
<td>X-Ways</td>
<td>Accessed</td>
<td>Created</td>
<td>Modified</td>
<td></td>
</tr>
<tr>
<td>Encase</td>
<td>Last Accessed</td>
<td>File Created</td>
<td>Last Written</td>
<td>Entry Modified</td>
</tr>
</tbody>
</table>

Table 41 - File Date and Time nomenclature: MFT Parser, FTK, X Ways, and Encase.

As part of the research, the process of downloading individual files via the browser was tested to determine the effect on the file dates and times. It was observed for all files downloaded as a single file from the three cloud service providers that the file timestamps (modified, written, accessed, and created) were the date and time of the download.

It is important to note that the contents of the files were not altered during the process of upload, storage and download. The nature of cloud storage is that files can be modified whilst stored in the account, by the owner and by others when shared. Google Docs and other providers allow for collaboration in relation to documents and other files, which is a feature of the service. Whilst the contents of the files were not altered during this research, it is possible in other instances for the contents of the files to be altered whilst stored in the account, which could result in differences between the original file when it was uploaded to an account, and the subsequent file downloaded from the account, depending on whether any changes were made whilst it was in cloud storage.

It is also important to understand the different timestamps associated with the files, depending on the method of access and download. As has been discussed, timestamps are easily altered and different in comparison with the original file, which should be
considered when forming a conclusion with regard to the relevance of a timestamp. For example, a conclusion reached about a timestamp may be incorrect if the circumstances are not apparent. Relying on a timestamp at face value may result in an incorrect assumption. This can have implications in an investigation and legal proceedings, as it could provide an alibi for a defendant to show there was no way they could create, modify, or access a file as they were in a meeting with colleagues.

The actual times in relation to an event could be different to what is assumed if the altered timestamps from cloud storage have not been considered. This could also have implications when conducting analysis on hard drives when a user has synchronised or downloaded files from an account prior to seizure, rather than copying a file via USB, or other transfer method. $MFT$ analysis and searching a hard drive for filenames may provide indications for cloud storage use, as will analysis in relation to the cloud service client software files previously mentioned in relation to usage logs (such as the Microsoft SkyDrive ‘syncdiagnostics.log’ file).

Forensic examiners are recommended to conduct their own tests to verify these findings, as new releases of client software may alter the way the files are downloaded in future, and cloud providers may alter the date and time information associated with files, and hence what information is downloaded from browser access to an account.

### 8.5 Data Collection: Summary

Current methodologies for forensic examiners to manually collect and preserve data from cloud storage services include accessing an account via a browser, or using client software provided by the service provider. This chapter examined the collection of data via these methods and determined that the file contents were not altered during the process of uploading, storing, and downloading files using Dropbox, Microsoft SkyDrive, and Google Drive. Also of note is that the associated file timestamps were different to those of the original file, and varied depending on the process undertaken and the service used. This can have implications in an investigation if incorrect assumptions are made based on the timestamp information, if the original method of obtaining the file is not considered. Relying on incorrect information could provide an alibi to a defendant which may be inaccurate.
The conclusion reached is that it is appropriate to use both methods; browser and client software, as the information and data from both methods provides a greater picture of the circumstances then either method alone would. There was additional information located within the actual accounts, and hence the preservation process should include accessing the account via a browser and recording the information associated with an account and each file. During this research, this was achieved using Microsoft Windows Expression Encoder 4 to record the VM screen.

It is also concluded that the process of preservation should be undertaken from a clean installation of Windows in a virtual machine. The account should be accessed from a browser first, the contents noted, and then the files downloaded as a zip file. The client software should then be installed and used to synchronise the files to the VM hard drive, as the dates and times associated with the synchronised file are different to the dates and times from the browser download, and include the modified date and time from the original file. The tables in this chapter outline the various time and date information for the cloud storage accounts being accessed by the browser and by the client software.
9 Summary

This chapter summarises the thesis and research. The first section is a summary of the chapters. Next, the research is outlined and the questions and hypotheses are listed, detailing the manner in which the questions are answered. The validity, findings and implications of the research are discussed, and the overall conclusion finalises the chapter.

Chapter 9 – Summary

9.1 Summary of chapters

This thesis commenced with an introduction and followed with chapters that described the research in detail. The thesis also included an overall summary at the beginning, acknowledgements, table of contents, list of figures, tables, and a glossary of common terms and abbreviations.

Chapter one introduced the thesis topic, providing information regarding cloud storage to enable the reader to understand the background of the topic. This also included background information regarding digital forensic analysis and issues for investigators.
The motivation for conducting the research was outlined, and the overall objectives listed. The structure of the thesis was then explained.

Chapter two examined current cloud forensic analysis literature with a focus on cloud storage. The first section outlined cloud computing, digital forensic analysis, and cloud storage implications. Issues relating to identification, preservation, analysis, and presentation were then discussed. Additional issues were considered, and a summary concluded the chapter.

Chapter three explained the research methodology used in this thesis. The research problem, purpose, questions and hypotheses were outlined. The two research questions and associated hypotheses were discussed, and the research methodology for each question detailed. Finally, the limitations of the research were outlined.

Chapter four explained the proposed Digital Forensic Analysis Cycle and how this can be applied to forensic analysis of cloud storage. Each step of the cycle was explained; Commence (Scope), Prepare and Respond, Identify and Collect, Preserve (Forensic Copy), Analyse, Present, Feedback, and Complete. The cyclic nature of the framework was then outlined.

Chapters five, six, and seven described the application of the framework to the analysis of popular cloud storage services; Dropbox, Microsoft SkyDrive, and Google Drive. In each case, the data remnants on a Windows 7 computer were first examined using the proposed framework. Next an Apple iPhone was examined to further assess the framework and to determine the data remnants. A Case Study was then presented to outline the application of the research findings. A summary for each chapter highlighted the relevant information for each provider.

