공학석사 학위논문

오픈소스를 이용한 포렌식 클라우드 시스템 프레임워크 설계 및 구현

2014년 2월

부 경 대 학 교 대 학 원

정보보호학 협동과정

김 헌

공학석사 학위논문

오픈소스를 이용한 포렌식 클라우드 시스템 프레임워크 설계 및 구현

지도교수 신 상 욱

이 논문을 공학석사 학위논문으로 제출함.

2014년 2월

부 경 대 학 교 대 학 원

정보보호학 협동과정

김 헌

김헌의 공학석사 학위논문을 인준함.

2014년 2월 21일

주심　　이학박사 이 경 현　(인)

위원　　이학박사 신　　원　(인)

위원　　이학박사 신 상 욱　(인)

차 례

그림 차례

표 차례

Design and Implementation of Open-source based
Forensics Cloud System Framework

Hun Kim

Interdisciplinary Program of Information Security, The Graduate School,
Pukyong National University

Abstract

With the growing use of digital devices such as computer, smart phone or portable data storages, digital information has become a greater priority as evidence for crime investigation. As a result, digital forensics is under the spotlight as the proper alternative to address these issues because it is a branch of forensic science encompassing the recovery and investigation of material which was found in digital devices, often in relation to computer crime. The process of digital forensics are generally divided into the two major tasks: data acquisition and analysis. This sounds so easy but it is not because there are many requirements closely related to the law. Specially, traditional forensic tools and techniques have various limitations considering the rapid changes of IT environment like Big data.

Therefore, this thesis proposes forensics-cloud framework for the forensics tool based on open sources. In this thesis, first, limitations of existing forensics tools is analyzed and considerations for forensics-cloud system is examined. And then, the framework and architecture of the forensics-cloud system is proposed. Also, the usage flow diagram of the proposed forensics cloud tool is described. Finally,

this thesis suggests the implementation result of SaaS model for the proposed system. The proposed framework has scalability and can improve functional aspects and efficiency of the existing forensics tools.

I. 서 론

1. 연구배경

디지털 포렌식은 컴퓨터 범죄가 일어난 대상에 대하여 디지털 증거를 수집 분석하는 일련의 조사 또는 수사과정을 일컫는다. 디지털 장치의 유형에 따라 컴퓨터 포렌식, 네트워크 포렌식, 포렌식 데이터 분석, 모바일 포렌식 등으로 나뉘어져 분류되기도 하며, 일반적으로 사이버 범죄수사, 침해사고 대응 등에 활용되고 있다. 디지털 포렌식을 통하여 수집된 증거는 분석을 통하여 사실 관계를 규명할 수 있으며, 주로 디스크 또는 네트워크를 중심으로 한 연구가 이루어지고 있다[2].

보통의 디지털 기기는 운영체제를 포함하고 있으며, 운영체제는 휘발성 저장매체와 비휘발성 저장매체를 사용한다. 증거 수집은 대상 매체의 운영체제 종료 여부에 따라서 데드 시스템과 라이브 시스템으로 나눈다. 데드 시스템은 증거 분석을 통해 원본 데이터가 변경되는 것을 막기 위해서 원본 데이터의 이미지를 만들게 되며, 원본 데이터의 무결성을 보장할 수 있는 이미징 기술을 요구한다. 라이브 시스템은 휘발성 저장매체에 있는 데이터들을 먼저 획득한 후에, 비휘발성 저장매체에 있는 데이터들을 획득하는 순으로 이루어진다. 이때 라이브 시스템의 증거 데이터를 얻기 위해 내부 명령어를 사용하기보다는 포렌식 도구를 사용해서 데이터를 획득해야 한다[5].

포렌식 도구는 디지털 포렌식 절차를 이해하고 효율적 또는 체계적으로 수행할 수 있도록 하는 독립 또는 통합 도구이다. 일반적으로

포렌식 수사를 위해 필요한 포렌식 도구는 한 가지가 아니기 때문에 수사 절차에 맞춰 각각 도구들의 용도에 맞게 사용해야 할 필요성이 있다. 대부분의 포렌식 도구는 다양한 디지털 저장매체별로 각각 존재하고 있다. 최근까지는 이러한 도구의 개별성이 별 문제가 되지 않았으나, 사건의 복잡성 및 급격한 IT 환경 변화와 신속한 분석을 위한 통합 포렌식 도구의 필요성이 제기되고 있다[3]. 특히 IT 환경의 변화는 기존 포렌식 도구에 많은 영향을 끼치고 있으며, 그 중 저장매체의 대용량화로 인한 디지털 포렌식의 조사 또는 수사시간의 증가 문제는 해결해야할 과제로 남아 있다. 따라서 기존 포렌식 도구에서 이슈가 되고 있는 클라우드 컴퓨팅 기술 요소를 추가하여, 단일 플랫폼으로 동작하는 포렌식의 한계점을 극복하고자 한다. 클라우드 컴퓨팅 환경은 클라우드 OS의 특성으로 디지털 포렌식 업무을 효율적이고 유연하게 처리될 수 있도록 하며, 통합적인 플랫폼으로 유지보수 및 관리가 용이하게 될 것이다. 따라서 디지털 포렌식 기술의 변화 및 발전에 따라 유연하게 대처할 수 있고 확장성까지 갖춘 포렌식 클라우드의 필요성이 대두된다.

본 논문에서는 기존 포렌식 도구를 이용한 조사 및 수사시간의 증가 문제와 새롭게 추가되거나 변화되는 포렌식 기법에 맞춰 오픈소스를 이용한 포렌식 도구의 기능 개선과 확장 가능한 클라우드 환경기반의 포렌식 도구 프레임워크를 제안하고자 한다.

2. 연구 내용 및 구성

본 논문에서는 먼저 디지털 포렌식의 동향에 대해 살펴보고 기존 포렌식 도구의 한계점에 대해서 분석한다. 그리고 디지털 포렌식 도구에서 개선할 수 있는 요소에 대해 살펴보고, 클라우드 컴퓨팅 시스템에 대한 장점을 접목시켰을 때의 고려사항을 알아본다.

기존의 디지털 포렌식 도구와 클라우드 컴퓨팅을 결합시킬 수 있는 기술적 요소에 대해 분석해보고 포렌식 도구의 기능을 클라우드 컴퓨팅과 결합시킨 새로운 포렌식 클라우드 프레임워크를 제안한다. 그리고 서비스될 수 있는 포렌식 클라우드의 구조에 대해 알아보고, 포렌식 클라우드의 서비스 모델별 흐름도를 제안한다.

제시한 포렌식 클라우드 프레임워크를 참조하여 포렌식 클라우드 도구의 프로토타입을 구현함으로서 차후 논의 되어야할 사항들을 기술한다.

본 논문의 구성으로는 2장에서는 관련 연구에 대해 알아보고 3장에서는 포렌식 도구 개선을 위한 기술적 요소에 대해 분석한다. 4장에서는 포렌식 클라우드가 가져야 되는 요소에 대해서 설명하고, 클라우드의 모델 3가지에 맞춰서 각각의 구조 및 흐름도를 제시한다. 5장에서는 기존 디지털 포렌식 도구와 포렌식 클라우드를 비교하여 장점들을 나타내고, 6장에서는 포렌식 클라우드의 핵심적인 요소를 이용한 프로토타입을 구현한 내용에 대해 기술하며 7장에서 결론을 맺는다.

II. 관련 연구

본 장에서는 국내 디지털 포렌식 동향에 대하여 간단히 알아보고 디지털 포렌식 절차에 대해 동작하고 있는 기존 디지털 포렌식 도구의 한계점과 클라우드 컴퓨팅 서비스 전환 시 고려사항에 대해 살펴본다.

1. 디지털 포렌식 동향

가. 국내 디지털 포렌식의 현황

최근 인터폴, 유로폴 등 국제 수사기관을 비롯한 각국 정부들은 디지털 포렌식에 대한 투자와 함께 기술지원에 노력하고 있다. 각 기관은 디지털포렌식전담팀을 신설하고 운영에 나서고 있으며, 미국의 경우 이미 법무부 사법연구원에서 주도적으로 디지털 포렌식에 대한 연구가 이루어지고 있다. 이는 매년 증가하고 있는 사이버위협에 대해 대응하기 위해서 디지털 포렌식 역량 강화는 필수적 요소이기 때문이다. 하지만 국내의 경우는 수사기관에 의해 디지털 포렌식이 연구되어지고 있다. 최근에는 기존에 디지털포렌식센터가 대검찰청 산하의 과학수사 지원 전문기관인 국가디지털포렌식센터(NDFC)로 개소되었지만, 객관적으로 수집 · 분석된 디지털 증거를 검증해줄 수 있는 중립적인 기관은 여전히 부족하며 디지털 포렌식 전문가 수급 또한 부족한 실정이다. <표 1>은 국내 디지털 포렌식 센터 구축 현황[3]을 나타내고 있다.

<표 1> 국내 디지털 포렌식 센터 구축 현황

구 분	도 입 기 관
수사기관	검찰청
	경찰청
사법수사기관	관세청
	전파관리소
조사기관	서울시청 조사과
	공정거래위원회
민간센터	딜로이트안진회계법인
	더존정보보호서비스
	경기대학교 F-Forensics
	고려대학교 센터

나. 디지털 포렌식 수사의 기술적 문제점

현재 일반적으로 사용하고 있는 디지털 포렌식은 과거 개인용 컴퓨터를 대상으로 설정된 절차이다. 따라서 현재 제시되고 있는 포렌식 기술을 그대로 적용하기에는 기술적 한계가 존재한다.

그럼에도 불구하고 국내에서는 아직 일관성 있는 포렌식 모델이나 분석도구 등이 마련되어지지 못하고 있다. 경찰청의 디지털 증거 표준 가이드라인에서도 알 수 있듯이 우리나라는 전문 포렌식 수사관 등의 부족으로 인하여 증거 수집과 분석이 별도로 이루어지고 있기에 우리나라 수사현실을 고려한 포렌식 도구 등이 개발되어야한다[6]. 뿐만 아니라 포렌식 기술이나 도구에 대한 신뢰성 있는 검증도 이루어지고 있지 않다. 만약 디지털 증거를 수집하는 수사관마다 자신의 경험과 기술에 따라 증거를 수집하는 방법과 절차가 제각각 다르다면 수집된 증거에 대한 신뢰성을 확보할 수 없을 것이다. 따라서 보다 체계적이

고 과학화된 디지털 포렌식의 모델과 도구가 개발·육성되어야 한다.

다. 기존 포렌식 도구 기능 비교 및 한계점

디지털 포렌식 도구의 효율적인 기능 개선을 위해서는 기존 포렌식 도구 서비스에 대한 문제점을 인식하고, 기존 디지털 포렌식의 수사 방법을 재분석할 필요성이 있다. <표 2>은 국내외 디지털 포렌식 전문 도구들을 분석하여 기능을 비교한 것이며[3], 이를 종합한 기존 포렌식 도구의 단점은 다음과 같이 4가지로 정리할 수 있다.

• 단일 플랫폼으로 동작하는 포렌식 도구는 머신에 장착된 스토리지, 메모리, CPU에 의존하여 그 이상의 처리 속도를 낼 수 없다.

