공학석사 학위논문

클라우드 컴퓨팅을 이용한 DFSaaS 프레임워크 연구

2012년 2월

부 경 대 학 교 대 학 원

정보보호학 협동과정

구 본 민

공학석사 학위논문

클라우드 컴퓨팅을 이용한 DFSaaS 프레임워크 연구

지도교수 신 상 욱

이 논문을 공학석사 학위논문으로 제출함.

2012년 2월

부 경 대 학 교 대 학 원

정보보호학 협동과정

구 본 민

구본민의 공학석사 학위논문을 인준함.

2012년 2월 24일

주심　　공학박사 송 하 주 (인)

위원　　이학박사 신　원 (인)

위원　　이학박사 신 상 욱 (인)

차 례

그림 차례

표 차례

A Study of DFSaaS Framework based on Cloud Computing

Bon Min Koo

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

Abstract

Existing digital forensic tools operate on a single platform such as Windows OS or Linux OS, which have been constantly improved in processing speed and supporting various media. Recently, the number of evidences in digital media is increasing and to process these data, high speed processing is needed. Enormous digital evidence will require more much processing time and resources to process them. However, a single platform based tool cannot satisfy this requirement.

Meanwhile, cloud computing provides the characteristics of rapid elasticity, measured service, on-demand self-service, ubiquitous network access and resource pooling. To deal with these limits of existing forensic tools, advantages and characteristics of cloud computing can be used to design an advanced forensic tool.

Therefore, this thesis proposes the structure of digital forensic software as a service(DFSaaS) on cloud computing to deal with drawbacks of existing tools and assist the efficient and

effective digital forensic procedures. And the use scenario of DFSaaS is presented. By using DFSaaS, digital forensics procedures can be performed by accessing the web or a dedicated client anywhere, anytime, if there is available network connection. Also, by the distributed/parallel processing of the proposed DFSaaS, the huge amounts of evidence data can be efficiently handled, and the relevant evidences can be obtained more quickly, through the collaboration of multiple analysts using DFSaaS.

I. 서 론

1. 연구배경

디지털 포렌식이란 과학적이거나 기술적인 기법을 사용하여 범죄 수사 또는 증거를 수집하는 행위이이다. 이를 위한 도구로 디지털 포렌식 도구가 존재한다. 이 도구는 법과 기술 간의 매개체가 될 수 있는 핵심 요소라 할 수 있다[1]. 디지털 포렌식은 증거 이미징, 분석, 검색, 보고서 작성 등의 일련의 절차를 요구한다. 기존의 디지털 포렌식 도구는 이러한 절차적 기능을 제공하는 것을 목적으로 개발되었다.

현존하는 대부분의 포렌식 도구들은 단일 플랫폼 상의 윈도우 운영체제에서 운용되는 통합 도구로 제공된다. 이동성을 위해 휴대형 하드디스크 드라이브에 저장되며 해당 매체 내에서 실행된다. 목적에 따라 전용 기능을 제공하는 하드웨어 형태의 도구로도 제작 된다. 추가적으로 압수한 데스크탑을 포렌식 연구실로 이동하기 위해서 특수한 장치가 필요할 수 있다.

현재 단일 플랫폼 형태의 포렌식 도구에서 2TB의 데이터를 이미징 하는데 7시간이 걸리며 비트와이즈 검색을 20MB/s 정도의 속도로 처리해 1TB 이미지를 검색 했을 때에는 14시간이 소요된다[2]. 이는 양적으로 증가하는 디지털 증거의 추세에 미루어 볼 때 향후 도구의 증거 처리 속도 문제를 야기 시킬 것이다. 또, 증거물을 포렌식 연구실로 이송을 하거나 고속 처리를 하기위해서는 기타 하드웨어 장치를 이용해야 하고, 이는 도구 사용에 있어 불필요한 번거로움을 초래한다.

따라서 기존의 도구를 아우를 수 있는 새로운 도구 개발이 시급하다. 프로세싱 속도 향상을 위해 단일 플랫폼의 한계를 극복해야하고,

이용의 번거로움을 피하기 위해 모든 장비들을 하나의 도구로 합쳐야 한다. 기존의 도구를 기능을 아우를 수 있고, 단점을 제거하기위해 클라우드 컴퓨팅 개념을 적용해 해결 방안을 모색하였다.