Chapter eight addressed the preservation of data which had been identified in the previous chapters. Firstly the forensic implications of accessing evidence were discussed, and then the process outlined. The research for each of the three popular cloud storage providers; Dropbox, Microsoft SkyDrive, and Google Drive, was outlined. This included findings in relation to accessing live data via a browser, and then using the client software of each provider, which were then compared for forensic suitability.
In the following sections, the original objectives are listed, and the method used to answer each question is explained. The findings in relation to the proposed framework, analysis of Dropbox, Microsoft SkyDrive, and Google Drive are listed. The research findings in relation to preservation of cloud stored data are then summarised. The benefits of the research are explained, and then areas for future research are summarised.

### 9.2 Research Objectives

The focus of this research is to discover cloud storage data remnants on popular client devices; a computer running Windows 7, and an Apple iPhone running iOS 4.2.1.

In the first chapter, the research objectives and deliverables were defined;

**Objective 1:** To examine current research published in literature relating to cloud storage and identified cloud storage analysis methodologies.

Chapter Two addressed this objective in detail, and concluded that; whilst there was some information directly relevant to answer some of the research sub questions, literature directly addressing the objectives was not located. Hence it was necessary to conduct primary research to explore the objectives.

**Objective 2:** To develop a digital forensic analysis framework that will assist practitioners, examiners, and researchers follow a standard process when undertaking forensic analysis of cloud storage services.

When commencing this research it was determined that there was a need to have a framework to guide the investigation process, and hence a framework was developed and assessed using the practical research to examine the benefits of using the framework when undertaking forensic analysis of cloud remnants. The proposed framework was outlined in Chapter Four, and built upon the guidelines for digital forensic analysis to provide for a process which follows a common digital forensic examination, enabling
Objective 3: To conduct research using popular cloud storage services; Dropbox, Microsoft SkyDrive, and Google Drive, to determine whether there are any data remnants which assist digital forensic analysis and investigations.

Chapters Five, Six and Seven covered the research of cloud storage services in depth. This research was undertaken using the framework developed for Objective Two. It was concluded that there a vast amount of data remnants resulting from the use of the examined cloud storage services, whether a browser is used, or client software is used, particularly on Windows 7 computer systems. Data remnants on an Apple iPhone were not as prevalent, but enough information was located to identify the cloud service used, and user and file information. Specific data sources in relation to this objective are outlined in the next section addressing Research Question One and the associated hypotheses.

The information that was located was shown to be applicable in law enforcement investigations. As explained in Chapter Five, the outcomes of this research proved to be...
beneficial for a current Police investigation. Whilst the investigation initially did not indicate cloud storage was potentially used, during the analysis process information was located which identified the use of Dropbox cloud storage to store illicit material. Unfortunately, the details of the investigation cannot be expanded as it is ongoing, but it is important to highlight the real-world application of this research is already achieving the goals of expanding the information and methodology available in forensic examinations to identify when cloud storage services are being used.

**Objective 4**: To examine the forensic implications of accessing and downloading cloud stored data from popular cloud storage services; Dropbox, Microsoft SkyDrive, and Google Drive.

Chapter Eight addressed this objective by explaining the research methodology and findings in relation to accessing cloud storage services, and the forensic implications associated with this. The specific findings are outlined in the next section, in answering Research Question Two.

### 9.3 Research findings

Analysis of the results of the literature review highlighted the issues in relation to forensic analysis of cloud storage as; whether there are any data remnants from cloud storage use, and what methods are available to preserve data stored in a cloud storage account.

The first primary research question was defined as:

\[ Q_1 \quad \text{What data remnants result from the use of cloud storage to identify its use?} \]

Chapters Five, Six and Seven covered the research of cloud storage services in depth. It was found that there were data remnants from cloud storage services on Windows 7 computer systems, and on the Apple iPhone 3G. These included the circumstances when the client software for the examined service was used, and also when the tested browsers were used. In addition, remnants were located when anti forensic processes
were undertaken, including erasing files, and cleaning browsing and file history information. No data remnants were located after a full erase was conducted though. In the research methodology chapter (Chapter Three), the first primary question was expanded into six sub questions, which are addressed further in this section with specific data remnants listed for each of the circumstances outlined in the sub questions. The findings for the sub questions serve to answer this main research question in depth.

\( Q_1 \) led to the following hypotheses;

\[ H_0 \] - There are no data remnants from cloud storage use to identify service the provider, username, or file details.

\[ H_1 – \text{There are remnants from cloud storage use which enable the identification of the service, a username, or file details.} \]

As outlined above, in this research experiment there were data remnants located after the cloud storage services examined were accessed. As a result, the second hypothesis (\( H_1 \)) is shown to be correct. After erasing software and cleaning software were used, \( H_1 \) was still correct, as there were data remnants able to be located to identify cloud storage use. If a full erase process was undertaken, the null hypotheses \( H_0 \) appears to be correct, but this would actually be an incorrect conclusion, as there were data remnants prior to the full erase, hence \( H_1 \) is still correct. The full erase process serves to require a further experiment and hypotheses in relation to whether data erasure removes the remnants from cloud storage use, and this research would conclude that the data remnants left by cloud storage use were able to be erased.

\( Q1a) \) What data remains on a Windows 7 computer hard drive after cloud storage client software is installed and used to upload and store data with each hosting provider.

The results of the analysis of the chosen cloud service providers are outlined in Chapters Five, Six and Seven. Focusing on the remnants when client software was used determines that it was possible to locate data remnants for each service. The existence of the software was able to be located by examining the $MFT in Windows 7 for the
service provider name; Dropbox, SkyDrive or Google Drive. Common locations were observed for the installation of client software, and were listed in the summary tables for each service, either in ‘AppData; Roaming or Local’, or in ‘Program Files’ locations. The username for each service was able to be located, either in database files on the hard drive, cookie files, or in memory captures. Of note was that the password for each service was able to be located, either in memory near common terms, or stored on the hard drive in plain text (representing a security risk).