• 기존 도구를 이용하여 디지털 포렌식을 수행할 시, 도구가 설치된 단일 컴퓨터에서 하나의 도구를 이용하여 업무를 진행하면 긴급 상황에 우선적으로 행해야 하는 작업에 있어 다른 전용의 고속 도구를 이용하여 처리해야 하는 번거로움이 있다.

• 새로운 형태의 데이터에 대한 도구 기능 확장 시, 도구 전용 웹사이트의 절차에 따라 지속적인 도구의 유지 보수 측면에서 일일이 업데이트를 실시해야 한다.

• 디지털 포렌식은 즉시 행해질 수 없고, 도구가 설치된 디지털 포렌식 센터 등 특정 장소에 이동해야 수행이 가능하다.

위 단점들을 보완하기 위해 기존 포렌식 도구에 대한 효율적인 기

술적 방안이 필요하며, 이에 따라 기존 포렌식 절차가 수정되어야 할 수도 있다.

<표 2> 디지털 포렌식 도구별 기능 비교[3]

기 능	G사 제품(해외)	F사 제품(해외)	P사 제품(해외)	F사 제품(국내)
분산처리 기능		•		
Database를 이용한 관리		•	•	
원격 데이터 획득 및 분석		•	•	•
한글 키워드 검색	•	•	•	•
메타데이터 복구(Meta Carving)	▲ 수동	•	•	•
데이터 복구(Data Carving)	▲ 수동	•	•	•
편리한 사용자 UI		•	•	•
인덱싱	•	•	•	
한글 보고서		•	•	•
파일 자동 분류(확장자, 카테고리 등)		•		•
한글 인터페이스	•	•	•	•
데이터베이스 분석		•		•
자동 작업 복구 기능		•		•
다양한 파일시스템 지원	•	•	•	▲
Mac OS X의 원활한 분석	▲	•	•	
암호 해독 기능		•		
XML 형태로 보고서 지원		•	•	
내부확장 개발 기능 지원	•	•	•	
외부확장 개발 기능 지원			•	

III. 디지털 포렌식 도구의 개선을 위한 요구사항 및 고려요소

본 장에서는 기존 디지털 포렌식 도구의 한계점을 극복하기 위해 요구사항들을 분석하고 이를 위한 고려해야 할 법적요소 및 기술적 요소에 대하여 분석한다. 그리고 각 기술적 요소에 대한 자세한 설명을 다룬다.

1. 디지털 포렌식 도구의 기능 요구사항

한국정보통신기술협회(Telecommunications Technology Assocication. 이하 TTA)는 2007년부터 현재까지 디지털 포렌식과 관련된 다양한 국가 표준을 제정하였다. 이를 참조하여 디지털 포렌식 도구의 요구사항에 대해 간단하게 알아본다. 디지털 포렌식 도구의 기능에 대한 요구사항과 관련된 표준은 다음과 같다.

- 컴퓨터 포렌식을 위한 디지털 데이터 수집도구 요구사항 (TTAS.KO-12.0057)[7]
- 컴퓨터 포렌식을 위한 디지털 데이터 분석도구 요구사항 (TTAK.KO-12.0081)[8]

가. 디지털 데이터 수집도구 요구사항

표준에는 디지털 데이터 수집도구의 일반적인 요구사항과 디스크 이미지 생성 기능을 갖는 수집도구의 요구사항 및 쓰기방지 장치가 없는 환경에서의 요구사항으로 구성되어 있다.

디지털 데이터 수집도구는 디지털 데이터 원본으로부터 복사 원본을 생성하는 도구로써 복사 원본의 형태로 디스크 이미지나 복제 디스크 둘중 하나를 생성할 수 있어야 한다. 또한, 디스크 이미지와 복제 디스크 생성 기능을 동시에 가질 수 있다.

<표 3> 디지털 데이터 수집도구 요구사항[7]

구분	요구사항
DA_MG_01	디지털 데이터 수집도구는 획득한 디지털 데이터 원본의 디스크 이미지나 복제 디스크를 생성할 수 있어야 한다.
DA_MG_02	디지털 데이터 수집도구는 디지털 데이터 원본 저장소의 전체 데이터를 수집할 수 있어야한다.
DA_MG_03	디지털 데이터 수집도구는 디지털 데이터 원본 저장소의 일정 부분 데이터를 수집할 수 있어야한다.
DA_MG_04	디지털 데이터 수집도구는 디지털 데이터 원본 저장소의 데이터를 완전하게 수집해야 한다.
DA_MG_05	디지털 데이터 수집도구는 디지털 데이터 원본 저장소의 데이터를 정확하게 수집해야 한다.
DA_MG_06	디지털 데이터 수집도구는 디지털 데이터 획득 과정에서 오류가 발생한다면 오류의 유형과 위치를 사용자에게 알려야 한다.
DA_MG_07	디지털 데이터 수집도구는 데이터 획득 과정에서 해결할 수 없는 오류가 발생한다면 복사 원본 저장소의 해당 위치에 분석 결과에 영향을 주지 않는 값으로 대체해야 한다.
DA_MG_08	디지털 데이터 수집도구는 디지털 데이터 복사 원본 작성시 저장소의 공간 부족을 포함한 기타 오류가 발생한다면 이를 사용자에게 알려야 한다.

*DA : data acquisition, M : mandatory, G : general

디지털 데이터 수집도구가 디스크 이미지 생성 기능을 제공한다면 하나 이상의 디스크 이미지 포맷을 지원해야하며, 디스크 이미지로부터 컴퓨터 파일 시스템으로 접근 가능한 디지털 데이터 원본 저장소와 동일한 형태의 디스크를 복원하는 수단을 제공해야 한다.

데이터 수집 과정에서 원본 데이터의 변경을 막기 위해 쓰기방지 장치(write bolcker)의 사용이 권장되나, 유닉스의 쓰기권한을 해제한 마운팅이나 DOS환경에서는 원본 데이터의 임의적 변경을 야기하지 않으므로 쓰기방지 장치의 사용이 강요되지는 않는다. 그러나 이러한 경우라도 사용자의 실수 등으로 인한 원본 데이터의 변경이 발생할 수 있으므로 쓰기방지 장치가 없는 환경에서 동작하는 디지털 데이터 수집도구는 수집 과정에서 디지털 데이터 원본의 변경이 발생했는지 확인하는 수단을 제공해야 한다.

나. 디지털 데이터 분석도구 요구사항

법정에서 디지털 증거가 효력을 가지기 위해서는 디지털 증거를 도출한 디지털 증거 분석도구의 신뢰성이 보장되어야 하며, 디지털 증거 분석도구가 정확하고 객관적인 결과를 일관성 있게 산출한다는 것을 보장하는 능력이 요구된다.

이를 위해 컴퓨터 포렌식을 위한 디지털 증거 분석도구가 만족해야 하는 일반적인 요구사항은 다음과 같다.[8]

- 유용성(usability) : 복잡도 문제를 해결하기 위해 분석 도구는 추상화 계층에서 데이터와 조사관에게 도움을 주는 포맷을 제공해야 한다. 최소한 조사관은 경계 계층으로 정의된 추상화 계층에 액세스해야한다. 분석도구는 조사관이 데이터를 부정확하게

해석하지 않도록 평문으로 그리고 정확한 포맷으로 데이터를 제시해야 한다.

- 포괄성(comprehensive) : 모든 증거를 식별할 수 있기 위해 조사관이 모든 출력 데이터에 접근할 수 있어야 한다.
- 정확성(accuracy) : 오류 문제를 해결하기 위해 분석 도구는 출력 데이터가 정확하고 결과가 적절하게 해석될 수 있도록 오류 한계가 계산된다는 것을 보장해야 한다.
- 동일성(deterministic) : 분석 도구의 정확성을 보장하기 위해 변형 규칙 집합과 입력이 주어지면 항상 동일한 출력을 산출해야 한다.
- 입증 가능(verifiable) : 분석 도구의 정확성을 보장하기 위해 결과를 입증하는 것이 가능할 필요가 있다. 이것은 수동으로 또는 보조의 독립적인 도구 집합을 사용하여 수행된다. 그러므로 출력이 입증될 수 있도록 각 계층의 입력과 출력에 액세스할 필요가 있다.

요구된 속성에 추가적으로 다음의 특성들이 권고된다.

- 읽기 전용(read-only) : 필수 요건은 아니지만 이 특성은 매우 권장되는 특성이다. 디지털 매체의 특성은 데이터의 정확한 사본을 쉽게 만들 수 있기 때문에, 원본을 수정하는 도구를 사용하기 전에 사본이 만들어질 수 있다. 결과를 입증하기 위해 입력의 사본이 필요하다.
- 건전도 검사(sanity check) : 모든 데이터 값은 추상화 계층에 입력으로 사용될 수 있다. 그러나 단지 몇몇 출력은 유효할 것

이다. 그러므로 조사관은 유효한 출력과 유효하지 않은 출력 사이에 구별할 수 있어야 한다. 데이터 조사관을 지원하기 위해 표현 도구는 건전도 검사를 수행해야 하고 유효한 지를 표시해야 한다.

2. 디지털 포렌식 도구의 법적 증명력 확보

디지털 포렌식 도구의 결과로 나온 증거의 증명력 판단은 결국 법관의 자유심증에 의하는 것이지만, 디지털 포렌식 절차가 일반화되고 신뢰할 수 있으며 정확한 분석이 이루어진다면 법관의 심증형성에 기여 할 수 있으므로 디지털 포렌식 기술과 비교하여 알아본다.

증명력 확보는 포렌식 절차 가운데, 증거물 보관 이송단계, 증거물 분석단계, 보고서 작성 단계가 해당된다.

증거물 보관 이송단계에서는 원본증거가 훼손되지 않아야 하고 보관 상황에 신뢰성이 담보 되어야 한다. 디지털 증거의 경우 대부분 자기의 형태로 저장된다. 따라서 전자기파에 노출되면 훼손될 가능성이 크다. 따라서 전자기파 방지할 수 있는 도구를 사용하여야 한다[4].

증거물 분석 단계에서는 디스크 포렌식의 경우에는 원본이 변경되지 않게 하기 위하여 원본을 사용하지 않고 디스크 이미징을 통하여 증거를 수집하고 분석한다. 원본데이터를 사용할 경우는 쓰기방지장치를 사용하며, 무결성 확보를 위해 해시 함수를 사용하여야 한다. 해쉬함수를 이용하여 원본데이터와 증거분서의 대상이 된 이미지에 대하여 해시값을 비교함으로써 무결성을 확보할 수 있을 것이다.

보고서 작성 단계에서는 법정에 제출하기 위한 것이므로 최대한 전문용어는 피하고 누구나 보아도 알아 볼 수 있을 정도의 쉽게 작성

되어야 할 것이다. 최소한의 신뢰성을 위해 검증된 도구를 사용하였고 전문가에 의해 행해졌으며 법적철자에 따라 행하여지고, 해시에 대한 자세한 설명 이후 해시 값이 동일하여 무결성이 확보된다는 것 등을 강조 하여야 할 것이다.

이러한 절차로 디지털 도구가 증거능력이 인정되었을 때 증거의 가치는 더욱 높아질 수 있고, 판사의 자유심증과정에서 증거의 정확성 및 신뢰성을 확보할 수 있어 최종적으로 범죄사실을 인정하는 증거로 사용될 수 잇을 것이다.