본 논문에서는 클라우드 컴퓨팅[3]을 이용하여 디지털 포렌식 절차에 따라 어디에서든 포렌식을 수행 할 수 있는 DFSaaS(Digital Forensic Software as a Service) 구조와 이에 대한 시나리오를 제시하고, 이들에 대한 프레임워크를 연구한다.

2. 연구 내용 및 구성

본 논문에서는 디지털 포렌식 절차와 클라우드 컴퓨팅에 대해서 설명하고 각 절차별 상세 설명 과 더불어 클라우드 컴퓨팅의 구조에 대해 살펴본다.

디지털 포렌식 절차에 따른 기능들을 도출하여 클라우드 컴퓨팅 기반의 포렌식 도구로써 가져야 할 기능들에 대한 구조를 그려보고 이들에 대한 내부 흐름도와 도구 사용 시나리오를 제시한다.

2장에서는 디지털 포렌식 절차에 대해 설명하고 각 절차에 대한 사항들을 살펴본다. 또한, 클라우드 컴퓨팅의 정의와 더불어 기본 구조를 살펴본다. 3장에서는 클라우드 컴퓨팅 기반 포렌식 도구인 Digital Forensic Software as a Service에 대한 구조와 내부 흐름도, 프레임워크를 제시하며 Digital Forensic Software as a Service 에 대한 시나리오를 그려보고 4장에서 결론을 맺는다.

II. 관련 연구

현재 디지털 포렌식 수사 절차는 그 과정이 정립되어 단계가 나뉘어져 있고, 각 단계를 아우르는 도구가 존재할 수 있고, 단계에 따른 고속 도구가 존재한다. 이에 따라 디지털 포렌식 수사 도구는 많은 종류가 배출되었다. 클라우드 컴퓨팅의 등장과 함께 이를 분석하여 서비스화 하고 있는 기업이 늘어나고 있다.

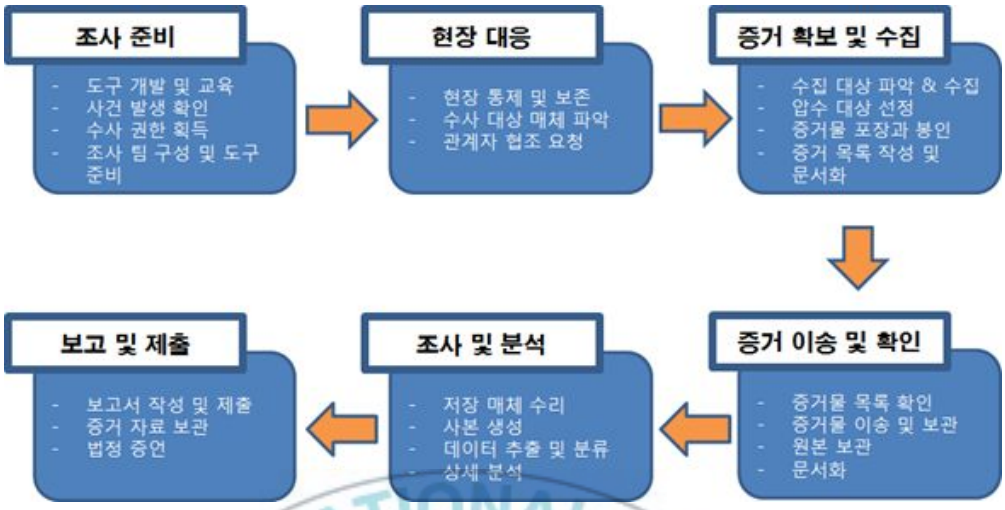
본 장에서는 디지털 포렌식 절차에 대해 살펴보고 각 절차에 관한 수사 방법을 살펴본다. 또한 클라우드 컴퓨팅의 정의 및 구조와 각 구성요소들을 살펴본다.