Files of significance for each service are listed in the following table (Table 42) and are files an examiner should look for to locate user history in relation to cloud storage.

<table>
<thead>
<tr>
<th>Service</th>
<th>Files of significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropbox</td>
<td>Versions prior to October 2011: <code>config.db</code> and <code>filecache.db</code>. From version 1.2.48 the files <code>config.dbx</code> and <code>filecache.dbx</code> are encrypted.</td>
</tr>
<tr>
<td>Microsoft SkyDrive</td>
<td><code>SyncDiagnostics.log</code> and <code>OwnerID.dat</code></td>
</tr>
<tr>
<td>Google Drive</td>
<td><code>sync_config.db</code> and <code>snapshot.db</code></td>
</tr>
</tbody>
</table>

Table 42 - Client software log files and databases

Internet history extracts also provided relevant information in relation to the use of cloud storage, as did operating systems and software files such as; Registry, Prefetch, Link files, $MFT, $Journal, Thumbscache, and other files listed in the analysis. The term ‘enabledogfood’ was observed in the Microsoft SkyDrive client software, and could be used to locate the presence of this software as the term was not located in the control base image files. Keyword search terms are listed in Table 43, and are terms an examiner can use to locate usernames, passwords, and the common URLs for the cloud storage services.

<table>
<thead>
<tr>
<th>Service</th>
<th>Keyword search terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropbox</td>
<td>u‘email’: free name periods login_email</td>
</tr>
<tr>
<td>Microsoft SkyDrive</td>
<td>&amp;login= &amp;passwd= enabledogfood</td>
</tr>
<tr>
<td>Google Drive</td>
<td>&amp;email= &amp;passwd= &amp;passwdagain=</td>
</tr>
</tbody>
</table>

Table 43 - Keyword search terms
Q1b) What data remains on a Windows 7 computer hard drive after cloud storage services are accessed via a web browser with each hosting provider?

When considering the data remnants after cloud storage accounts were accessed via the four browsers, it was possible to determine the username for the account from Internet Explorer, Mozilla Firefox and Google Chrome; Temporary Internet Files, ‘formhistory.sqlite’, and ‘Web Data’ files, and Apple Safari in Pagefile.sys. The username was also located in the memory captures. With SkyDrive, the OwnerID was observed in the URL for the skydrive account access. The password was not observed in any of the VM’s. The URL for the cloud storage accounts was observed in Internet history extracts, and the files and file contents were able to be located on the hard drive in a range of locations, such as System Volume Information and Unallocated space.

Q1c) What data is observed in network traffic when client software or browser access is undertaken?

Analysis of the network PCAP capture files was able to determine the common URL’s and IP’s used for each of the cloud storage services examined, which are listed in each chapter. The analysis was able to determine that Dropbox is hosted on the Amazon AWS service. All the cloud storage services utilised the Verisign/Thawte digital certificates. With Dropbox, there were different URL’s observed when using the client software in comparison with when using a browser; ‘www’ or ‘web’ was observed when a browser was used, and ‘client’ was observed when the software was used.

The username was observed in cleartext when Google Drive was accessed with the Internet Explorer browser, but not with the other browsers. No file contents were observed in the network traffic, indicating the data is encrypted. The filenames were able to be read in cleartext in the network traffic when SkyDrive was accessed, which highlights the importance of capturing network traffic when possible, and also presents a security issue, as it is expected that this data would be encrypted.
Q1d) What data remains in memory when client software or browser access is undertaken?

Analysis of the memory files was able to locate a wide range of information. Of note was the ability to locate the password for all the services in freetext. These were able to be located near common terms; ‘free name periods’ for Dropbox, ‘&passwd=’ for SkyDrive, and ‘&passwd=’ or ‘&passwdagain=’ for Google Drive. The username was also able to be located in memory, near the text; ‘u’email’:’ for Dropbox, ‘&login=’ for SkyDrive, and ‘? –Email’, ‘?? Email’ or ‘<email>’ for Google Drive. Filename references and the contents of the test files were able to be located in memory captures, including when anti-forensic procedures had been undertaken. Some information was also able to be located in pagefile.sys and System Volume Information remnants.

Q1e) What data remains on an Apple iPhone 3G after cloud storage services are accessed via a web browser with each hosting provider?

Analysis of the extracted data from the Apple iPhone 3G for Dropbox was not able to locate the username or password for the account. In relation to SkyDrive, the OwnerID was located in a URL with ‘https://skydrive.live.com’ in the ‘history.plist’ file, and also in a cookies file. With Google Drive, the username was located in the file ‘cookies.binarycookies.’ The contents of the test files were not able to be located in the extracted data. Filenames of the test data was able to be located in the history extracts for all the cloud services, in the history.plist file and the XRY Web-History.txt file.

Q1f) What data remains on an Apple iPhone 3G after cloud storage services are accessed via an installed application from each hosting provider?

The username for the Dropbox and the SkyDrive accounts were able to be extracted when the applications was used on the Apple iPhone 3G. With Dropbox the username was located in the ‘com.getdropbox.Dropbox.plist’ file near the text; ‘=]DBAccountInfo’ and in the ‘keychain-backup.plist’ file. For SkyDrive the username was located in the ‘keychain-backup.plist’ near the text; useridentity, windows live, and ‘.com.microsoft.skydrive’. This information appears to be parsed by XRY in the
Information in iTunes sqlite file indicated the Dropbox application was present, and the 'com.apple.AppStore.plist' file lists ‘http://itunes.apple.com/au/app/skydrive’ application. No file contents or filenames were located in the extract. The application for Google Drive was unable to be installed on the test device, and hence was not included in the scope of the research.

The inability to extract the file contents in relation to the Dropbox application supports the findings of Zhu (2011), who was also unable to extract the file contents from the Dropbox application.