3. 효율적인 디지털 포렌식을 위한 기술적 요소

효율적인 디지털 포렌식을 제공하기 위해서는 다음과 같은 기술적 요소가 제공되어야 한다. 첫째, 대용량 디지털 포렌식 서비스가 가능해야 한다. 이는 디지털 저장매체의 발전으로 인해 1TB 이상의 하드디스크가 보편화되어 있기 때문에 실제 디지털 포렌식 수사에 있어 수집 및 분석시간의 과다소요로 업무 진행에 큰 영향을 끼치기 때문이다. 따라서 디지털 포렌식 도구는 수집 및 분석시간을 줄이기 위해 분산 색인 기능 및 검색을 지원해야 할 것이다.

둘째, 효율적인 색인을 위해 각각의 어플리케이션이 만들어내는 문서의 포맷에 대한 파서가 필요하다. 문서의 포맷은 지속적으로 변화하고 계속적으로 새로운 규격의 포맷이 개발되어지기 때문에 이러한 변화에 맞춰 확장 가능한 형태의 파서가 색인 작업을 보조해야 할 것이다.

셋째, 가상화 기술을 이용하여 원활한 포렌식 환경과 사용자가 원하는 기능으로 이루어진 플랫폼 또는 응용프로그램을 제공할 수 있어

야 한다. 또한 기존 포렌식 도구를 활용하거나 응용프로그램을 개발, 확장할 수 있는 환경이 구축되어야 한다.

넷째, 클라우드 컴퓨팅 기술을 이용하여 일관성 있고 통합된 디지털 포렌식 환경을 제공해야 한다. 다음은 각각의 기술적 요소에 대한 자세한 설명이다.

가. 분산 색인 및 검색, 파싱

분산 색인 기술은 기존의 색인 시간을 줄이기 위해 분산 처리 환경 시스템을 적용한 것을 말한다. 문서의 내용을 색인함으로써 기존 검색 시간을 줄일 수 있지만, 빅데이터의 출현으로 데이터의 규모가 커져 색인 시간의 오버헤드가 발생하는 경우가 생기고 있다. 분산 색인 기술은 이에 대한 방안으로서, 여러 대의 서버에 걸쳐 문서 색인 작업을 나누어 처리해 규모가 큰 데이터도 색인 및 검색 속도를 빠르게 할 수 있는 기술이다.

파싱이란 자료를 원하는 형태로 재가공하는 방식이며 데이터 색인에 있어서 여러 가지 자료 포맷들을 하나의 공통된 형태로 변환시키기 위해 필요한 기술이다. 따라서 기술의 발전으로 새로운 포맷이 생겨나게 되면 파싱의 기능 또한 마찬가지로 추가되어야 한다.

나. 가상화 기술

컴퓨터 운영체제(Operationg System)를 시스템 구조나 하드웨어에 영향을 받지 않고 설치, 사용할 수 있도록 하는 기술이다. 일반적으로 운영체제는 특정 시스템 구조나 하드웨어에 특화되어있어 운영체제의 교체가 쉽지 않으며, 하나의 시스템에서 여러 운영체제를 동시에

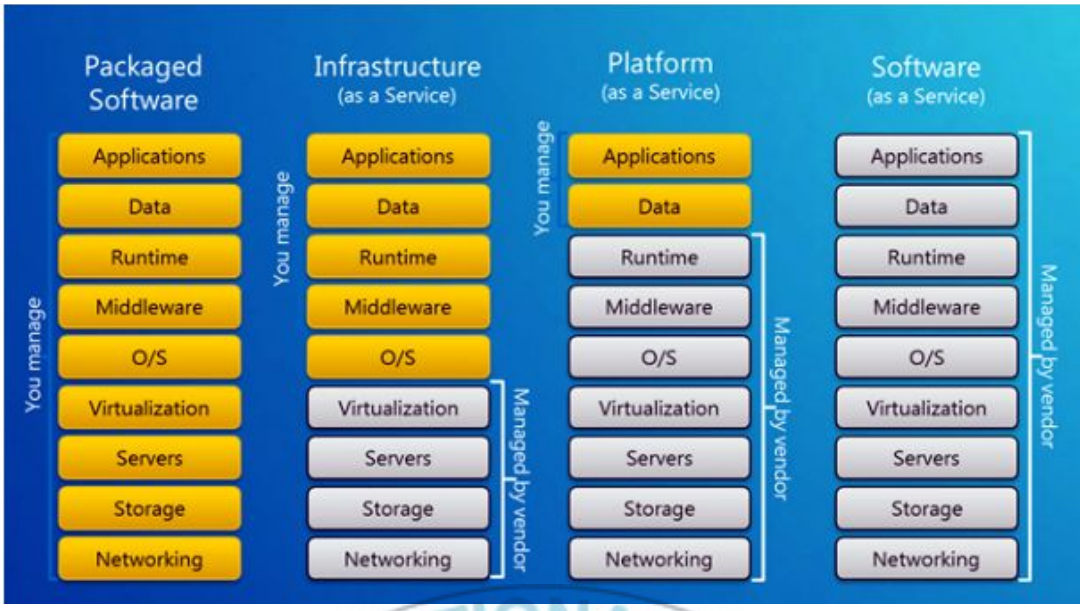
운영하는 것도 거의 불가능하다. 그러나 다양한 업무 수행을 위해서는 하나의 시스템에 여러 운영체제를 얹거나 운영체제를 교체하여 낡은 컴퓨터를 재활용하는 기술이 필요하며, 특히 최근에는 서버나 PC 수준에서도 이러한 기능이 절대적으로 요구되고 있다. 대표적으로는 VM Ware나 MS의 가상서버 기능이 있으며, 인텔 및 AMD 등이 가상화 기술을 지원하는 칩을 개발하고 있다. 컴퓨터에서 가상화 기술을 사용하는 방법은 운영체제 위에 가상머신 지원 프로그램을 사용하는 방법과 가상머신 위에 운영체제를 올리는 방법 등이 있다.

다. 클라우드 컴퓨팅

클라우드 컴퓨팅 기술은 네트워크로 연결된 서버 / 클라이언트 방식의 정보처리에서 클라이언트의 리소스를 사용하지 않고, 서버의 리소스를 사용하여 처리하는 기술을 의미한다. 클라우드 서비스는 IT의 모든 것을 서비스 형태로 제공하는 것이며, 서비스 제공 형태에 따라서 SaaS, PaaS, IaaS로 분류할 수 있다[1].

- SaaS(Software as a Service)

클라우드 환경에서 동작하는 응용프로그램을 서비스 형태로 제공하는 것을 말한다. 사용자는 사용하는 시스템의 구조와 필요한 기능을 모두 Contents Service Provider(이하 CSP)에게 제공받고, 서비스에 대한 비용을 지불하는 것이 일반적인 형태이다. 사용자는 CSP가 제공하는 응용프로그램만 이용할 수 있으며, CSP는 서비스 제공에 대한 데이터 관리, 백업 등을 담당한다.



[그림 1] 클라우드 컴퓨팅의 서비스 모델

- PaaS(Platform as a Service)

서비스를 개발할 수 있는 안정적인 환경(Platform)과 그 환경을 이용하는 응용프로그램을 개발할 수 있는 API까지 제공하는 형태이다. 사용자는 CSP가 제공하는 환경을 이용하여 구축할 수 있지만, CSP가 정의한 방식으로 개발이 한정된다.

- IaaS(Infrastructure as a Service)

서버를 운영하기 위해서는 IP, Network, Storage, 전력 등의 인프라 구축이 필요하다. IaaS 서비스는 이러한 인프라를 가상의 환경에서 쉽고 편하게 이용할 수 있도록 서비스 형태로 제공하는 것이다. IaaS를 서비스로 제공하기 위해서는 기존 서버 호스팅보다 H/W 확장성과 신속성을 요구하게 되는데, 이는 가상화 기술을 이용하여 제공된다.

라. Open Source

- Lucene & Solr[9][10]

가장 널리 사용되는 오픈소스 검색 엔진 중 하나인 Apache Lucene과 Solr는 대용량 데이터를 지원하며 분산 색인 아키텍처, 실시간 결과 출력 기능을 가지고 있다. 최근 4.x 업데이트로 인하여 빅데이터에 관한 기술이 추가되었고, Solr 컴포넌트에는 분산 인덱싱(distributed indexing)이라는 새로운 기술이 추가되었다. 분산 인덱싱은 여러 대의 서버에 걸쳐 문서 색인 작업을 나누는 기능으로, 데이터 규모가 커져도 검색 속도를 빠르게 할 수 있다. Solr는 이를 위해 정보 색인 작업을 다중 쓰레드로 처리하고 동시에 디스크에 기록도 할 수 있다.

Solr Cloud는 이러한 Solr Core들을 클러스터 서버 환경에 설치하여 분산 색인을 가능하게 한다.

- Tika[11]

Apache 프로젝트 중 하나로서, 기존 파서들을 이용하여 다양한 문서로부터 텍스트나 메타 데이터를 추출하는 툴킷이다. 비슷한 관련 프로젝트로는 POI가 있으며, 지원하는 모든 파일 형식들을 위한 문자 추출 라이브러리를 제공한다. Tika와 POI 라이브러리는 밀접하게 동작하며 문서 파싱에 대한 기능들을 활용할 수 있다.

● Cloud Platform

(1) Open Stack[12]

오픈스택은 IaaS(Infrastructure as a Service)형태의 클라우드 컴퓨팅 오픈소스 프로젝트이다. 2012년 창설된 비영리 단체인 OpenStack foundation에서 유지 보수하고 있으며 아파치 라이센스 하에 배포된다. 오픈스택의 구조는 프로세싱, 저장 공간, 네트워킹의 가용자원을 제어하는 목적의 여러 개의 하위 프로젝트로 이루어져 있다. 대시 보드 프로젝트는 다른 하위 프로젝트의 운영 제어를 웹 인터페이스를 통해 담당한다.

(2) Cloud Stack[14]

오픈스택과 마찬가지로 IaaS 솔루션으로써 Public, Private, Hybrid 클라우드 인프라 스트럭처를 관리하기 위하여 시트릭스 사에서 제공되는 오픈소스 소프트웨어이다. 클라우드 스택은 클라우드 인프라 스트럭처를 구성하는 네트워크, 스토리지, 컴퓨터 노드를 관리한다.

4. 클라우드를 접목한 디지털 포렌식 도구의 요구사항

디지털 포렌식 도구를 클라우드화 하였을 경우, 포렌식 목적에서 고려해야할 요구사항들은 기존 도구와 차이가 없을 것이다. 기존의 디지털 포렌식 도구의 기능들을 클라우드 요소에 올려서 사용하여 단일 플랫폼 기반의 한계점을 극복하기 위한 것으로 기능 자체의 변화는 없을 것이며, 클라우드화 하였을 경우의 디지털 증거의 수집에 대한 부

분에 대해 새로운 요소가 추가되어 이에 대한 검증이 필요하게 될 것이다. 디지털 증거 확보를 위해 증거 이미지를 생성하는 경우 기존 도구와 마찬가지로 클라우드 상에서 해쉬값을 이용한 원본과의 무결성 검사가 이루어질 것이기에 증거능력이 보장될 것이며, 이에 대한 분석 결과 또한 마찬가지 일 것이다. 따라서 기존 포렌식 도구처럼 요구사항을 만족하며, 보고서 작성 단계에서 법관과 소송 당사자들이 이해하기 쉽게 작성되고, 동일성, 무결성 등이 확보되었다는 것을 강조한다면 포렌식 클라우드의 효용성 및 필요성이 대두될 수 있다.

