한국항행학회 논문지 제17권 제1호 2013년 2월 (JKONI 17(1): 63-68, Feb. 2013)

클라우드 환경을 고려한 디지털 포렌식 프레임워크

Digital Forensics Framework for Cloud Computing

이창훈*

Chang-Hoon Lee*

요 약

최근 세계적인 경제위기 속에서 국내외의 기업들이 IT 투자를 보류하거나 예산을 대폭 삭감하고 있다. 이에 기업들은 IT 부문에 있어서의 비용 절감을 통한 위기 극복 방안을 모색하고 있으며, 이러한 상황에서 클라우드 컴퓨팅(Cloud Computing)은 위기 극복을 위한 최적의 솔루션으로 빠르게 부상하고 있다. 또한 디지털 포렌식 조사과정에서 조사 대상 시스템의 사용자가 클라우드 서비스를 사용했는지 여부는 추가적은 조사 대상의 선정에 매우 중요한 요소이다. Daum Cloud, Google Docs와 같은 클라우드 서비스를 사용하였을 경우, 로그인 정보를 획득하여 원격지의 클라우드 서비스에 접속이 가능한 경우가 있다. 이러한 경우에는 원격지의 증거 데이터를 수집할 수 있는다. 따라서 다양한 클라우드 서비스에서 데이터를 수집하고 분석하는 방안에 대하여 연구가 필요하다. 이에 본 연구에서는 서비스별 데이터 수집 및 분석 기법에 대해 연구하여 클라우드 환경을 고려한 디지털 포렌식 프레임워크를 제안한다.

Abstract

Recently, companies seek a way to overcome their financial crisis by reducing costs in the field of IT. In such a circumstance, cloud computing is rapidly emerging as an optimal solution to the crisis. Even in a digital forensic investigation, whether users of an investigated system have used a cloud service is a very important factor in selecting additional investigated subjects. When a user has used cloud services, such as Daum Cloud and Google Docs, it is possible to connect to the could service from a remote place by acquiring the user's log-in information. In such a case, evidence data should be collected from the remote place for an efficient digital forensic investigation, and it is needed to conduct research on the collection and analysis of data from various kinds of cloud services. Thus, this study suggested a digital forensic framework considering cloud environments by investigating collection and analysis techniques for each cloud service

Key words : Digital Forensics, Cloud Computing, Digital Evidence

I. 서 론

최근 클라우드 컴퓨팅 서비스의 사용이 급격하게 증가하고 있다. 기업에서 뿐만 아니라 개인 사용자들도 클라우드 컴퓨팅 서비스를 이용하고 있다. 이렇게 클라우드 컴퓨팅 서비스의 사용은 증가하고 있지만, 클라우드 컴퓨팅에 대응할 디지털 포렌식은 현재 학문적, 실용적 체계화 수준이 미비한 실정이다. 따라

* 서울과학기술대학교 컴퓨터공학과(Department of Computer Science and Engineering, Seoul National University of Science and Technology)
· 제1저자 (First Author) : 이창훈(Changhoon Lee, tel:+82-2-970-6712, email : chlee@seoultech.ac.kr)
· 접수일자 : 2013년 1월 21일 · 심사(수정)일자 : 2013년 1월 25일 (수정일자 : 2013년 2월 23일) · 게재일자 : 2013년 2월 28일
http://dx.doi.org/10.12673/jkoni.2013.17.01.063

서 상용화가 진행되고 있는 클라우드 시대를 대비하기 위해서 클라우드 컴퓨팅에 대한 이해를 바탕으로 체계적인 디지털 포렌식을 준비할 필요가 있다.

최근 언론에서 가상화 기술과 클라우드 컴퓨팅이라는 용어가 자주 등장한다. 그만큼 클라우드 컴퓨팅 기술은 21세기 세계 IT시장의 최대 화두이며, 국내에서도 이러한 IT 시장 흐름의 변화에 발맞추기 위해 기업과 정부 모두 분주하게 움직이고 있다. 미국 NIST(National Institute of Standards and Technology)는 클라우드 컴퓨팅을 “네트워크, 서버, 스토리지, 응용프로그램 등과 같은 컴퓨팅 자원들의 공유된 풀에 언제든지 편리하게 네트워크로 접근가능한 방식의 모델”로 정의하고 있다. 즉, 사용자는 단말기를 사용하여 유선, 무선통신으로 인터넷에 접속만하면, IT 자원들을 필요한 때 필요한 만큼 빌려 쓰고 사용한 만큼 요금을 지불하는 형식의 서비스 형태이다.