The second primary research question was defined as:

\[ Q_2 \text{ What forensically sound methods are available to preserve data stored in a cloud storage account?} \]

Chapter Eight detailed the research of preservation of data from cloud storage services in depth. It was found that the process of uploading then downloading the test files did not alter the data within the files, as the MD5 hash values of the files remained the same. The file date and time stamps were, however different to the original test files, and this needs to be considered when forming conclusions in relation to the times and dates of the files within cloud storage accounts. There was additional information located within the actual accounts, and hence the preservation process should include accessing the account via a browser and noting the information associated with an account, and also associated with each file. This process can be recorded using software such as Microsoft Windows Expression Encoder 4. It is concluded that the process of preservation should be undertaken from a clean installation of Windows in a virtual machine, which is recorded. The account should be accessed from a browser first, the contents noted, and then the files downloaded as a zip file. The client software should then be installed and used to synchronise the files to the VM hard drive, as the dates and times associated with the synchronised file are different to the dates and times from the browser download, and include the modified date and time from the original file. The tables in Chapter Eight outline the various time and date information for the cloud storage accounts being accessed by the browser and by the client software.
$Q_2$ led to the following hypotheses;

$H_0$  the process of downloading files from cloud storage does not alter the internal data or the associated file metadata.

$H_1$  the process of downloading files from cloud storage alters the internal file data and the associated file metadata.

$H_2$  the process of downloading files from cloud storage does not alter the internal data, but does alter the file metadata.

$H_3$  the process of downloading files from cloud storage alters the internal data, but not the associated file metadata.

As outlined above, the process of downloading files did not alter the contents of the files; hence $H_1$ and $H_3$ were not valid, leaving the null hypotheses ($H_0$) and $H_2$ to be considered. As the file date and time associated with the downloaded files was altered from the original file, $H_0$ was not valid. This leaves $H_2$ as the valid hypotheses. The information observed in the time and date stamps after the files were downloaded from the cloud storage accounts proved to be useful for an investigation, even though the data is different to the original file data, as the different timestamps associated with the different methods of downloading the files provides additional information which may be relevant to an investigation.

$Q_2$ a

What data can be acquired and preserved from a cloud storage account using existing forensic tools, methodologies, and procedures when applied to cloud storage investigations?

As stated in Chapter Eight, there is currently no software available to preserve data from cloud storage accounts; Dropbox, Microsoft SkyDrive, and Google Drive. The literature review was also unable to locate any methodology or procedures specifically addressing the preservation of data from cloud storage accounts. Dykstra and Sherman (2012) researched the process of preserving data from remote IaaS systems via Amazon EC2,
but this is not applicable for file storage as the process involves running software within the cloud environment, which is not possible with Dropbox, SkyDrive or Google Drive.

Hence there are no existing forensic tools, methodologies or procedures associated with preserving data from cloud storage accounts. Abiding by the principles of digital forensic acquisition (SAI 2003; ACPO 2006; NIJ 2008), the methodology outlined in Chapter Three for Research Question 2 was prepared and used in the process outlined in Chapter Eight. This served to outline and test a methodology to preserve data from cloud storage accounts in a forensically sound manner, resulting in the data from the test accounts being able to be acquired and preserved. The data consisted of the file contents, and a range of associated file information, such as; size, hash values, various time and date information, and owner, user and modifier details. This process applies for those jurisdictions with the legal provision to collect data via accessing an account when under a warrant (such as Australia and Norway), and will vary amongst different jurisdictions. Legal processes for each jurisdiction would need to be met prior to undertaking any acquisition or preservation processes.

9.4 Research validity

This research was undertaken using popular web browsers, operating systems, software and processes. At a general level, results indicate there are data remnants which enable the identification of cloud storage accounts, user details, and file information. Data can be preserved from cloud storage accounts without modifying the hosted data, however changes were observed in relation to the date and time information associated with files. Control media was used to ensure the findings were based on newly installed and known media, hence the changes that were made were observed.

9.5 Future research opportunities

It has been identified that conducting research into physical extracts of data from an iPhone would be of relevance, when compared to the information obtained from a logical extract. In addition, future research opportunities include conducting research in
to the remnants of other cloud storage services such as Amazon S3, iCloud, and UbuntuOne. This will assist to determine data remnants from other cloud service providers and encompass a methodology that can be used to identify cloud storage providers and cloud computing providers. This should also encompass any future developments in the field of remote storage, and developing a consistent digital forensic framework, such as the one proposed and examined in this research.

9.6 Conclusion

When investigating the storage of data using cloud service providers, the initial stages of an investigation will be to identify the service provider and particular user account details. This will enable examiners to identify the potential location of data, and act to secure this data in a timely manner. It was demonstrated that an examiner can identify cloud storage account use by undertaking keyword searches, hash comparison, and examine common files and file locations to locate relevant information. Cloud storage username details can be determined from forensic images.

Of great interest is that account passwords can, in some instances, be located in plain text in memory captures and stored on a hard drive. In addition, it was possible to gain full access to a Google Drive account when the client software was installed, by using the ‘visit Google Drive on the web’ option from the icon in the System Tray, without knowing the username or password. In previous versions of Dropbox it was possible to copy the software files to another computer and use them to synchronise to an account, however, in the versions examined in this research, this was not possible. In addition, it was possible to gain full access to a Dropbox account through the client software from an icon in the System Tray; ‘launch Dropbox website’ without requiring a username or password. With SkyDrive, it was possible to synchronise to an account if the client software was installed on a PC with an associated username and password by running the PC within a virtual machine connected to the Internet. Attempting to connect to the user account with the Microsoft SkyDrive client software system tray icon required knowledge of the username and password to gain access to an account.