기존의 어플리케이션을 클라우드 컴퓨팅으로 이동시키기 위해서 어플리케이션 자체가 클라우드 플랫폼으로 이동 및 개발 또는 호스팅하기 적합한 어플리케이션이 있는가 하면, 클라우드를 사용하기에 부적합한 어플리케이션도 있다. 따라서 특정 어플리케이션을 클라우드로 운영하는 것에 대해 실용적인가를 먼저 결정한 후 클라우드를 운영한다. 대부분 EnCase나 FTK와 같은 도구들을 클라우드 컴퓨팅 서비스의 VM에 바로 접목시키면 별다른 수정 없이 클라우드 서비스에서 안정적이고, 확장성을 가지며 손쉽게 비용 절감과 같은 특징들을 이룰 수 있을 것이라 생각하지만, 클라우드에서 어플리케이션을 수정 없이 구동하는 것은 아무런 특징을 가질 수 없다. 일반적으로 단일 컴퓨터에서 동작하도록 디자인된 어플리케이션은 여러 대의 컴퓨터로 된 인프라에서 사용 가능한 확장성이 배제되어 있다[13].

본 논문에서 설계하는 포렌식 클라우드는 이러한 법적요소 및 보안적인 요소에 대하여 미리 정의되어 있고 가정된 상황에서의 디지털 포렌식 클라우드의 기술적인 요소에 대해 프레임워크를 설계하며 다룬다.

IV. 오픈소스 기반의 포렌식 클라우드 프레임워크 설계

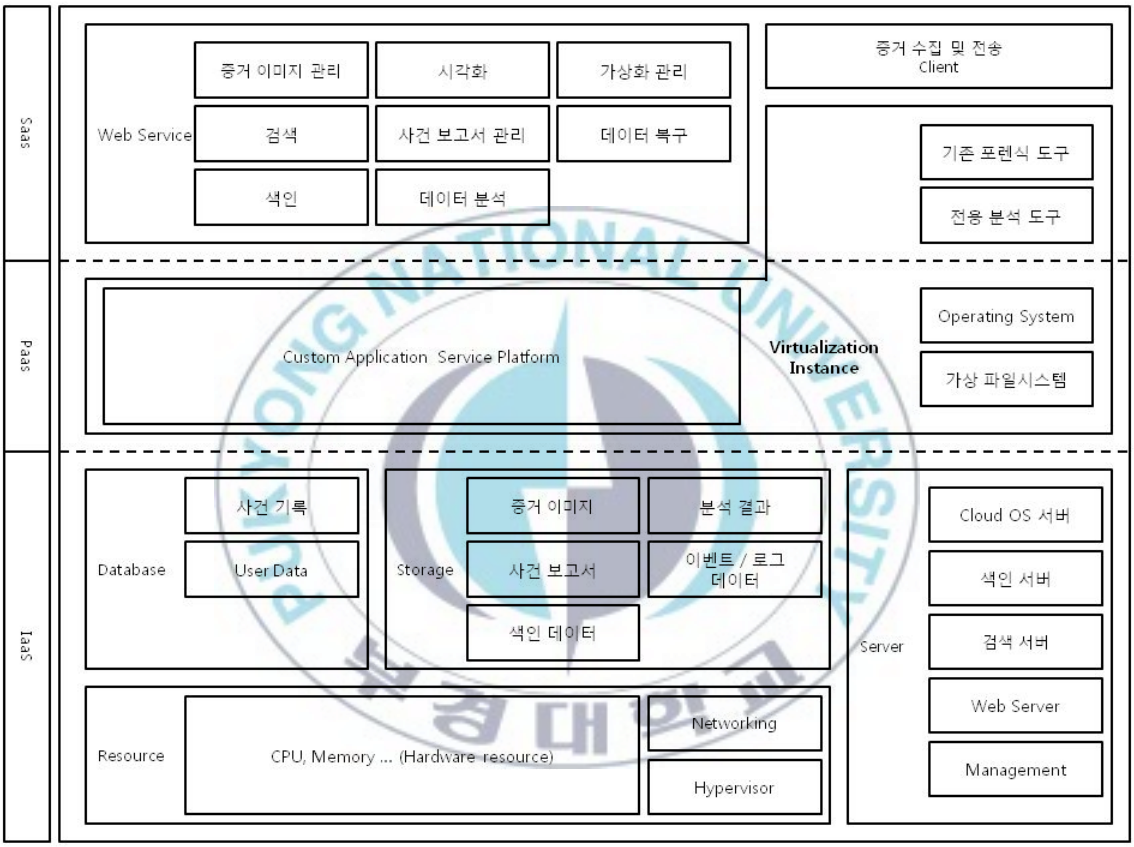
본 장에서는 오픈소스 기반의 포렌식 클라우드 프레임워크를 설계하기 위해, 기존 클라우드 컴퓨팅 모델에서 포렌식 서비스를 접목시켰을 경우 먼저 고려해야 할 사항들을 살펴본다. 그리고 디지털 포렌식 절차에 따른 기능들을 축약하여 포렌식 클라우드의 프레임워크를 제시하며, 오픈소스를 이용하여 포렌식 클라우드 서비스로 만들 수 있는 도구의 구조와 흐름도를 설명한다.

1. 포렌식 클라우드 정의

포렌식 클라우드는 기존 디지털 포렌식에 필요한 모든 기능 또는 환경을 서비스로 제공하는 형태를 말한다. 포렌식 클라우드가 제공하는 서비스는 사건 Case별로 사건 관리, 증거 수집, 증거 분석, 보고서 작성 등 디지털 포렌식 절차에 맞춰 제공한다. 이는 각 조사관, 분석관, 수집관, 기타 관련자들이 단일 시스템으로 업무를 수행하는 것보다 더 빠르고 효율적으로 수행하기 위한 통합적인 디지털 포렌식 환경을 제공하는 것을 기대할 수 있다.

기존 클라우드 서비스 모델이 각 목적에 따라 3가지로 분류되어 제공되는 형태처럼 포렌식 클라우드도 디지털 포렌식의 기능 및 환경에 따라 분류되어야 할 필요성이 있다. 사용자는 단순하게 디지털 포렌식 기능을 클라우드로 제공받기 위해서는 SaaS 모델이 될 것이고, 사용자가 커스터마이징하여 원하는 디지털 포렌식 기능 및 환경들을

제공하기 위해 PaaS의 형태가 되어야 할 것이다. 또한 포렌식 클라우드 인프라 자체를 서비스 받는 경우도 있을 수 있다. 따라서 포렌식 클라우드는 기존 클라우드 컴퓨팅의 안정적이고, 확장성을 지닌 특성과 함께 더 나은 디지털 포렌식 기능 및 환경을 모두 충족 할 수 있어야 할 것이다.



[그림 2] 포렌식 클라우드 프레임워크

2. 포렌식 클라우드 프레임워크

포렌식 클라우드 정의에 따라서 설계한 포렌식 클라우드 프레임워크[그림 2]는 디지털 포렌식 절차 및 기능을 축약하고 클라우드를 접목시킨 프레임워크이다. 그리고 앞 절에서 설명했다시피 기존 클라우

드의 정의에 따라 포렌식 클라우드 프레임워크를 IaaS, PaaS, SaaS로 나누어 분류를 한다. 다음은 각 부분들에 대한 세부적인 설명이다. 제안된 포렌식 클라우드에서는 보안요소에 대한 것들이 빠져 있으며, 3.4절에서 설명했다시피 보안요소에 대한 것들은 미리 정의되어 있다고 가정한 후 기능적 요소에 대한 부분들만 고려하여 포렌식 클라우드 프레임워크를 제안한다.

가. 포렌식 클라우드의 IaaS 부분

프레임 워크의 IaaS 부분은 포렌식 클라우드 구축을 위해서 필요한 인프라 서비스이다. 아래 <표 3>은 포렌식 클라우드의 IaaS 서비스를 위해서 필요한 기본적인 서비스에 대한 설명이다.

Resource는 클라우드 컴퓨팅 환경이 구성되기 위한 기본적인 서버 하드웨어 자원 및 기타 자원 등으로 이루어져 있다. 포렌식 클라우드는 원활한 업무 환경 및 확장성을 고려하여 여러 대의 서버 컴퓨터로 이루어져야 할 것이다. CSP는 각 서버 컴퓨터의 하드웨어 자원에 따라 역할을 정하고 관리하여야 할 것이다. 네트워킹 자원과 가상화 자원은 클라우드 서비스에 있어서 가상화 인스턴스를 생성하고 관리하기 위해 필수적인 요소이다. 가상화 자원은 다수의 OS를 하나의 컴퓨터 시스템에서 가동할 수 있게 하는 자원이며, 가상화 자원을 이용하여 생성된 가상화 인스턴스는 네트워크 자원을 이용하여 클라우드 서버의 내외부와 통신을 가능하게 한다.

Server 부분은 포렌식 클라우드 서비스를 제공하기 위한 관련 서버의 그룹이다. 클라우드 OS는 클라우드 컴퓨팅을 가능케 하는 운영체제로써 가상화 자원 및 생성된 가상화 인스턴트들을 관리하는 운영 소프트웨어이다. 이외에도 네트워크 자원을 이용하여 내부 사설망을

구축할 수 있도록 하며, 웹 환경을 통하여 서비스를 이용 또는 관리하기 위한 대시보드 및 웹 서버 기능을 내포하고 있다. 색인 서버와 검색 서버는 디지털 포렌식 수사에 있어서 포렌식 클라우드에 업로드 된 증거 이미지 또는 증거 자료들을 색인하여 빠르고 효율적이 검색 할 수 있도록 전문화 된 서버이다. 색인 서버는 색인 데이터를 별도의 스토리지에 저장하여 사건 Case별로 관리하게 되며, 스토리지에 저장된 색인 데이터를 사건 관련 사용자가 검색 서버에 접속하여 정보를 검색할 수 있도록 한다.

Storage는 증거 이미지 또는 자료, 색인 데이터, 분석 결과, 사건 보고서 등 포렌식 클라우드를 이용하여 생성된 모든 자료들이 저장되는 곳이며, 모든 사용자의 이벤트 및 로그 기록을 보관되어야 할 것이다.

Database는 클라우드 시스템의 사용자의 권한이나 정보를 저장하거나 기타 반영구적으로 저장되어야 하는 정보들을 가지고 있어야 한다.