IT 자원을 인프라로 사용하는 클라우드 컴퓨팅은 포털 사이트에서 제공하는 웹메일이나 블로그, 웹하드 서비스나 웹호스팅 서비스를 통해 이미 개념적으로 사용되고 있었다. 그러나 예전에는 소프트웨어 기술의 한계와 네트워크가 전달할 수 있는 물리적인 정보량의 한계로 인해 인터넷을 통해 제공 가능한 서비스의 수준과 범위가 제한적이었고, 클라우드 컴퓨팅의 시장 가치가 낮았다. 하지만 최근에는 네트워크의 고도화와 함께 가상화와 같은 소프트웨어 기술이 발전되면서 광범위한 분야의 소프트웨어와 IT 자원들이 인터넷을 통해 제공될 수 있는 환경이 마련되었다. 기존의 그리드 컴퓨팅, 분산 컴퓨팅, 유틸리티 컴퓨팅, 웹 서비스, 서버 및 스토리지의 가상화 기술과 공개 소프트웨어 등과 같은 기반 기술들을 융합하여 하나의 커다란 컴퓨팅 환경을 구성한 것이다. 클라우드 컴퓨팅이라는 용어가 광범위하게 사용되고 있지만, 모든 클라우드 모델들이 다 똑같지는 않다. 클라우드 컴퓨팅 서비스는 SaaS (Software as a Service), PaaS (Platform as a service), IaaS (Infrastructure as a Service)의 세 가지로 나눌 수 있다. 클라우드 컴퓨팅 환경 적용할 때에는 이러한 세 가지 모델 사이의 차이점과 유사점을 사전에 충분히 고려해야 한다[1].

Ⅱ. 클라우드 컴퓨팅 환경에서의 디지털 포렌식

클라우드 컴퓨팅은 서비스의 형태에 따라 크게 3가지로 분류할 수 있다. 먼저 SaaS(Software as a Service)로 특별한 설치 과정 없이 웹브라우저로 응용프로그램을 사용할 수 있는 서비스이다. 현재 문서작성, 그림 편집 등 다양한 기능의 SaaS가 존재하며, 인터넷 연결이 가능한 환경에서 웹브라우저만 있으면 편리하게 사용할 수 있다. 다음으로 PaaS(Platform as a Service)는 응용프로그램 개발 환경을 제공하는 서비스로 Microsoft, Google 등의 기업에서 다양한 종류의 개발 환경을 제공하고 있다. 마지막으로 IaaS(Infrastructure as a Service)는 인프라(서버, 스토리지 등)를 제공하는 것으로 Amazon, IBM, Oracle, KT, SK, NHN 등 국내외 여러 기업에서 서비스하고 있다.

이와 같이 클라우드 컴퓨팅 시장이 증가함에 따라 다양한 기업들에 의해 여러 가지 종류의 서비스가 제공되고 있다. 클라우드 컴퓨팅 서비스를 이용하는 경우에는 기존 데스크톱 컴퓨터 환경과는 달리 사용자의 행위에 의한 데이터가 로컬 시스템에 거의 남지 않으며, 심지어 원격지의 서버 시스템에 모든 데이터가 저장되기도 한다. 따라서 클라우드 컴퓨팅 서비스를 사용하는 용의자의 시스템을 조사할 때 기존의 디지털 포렌식 기술을 활용하는데는 한계가 있으며, 클라우드 컴퓨팅 환경에 대응하기 위한 포렌식 기술에 대한 연구가 필요하다. 클라우드 컴퓨팅에 대응하기 위한 디지털 포렌식 기술은 현재 학문적, 실용적 체계화 수준이 미비하여 실제 수사 환경을 고려한 연구가 필요한 실정이다. 이를 위해서는 다양한 클라우드 컴퓨팅 서비스 별 구성 형태에 대한 이해가 필요하며, 각 서비스의 형태에 따라서 법적효력을 지니는 증거 데이터 수집 체계를 확립해야 한다. 기술적으로 보았을 때, 용의자의 시스템에서 클라우드 서비스의 사용 기록을 탐지하는 기술이 필요하며 이를 통해 추가적인 분석 대상의 선별과 대응이 가능하도록 해야 한다. 또한 클라우드 컴퓨팅의 기본이 되는 서버 가상화 솔루션에 대한 이해를 통해 서버 가상화 환경 대응을 위한 포렌식 기법을 연구해야 한다[3].