Also of note was the identification of the log files for SkyDrive and Google Drive, which provide file and use history which may be important in an investigation. Also of
note was the discovery that Dropbox log files are now encrypted, and information may not be available, depending on the version a user had installed. Browser information was important, and there is software and procedures for this to be extracted for analysis. The use of anti-forensic software, Eraser and CCleaner, did not remove data remnants; however, the use of a full erase process (DBAN) removed all data from the hard drives.

As outlined, there is a wide range of investigation points for an examiner to determine the use of cloud storage, such as; directory listings, prefetch files, link files, thumbnails, registry, browser history, and memory captures. By determining the data remnants, this research provides a better understanding of the type of artifacts that are likely to remain, and the access point(s) for digital forensics examiners to assist an investigation.

The real world application of this research was demonstrated when Dropbox information was located in a current Police investigation, indicating illicit data hosted within a user account, as mentioned in Chapter Five. The use of the proposed framework was also beneficial to guide the research, and would be applicable in a digital forensic investigation to lead the investigation from commencement to completion. Of note was the ability to locate username information, and passwords for accounts in plain text. Whilst this is a security concern for users, for an examiner this information can be used to gain access to evidence which may not be possible without knowledge of the username or password. The appropriate legal considerations would need to be reviewed for each jurisdiction to ensure this was valid.
References

All URLs were last accessed (and correct on) 21 October 2012


Brezinski, D & Killalea, T (2002), 'Rfc 3227 - Guidelines for Evidence Collection and Archiving'.


Karyda, M (2009), 'The Socioeconomic Background of Electronic Crime', *Socioeconomic and Legal Implications of Electronic Intrusion*, p. 1.


184
Reilly, D, Wren, C & Berry, T (2010), 'Cloud Computing: Forensic Challenges for Law
Enforcement', paper presented at the International Conference for Internet
Technology and Secured Transactions (ICITST).

Forensic Investigations', *International Journal of Multimedia and Image
Processing (IJMIP)*, vol. 1, pp. 26-34.

Roach, M (2012), *Cumulus Data Launches Cloud Collection Service*, updated 23
August 2012, Law.com,
<http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=12025685
9159&Cumulus_Data_Launches_Cloud_Collection_Service&slreturn=201208
08064047>.

for Large-Scale Forensic Computing', *Advances in Digital Forensics V*, pp. 201-
214.

Ruan, K, Carthy, J & Kechadi, T (2011), 'Survey on Cloud Forensics and Critical
Criteria for Cloud Forensic Capability: A Preliminary Analysis'.

Ruan, K, Carthy, J, Kechadi, T & Crosbie, M (2011), 'Cloud Forensics', *Advances in
Digital Forensics VII*, pp. 35-46.

Australia, Sydney, Australia.

Sandoval, G (2012), *U.S. Says Megaupload's Hosting Service Is No Innocent Bystander*,
1023_3-57414104-93/u.s-says-megauploads-hosting-service-is-no-innocent-
bystander/>.!

Slay, J, Lin, Y, Turnbull, B, Beckett, J & Lin, P (2009), 'Towards a Formalization of
37-47.

Models of Cloud Computing', *Journal of Network and Computer Applications*,
vol. 34, no. 1, pp. 1-11.

Sunde, IM (2006), *Lov Og Rett I Cyberspace (Law and Order in Cyberspace)* Chapter
Internet Crime and Investigation Methods, Fagbokforlaget.


Wolthusen, SD (2009), 'Overcast: Forensic Discovery in Cloud Environments', paper presented at the Fifth International Conference on IT Security Incident Management and IT Forensics.


Use of electronic equipment at premises

(1) The executing officer of a warrant in relation to premises, or a constable assisting, may operate electronic equipment at the warrant premises to access data (including data not held at the premises) if he or she suspects on reasonable grounds that the data constitutes evidential material.

Note: A constable can obtain an order requiring a person with knowledge of a computer or computer system to provide assistance: see section 3LA.

(1A) If the executing officer or constable assisting suspects on reasonable grounds that any data accessed by operating the electronic equipment constitutes evidential material, he or she may:

(a) copy any or all of the data accessed by operating the electronic equipment to a disk, tape or other associated device brought to the premises; or

(b) if the occupier of the premises agrees in writing--copy any or all of the data accessed by operating the electronic equipment to a disk, tape or other associated device at the premises;

and take the device from the premises.

(1B) If:

(a) the executing officer or constable assisting takes the device from the premises; and

(b) the Commissioner is satisfied that the data is not required (or is no longer required) for a purpose mentioned in section 3ZQU or for other judicial or administrative review proceedings;

the Commissioner must arrange for:

(c) the removal of the data from any device in the control of the Australian Federal Police; and
(d) the destruction of any other reproduction of the data in the control of the Australian Federal Police.

(2) If the executing officer or a constable assisting, after operating the equipment, finds that evidential material is accessible by doing so, he or she may:
   (a) seize the equipment and any disk, tape or other associated device; or
   (b) if the material can, by using facilities at the premises, be put in documentary form--operate the facilities to put the material in that form and seize the documents so produced.

(3) A constable may seize equipment under paragraph (2)(a) only if:
   (a) it is not practicable to copy the data as mentioned in subsection (1A) or to put the material in documentary form as mentioned in paragraph (2)(b); or
   (b) possession by the occupier of the equipment could constitute an offence.

(4) If the executing officer or a constable assisting suspects on reasonable grounds that:
   (a) evidential material may be accessible by operating electronic equipment at the premises; and
   (b) expert assistance is required to operate the equipment; and
   (c) if he or she does not take action under this subsection, the material may be destroyed, altered or otherwise interfered with;

he or she may do whatever is necessary to secure the equipment, whether by locking it up, placing a guard or otherwise.

(5) The executing officer or a constable assisting must give notice to the occupier of the premises of his or her intention to secure equipment and of the fact that the equipment may be secured for up to 24 hours.

(6) The equipment may be secured:
   (a) for a period not exceeding 24 hours; or
   (b) until the equipment has been operated by the expert;

whichever happens first.