1. 디지털 포렌식

가. 기본적인 디지털 포렌식 절차

범죄 사실을 규명하기 위해 각종 증거를 과학적으로 분석하는 분야를 법과학(forensic science)라 한다. 여기서 포렌식(forensic)이란 '법정의', '공개토론이나 변론에 사용되는', '수사와 법정에서의 증거 또는 사실관계를 확정하기 위하여 사용하는 과학이나 기술에 관한', '범죄와 관련된 증거물을 과학적으로 조사하여 정보를 찾아내기 위한' 이라는 의미[4]가 있다. 법과학은 지문, 모발, DNA 감식, 변사체 검시 등과 같은 전통적인 법의학(forensic medicine) 분야를 의미하였는데, 점차 발전하여 범죄 수시 및 민·형사 소송 등 법정에 사용되는 증거의 수집, 보존, 분석을 위한 응용과학 분야를 통칭하는 용어로 사용되고 있다. 컴퓨터가 도입됨에 따라 컴퓨터를 대상으로 하는

범죄와 컴퓨터를 도구로 사용하는 범죄가 발생하고 있으며, 범죄 자료의 저장소로 컴퓨터가 사용되고 있다. 또한 각종 디지털 기기가 보급되면서 디지털 기기에 저장되어 있는 데이터가 범죄 사실을 입증하는 증거로 사용되거나 사건의 실체를 규명하는 실마리로 사용되는 사례가 증가하고 있다. 이와 같이 디지털 기기에 남아 있는 각종 데이터를 조사하여 사건을 규명하는 법과학 분야를 법전산(forensic computing) 또는 디지털 포렌식(digital forensic)이라 말한다. IT 기술의 발전 및 급격한 정보화 사회로의 변화는 정보의 디지털화를 가속시켜서 컴퓨터 관련 범죄 뿐만 아니라 일반 범죄에서도 중요 증거 또는 단서가 컴퓨터를 포함한 디지털 정보 기기 내에 보관하는 경우가 증가함에 따라 증거 수집 및 분석을 위한 전문적인 디지털 포렌식 기술이 요구 된다. 디지털 포렌식은 조사 대상에 따라 컴퓨터 포렌식, 모바일 포렌식이라는 용어로 나뉘어 사용되고 있으며, 또한 디지털 데이터의 유형에 따라 활성 데이터, 파일 시스템, 데이터베이스, 응용 프로그램 사용 흔적 데이터, 각종 로그 데이터, 악성 코드, 암호 및 은닉 데이터 등을 조사 대상으로 한다. 이외에도 모든 디지털 데이터가 조사 대상이 되며, 사건에 따라 조사해야 하는 대상이 달라진다. 디지털 포렌식의 기본 절차[4][5][6][7][8]는 아래 [그림 1]과 같다.



[그림 1] 디지털 포렌식 수사 절차

나. 절차 별 상세 기능

(1) 조사 준비

디지털 포렌식 조사는 일반 범죄 수사와 동일하게 피해자의 신고나 자체 조사를 통해 사건을 인지하고 원인 파악을 할 필요성이 제기되면 시작한다. 디지털 기기의 다양성, 디지털 데이터의 복잡한 특성 때문에 디지털 포렌식 수사를 진행하는 조직은 본격적인 조사를 시작하기 전에 사전 준비 과정을 거치게 된다. 조사 준비 과정은 도구 개발 및 교육, 사건 발생 확인, 수사 권한 획득, 조사 팀 구성 및 도구 준비 단계로 나뉜다.

- 도구 개발 및 교육 : 디지털 기기는 계속해서 신제품이 출시되고, 지속적으로 개량되고 있다. 따라서 해당 기기에 대한 연구가 선행되어 있지 못하면 조사 자체가 어려울 수 있다. 디지털 데이

터는 한번 손상되면 회복이 불가능하기 때문에 조사 과정에서 발생한 실수는 사건 해결에 치명적인 영향을 미칠 수 있다. 따라서 디지털 기기의 취급 방법, 조사 원칙 등에 대한 사전 교육 및 교육 과정이 있어야 한다.

- 사건 발생 확인 : 사건 발생을 인지하면, 조사할 필요가 있는지 확인해야 한다. 확인 과정을 거쳐 주요 조사 대상을 결정하고 전반적인 수사 계획을 수립한다. 조사 대상의 선정은 전체 조사 전략을 수립하고 진행하는데 있어 결정적인 역할을 하므로 조사 책임자는 사건의 유형을 잘 파악하고 주요 수집 대상을 선정할 수 있는 능력이 필요하다.
- 수사 권한 획득 : 보통 사건을 조사하는 과정에서 불가피하게 알게 되는 정보가 있다. 특히 기업 비밀이나 프라이버시 관련 사항 등 외부에 노출되지 않아야 할 정보가 포함 될 수 잇는데, 이에 접근할 권한이 없다면 관련 내용에 대한 조사를 하지 말아야 하며, 조사가 꼭 필요하다면 그에 적합한 권한 확보가 선행되어야 한다.
- 조사 팀 구성 및 도구 준비 : 사건의 유형, 조사자의 전문성, 압수 대상 장소를 고려하여 조사 팀을 구성하고, 현장 도착 후 수색 절차, 증거 수집 방법과 범위, 각 조사자의 역할 등을 분담한다. 도구는 크게 증거 수집 장비, 증거 봉인 및 포장 장비, 운반 장비로 나눌 수 있다.