<표 4> 포렌식 클라우드의 IaaS 부분

구분	서비스 내용	상세 설명
Resource	기본적인 서버 자원	클라우드 시스템을 위한 기본적인 하드웨어 자원.
	네트워킹 자원	클라우드 시스템 내의 가상 네트워크 형성을 위한 자원.
	가상화 자원	클라우드의 가상화 인스턴스를 생성하기 위한 자원.
Server	Cloud OS	클라우드 컴퓨팅 및 시스템 구축을 위한 운영체제.
	색인 서버	클라우드 환경에서 동작하는 색인 및 검색 서버. 처리할 데이터의 크기에 따라 단일 또는 분산으로 수행됨.
	검색 서버	
	웹 서버	클라우드 접속 GUI를 제공하기 위한
	관리 서버	PaaS의 Service들과 기타 서버들을 통괄하기 위한 관리 서버
Storage	증거 이미지	증거 이미지 저장
	색인 데이터	증거 데이터의 색인 결과 저장
	분석 결과	증거 분석 결과 저장
	사건 보고서	사건 보고서를 목록 화하여 저장
	이벤트 / 로그 데이터	클라우드 서버의 이벤트 / 로그 데이터 저장
Database	사건 기록	사건 Case 별로 정보 저장
	사용자 데이터	클라우드 시스템 사용자 정보 및 접속 권한 등을 저장

<표 5> 포렌식 클라우드의 PaaS 부분

구분	서비스 내용	상세 설명
가상화 인스턴스 플랫폼	운영체제	가상화 인스턴스에 생성될 수 있는 운영체제(ex. windows, linux 등)
	가상 파일 시스템	업로드 된 증거 이미지를 바로 검색하지 않고 복구하여 데이터를 색인하기 위한 이미지 복구용 파일 시스템. IaaS의 색인 서버는 이미지가 복구된 가상 파일시스템을 빠르게 색인하여 사용자가 검색 및 데이터를 분석할 수 있도록 함.
사용자 정의 폴랫폼	사용자 정의 어플리케이션	응용프로그램 개발 API 및 개발 환경 제공. 응용프로그램 커스터마이징. 사용자가 직접 플랫폼을 구성하여 디지털 포렌식 업무를 수행할 수 있도록 함.

나. 포렌식 클라우드의 PaaS 부분

포렌식 클라우드의 PaaS 부분은 사용자가 포렌식 클라우드에서 제공되는 서비스를 커스터마이징할 수 있도록 플랫폼을 서비스로 제공한다. 사용자는 포렌식 클라우드의 SaaS 부분에서 클라우드 API를 제공 받아 사용자에게 필요한 새로운 응용프로그램을 개발 할 수 있는 환경을 구성하거나 개발한다. 이러한 작업은 IaaS에서 제공되는 인프라를 기반으로, 생성된 가상화 인스턴스에서 작업을 수행하게 된다. <표 4>는 PaaS부분의 기능을 요약한 것이다.

가상화 인스턴스는 사용자의 목적에 따라 운영체제가 인스톨 되거나 증거 이미지를 복원한 하나의 가상 파일 시스템이 될 수 있다. 운영체제가 인스톨 된 인스턴트는 PaaS가 제공하는 환경 안에서 개발 또는 디지털 포렌식을 수행할 수 있다. 클라우드의 PaaS 모델은 사용자에게 좀 더 유연한 디지털 포렌식 환경을 제공할 수 있도록 설계 되어야 할 것이다.

다. 포렌식 클라우드의 SaaS 부분

포렌식 클라우드의 SaaS 부분은 클라우드의 웹 환경을 이용하여 디지털 포렌식 도구의 기능을 사용자에게 제공하는 것이다. 사용자는 원하는 디지털 포렌식 도구의 기능을 웹 환경에서 바로 이용할 수 있거나, 가상화 인스턴트에 접속하여 디지털 포렌식을 수행할 수 있다. 따라서 CSP는 사용자에게 기본적인 디지털 포렌식 도구의 기능을 웹 또는 어플리케이션을 제공할 수 있어야 할 것이다. <표 6>는 SaaS 부분의 기능을 요약한 표이다.

웹 서비스는 포렌식 클라우드를 이용하는 사용자가 수사 및 조사, 분석에 관한 모든 작업을 할 수 있도록 서비스 형태로 제공한다. 포렌식 클라우드는 여러 개의 사건들을 통괄하므로 각 사건 Case별로 관리할 수 있는 기능을 가지고 있어야 한다. 기능별로는 증거 이미지 업로드, 관리, 데이터 복구 및 분석, 데이터 색인 및 검색, 시각화, 사건 보고서 관리 등 디지털 절차별 기능이 모두 CSP에 의해서 제공되어야 할 것이다. 필요에 따라서 사용자는 CSP가 제공하는 서비스가 아닌 기존 포렌식 도구를 사용하여야 하는 상황이 발생할 수 있다. 이때는 사용자가 원하는 운영체제가 인스톨된 가상화 인스턴트를 생성해 기존 포렌식 도구를 사용할 수 있는 환경을 서비스 받게 될 것이다.

전용 분석 도구는 웹 환경에서 표현하기 어렵거나 제공하기 힘든 기능들을 가상화 인스턴트 내의 어플리케이션 형태로 제공하는 도구를 말한다. 이로써 사용자는 웹 환경과 가상화 환경을 디지털 포렌식 업무 특성에 맞추어 포렌식 클라우드를 사용하여야 할 것이다.

포렌식 클라우드는 클라우드 특성상 서버의 자원을 사용하여 디지털 포렌식 업무가 수행되지만 필요에 따라서 별도의 Client부분이 필요할 수도 있다. 예를 들어 증거이미지를 클라우드에 전송하는데 있어서 웹 환경의 접속이 원활하지 않거나 가상화 인스턴트에 접속하기 위한 별도의 어플리케이션이 필요한 상황이거나, 기타 예외적인 상황을 고려할 수 있다. 따라서 포렌식 클라우드의 기능을 보조하는 역할의 Client가 존재하여야 하며, CSP 및 포렌식 클라우드 사용자의 필요에 의하여 기능 및 구조는 바뀔 수 있다.

<표 6> 포렌식 클라우드의 SaaS 부분

구분	서비스 내용	상세 설명
Web Service	색인 및 검색	증거 이미지를 복구하여 색인 후 사용자가 검색할 수 있는 UI를 제공
	증거 이미지 관리	클라우드에 업로드 되는 증거 이미지들을 리스트를 작성하여 관리
	데이터 분석	색인된 데이터를 웹 인터페이스에서 분석할 수 있도록 UI를 제공
	데이터 복구	추출된 데이터를 복구 하거나 증거 이미지를 가상 파일시스템에 복구
	시각화	데이터 시각화 및 기타 기능 제공
	사건 보고서 관리	사건 Case 별로 보고서를 작성 및 관리하는 기능 제공
	가상화 관리	분석관, 수집관이 필요시 가상화 인스턴스를 사용하여 작업을 수행할 수 있도록 가상화 인스턴스 생성 및 삭제, 또는 관리할 수 있는 기능 제공
가상화 인스턴스	기존 디지털 포렌식 도구 사용	OS가 인스톨되어 있는 가상화 인스턴스에서 기존에 사용하던 포렌식 도구를 활용하여 디지털 포렌식을 수행할 수 있도록 함.
	전용 분석 도구	웹 서비스에서 제공할 수 없는 분석 기능을 구현하여 가상화 인스턴스 내의 어플리케이션 형태로 제공
기 타 Client	증거 수집 및 전송	클라우드 내의 스토리지로 디지털 포렌식 데이터 및 기타데이터를 전송하기 위한 전용 클라이언트 어플리케이션. 웹 UI를 이용하여 전송 할 수 있으나, 상황 또는 취급하는 데이터에 따라 활용 될 수 있도록 어플리케이션 형태의 클라이언트를 제공

3. 오픈소스 기반의 포렌식 클라우드 구조

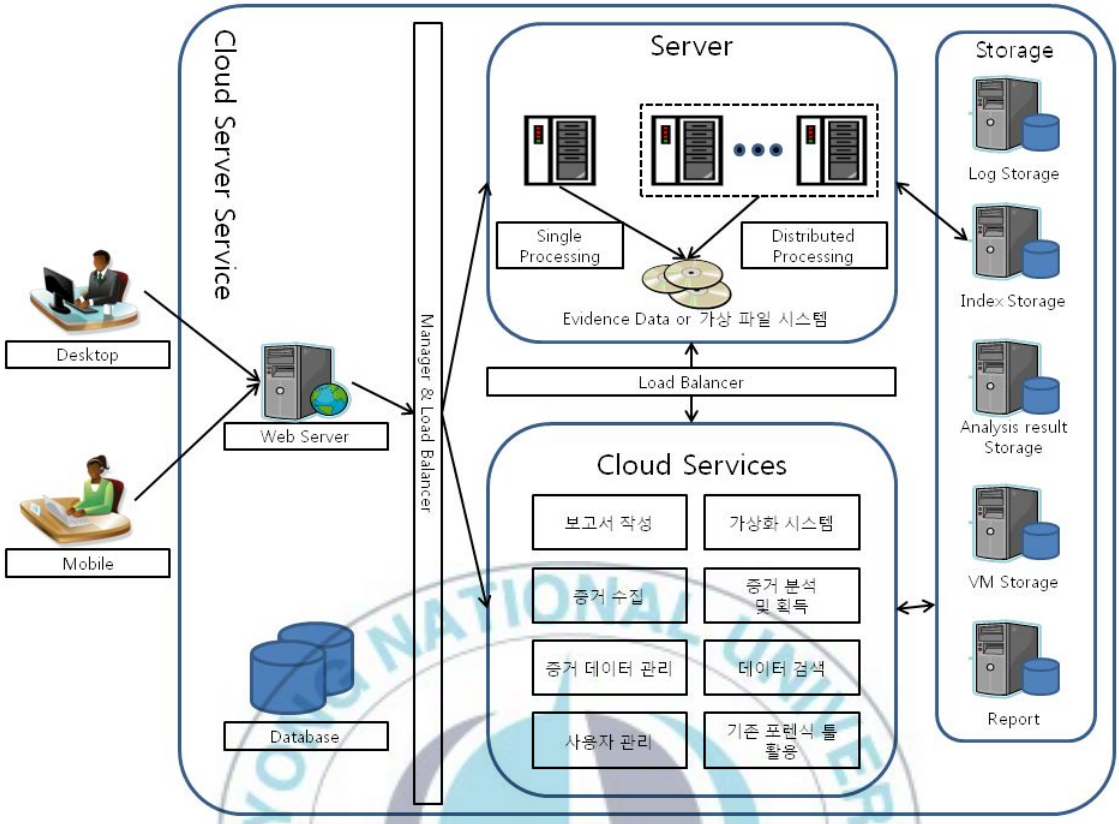
포렌식 클라우드 서비스는 웹 서버를 이용하여 인증된 사용자의 직무에 맞게 웹 환경에서 클라우드 서비스를 제공 받는다. 포렌식 클라우드는 접속한 사용자의 디지털 포렌식 수행 규모에 따라서 자원을 할당하여 준다. 따라서 대용량 디지털 포렌식 수행 업무에 있어서 포렌식 클라우드는 클라우드 기능 외의 분산 처리 시스템을 가지고 있어야 하며, 주로 분산처리에 이용되는 부분은 색인 부분과 검색 부분, 데이터 분석 부분이 되어야 할 것이다. [그림 3]은 포렌식 클라우드 프레임워크를 기반으로 구성된 도구의 전체적인 구조를 나타낸다. 도구의 기본 구조는 기존 클라우드 컴퓨팅 시스템 구성요소를 제외한 5가지의 요소로 나눌 수 있다.

가. 웹 서버 군

클라우드에 접속하는 사용자는 데이터베이스에 저장된 정보를 이용하여 사용자 인증 또는 공인인증서를 이용한 인증, 기타 인증 모듈 등을 이용하여 접속할 수 잇는 기능이 있으며, 매니저 및 로드밸런서에 의하여 접속 된 사용자의 관리 및 사용자의 직무에 맞는 기능을 제공하게 된다.