(7) If the executing officer or a constable assisting believes on reasonable grounds that the expert assistance will not be available within 24 hours, he or she may apply to an issuing officer for an extension of that period.

(8) The executing officer or a constable assisting must give notice to the occupier of the premises of his or her intention to apply for an extension, and the occupier is entitled to be heard in relation to the application.

(9) The provisions of this Division relating to the issue of warrants apply, with such modifications as are necessary, to the issuing of an extension.
Person with knowledge of a computer or a computer system to assist access etc.

(1) A constable may apply to a magistrate for an order requiring a specified person to provide any information or assistance that is reasonable and necessary to allow a constable to do one or more of the following:

(a) access data held in, or accessible from, a computer or data storage device that:
   (i) is on warrant premises; or
   (ii) has been moved under subsection 3K(2) and is at a place for examination or processing; or
   (iii) has been seized under this Division;
(b) copy data held in, or accessible from, a computer, or data storage device, described in paragraph (a) to another data storage device;
(c) convert into documentary form or another form intelligible to a constable:
   (i) data held in, or accessible from, a computer, or data storage device, described in paragraph (a); or
   (ii) data held in a data storage device to which the data was copied as described in paragraph (b); or
   (iii) data held in a data storage device removed from warrant premises under subsection 3L(1A).

(2) The magistrate may grant the order if the magistrate is satisfied that:
(a) there are reasonable grounds for suspecting that evidential material is held in, or is accessible from, the computer or data storage device; and
(b) the specified person is:
   (i) reasonably suspected of having committed the offence stated in the relevant warrant; or
   (ii) the owner or lessee of the computer or device; or
   (iii) an employee of the owner or lessee of the computer or device; or
   (iv) a person engaged under a contract for services by the owner or lessee of the computer or device; or
   (v) a person who uses or has used the computer or device; or
   (vi) a person who is or was a system administrator for the system including the computer or device; and
(c) the specified person has relevant knowledge of:
   (i) the computer or device or a computer network of which the computer or device forms or formed a part; or
   (ii) measures applied to protect data held in, or accessible from, the computer or device.

(3) If:
(a) the computer or data storage device that is the subject of the order is seized under this Division; and
(b) the order was granted on the basis of an application made before the seizure;
the order does not have effect on or after the seizure.

Note: An application for another order under this section relating to the computer or data storage device may be made after the seizure.

(4) If the computer or data storage device is not on warrant premises, the order must:
(a) specify the period within which the person must provide the information or assistance; and
(b) specify the place at which the person must provide the information or assistance; and
(c) specify the conditions (if any) determined by the magistrate as the conditions to which the requirement on the person to provide the information or assistance is subject.

(5) A person commits an offence if the person fails to comply with the order.
Penalty for contravention of this subsection: Imprisonment for 2 years.
Access to and downloading of data from computers (including access to computers outside premises the subject of a warrant)

75B Access to and downloading of data from computers (including access to computers outside premises the subject of a warrant)
(1) The person executing or assisting in the execution of a warrant to which this Division applies may operate equipment at the premises the subject of the warrant to access data (including data held at premises other than the subject premises) if the person believes on reasonable grounds that the data might be data that could be seized under the warrant.

Note: Under section 49, data may be seized under a warrant if connected with an offence. Section 46 (3) provides that a thing is connected with an offence if, for example, it will provide evidence of the commission of the offence.

(2) The person executing or assisting in the execution of the warrant may:
(a) copy any accessed data to a disk, tape or other data storage device brought to the premises, and
(b) with the approval of the occupier of the premises, copy any accessed data to a disk, tape or other data storage device already at the premises, and
(c) take the disk, tape or other data storage device from the premises to examine the accessed data to determine whether it (or any part of it) is data that could be seized under the warrant.

(3) The person executing or assisting in the execution of the warrant may operate the equipment to put any data that could be seized in documentary form and seize the document so produced.

(4) The person executing or assisting in the execution of the warrant may seize the equipment and any disk, tape or other data storage device:
(a) if it is not practicable to exercise the powers referred to in subsection (2) or (3) in relation to the data, or
(b) if possession by the occupier of the equipment or device could constitute an offence.

(5) This section does not authorise the operation of equipment already at the premises the subject of the warrant to access data unless the person operating the equipment has reasonable grounds to believe that the equipment can be operated without damaging the equipment or the data.

(6) The responsible officer for an authority must arrange for the removal of any data obtained by the exercise of a power referred to in this section by a member of the authority from any device under the control of the authority and the destruction of any other reproduction of the data in the control of the authority if the responsible officer is satisfied that the data is data that could not be seized under the warrant.

(7) Subsection (6) does not require the destruction of court records.

(8) In this section, “responsible officer for an authority” means the following:
(a) in relation to data obtained by a police officer-the Commissioner of Police,
(b) in relation to data obtained by a member of staff of the Police Integrity Commission-the Commissioner for the Police Integrity Commission,
(c) in relation to data obtained by a member of staff of the New South Wales Crime Commission-the Commissioner for the New South Wales Crime Commission,
(d) in relation to data obtained in execution of a search warrant issued under a provision of an Act specified in Schedule 2-the person prescribed by the regulations.
67—General search warrants
   (1) Despite any law or custom to the contrary, the Commissioner may issue
general search warrants to such police officers as the Commissioner thinks fit.
   (2) Every such warrant must be in the form in the Schedule, or in a form to
the same effect, and must be signed by the Commissioner.
   (3) Every such warrant will, subject to prior revocation by the Commissioner,
remain in force for six months from the date of the warrant, or for a shorter period
specified in the warrant.
   (4) The police officer named in any such warrant may, at any time of the day
or night, exercise all or any of the following powers:
      (a) the officer may, with such assistants as he or she thinks necessary, enter
into, break open and search any house, building, premises or place where he or she has
reasonable cause to suspect that—
           (i) an offence has been recently committed, or is about to be committed;
or
           (ii) there are stolen goods; or
           (iii) there is anything that may afford evidence as to the commission of
an offence; or
           (iv) there is anything that may be intended to be used for the purpose of
committing an offence;
      (b) the officer may break open and search any cupboards, drawers, chests,
trunks, boxes, packages or other things, whether fixtures or not, in which he or she has
reasonable cause to suspect that—
           (i) there are stolen goods; or
           (ii) there is anything that may afford evidence as to the commission of
an offence; or
           (iii) there is anything that may be intended to be used for the purpose of
committing an offence;
      (c) the officer may seize any such goods or things to be dealt with
according to law.
176—Person with knowledge of a computer or a computer system to assist access etc

