클라우드 컴퓨팅 플랫폼에서 디지털 증거 수집 절차†

한수빈*, 이태림*, 신상욱**
*부경대학교 대학원 정보보호학(협)
**부경대학교 IT 융합응용공학과
e-mail : subin4853@gmail.com

Procedure for the acquisition of digital evidence on a cloud computing platform

Su bin Han*, Tae-Rim Lee*, Sang Uk Shin**
*Dept. of Computer Science, Pukyong National University
**Dept. of IT Convergence and Application Eng., Pukyong National University

요 약

클라우드 컴퓨팅은 IT 자원의 효율적인 관리와 비용 대비 양질의 서비스 제공을 위한 새로운 패러다임으로써, 국내외의 기업뿐만 아니라 많은 사용자들에게 주목 받고 있다. 하지만 관련 시장의 빠른 성장과 함께 다양한 사이버 범죄에 노출될 수 있는 위험이 높아졌음에도 불구하고, 클라우드 컴퓨팅에 대한 디지털 포렌식은 실질적인 역할을 수행하기에 아직 미비한 실정이다. 클라우드 컴퓨팅은 증거 데이터가 물리적으로 분산되어 있고, 자원이 가상공간에 존재할 수 있기 때문에 기존의 디지털 포렌식 수사와는 다르게 접근해야 한다. 이에, 본 논문에서는 추상화된 클라우드 계층에 따른 기존 포렌식 절차 상의 데이터 수집 방법에 관한 한계를 분석하고, 확보한 증거 데이터의 신뢰성 보장 및 다양한 클라우드 환경에 보다 유연하게 적용할 수 있는 디지털 증거 수집 절차를 제안한다. 해당 절차는 클라우드 구성 요소들 중 물리적인 자원들을 가상화하여 논리적으로 구성할 수 있도록 하며, 가상화된 자원들을 서비스 목적에 따라 폭넓게 활용할 수 있도록 관리 체계를 제공해주는 클라우드 플랫폼을 기반으로 한다.

1. 서론

최근 국내외 기업뿐만 아니라 개인 사용자들의 클라우드 컴퓨팅 서비스 사용이 급격하게 증가하고 있다. 미국의 정보 기술 연구 및 자문 회사인 가트너(Gartner)는 “2014 년 개인용 클라우드(Personal Cloud)가 개인 PC 를 대신해 디지털 라이프의 새로운 허브 역할을 할 것”이라고 전망하며 “현재의 단말 의존적인 IT 환경이 클라우드 기반 서비스 환경으로 전환될 것”이라고 발표했다. 이러한 클라우드 컴퓨팅의 성장은 데이터 및 트래픽의 급격한 증가와 함께 인터넷 업체의 성공과 기술 공유, 오픈소스의 활성화, 가상화 기술의 발전을 바탕으로 이루어졌다[9].

하지만 관련 시장의 빠른 성장과 함께 다양한 사이버 범죄에 노출될 수 있는 위험이 높아졌음에도 불구하고, 클라우드 컴퓨팅 시스템에 대한 디지털 포렌식은 아직 미비한 실정이며, 새로운 기술 및 법적 문제를 발생시키고 있다. 특히 포렌식 절차 상에서 증거 수집의 경우, 클라우드 서비스의 가용성 문제로 인해 물리적인 장비들의 압수가 현실적으로 불가능하며, 사용되는 클라우드 플랫폼 및 가상화 기술들에 의해 그 구성이 매우 다양하고 복잡하게 나타나지만 수사자의 접근 가능 범위는 관리 시스템 정도로 제한적일 수 밖에 없다. 이는 클라우드 환경에서 효과적인 증거 수집을 위해서는 기존의 포렌식 수사 방식과 다르게 접근해야 함을 의미하지만, 이를 지원하는 뚜렷한 포렌식 도구, 실질적인 정책 또는 절차들이 전무한 상황이다[4]. 또한 포렌식적으로 의미 있는 클라우드 사용자들의 서비스 이용과 관련된 데이터들이 대부분 가상화된 영역에 저장되어 있음을 고려해볼 때, 이를 증거로 활용하기 위해서는 해당 데이터를 획득하는 절차에 대한 신뢰성 문제를 우선적으로 해결해야 한다.