나. 매니저 및 로드 밸런서 군

웹 인터페이스 또는 가상화 인스턴스에 접속하기 위한 전용 클라이언트로 접속한 사용자들에게 기능을 원활화게 제공하기 위한 조정 및 관리를 담당한다.



[그림 3] 포렌식 클라우드 서비스의 구조

다. Server 군

포렌식 클라우드의 핵심 기능으로 증거 수집 및 분석을 위한 서버이다. 주로 색인 및 검색, 분석, 처리 등을 담당하며, 서버는 여러 대의 분산 클러스터 환경으로 구성되어 있다. 서버는 사건 Case의 데이터 규모에 따라서 단일 작업이 이루어지거나, 규모가 큰 경우 분산 처리 작업이 이루어질 수 있다. 또한 데이터 규모가 큰 색인의 경우 색인 데이터가 나누어져 스토리지에 저장될 수 있다. 따라서 색인 데이터의 관리를 효율적으로 수행하기 위해 클라우드 서버가 유휴상태일 경우 여러 개로 분할된 색인 데이터를 병합하여 사건 Case별로 관리되어질 수 있도록 한다.

라. 클라우드 서비스 군

기존의 디지털 포렌식 도구들의 기능들을 클라우드용 어플리케이션 형태로 서비스하기 위한 부분이다. 대표적으로 디지털 포렌식 절차에 따라 필요한 기능 및 클라우드 자원을 관리하는 기능이 존재하며, 기능 추가 및 업데이트가 효율적으로 이루어질 수 있도록 각 기능들은 어플리케이션 단위로 나누어져 있다. 클라우드 서비스 군의 기능들은 기본적으로 웹 형태로 제공되며, 웹에서 표현하기 힘든 기능들은 가상화 인스턴스를 이용하여 가상화 OS 내에서 사용될 수 있도록 환경이 구성되어야 할 것이다.

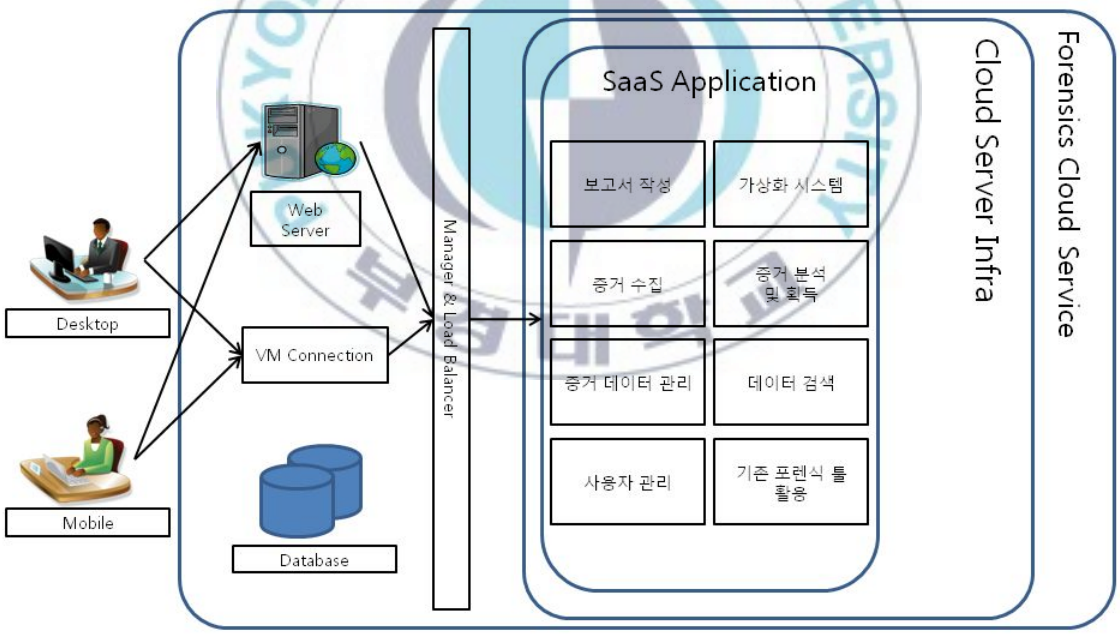
마. 스토리지 군

포렌식 클라우드의 저장 공간은 디지털 포렌식 대상의 취급 데이터 종류에 따라서 분할되어 관리되어야 한다. 데이터의 종류 별로 저장되지만 하나의 스토리지 공간 안에 여러 사건 Case들의 데이터가 모두 저장되어 있기 때문에 각 Case별로 관련 사용자만 접근 할 수 있도록 접근 권한 관리 및 보안 기능을 가지고 있어야 할 것이다.

4. 포렌식 클라우드 사용 흐름도

가. SaaS 모델의 흐름도

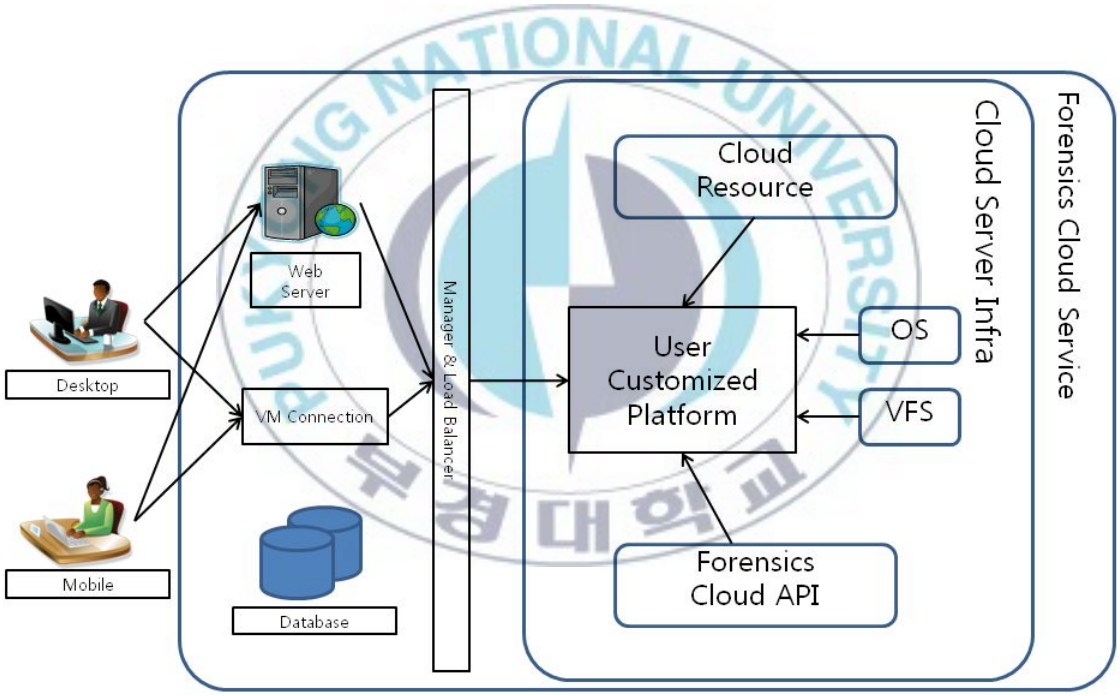
[그림 4]는 SaaS 모델의 흐름도를 나타낸다. 포렌식 클라우드의 SaaS 이용자는 웹 서버 또는 가상화 인스턴스 접속 클라이언트를 통하여 사용자 인증을 거치게 된다. 사용자 인증은 데이터베이스에 저장된 사용자 정보 또는 별도의 인증 모듈을 이용한다. 접속한 사용자는 매니저 및 로드 밸런서에 의해서 CSP에 의해 제공되는 포렌식 클라우드 어플리케이션 서비스를 그대로 제공 받는 흐름이 되겠다.



[그림 4] 포렌식 클라우드 SaaS 흐름도

나. PaaS 모델의 흐름도

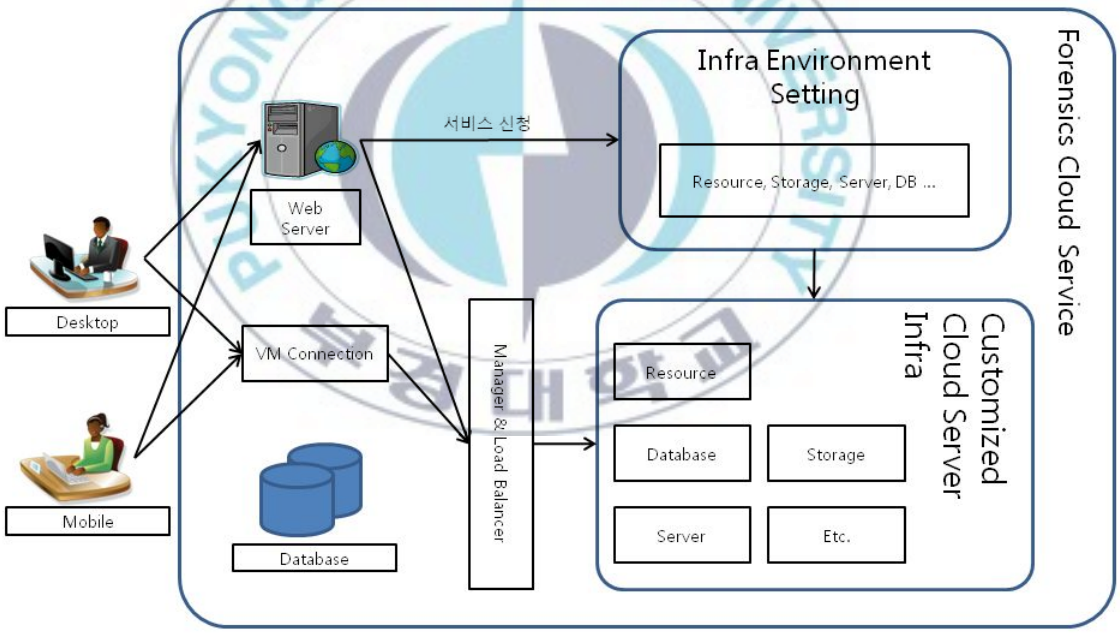
포렌식 클라우드 PaaS 모델 흐름도는 SaaS 모델과 동일하게 사용자 인증을 거친 다음, 포렌식 클라우드에서 사용자가 원하는 디지털 포렌식 플랫폼을 서비스로 제공 받는다. 이러한 사용자 정의 플랫폼은 클라우드 리소스, 클라우드용 디지털 포렌식 도구 개발 API를 제공하며, 운영체제 및 가상 파일시스템을 이용하여 다양한 디지털 포렌식 개발 도구 및 환경을 구성할 수 있다. [그림 5]은 PaaS 모델의 흐름도를 나타낸다.



[그림 5] 포렌식 클라우드 PaaS 흐름도

다. IaaS 모델의 흐름도

포렌식 클라우드 IaaS 모델은 사용자가 포렌식 클라우드의 인프라를 이용하기 위해서 서비스 신청 단계가 필요할 것이다. 사용자가 원하는 인프라 환경을 구성하고, 구성된 인프라를 기반으로 포렌식 클라우드를 구축한다. CSP는 포렌식 클라우드 구성에 필요한 인프라 및 클라우드에서 필요한 디지털 포렌식의 기능들을 제공해야 할 것이며, 인프라의 전반적인 유지 보수 및 관리를 담당한다. [그림 6]은 IaaS 모델의 흐름도를 나타낸다.