(2) 현장 대응

사건을 인지하고 조사를 진행하기로 결정하면 면밀한 계획을 수립한 후, 현장에 출동하여 조사에 필요한 조치를 취한다. 즉, 현

장을 통제 및 보존하고, 관계자에게 협조를 요청하며, 조사 대상 매체를 파악한다. 이 단계의 목표는 본격적인 증거 수집을 시작하기 이전에 현장을 통제하고 이를 보존하여 필요한 모든 증거가 수집되도록 준비한다.

- 현장 통제와 보존 : 현장에 도착하여 피조사자에게 부여된 권한의 범위를 명확히 설명하고 조사할 대상을 신속히 파악한 후, 현장 보존과 통제를 실시함으로써 현장 조사가 시작된다. 다이어리 노트, 일지, 포스트-잇 메모지, 컴퓨터 출력물 등에서 유용한 단서를 찾을 수 있으므로 물리적 증거물을 함께 보존한다.
- 수사 대상 매체 파악 : 현장에 있는 모든 것이 수사 대상이라고 할 수 있다. 물리적인 증거와 디지털 기기, 저장매체, 네트워크 구성을 파악한다.
- 관계자 협조 요청 : 피조사자가 협조하면 상당히 쉽게 증거를 수집할 수 있다. 전산 관리자, 데이터베이스 관리자와 같이 포렌식 관점에서 중요한 임무를 책임지고 있는 인원을 파악하여 분석에 대한 협조를 구하고, 범죄 혐의가 없는 자를 설득하여 협조를 이끌어 내어 이를 토대로 핵심 용의자와 면담함으로써 원활한 조사가 이루어지도록 유도한다.

(3) 증거 확보 및 수집

현장 대응 과정에서 기본적인 조치를 하였으면, 조사 대상자의 증거물을 확보하고 어떤 종류의 데이터를 어떤 방법으로 수집할 것인지를 결정한다. 영장의 기재 내용에 의거하여 압수 수색을 진행하고 필요한 증거물을 압수하는 과정이다.

- 수집 대상 파악 및 수집 : 주요 확보 대상인 컴퓨터나 서버 시스템 등의 위치가 파악되었으면, 증거 수집 방법을 신중히 결정한다. 시스템 및 네트워크 상황을 최대한 고려하여 어떠한 방법으로 증거 자료를 획득할 것인지를 결정하고, 시스템 압수 가능 여부, 디스크 이미징을 통한 압수 여부, 활성 데이터의 수집 필요성을 고려하여 증거물을 확보한다.
- 압수 대상 선정 : 증거 확보는 크게 시스템 확보, 원본 하드 디스크 드라이브 확보, 하드 디스크 복제, 선별 데이터 수집으로 나눌 수 있다. 시스템 확보는 하드 디스크에 저장된 데이터 및 시스템 전체에 대한 정밀 분석이 요구 될 때 실시하며, 필요한 경우 컴퓨터 본체와 모니터, 프린터, 케이블 등 주변 장치도 확보한다. 원본 하드 디스크 드라이브 확보는 하드 디스크에 저장된 데이터의 복구 및 분석이 필요하고, 원본을 보존할 필요가 있는 경우 실시하며, 하드 디스크 복제는 저장된 데이터를 복구하고 분석할 필요가 있지만, 원본을 압수 할 수 없는 경우 실시한다. 비활성 시스템의 경우는 전원이 꺼진 상태 그대로 두며, 활성 시스템인 경우는 전원을 차단하기 전에 휘발성 데이터의 수집 필요성을 판단한 후 필요하다면 휘발성 데이터를 수집한다.
- 증거물 포장과 봉인 : 확보한 디지털 저장 매체는 물리적인 충격에 의한 증거 훼손을 막기 위해 충격 방지제로 포장하고, 수집 과정에서 증거물에 대한 위·변조가 없었음을 증명하기 위해서 봉인한다. 증거물의 조작을 방지하기 위해 밀봉전용 특수 테이프(Evidence Tape)를 이용하여 포장한다. 정전기에 약한 매체라면 반드시 정전기방지 봉투를 이용해 먼저 포장을 한 후 밀봉해야 한다.