따라서, 본 논문에서는 다양한 클라우드 환경에 유연하게 적용할 수 있는 디지털 증거 수집 절차를 제안하기 위해, 먼저 일반적인 클라우드 컴퓨팅 시스템 구성 요소의 추상화된 계층에 따른 역할과 수집 가능한 데이터에 대해 분석한다. 또한 클라우드 환경에서 기존 포렌식 절차 상의 데이터 수집 방법에 관한 한계를 분석하고, 확보한 증거 데이터의 신뢰성 보장을 위해 클라우드 컴퓨팅 플랫폼 기반의 디지털 증거 수집 절차를 제안한다.

† 이 논문은 2011 년도 정부(미래창조과학부)의 재원으로 한국연구재단-차세대정보컴퓨팅기술개발사업의 지원을 받아 수행된 연구임(No. 2011-0029927).

2. 관련연구

2.1 클라우드 컴퓨팅(Cloud Computing)

클라우드 컴퓨팅은 광범위하게 사용하는 용어로 많은 업체나 기관에서 다양하게 정의되는데, Cloud Computing Use Cases group 에서는 “다양한 클라이언트 디바이스에서 필요한 시점에 인터넷을 이용하여 공유풀에 있는 서버, 스토리지, 애플리케이션, 서비스 같은 IT 리소스에 쉽게 접근할 수 있게 하는 모델” 이라고 클라우드 컴퓨팅을 정의하고 있다. 클라우드 컴퓨팅을 구분하는 기준은 여러 가지가 있지만, 일반적으로는 제공 서비스와 배치 방식을 기준으로 한다[9].

클라우드 컴퓨팅은 제공 서비스에 따라 IaaS(Infrastructure as a Service), PaaS (Platform as a Service), SaaS (Software as a Service)로 나뉘며, 먼저 SaaS 는 소프트웨어를 설치하지 않고 서비스 제공자가 제공하는 소프트웨어를 사용하며, 대표적인 서비스 사례로 Salesforce.com 과 구글 앱스를 들 수 있다. PaaS 는 응용프로그램이 실행되는 환경을 서비스 형태로 제공하는 플랫폼 서비스로, 대표적인 서비스 사례로 구글 앱엔진, Microsoft 의 Azure 등이 있다. 마지막으로 IaaS 는 서버, 스토리지 등의 서비스를 제공하는 것으로 인프라 서비스를 구축할 때 필요한 환경을 제공해주며, 대표적인 서비스는 아마존의 AWS, IBM, KT 등 국내외 여러 기업에서 제공하고 있다[9].

배치 방식에 따라 public, private, community, hybrid 로 나눌 수 있다. 공용 클라우드(public cloud)는 기업 내부가 아닌 외부의 사용자에게 자원을 제공하는 방식으로, 대중 사용자를 위한 서비스다. 사설 클라우드(private cloud)는 기업의 내부 구성원을 대상으로 하는 서비스로, 별도로 서비스를 구축해야 하기 때문에 서비스 구축 비용과 전문 인력이 필요하며 일정 규모 이상이 되어야 비용 절감 효과가 크다[9]. 하이브리드 클라우드는 공용과 사설을 혼합한 서비스 개념이다.

2.2 클라우드 포렌식(Cloud Forensic)

J. Dykstra and A. T. Sherman(2012)[1]는 클라우드 포렌식에서 수집을 위한 도구를 평가하기 전에, 클라우드 환경에서 신뢰를 이해하는 것에 대한 중요성을 언급하면서, IaaS 클라우드 컴퓨팅에서의 6 개의 신뢰모델을 제시한다. 또한, 클라우드 컴퓨팅 환경에서 포렌식 조사를 수행할 때 선택할 수 있는 방법에 대해 초점을 맞추고 있고, 클라우드 포렌식에서 사용 가능한 도구(Encase, FTK, etc..)들로 데이터를 확보하고 이를 바탕으로 도구에 관하여 측정하고 평가했다.