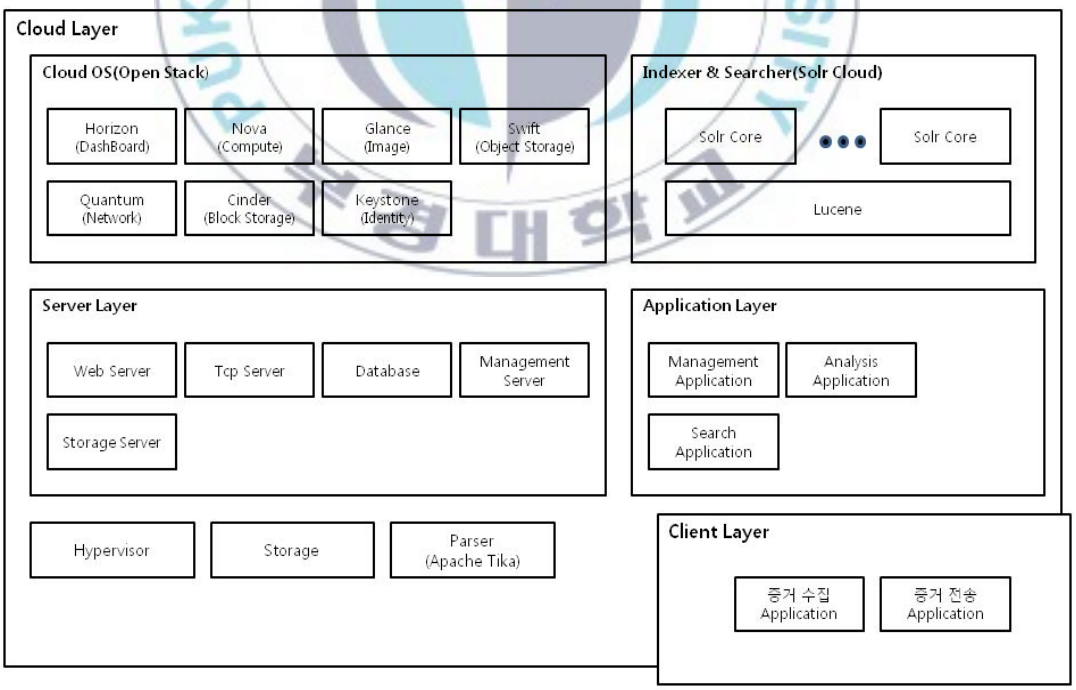
[그림 6] 포렌식 클라우드 IaaS 흐름도

V. 오픈소스 기반의 디지털 포렌식 클라우드 구현

본 장에서는 제안한 포렌식 클라우드의 프레임워크를 참조하여 오픈소스를 이용한 소프트웨어 아키텍처를 설계하고 테스트 환경에서 구현한다. 구현되는 포렌식 클라우드는 제안한 프레임워크에서 필수적인 요소만 추린 핵심 구현이다.

1. Software Architecture 및 개발환경

가. Software Architecture



[그림 7] 포렌식 클라우드 Software Architecture

3장에서 제시한 프레임워크 및 각 서비스별 절차를 참조하여 포렌식 클라우드 서비스에서 필요한 최소한의 기능을 선별한 후 설계한 아키텍처는 [그림 7]과 같다. 가상화 환경 및 전반적인 클라우드 환경 구축을 위해 OpenStack을 이용하며, 다중 색인 및 검색 서비스를 위해 Solr Core를 이용한 Solr Cloud를 구축 한다. 그리고 포렌식 서비스를 위해 웹 환경의 색인 및 검색 기능 구현 및 기타 사용자 관리 서비스도 포함하며, 이를 관리하기 위한 관리 서버를 구현 한다. 보안적인 요소는 3.4절에서 설명했다시피 정의 및 구현되어 있다고 가정한다.

나. 개발 환경

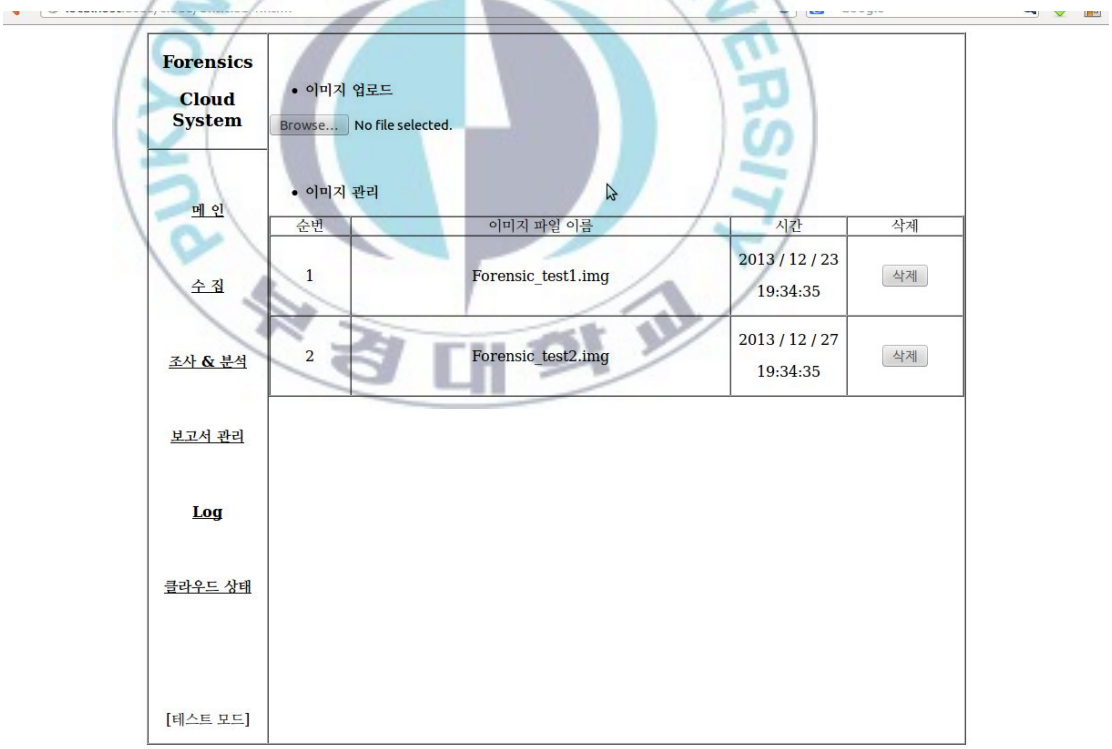
- 프로세서 : Intel i7-2600 CPU @ 3.4GHz x 8
- 메모리 : DDR3 - 4G x 2, DDR3- 2G x 2, Total 12G ram
- 하드디스크 : 320G SATA3
- 운영체제 : Ubuntu Server 64bit 12.04 LTS
- 개발 언어 : Java, Jsp, node.js
- 오픈 소스 :Solr, OpenStack(devstack), Tika, Zookeeper
- 어플리케이션: Tomcat, Jutty, eclipse.

2. 구현 결과

포렌식 클라우드 구현에 앞서 제시한 프레임워크의 3가지 서비스 모델을 모두 구현하기에는 구현 규모가 너무 크며, 개발 환경 또한 많이 부족한 부분이 많아 PaaS와 IaaS 구조를 모두 포함하고 있는

SaaS 기반으로 포렌식 클라우드 시스템의 프로토타입을 구현한다. SaaS의 내부 구조는 크게 3부분으로 색인과 검색을 담당하는 Solr Cloud와 가상화 시스템을 관리하는 OpenStack Cloud, 사용자 접속 UI 및 서비스를 제공하는 웹서버로 구성되어 진다.

클라우드 컴퓨팅에서 SaaS 사용자는 대부분 웹 페이지를 통하여 서비스를 받게 된다. 따라서 앞서 제시한 포렌식 클라우드 시스템의 프로토타입 구현을 위해 OpenStack과 Solr Cloud를 설치하여 서비스를 구성하고 이를 관리하고 접속하는 사용자에게 서비스하기 위한 웹 서버를 구현하였다. 웹페이지에서 제공되는 기능들은 차후 PaaS 서비스 구성을 위해 각 기능별로 분할되어 있다.



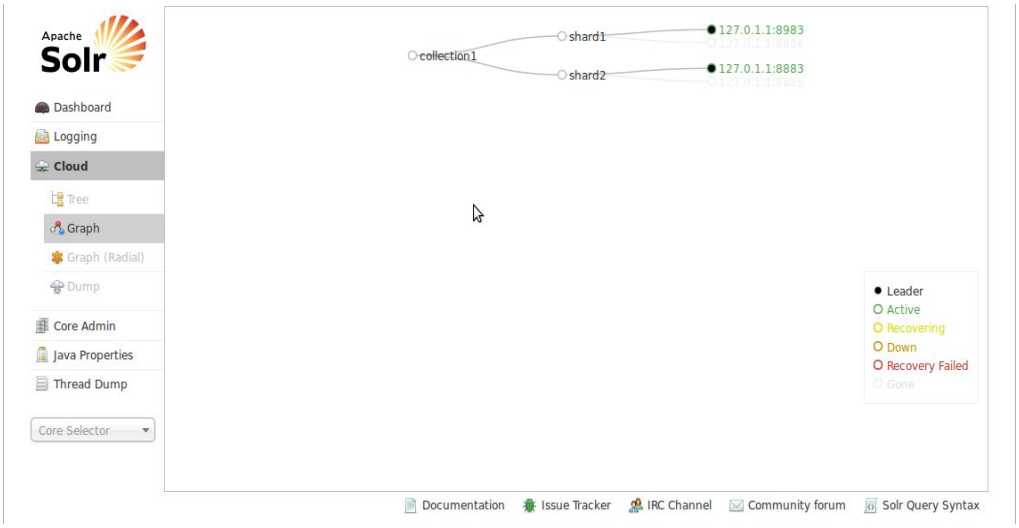
[그림 8] 포렌식 클라우드 웹 서비스 - 이미지 업로드

[그림 8]은 포렌식 클라우드의 서비스 중 이미지를 업로드 및 관

리하는 페이지를 나타낸다. 사용자는 클라우드 웹페이지에 접속하여 이미지를 업로드 하고, 업로드 된 이미지 파일에 대해 색인 및 검색을 수행한다. [그림 9]는 업로드 된 이미지를 색인하고 검색하는 화면이다. 색인은 먼저 업로드 된 이미지가 가상 드라이브로 마운트가 되며, 웹페이지에서 전달된 명령이 Solr Cloud로 전달되어 이미지가 마운트 된 경로로 색인 작업을 수행하게 된다.

[그림 9] 포렌식 클라우드 웹 서비스 - 색인 및 검색

Solr Cloud는 여러 개의 Solr Core를 Zookeeper에 연결 시켜 하나의 Cloud로 구성한다. 본 프로토타입 구현에서는 하나의 단일서버에 Standard 모드로 Zookeeper를 구성하여 2개의 Solr Core를 연결 시켰고, Shard의 개수를 2개로 설정하였다.