(1) An authorised officer responsible for executing a warrant may apply to a magistrate for an order requiring a specified person to provide any information or assistance that is reasonable or necessary to allow an authorised officer to do one or more of the following:

(a) access data held in or accessible from a computer that is on the premises specified in the warrant;
(b) copy the data to a data storage device;
(c) convert the data into documentary form.

(2) The magistrate may make an order if satisfied that—

(a) there are reasonable grounds for suspecting that material liable to seizure under this Act is accessible from the computer referred to in subsection (1)(a); and

(b) the specified person is—

(i) reasonably suspected of possessing, or having under his or her control, material liable to seizure under this Act; or
(ii) the owner or lessee of the computer; or
(iii) an employee of the owner or lessee of the computer; and

(c) the specified person has knowledge of—

(i) the computer or a computer network of which the computer forms a part; or
(ii) measures applied to protect data held in or accessible from the computer.

(3) A person who fails to comply with an order under this section is guilty of an offence.

Maximum penalty: $2 500 or imprisonment for 6 months.
Road Traffic Act 1961 (SA) - Sect 40q

40Q—Power to inspect vehicle on road or certain official premises

(1) This section applies to a vehicle located at a place—
   (a) on a road; or
   (b) in or on premises occupied or owned by a public authority, whether or not the vehicle is unattended.

(2) An authorised officer or police officer may inspect a vehicle for compliance purposes.

(3) The officer may enter the vehicle for the purpose of or in connection with conducting the inspection.

(4) The officer may exercise powers under this section at any time, and without the consent of the driver or other person apparently in charge of the vehicle or any other person.

(5) Without limiting the above, the power to inspect a vehicle under this section includes any or all of the following:
   (a) the power to weigh, test, measure or take photographs of the vehicle or any part of it or its equipment or load;
   (b) the power to check the existence or details of, or take photographs of, placards or other information required under an Australian road law or under an approved road transport compliance scheme to be displayed in or on the vehicle, including placards or other information relating to its specifications, capabilities or legal entitlements;
   (c) the power to inspect and take copies of or extracts from any records that are located in or on the vehicle and that are required to be carried in or on the vehicle under an Australian road law or under an approved road transport compliance scheme;
   (d) the power to access or download information that is required to be kept under an Australian road law or under an approved road transport compliance scheme and that is—
      (i) stored electronically in equipment located in or on the vehicle; or
      (ii) accessible electronically from equipment located in or on the vehicle.

(6) This section does not authorise the use of force, but the officer may under this section do any or all of the following:
   (a) open unlocked doors and other unlocked panels and objects;
   (b) inspect anything that has been opened or otherwise accessed under the power to use reasonable force in the exercise of a power to enter or move a vehicle under Subdivision 3;
   (c) move but not take away anything that is not locked up or sealed.
Appendix B – Example of Examination Report

(adapted from NIJ (2004))

REPORT OF MEDIA ANALYSIS
MEMORANDUM FOR: Police Investigator Jones

SUBJECT: Forensic Media Analysis Report
SUBJECT: SUSPECT
Case Number: 012345

1. Status: Complete.

2. Summary of Findings:
   - 3 document files containing information relating to importation and sales of steroids.
   - 14 shortcut files that pointed to files stored in Google Drive account with the username ‘suspect@badstuff.com’
   - Internet History containing links to files stored in Google Drive

3. Items Analysed:

<table>
<thead>
<tr>
<th>TAG NUMBER</th>
<th>ITEM DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>012345</td>
<td>One Generic laptop, Serial # 123456789</td>
</tr>
<tr>
<td>012346</td>
<td>One Tower PC, Serial # not listed</td>
</tr>
<tr>
<td>012347</td>
<td>Apple iPhone 3G, Serial # 123456780</td>
</tr>
</tbody>
</table>

4. Details of Findings for item 012345:

Findings in this paragraph related to the Generic Hard Drive, Model ABCDE, Serial # 3456ABCD, recovered from Tag Number 012345, One Generic laptop, Serial # 123456789.

1) The examined hard drive was found to contain a Microsoft Windows 7 operating system.

2) The directory and file listing for the media was saved to ‘012345 LT HD File Listing.csv’.

3) The Internet History at C:\USERS\SUSPECT\APPDATA\LOCAL\MICROSOFT\WINDOWS\INTERNET EXPLORER\HISTORY\INDEX.DAT, was found to contain URL references to files stored in a Google Drive account with the filenames; “steroid customers.docx”, “steroid importations.docx”, and “steroid sales.docx”. The last accessed dates for the references disclosed the date and times as 5 July 2012 between 11:23 p.m. and 11:57 p.m.
5. Details of Findings for item 012346:
Findings in this paragraph related to the Generic Hard Drive, Model ABCDF, Serial # 3456ABCE, recovered from Tag Number 012346, One Tower PC, Serial # not listed.