J. Dykstra and A. T. Sherman(2013)[2]는 위의 논문을 조금 더 구체적으로 구현한 논문으로 OpenStack 클라우드 컴퓨팅 플랫폼에 대한 디지털 포렌식 도구를 구현했다. 이 논문에 따르면 저자는 도구는 게스트 가상 머신 (VM) 또는 하이퍼바이저를 신뢰할 필요 없이 클라우드 공급자의 지원을 필요로 하지 않고 포렌식 데이터에 대한 액세스를 제공하기 때문에 관리 면에서 사용자 중심의 포렌식 기능을 위한 솔루션이라고 말한다. 그러나 클라우드에 있는 데이터의 보존 문제와 OpenStack 의 업데이트에 따른 문제 등을 가지고 있다.

T. Rubsamen and C. Reich(2013)[3]는 클라우드 컴퓨팅 서비스에서 취득할 수 있는 증거를 획득하는 방법과 함께 로깅과 증거를 획득하는 것에 초점을 맞추고 있다.

C. H. Lee(2013)[4]는 서비스별 데이터 수집 및 분석 기법에 대해 연구하여 클라우드 환경을 고려한 디지털 포렌식 프레임워크를 제안하고 있다. 또한 클라우드 컴퓨팅 환경에서의 데이터는 현재 4 개의 기업에서 서비스되고 있는 클라우드 서비스를 분석대상으로 수집했다. 이러한 점은 IaaS 계층으로 제한하고 Cloud Service Provider(CSP)에게 의존하지 않고 조사하려는 우리의 관점과 상이하다.

S. Zawoad and R. Hasan(2013)[5]는 디지털 포렌식의 수사 과정(그림 1)에 대해 설명하면서 클라우드 포렌식에 대한 전체적인 개요를 명시하면서 클라우드 환경에서의 포렌식이 가지는 취약점과 요구사항에 관하여 설명하고 있었다. 이 외에도 클라우드 환경에서의 신뢰성 확보[6]와 클라우드 컴퓨팅 환경의 신뢰성 증가를 위한 보안 로깅[7]에 관하여 많은 연구자들이 언급하고 있었다.

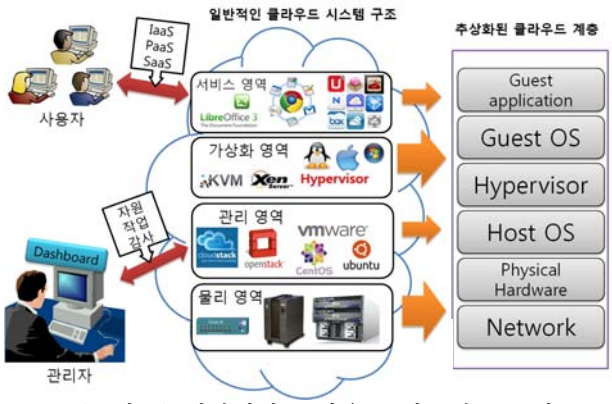
(그림 **1**) 디지털 포렌식 절차

따라서, 본 논문에서는 위와 같은 클라우드 환경의 신뢰성에 대한 이해를 바탕으로 일반적인 클라우드 컴퓨팅 시스템 구성 요소의 추상화된 계층에 따른 역할과 수집 가능한 데이터에 대해 분석한다. 또한, 기존의 포렌식 절차 상의 데이터 수집에 따른 한계를 기술하고, 이를 해결하는 방안으로 클라우드 플랫폼을 고려한 디지털 포렌식 절차에 대해 제안한다.

3. 클라우드 플랫폼을 이용한 디지털 포렌식

3.1 클라우드 환경에서의 수집 가능한 데이터

그림 2 는 일반적인 클라우드 컴퓨팅 구조와 그에 따른 추상화된 클라우드 계층을 보여준다. 각 계층에서는 다른 포렌식 수집 활동이 이루어지며, 계층에 대한 신뢰성과 안정성의 정도가 각각 다르게 요구되



(그림 2) 일반적인 클라우드 시스템 구조와 추상화된 클라우드 계층

기 때문에 계층을 나누어서 분석해야 하고 상위로 올라 갈수록 누적된 신뢰가 필요하다[1]. 공용 클라우드에서 모든 계층은 제공자에 일부 신뢰를 요구하는데, 특히 악의적인 내부자에 대비한 신뢰를 요구한다.