[그림 10] Solr Cloud 구성

업로드 된 이미지 데이터 및 색인 데이터는 사용자가 생성한 케이스별 단위로 저장되며, 검색은 케이스 내에서 생성된 모든 색인들을 통합적으로 검색 할 수 있도록 하였다. 색인 및 검색 기능은 기본적인 것만 구현하였으며, 목적에 따라서 색인 데이터 관리 및 검색 방법이 변경 될 수 있을 것이다.

3. 기존 디지털 포렌식 도구와 포렌식 클라우드 비교

기존 디지털 포렌식 도구와 설계한 포렌식 클라우드를 비교 분석하여 포렌식 클라우드를 도입으로 얻을 수 있는 특징들을 알아본다.

다음 <표 7>는 기존 디지털 포렌식 도구의 단점과 클라우드 포렌식으로 얻을 수 있는 장점을 나열한 것이다.

<표 7> 기존 포렌식 도구와 포렌식 클라우드 도구의 비교

항목	기존 포렌식 도구	포렌식 클라우드
동작 환경	단일 플랫폼 및 단일 프로세서에 의존한 디지털 포렌식 도구의 처리	한 대의 서버 또는 분산 클러스터로 구성된 서버의 자원으로 인한 처리
작업 처리	각 디지털 포렌식 업무별로 디지털 포렌식 도구 필요	로드밸런서에 의한 여러 명의 사용자 작업 분할 분산 클러스터 서버의 자원을 통합적으로 관리
유지 보수 및 확장성	도구의 개발사를 통한 업데이트 및 유지보수	CSP에 의해 통합적인 클라우드 서비스 관리
제약성	디지털 포렌식 도구가 설치된 자리에서 포렌식 업무를 수행해야만 하는 공간상의 제약	웹 환경 및 네트워크를 통하여 포렌식 클라우드 서비스를 이용할 수 있음. 네트워크가 연결되어 있지 않으면 서비스 이용 불가

따라서 포렌식 클라우드 구축 시 얻을 수 있는 효과들을 정리하면 다음과 같다.

● 디지털 포렌식 도구의 통합적인 솔루션 및 환경 제공

포렌식 클라우드는 단순하게 서비스 제공뿐만 아니라 업무의 특성에 맞게 기능들을 사용자가 원하는 대로 구성 및 개발할 수 있는 환경을 제공한다. 따라서 포렌식 클라우드를 이용하고자 하는 수사기관 및 관련 기업들의 특성에 맞춰 서비스 모델을 선택할 수 있어 시간 및 비

용을 절감 할 수 있다.

● 디지털 포렌식 업무 간의 협업 가능

기존 디지털 포렌식 업무가 각각의 수집관, 조사관, 분석관리관 등 사용하는 포렌식 도구나 업무가 달라 협업이 이루어지지 않았다면, 포렌식 클라우드로 인한 통합 포렌식 환경 제공으로 업무 절차에 따라 진행도 확인 및 사용자간의 커뮤니케이션이 가능하다.

● 도구의 유지 보수 및 확장성

디지털 포렌식 도구가 수행 작업에 따라 종류가 다양하고 제공하는 기능도 도구에 따라 차이가 크다. 포렌식 클라우드가 서비스하는 기능들이 각각의 어플리케이션으로 이루어져 있어 유지보수가 용이하고 필요한 기능들을 추가하기 쉬워 확장성 또한 보장된다.

● 디지털 포렌식 환경에 대한 공간적 제약 사항 해결

클라우드의 공통적인 특성으로 포렌식 클라우드 또한 웹 접속 환경 및 네트워크 접속이 가능한 환경에서 어디서든 사용자 인증 후 접속할 수 있으며 공간적 제약 사항을 해결할 수 있다.

● 디지털 포렌식 도구의 처리 속도 한계 극복

기존 단일 플랫폼에서 사용되는 도구의 한계는 단일 컴퓨터의 성능에 좌우된다. 하지만 포렌식 클라우드는 클라우드 특성상 클라이언트의 자원이 아닌 서버의 자원을 사용하기 때문에 서버의 클러스터 환경 또는 인프라가 확장된다면 성능 또한 크게 증가될 수 있다.

VI. 결론

본 논문에서는 기존 디지털 포렌식 기술의 동향에 대해 알아보고 디지털 포렌식 절차에서 현재 쓰이고 있는 디지털 포렌식 도구의 문제점에 대해 분석했다.

기존 디지털 포렌식 도구는 단일 플랫폼에서 동작하는 한계점을 극복하기 위해 새로운 기술적 요소의 필요성을 제시하고, 미래 IT 환경 변화에 대해 따라 기존의 포렌식 기술이 가질 수 있는 한계를 뛰어넘기 위한 대안으로써 클라우드 컴퓨팅 구조를 분석하였다. 그리고 클라우드 컴퓨팅과 디지털 포렌식 도구의 기술적 요소 결합에 대해 제시했다. 이를 위해 일반화 된 클라우드 컴퓨팅 구조와 서비스 형태들을 분석하였으며, 특징 요소들을 파악하였다. 또한 여러 가지 관점에서 필수적이고 실현 가능한 포렌식 절차 별 기능들을 선별하고 이들을 효과적으로 포함할 수 있는 포렌식 클라우드 프레임워크를 설계 하였다.

제안한 포렌식 클라우드 프레임워크의 구조를 제안하고, 포렌식 클라우드 기본적으로 가져야하는 구성 요소에 대해 설명하였다. 그리고 사용자가 포렌식 클라우드를 이용했을 때 이루어지는 모델별 흐름도에 대해 설명하였다.

본 논문에서 제안한 포렌식 클라우드는 이슈화되어 있는 클라우드 OS를 기반으로 한 프레임워크로써 향후 클라우드 컴퓨팅 기술이 발전함으로 프레임워크에 대한 내용 또한 추가되거나 변경될 수 있다. 하지만 기존 포렌식 도구가 가지는 한계점을 극복하기 위해서는 클라우드 컴퓨팅 기술과 같은 새로운 기술이 접목되어 포렌식 도구가 개별화되지 않고 통합적으로 이루어져 디지털 포렌식 업무 효율성 증대 및 발전이 필요할 것이며 지속적으로 연구 할 필요가 있다.

참고 문헌

[1] 민옥기, 김학영, 남궁한(2009. 8), “클라우드 컴퓨팅 기술 동향”, 전자통신통향분석, 제 24권 제 4호, 한국전자통신연구원

[2] 손정환, 김귀남(2005. 3), “국내 디지털 포렌식 기술 현황과 발전 방안”, 정보보안논문지, 제5권 제1호 pp.11-18, 한국융합보안학회

[3] 신용녀, 신승목(2010. 11), “대용량 디지털포렌식 서비스에 대한 실증적 연구”, Internet and Information Security, 제1권 제 2호 pp.83-100, 한국인터넷진흥원

[4] 이상진(2007. 11), “디지털 포렌식스 기술 발전방안”, 디지털 포렌식 연구, 제1권 제1호 pp.1-22, 한국디지털포렌식학회

[5] 한지성, 이상진(2011. 04), “라이브 포렌식을 위한 윈도우즈 물리 메모리 분석 도구”, 한국정보보호학회논문지, 제21권 제 2호 pp.71-82, 한국정보보호학회

[6] 한국정보통신기술협회(2007. 12), “컴퓨터 포렌식 가이드라인”, 표준번호:TTAS.KO-12.0058

[7] 한국정보통신기술협회(2007. 12), “컴퓨터 포렌식을 위한 디지털 데이터 수집도구 요구사항.”, 표준번호:TTAS.KO-12.0057

[8] 한국정보통신기술협회(2007. 12), “컴퓨터 포렌식을 위한 디지털 데이터 분석 도구 요구사항.”, 표준번호:TTAK.KO-12.0081

[9] Apache Lucene, http://lucene.apache.org/

[10] Apache Solr, http://lucene.apache.org/solr/

[11] Apache Tika, http://tika.apache.org/

[12] CloudStack, http://cloudstack.apache.org/

[13] Darryl Chantry(2010. 5), “Mapping Applications to the

Cloud", The architecture journal, pp.2-9
[14] OpenStack, http://www.openstack.org/

감사의 글

제가 LACUC에 학부 2학년 때부터 와서 지금까지 5년 가까이란 시간이 지났습니다. 복학하고 NAN이라는 동아리에 들어갔다가 만난 찬두선배가 저에게 갑자기 전화 와서 연구실에 와라고 했을 때가 엊그저께 같은데, 이젠 집보다 더 익숙한 연구실의 제자리가 마냥 떠난다고 생각하니 많이 아쉽고 또 뒤를 많이 돌아보게 되며 감사의 글을 올립니다.

먼저, 부족했던 저를 연구실에 받아주시고 공부시켜 주신 신상욱 교수님께 정말 감사합니다. 별로 공부는 잘하는 편은 아니었지만, 열정을 높게 봐주시고 기대해주셨던 것이 항상 기억에 남습니다. 마지막까지 신경써주시고 무사히 졸업할 수 있도록 신경써주신 것 다시 한번 감사합니다. 다음으로 바쁘시지만 논문 완성을 위해 조언을 해주시고 신경 써주신 이경현 교수님과 신원 교수님께 감사드립니다.

그리고 대학생활을 해오면서 가장 기억에 남는 것은 LACUC연구실 멤버와 NAN 선후배동기들입니다. 저에게 NAN이란 동아리는 선택한 것은 정말 잘했다라고 생각합니다. 처음 만났던 기현선배와 찬두선배, 성지, 희경이, 혜진이와 동아리 생활을 하며, LACUC에 들어와서 만난 수완선배, 태림선배, 주영선배, 본민선배는 저에게 잊지 못할 선배들입니다. 저에게 많은 얘기를 건네주시고, 가끔 조언도 해주신 이혜주 교수님. 그리고 학부 때 같이 연구실 생활했던 중원이. 지금 생각해보지만 너랑 참 성격은 달랐어도 같이 공부하면서 재미는 있었다. 그리고 석사 들어오면서 같이 석사생활을 하자고 꼬셨던 내 고등학교 동창 헌민이, 학부생활을 같이 보냈던 상민이, 믿음직 스러운 후배 명완

이, 정환이, 나랑 미운 정 고운 정 다든 기웅이, 연구실 안방마님 된 수빈이, 석사 후배 창수, 병도, 다른 연구실이지만 도움을 많이 주신 요섭선배, 오석이 석사 후기 때 만난 원준 선배 그리고 나를 도와주고 친해졌던 모두들. 나에게 와서 전공에 대해 공부해갔던 많은 후배들. 그리고 마지막 후배이자 지금 열심히 나에게 배우고 있는 혜인이. 정말 학교에 오래 있었던 만큼 감사해야할 사람들이 많습니다.

석사 들어오고 여러 감사의 글을 보면서 나는 언제 이렇게 적어보지 했던 지가 엊그저께 같습니다. 제 밑으로 한참 공부하고 있는 후배들에게 끊임없이 노력하고, 학교 다닐 때 이것저것 많이 경험해보라고 말하고 싶습니다. 비록 감사의 글에 이름은 없더라도 저에겐 이 학교에서 만난 모든 사람들이 소중하고 감사합니다.

마지막으로 저를 이렇게 무사히 졸업할 수 있도록 저를 믿어주시고 고생하시는 우리 아버지, 어머니께 이 논문을 바칩니다. 앞으로 더 열심히 해서 큰 사람이 되겠습니다.