Ⅲ. 클라우드 컴퓨팅 환경에서의 데이터 수집

클라우드 컴퓨팅 서비스를 사용하기 위해서는 어떠한 형태를 가지든지 클라이언트가 필요하며, 일반 사용자들은 데스크톱, 노트북, 태블릿 PC, 스마트폰 등과 같이 다양한 클라이언트 시스템을 사용하고 있다. 다양한 시스템에는 다양한 형태의 흔적이 존재하게 된다. 클라우드 컴퓨터를 사용했다는 것을 확신할 수 있는 흔적이 존재하고 이 외에 클라우드 컴퓨팅 서비스에 따라 로그인 정보를 추가적으로 획득할 수 있다. 클라이언트 관점에서 클라우드 서비스는 인터넷 서비스를 기반으로 제공되고 있다[2]. 사용자는 주로 웹브라우저를 통해 클라우드 서비스를 이용하기 때문에 웹브라우저의 사용 흔적을 중점적으로 분석한다. Internet Explorer를 비롯한 FireFox, Chrome, Safari 등 다양한 웹브라우저에서는 인터넷 사용 내역을 웹 브라우저 생성 파일에 기록한다. 이러한 파일들에는 임시 인터넷 파일, 웹 캐쉬, 웹 히스토리 등이 포함되고 이를 분석해 서비스 사용 여부를 알 수 있다. 클라우드 서비스들 중에 응용프로그램을 제공하는데 사용 여부를 확인하기 위해 레지스트리를 분석한다. 일부 클라우드 서비스는 로그인 정보 등이 남는 경우도 있기 때문에 레지스트리는 필수적으로 조사해야한다[4].

표 1. 분석 대상
Table 1. Target Cloud Service

제공사	서비스 이름
KT	Ucloud
NHN	N Drive
Daum	Daum Cloud
Google	Google Docs

현재 웹을 통해 서비스되고 있는 형식은 대부분 클라우드 형태로 다수의 대중을 위하여 인터넷 기반으로 운영되고 있다[5]. 본 장에서 윈도우(Windows) 시스템에서 대표적인 클라우드 서비스를 사용할 시에 생기는 흔적을 분석한다. 분석 대상 클라우드 서비스 목록은 표 1.과 같다.

3-1 KT Ucloud

KT의 Ucloud 서비스를 이용하는 경우에 웹브라우저 로그에 해당 접속 기록이 저장된다. 만약 사용자가 '문서 미리 보기' 기능을 사용하였다면 'http://docview.ucloud.olleh.com/view'가 포함된다. 또한 '미리 듣기 기능'을 사용하여 음악 파일을 재생했다면 'http://dosirak.ucloud.olleh.com'이 포함된다. 이와 같이 인터넷 접속 기록을 통해서 Ucloud 서비스의 사용 내역을 확인할 수 있다[6].

그림 1. Ucloud 웹 사이트 접속 기록 분석
Fig 1. Ucloud web history

3-2 Naver N Drive

Naver N ucloud 서비스를 이용하여 파일을 올리거나 다운로드 받는 경우 웹브라우저 로그에 해당 접속 기록이 저장된다. 만약 사용자가 '네이버 워드' 기능으로 문서 파일에 대한 열람 또는 편집을 시도했다면, 'http://word.office.naver.com'이 포함된 URL을 확인할 수 있다. 또한 이미지 파일을 열어본 경우에는 'http://nphoto.naver.com'이 포함되며, 파일을 메일로 보낸 경우에는 'http://mail2.naver.com'이 포함된다. 또한 관련된 레지스트리 흔적은 크게 두 가지로 구분할 수 있다. 첫 번째는 설치와 관련된 정보로 윈도우 기본 서비스에 등록이 되어서 시스템이 시작될 때 자동으로 실행되도록 한다.