- Network :

클라우드와 같은 복잡한 컴퓨팅 모델에서는 다양한 목적을 가진 사용자들이 참여하고 있기 때문에, 특정 사용자에 의해 이용되는 네트워크 리소스를 모니터링 할 수 있어야 한다. 네트워크 자원은 물리적 또는 가상이 될 수 있는데 이러한 자원들은 사용자 간에 공유가 될 수 있다. 또한, IaaS 에서는 호스트 시스템 하나의 네트워크 카드는 여러 VM 에 의해 이용되고, 다양한 사용자를 포함하고 있다. 각 사용자의 트래픽을 구별하는 것은 책임추적성에 대한 핵심적인 문제이며, 사용자의 네트워크 활동의 흔적이 증거로 사용될 수 있어야 클라우드의 다른 서비스 모델에도 적용 할 수 있다.

- Physical Hardware :

물리적 하드웨어에서는 원본 하드디스크 확보 또는 하드디스크 복제를 통해 저장된 데이터의 복구 및 분석이 필요하고 원본을 보존할 필요가 있을 때 수행된다. 이 때 획득할 수 있는 데이터는 파일 시스템의 메타정보(마지막 수정 시간, 마지막 접근 시간, 생성 시간, 변경 상태 등), 시스템과 응용 프로그램의 로그를 확인할 수 있는데 여기에는 에러 로그, 설치 로그, 네트워크 연결 로그, 보안 로그 등이 있다[8].

- Host OS :

Host OS 에는 클라우드 인프라의 중앙제어에 관한 구성요소, 클라우드 서비스 사용, 액세스 권한, 환경구성, 자원 프로비저닝, 정책, 사용자 로그인 등의 정보를 제공하는 클라우드 플랫폼이 포함되어 있으며 클라우드 관리 시스템이라는 용어로도 사용된다. 가상 디스크에 대한 접근으로 데이터를 획득할 수 있고, 클라우드 플랫폼에서 제공하는 로그파일과 VM 의 스냅샷은 관심 있게 볼 수 있는 정보이다. 또한 관리자는 클라우드 플랫폼을 통해서 자원관리, 작업관리, 감사 등을 수행할 수 있다.

- Hypervisor :

하이퍼바이저에서 데이터의 사용은 IDS(Intrusion Detection System)의 동작을 통해 증거를 획득할 수 있다[3]. 이러한 조사는 하이퍼바이저에 대한 액세스 권한이 필요하기 때문에 IaaS 클라우드 조사에 적합하다.

- Guest OS/application:

인스턴스 내부에서 정보를 얻기 위해서는 원격 포렌식 소프트웨어를 통해 증거를 획득하는 방법이 있으며, VM 내부에 추가로 소프트웨어를 설치해서 VM 내부정보를 획득할 수 있다.

어플리케이션에서는 어플리케이션 로그, 인증 로그를 얻을 수 있으며 multi-tenant 로그 데이터에 대해서는 다중 리소스로부터 분리와 병합이 함께 되어야 한다.

3.2 클라우드 환경에서의 데이터 수집의 한계

데이터 수집에서의 취약점은 포렌식 수사과정과 증거 획득 여부에 결정적인 영향을 끼치기 때문에 취약점을 분석하고 이에 따른 해결방안에 대한 연구가 필요하다.

표 1 은 클라우드 계층에 따른 데이터 수집의 취약점을 보여준다. 클라우드 플랫폼은 서버, 스토리지, 네트워크와 같은 물리 자원을 가상화하여 논리적으로 구성할 수 있도록 하며, 스토리지는 논리적이며 할당된 공간에 초점이 맞추고 있기 때문에 물리적 접근과는 다르게 데이터를 확보해야 한다. 이러한 점은 물리적 디스크를 압수하는 포렌식 수사와는 달리 복잡한 과정을 거쳐야 하며, 신속하게 획득해야 하는 휘발성 데이터에 대한 수집이 어렵고 수사 과정에서 지연을 발생시킨다[4].