● 증거 목록 작성 및 문서화 : 조사 과정의 기록은 모든 단계에 적용되는 것이지만, 현장 조사 과정에서 더욱 중요하다. 증거물 목록은 증거 처리 과정의 첫 시작이므로 상세히 작성되어야 한다.

(4) 증거 이송 및 확인

디지털 증거 운반에서 가장 주의해야할 사항은 획득한 증거물의 진정성 유지와 훼손 방지이다. 또한 증거물의 누락 및 도난이 없도록 연계보관 원칙을 철저히 유지하고, 반복 확인 과정을 거치는 일 또한 중요하다. 모든 증거물들이 무결하다는 것이 확인되면 이 과정을 지켜본 인수 책임자는 증거물 목록에 이상 없이 전달 받았다는 서명을 하고 이를 문서화 한다.

(5) 조사 및 분석

점차 사건 조사에서 다루는 디지털 데이터의 양이 증가하고 있으며, 범죄 유형에 따라 조사해야 할 데이터가 서로 다르다. 따라서 효과적인 조사를 위해서는 수집한 데이터를 체계적으로 분류하고 사건 특성에 맞는 데이터를 선별하는 과정이 필요하다.

● 저장 매체 수리 : 저장 매체의 데이터를 분석하기 위해서는 데이터에 접근 할 수 있어야 한다. 고장이 있거나 증거 인멸을 위해 인위적으로 파괴한 경우는 수리 과정을 거친다.
● 사본 생성 : 분석을 시작하기 전에 수집한 디지털 증거물을 복제한다. 증거물 원본에서 직접 분석을 시행하게 되면 무결성에 손상을 줄 수 있으므로 원본과 동일하게 복제한 사본을 생성한다. 일반적으로 2개의 사본을 생성하여, 하나는 보관용, 다른 하

나는 분석용으로 사용한다. 증거물을 복제할 때는 원본의 모든 데이터를 포함해야 하며, 원본에 없는 내용이 포함되지 않아야 한다.

- 데이터 추출 및 분류 : 증거물 복제가 완료되면 데이터 분류를 위해 분석용 증거물에서 데이터를 추출한다. 정상적으로 파일 시스템을 인식 할 수 있으면, 정상 파일을 추출한 후 메타 데이터로부터 삭제된 데이터를 복구할 수 있는 파일을 추출한다. 그 후 미할당 영역을 대상으로 파일 카빙(Carving)을 실시한다. 파일 시스템을 인식하지 못하면 저장 매체 전체를 대상으로 파일 카빙을 실시한다. 파일 카빙이 종료된 후에 인식하지 못하고 남는 데이터는 따로 모아 문자열을 추출하고, 상세 분석을 위해 보관한다.
- 상세 분석 : 데이터 분류가 완료되면 분석 과정의 마지막 단계인 상세 분석을 수행한다. 상세 분석의 예로 인터넷 사용 흔적, 사용자 활동 정보 분석, 시스템 사용 정보 분석, 응용 프로그램 사용 흔적 분석, 파일 분석이 있다. 인터넷 사용 흔적 분석은 웹 브라우저, 메신저, 전자 메일 분석으로 나뉘며, 시스템 사용 흔적 분석은 사용자 정보, 응용 프로그램 및 하드웨어 설치 정보 등에 관한 내용을 분석한다. 윈도우 시스템에서는 시스템 설치, 사용 정보, 사용자 활동 정보, 응용 프로그램 사용 정보 등이 모두 레지스트리에 저장되므로, 이를 분석하면 유용한 정보를 얻을 수 있다.

(6) 보고 및 제출

결과 보고서는 조사, 분석자의 모든 행동과 관찰 내역, 분석 과

정 등이 정확히 기록되어야 하고 각 단계의 결과와 완벽히 일치해야 한다. 보고서의 내용은 쉽게 이해할 수 있는 용어를 사용하여 정확하고 간결하며 논리 정연하게 작성한다. 결과 보고서에 포함되어야 할 항목은 다음과 같다.