두 번째는 사용자의 계정과 관련된 정보이다. N 드라이브에 접속하기 위해서 사용자가 자신의 계

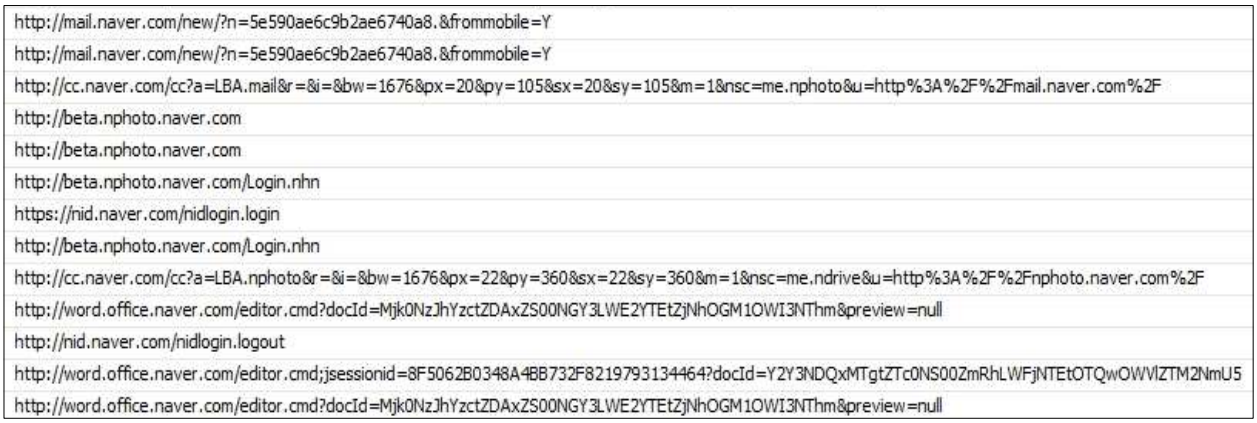

그림 2. Naver N Drive 접속 기록
Fig. 2. Naver N Drive connection history

정으로 로그인하면 레지스트리 키 'HKCU\Software\NHN orporation\NaverNDrive\LoginOpt'에 2개의 값(Recent Drive letter, recent_id0)이 생성된다. 이를 통해 계정정보 확인이 가능하다.

3-3 Daum cloud

다음 cloud를 사용한 경우에는 웹브라우저 로그에서 해당 기록을 확인할 수 있다. 사용자가 문서 파일을 '미리 보기' 한 경우에는 'http://cloud.daum.net/disk/viewer'가 포함된 URL을 확인할 수 있다. 또한 파일을 메일로 보낸 경우에는 'http://mail2.daum.net/hanmailex'가 포함되며, 'from' 파라미터를 통해 Daum 클라우드 서비스를 이용했다는 것을 확인할 수 있다.

3-4 Google Docs

Google Docs의 경우, 임시파일 중에 사용 흔적이 남는 파일은 크게 문서목록, 문서편집, 문서열람 했을 때의 사용자 행위로 구분할 수 있다. 사용자가 웹브라우저를 통해 'Google docs'에 접속했을 때 메인화면에 문서목록을 저장하고 있는 'docs_google_com[숫자].htm' 파일이 임시로 생성된다. 문서목록 중 한글 문자열의 경우 유니코드로 변환된 값이 저장된다.