저장된 이미지와 스냅샷을 통해 내부 정보를 얻을 수 있지만 CSP 에서 영구저장소를 제공하지 않으면 인스턴스의 종료와 함께 사라질 수 도 있다. 또한, 악의적인 사용자의 강제 종료나 간단한 명령어로 스냅샷을 삭제 시킬 수 있기 때문에 이러한 경우 수집에 어려움이 따른다. 어플리케이션 계층에서는 데이터를 획득할 때 사용자 이벤트가 클라우드 공급자 측에 위치되어 있기 때문에 공급자측의 협조 또는 수사기관의 영장을 통해서만 획득 할 수 있으므로 수사 과정의 지연뿐만 아니라 공급자와 어플리케이션을 신뢰해야 한다.

<표 1> 클라우드 계층에 따른 데이터 수집의 취약점

Cloud Layer	취약점
Guest application	(1) 사용자 이벤트가 공급자측에 위치 (2) 휘발성 데이터
Guest OS	(1) 스냅샷의 보존 문제
Hypervisor	(1) 종류에 따라 데이터에 대한 보존과 형식이 다양
Host OS	(1) 공용 클라우드는 라이브 포렌식 및 휘발성 데이터에 대한 액세스를 허용하지 않음 (2) 클라우드 플랫폼에 따라 로그파일 및 리소스가 분산
Hardware	(1) 서버 시스템의 특성상 하드웨어를 압수할 수 없는 경우가 대부분이기 때문에 원본 획득이 어려움.
Network	(1) 기존의 네트워크 장치 또는 모니터링 솔루션은 multi-tenant 환경에서 증거를 획득하는데 어려움.

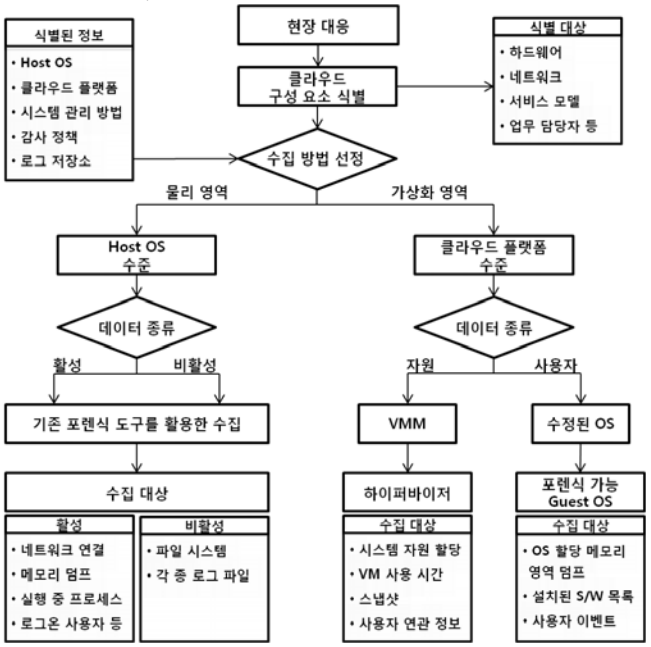
이러한 취약점들은 수사과정에서 매우 중요한 영향을 끼치고 있으며, 증거 획득 여부에 결정적인 역할을 하고 있다. 그러므로 이러한 취약점에 대한 지속적인 연구와 이에 대응하는 해결방안에 관한 연구도 함께 이루어져야 한다.

3.3 클라우드 플랫폼을 고려한 디지털 포렌식

클라우드 환경에서의 기존 디지털 포렌식 수사는 앞서 살펴본 바와 같이 데이터 수집에 따른 취약점이 존재한다. 이러한 취약점을 해결하기 위해 가상 자원들을 서비스화하여 관리 체계를 제공하는 클라우드 플랫폼을 이용해서 접근한다.

클라우드 플랫폼에서 모니터링을 통해 각 사용자에 대한 네트워크 트래픽을 구별하여 이를 통해 특정 사용자에 관한 증거를 획득하고 데이터를 통합 관리하는 데이터베이스에 수집하고 보존한다면 분산되어 있는 데이터에 대한 수집의 한계를 해결할 수 있다. 또한, 영구저장소의 부재에 따른 보존문제와 사용자의 조작, 부팅 또는 강제 종료에 의한 데이터 손실에 대해서는 클라우드 플랫폼에서 영구 저장소에 스냅샷을 보존하여 데이터 손실이 발생하는 것을 방지한다. 사용자 이벤트, S/W 목록과 같은 데이터에 대한 수집은 공인된 포렌식 도구가 설치된 OS 를 사용자에게 제공해서 데이터를 수집할 수 있다.