1) The examined hard drive was found to contain a Microsoft Windows 7 operating system.
2) The directory and file listing for the media was saved to ‘012346 PC HD File Listing.csv’.
3) Google Drive client software was installed on the PC, located at C:\Program Files\Google\Drive\. During an examination on scene, the client software was used to connect to the stored client account details using the link option ‘visit Google Drive on the web’. The following files were observed stored in the Google Drive account;
   - File 1 “steroid customers.docx” the files’ modified date and times are 5 July 2012 11:33 p.m. (ACST).
   - File 2 “steroid importations.docx” the files’ modified date and times are 5 July 2012 11:35 p.m. (ACST).
   - File 3 “steroid sales.docx” the files’ modified date and times are 5 July 2012 11:57 p.m. (ACST).
   These files were downloaded at the time of on scene access to a compressed ZIP file; ‘012346 GD.ZIP’
4) These 3 files were also located in the default Google Drive synchronisation folder on the hard drive, located at ‘C:\Users\Suspect\Google\Drive\’
5) The Internet History at C:\\Users\\Suspect\\AppData\\Local\\Microsoft\\Windows\\Internet Explorer\History\index.dat, was found to contain URL references to files stored in a Google Drive account with the filenames; “steroid customers.docx”, “steroid importations.docx”, and “steroid sales.docx”.

6. Details of Findings for item 012347:
Findings in this paragraph related to the Apple iPhone 3G logical evidence extract, recovered from Tag Number 012347, Apple iPhone 3G, Serial # 123456780.

1) The examined iPhone was found to contain iOS 4.2.1 operating system.
2) The Microsystemation XRY extract file was saved to ‘012347 Apple iPhone.xry’.
3) The Internet History at ‘Web History.txt’ was found to contain URL references to files stored in a Google Drive account with the filenames; “steroid customers.docx”, “steroid importations.docx”, and “steroid sales.docx”.

6. Items Provided: In addition to this hard copy report, one compact disk (CD) was submitted with an electronic copy of this report. The report on CD contains hyperlinks to the above-mentioned files and directories.

EXAMINER PERRY
Computer Forensic Examiner

Released by_______________________

**Principle 1:** No action taken by law enforcement agencies or their agents should change data held on a computer or storage media which may subsequently be relied upon in court.

**Principle 2:** In circumstances where a person finds it necessary to access original data held on a computer or on storage media, that person must be competent to do so and be able to give evidence explaining the relevance and the implications of their actions.

**Principle 3:** An audit trail or other record of all processes applied to computer-based electronic evidence should be created and preserved. An independent third party should be able to examine those processes and achieve the same result.

**Principle 4:** The person in charge of the investigation (the case officer) has overall responsibility for ensuring that the law and these principles are adhered to.

NIJ First Responders (2008)

Handling Digital Evidence at the Scene

Precautions should be taken in the collection, preservation, and transportation of digital evidence. First responders may follow the steps listed below to guide their handling of digital evidence at an electronic crime scene:

- Recognize the investigative value of digital evidence.
- Assess available resources.
- Identify the equipment and supplies that should be taken to electronic crime scenes.
- Designate the assignments, roles, and responsibilities of personnel involved in the investigation.
- Recognize, identify, seize, and secure all digital evidence at the scene.
- Document the entire scene and the specific location of the evidence found.
- Collect, label, and preserve the digital evidence.
- Package and transport digital evidence in a secure manner.
- Ensure that officer safety and the safety of others remain the highest priority.
NIJ Forensic Examination (2004)
1. Policy and Procedure Development
   **Principle:** Computer forensics as a discipline demands specially trained personnel, support from management, and the necessary funding to keep a unit operating. This can be attained by constructing a comprehensive training program for examiners, sound digital evidence recovery techniques, and a commitment to keep any developed unit operating at maximum efficiency.

2. Evidence Assessment
   **Principle:** The digital evidence should be thoroughly assessed with respect to the scope of the case to determine the course of action.

3. Evidence Acquisition
   **Principle:** Digital evidence, by its very nature, is fragile and can be altered, damaged, or destroyed by improper handling or examination. For these reasons, special precautions should be taken to preserve this type of evidence. Failure to do so may render it unusable or lead to an inaccurate conclusion.

4. Evidence Examination
   **Principle:** General forensic principles apply when examining digital evidence. Different types of cases and media may require different methods of examination. Persons conducting an examination of digital evidence should be trained for this purpose.

5. Documenting and Reporting
   **Principle:** The examiner is responsible for completely and accurately reporting his or her findings and the results of the analysis of the digital evidence examination. Documentation is an ongoing process throughout the examination. It is important to accurately record the steps taken during the digital evidence examination.


I, Darren Paul Quick, make the following declarations regarding the publications submitted for consideration titled:

- ‘Dropbox Analysis: Data Remnants on User Machines’.
- ‘Digital Droplets: Microsoft SkyDrive forensic data remnants’.
- ‘Forensic Collection of Cloud Storage Data from a Law Enforcement Perspective’.
- ‘Google Drive: Forensic Analysis of data remnants’.

1. I am the primary author of the papers.

2. I was primarily responsible for the research undertaken for the papers.

3. I was primarily responsible in determining the content and focus of the papers.

4. I authored the papers and was provided input and assistance in relation to background material, feedback and editing from the co-author and my thesis supervisor Dr Kim-Kwang Raymond Choo.

I declare in reference to the above that the contribution towards the research presented in the papers is substantially my own.

Signed: Darren Quick Date: 28 October 2012
I, Dr Kim-Kwang Raymond Choo, make the following declarations regarding the publications submitted for consideration titled:

- ‘Dropbox Analysis: Data Remnants on User Machines’.
- ‘Digital Droplets: Microsoft SkyDrive forensic data remnants’.
- ‘Forensic Collection of Cloud Storage Data from a Law Enforcement Perspective’.
- ‘Google Drive: Forensic Analysis of data remnants’.

1. I am the secondary author of the papers.

2. I provided input and assistance in relation to background material, feedback, editing and guidance to the primary author of the paper.

I declare in reference to the above that the contribution towards the research presented in the papers is substantially the primary author, Darren Paul Quick.

Signed: Kim-Kwang Raymond Choo Date: 30-Oct-2012