Ⅳ. 클라우드 환경에서의 디지털 포렌식 프레임워크 제안

클라우드 컴퓨팅 환경에서 디지털 포렌식의 문제점은 클라우드 컴퓨팅 서비스로부터 사건 관련 데이터를 얻기 어렵다는 것이다. 사용자 데이터가 세계 도처에 물리적으로 분산되어 존재하기 때문에 사법 관할권이 다른 클라우드 벤더의 스토리지에 존재할 경우 데이터 확보의 문제점이 발생한

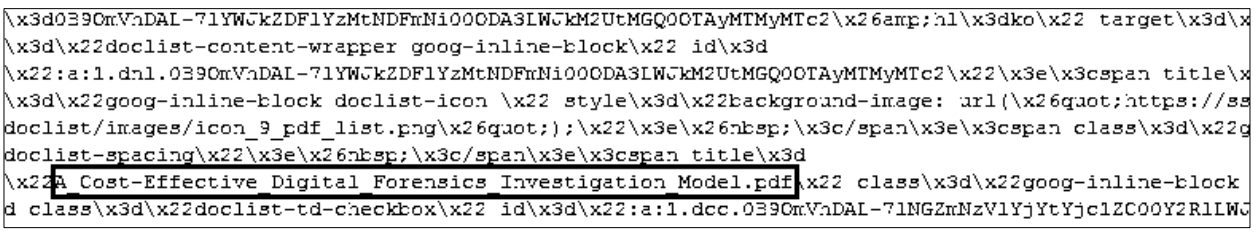

그림 3. Google Docs의 사용자 문서목록 파일 정보
Fig. 3. Google Docs User documents information

다. 또한 클라우드 컴퓨팅 서비스 종류(SaaS, PaaS, IaaS)나 클라우드 컴퓨팅 배치 모델(public, private, community, hybrid)에 따라 증거 수집 및 분석 방안이나 수사 절차 등이 상이하기 때문에 조사 방법을 사전에 숙지하고, 상황에 맞는 수사 업무에 임해야 한다는 것이다[7]. 예를 들면, 수사대상 클라우드 컴퓨팅 환경이 PaaS나 IaaS인 경우, 클라우드 컴퓨팅 시스템을 통해 데이터에 접근하는 경우, 레지스트리 또는 임시 인터넷 파일 등 운영체제에 기록되는 데이터들이 가상환경에 존재하거나 저장된다. 전통적 포렌식 수사에서 물리적인 컴퓨터 장치들을 압수하는 것은 상대적으로 간단한 업무이지만, 공용 클라우드 컴퓨팅 시스템에 저장된 데이터에 접근하는 일은 법적으로 한층 복잡하며, 신속하게 증거 자료를 복구해야 하는 수사 업무에서 시간을 지연시키는 원인이 될 수 있다. 디스크, 메모리, 네트워크 등은 공유하는 가상환경에서 개체를 물리적으로 수집하기 어려우며 전통적인 소유권 경계가 흐려지고 있는 국외 클라우드 서비스 업체에 대한 조사 관할권 문제가 존재한다. 이러한 현재의 상황에서 디지털 포렌식 관점에서의 클라우드 컴퓨팅 서비스 연구 결과 그림 4.와 같이 클라우드 컴퓨팅 서비스 환경에 대한 디지털 포렌식 프레임워크를 수립하였다.

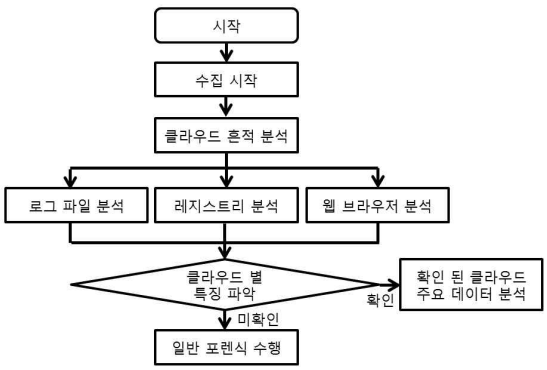

그림 4. 클라우드 환경을 고려한 디지털 포렌식 프레임워크

Fig. 4. Digital forensics framework for Cloud Computing

해당 프로세스에서 가장 중요한 부분은 해당 시스템에 대한 디지털 포렌식 조사 시에 클라우드 컴퓨팅 서비스의 사용 유무를 확인하는 내용이다.