위에서 언급한 방안들을 반영한 포렌식 절차는 그림 3 과 같이 도식화할 수 있다. 먼저 클라우드 구성요소 식별을 통해 식별대상을 얻고, 식별된 정보를 통해 물리 영역 또는 가상화 영역을 선정한다. 물리 영역은 Host OS 수준으로 일반적인 OS 와 마찬가지로, 기존 포렌식 도구를 이용하여 활성, 비활성에 따라 수집 대상을 수집한다. 가상화 영역의 경우 클라우드 플랫폼 수준에서 데이터의 종류는 자원과 사용자로 나뉘며, 자원과 관련된 데이터는 Virtual machine monitor(VMM) 즉, 하이퍼바이저에서 시스템 자원 할당, VM 사용 시간, 스냅샷을 수집할 수 있으며 이러한 정보와 클라우드 플랫폼의 사용자 관리 정보를 통해 사용자 연관 정보를 수집할 수 있다. 사용자에 대한 데이터는 사용자에게 공인된 포렌식 도구를 설치한 OS 를 제공하여 OS 할당 메모리 영역 덤프, 설치된 S/W 목록, 사용자 이벤트를 수집할 수 있다.



(그림 **3**) 클라우드 플랫폼을 고려한 포렌식 절차

4. 결론 및 향후 연구

본 논문에서는 다양한 클라우드 환경에 유연하게 적용할 수 있는 디지털 증거 수집 절차를 제안하기 위해 추상화된 클라우드 계층에 따라 수집 가능한 데이터를 분류하고, 기존 포렌식 절차 상의 데이터 수집 방법에 관한 한계를 분석하고, 확보한 증거 데이터의 신뢰성 보장을 위해 클라우드 플랫폼 기반의 데이터 수집 절차를 도식화하여 제안했다.

차후에는 앞서 제안한 해결방안을 바탕으로 오픈소스 클라우드 플랫폼인 OpenStack 을 구축하고, 데이터 수집과 보존문제를 해결하기 위해 분산된 자원과 로그를 통합 관리하는 데이터베이스를 구축할 예정이며, 하이퍼바이저에서의 스냅샷과 내부정보에 대한 보존문제에 대해서도 영구 저장소를 통해 해결할 수 있도록 세부사항에 대해서 연구할 예정이다.

참고문헌

[1] J. Dykstra , A .T. Sherman, “Acquiring Forensic Evidence from Infrastructure-as-a-Service Cloud Computing: Exploring and Evaluating Tools, Trust, and Techniques”, *Digital Investigation*, Vol 9, pp.90–98, 2012.

[2] J. Dykstra , A .T. Sherman, “Design and implementation of FROST: Digital forensic tools for the OpenStack cloud computing platform”, *Digital Investigation*, Vol 9, pp.87–95, 2013.

[3] T. Rubsamen1, C. Reich, “Evidence for Accountable Cloud Computing Services”, *Pre-Proceedings of International Workshop on Trustworthiness*, Accountability and Forensics in the Cloud (TAFC), 2013.

[4] C. H. Lee, “클라우드 환경을 고려한 디지털 포렌식 프레임워크”, *한국항행학회 논문지*, Vol 17, No.1, pp.63-38, 2013.

[5] S. Zawoad, R. Hasan, “Digital Forensics in the Cloud”, *The Journal of Defense Software Engineering*, Vol 26, No.5, pp.17-20, 2013.

[6] I. M. Abbadi, J. Lyle, “Challenges for Provenance in Cloud Computing”, *3rd USENIX Workshop on the Theory and Practice of Provenance*, USENIX Association, 2011.

[7] M. M. Potey, D. D. Nikumbh, “Achieving Accountability and Secure Logging to Increase Trust in Cloud Environment”, *International Journal of Computer Applications*, Vol 73, No.17, 2013.

[8] 이상진, “디지털 포렌식 개론”, 이룬, pp.105-137, 2010.

[9] 공용준, 오영일, 심탁길, “실전 클라우드 인프라 구축 기술”, 한빛미디어, pp.22-33 , 2014.