- 사건(Case) 및 보고서 번호
- 증거 수집 일시, 보고서 작성 일시
- 조사, 분석자, 보고서 작성자
- 조사, 분석에 사용된 장비 및 환경
- 각 절차에 대한 개략적 설명
- 사진 및 인쇄물 등과 같은 첨부 자료
- 추출 및 분석된 증거 데이터의 상세 설명
- 분석 결과 및 결론

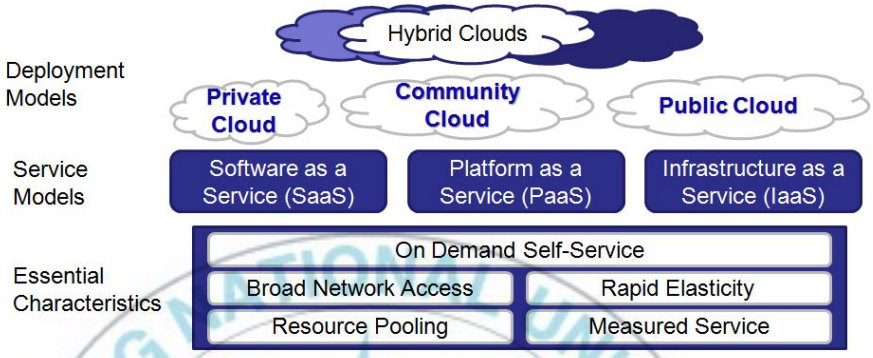
  분석 보고서에 최종 날인 하고 확인한 조사 책임자는 향후 법정에 출두하여 분석 결과에 대해 증언할 가능성이 있다. 따라서 법정 증언을 위해 분석 결과가 법정에 있는 일반인이 이해할 수 있도록 기술되어야 한다.

2. 클라우드 컴퓨팅

가. 클라우드 컴퓨팅의 정의

NIST[3]에서 정의한 클라우드 컴퓨팅은 빠르게 변화할 수 있고, 최소한의 관리 노력으로 이루어지며, 서비스 제공자의 상호 작용과 릴리스 구성을 받아들여 컴퓨팅 자원을 공유하고, 유비쿼터스와

On-Demand 형식의 접근을 활성화 하는 모델이다. 이러한 클라우드 컴퓨팅은 5개의 필수적 특성을 지니며 3개의 서비스 모델을 제시하고 있고 4개의 배포 모델로 구성되어 있다.



[그림 2] 클라우드 컴퓨팅 기본 구조

(1) 5가지 필수적 특징

- On-Demand Self-Service : 소비자가 서비스 제공자와 상호작용(서버 시간, 네트워크 스토리지 접근)을 고려하지 않고 서비스를 이용할 수 있어야 한다.
- Broad Network Access : 표준 메커니즘을 통하여 많은 클라이언트 플랫폼에서 네트워크 접근이 가능해야 한다.
- Resource Pooling : 공급자의 컴퓨팅 리소스가 동적으로 할당된 멀티 모델들을 사용하여 여러 소비자의 수요에 따라 고객이 사용하는 자원의 정확한 위치를 알 수 없더라도 서비스 할 수 있어야 한다.
- Rapid Elasticity : 클라우드 컴퓨팅 서비스를 사용하는 사용자는 자원을 무한대로 확장할 수 있거나 필요한 만큼의 수준으로 줄일 수 있어야 하며, 어떤 경우에든 신속하고 빠르게 소비자에

게 서비스를 제공할 수 있어야 한다.

- Measured Service : 클라우드 시스템은 자동 제어와 추상화 능력을 활용하여 최적의 리소스를 사용할 수 있어야 하며, 리소스 사용량의 모니터링을 통해 시스템 제어, 공급자 및 소비자 서비스의 활용에 대한 요금 및 이에 대한 투명성을 보장해야 한다.