만약 클라우드 상에 중요 증거 데이터가 존재하는 상황에서 디지털 증거 분석 시 클라우드 컴퓨팅 서비스의 사용 여부를 고려하지 않는다면, 아무리 정밀한 분석 절차를 진행하더라도 정작 핵심적인 증거 데이터들은 획득할 수 없을 것이다. 또한 최초 클라우드 컴퓨팅 서비스의 사용여부를 확인하지 않은 채 기존의 디지털 포렌식 분석 절차를 수행한 후 뒤늦게 클라우드 컴퓨팅 서비스 사용을 확인했다면, 중요한 증거가 사라졌을 가능성이 존재한다. 따라서 디지털 증거 분석의 최초 과정에서 클라우드 컴퓨팅 서비스의 사용 유무를 확인한 후에 앞서 제시했던 클라우드 컴퓨팅 포렌식 조사 절차와 기존의 디지털 포렌식 조사 절차를 병렬로 진행하는 방안이 필요하다.

Ⅴ. 결 론

클라우드 컴퓨팅 서비스를 이용하는 경우에는 기존 데스크톱 컴퓨터 환경과는 달리 사용자의 행위에 의한 데이터가 로컬 시스템에 거의 남지 않으며, 심지어 원격지의 서버 시스템에 모든 데이터가 저장되기도 한다. 이러한 클라우드 컴퓨팅의 특징은 디지털 포렌식 조사를 어렵게 할 수 있지만, 클라우드 컴퓨팅 환경이 갖는 특징을 최대한 이용한다면 조사를 하는 것이 불가능하지 않다.

클라우드 컴퓨팅 서비스를 사용하는 용의자의 시스템을 조사하기 위해 기존의 디지털 포렌식 기술을 그대로 적용하는 것에는 한계가 있다. 그러므로 클라우드 컴퓨팅 환경에서 디지털 포렌식 조사를 수행하기 위해서는 기존의 디지털 포렌식 절차에 앞서 클라우드 컴퓨팅 서비스의 사용 유무를 파악하는 과정이 필요하다. 클라우드 컴퓨팅 서비스를 사용한 흔적이 발견되면, 용의자가 클라우드 컴퓨팅 서비스에 접근하기 위해 사용했던 모든 단말 기기를 압수하여 데이터를 수집 및 분석해야 한다. 또한 용의자가 사용한 클라우드 컴퓨팅 서비스의 계정 정보를 알 수 있는 경우에는 해당 서비스의 원격지 서버에 접속하여 데이터를 수집할 수 있을 것이다.

감사의 글

이 연구는 서울과학기술대학교 교내 학술연구비 지원으로 수행되었습니다.

Reference

[1] Emma Webb Hobsosn, “Digital Investigations in the Cloud”

[2] Davic Mitchell Smith, David W. Cearley, Daryl C. Plummer, “Key Issues for Cloud Computing, 2009”

[3] Enrique Castro-Leon, Bernard Golden and Miguel Gomez, “Cloud Computing Basics”

[4] IAnewsletter, “Cyber Forensics in the Cloud”

[5] Keyun Ruan, Ibrahim Baggili(PhD), Prof Joe Carthy, Prof Tahar Kechadi, University College Dublin, Zayed University, “Surbey on Cloud forensics and critical criteria for cloud forensic capability: A preliminary analysis”

[6] Keith J. Jones, Forensic Analysis of Internet Explorer Activity Files. Foundstone, http://www.foundsrone.com/us/pdf/ wp_index_dat.pdf, 2003.

[7] D.Reilly, C Wren, T.Berry, “Cloud Computing: Forensic Challenges for Law Enforcement”, *School of Computing and Mathematical Sciences, Liverpool John Moores University,* UK.

이 창 훈 (Chang-Hoon Lee)

2001년 2월 : 한양대학교 수학과 (이학사)

2003년 2월 : 고려대학교 정보보호 대학원(공학석사)

2008년 2월 : 고려대학교 정보보호 대학원(공학박사)

2009년 3월 ~ 2011년 2월 : 한신대학교 컴퓨터공학부 전임강사

2011년 3월 ~ 2012년 2월 : 한신대학교 컴퓨터공학부 조교수

2012년 3월 ~ 현재 : 서울과학기술대학교 컴퓨터공학과 조교수

관심분야 : 정보보호, 암호학, 디지털포렌식, 컴퓨터이론