(2) 3가지 서비스 모델

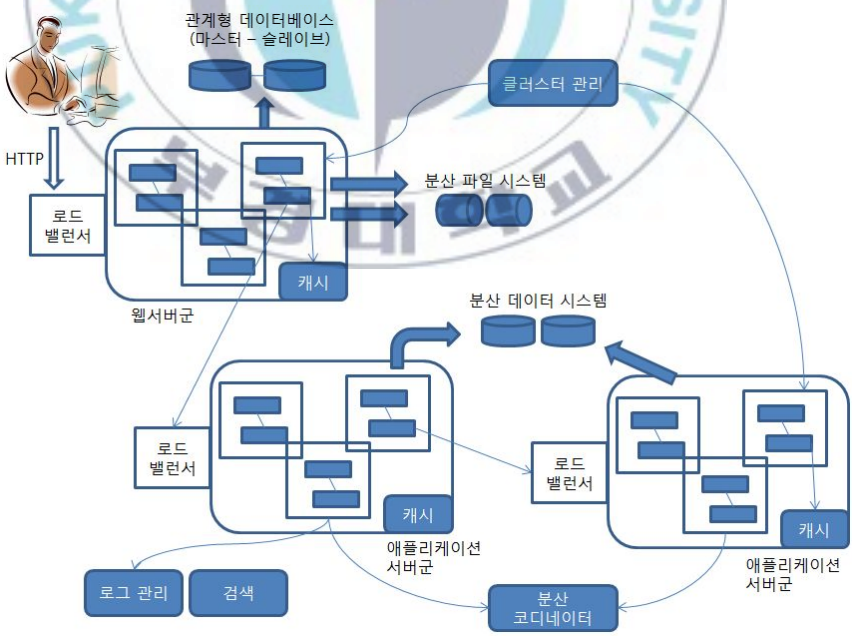
- SaaS(Software as a Service) : 클라우드 인프라에서 실행되는 공급자의 응용프로그램을 이용하여 서비스를 제공한다. 네트워크, 서버, 시스템, 스토리지 또는 개별 응용 프로그램 기능 등 오로지 클라우드 서비스 제공업체가 제공하는 어플리케이션만을 이용할 수 있다.
- PaaS(Platform as a Service) : 사용자가 서비스 공급자가 지원하는 도구를 사용하여 어플리케이션을 구축할 수 있다. 하지만 PaaS는 서비스 제공업체가 정의한 방식으로 한정된다.
- IaaS(Infrastructure as a Service) : 제공업체가 컴퓨터 하드웨어를 제공하고, 사용자들은 여기에 데이터를 저장하거나 원하는 모든 형태의 어플리케이션을 개발해 구동시킬 수 있는 클라우드 서비스이다. 소비자가 기본 클라우드 인프라를 제어하지만 OS, 스토리지 등 배포된 어플리케이션만 제어할 수 있으며, 일부 네트워킹 구성 요소(호스트 파이어월)를 제한된 조건하에 제어가 가능하다.

(3) 4가지 배포 모델

- Private Cloud : 기업에서 전적으로 운영하는 클라우드 인프라

를 말한다. Public Cloud의 개념 중 일부를 제한된 네트워크상에서 특정 기업이나 특정 사용자만을 대상으로 한다.

- Community Cloud : 여러 기업에서 정책이나 보안 요소 등 공유가 필요한 특정 커뮤니티를 클라우드 인프라로 지원한다.
- Public Cloud : 일반 기업이나 대규모 산업그룹에서 사용할 수 있는 클라우드 인프라로 기업에서 클라우드 서비스를 판매할 수 있다. 서비스 내부에 저장된 데이터나 기능, 서버 같은 자원은 각 서비스에서 사용자 별로 권한 관리 되거나 격리돼 서비스 사용자 간에는 전혀 간섭이 없다.
- Hybrid Cloud : 클라우드 배포 모델이 2개 이상 묶여 있는 클라우드 인프라로 보통 Public Cloud와 Private Cloud를 병행 사용하는 방식이다.



[그림 3] SaaS의 일반적인 구조

위 [그림 3]는 SaaS의 일반적인 구조를 나타낸다. 시스템은 서브 시스템 단위별로 분산돼 있으며, 각 서브 시스템은 부하 정도에 따라 여러 대의 서버 풀을 구성해 운영된다. 서버 풀의 부하를 분산해주는 로드밸런서가 배치되어 운영되며, 분산된 환경을 관리해주는 클러스터 관리 도구와 분산 코디네이터가 추가된다. 각 서버 군은 데이터 서버의 부하나 서버 간 부하를 줄이기 위해 캐시 서비스를 이용한다. 시스템 구성의 주요 특징은 애플리케이션 서버, 파일 시스템, 데이터베이스, 캐시 같은 요소들이 내부적으로 분산 구조로 구성돼 있으며, 시스템 전체 구성도 분산 구조로 되어 있다.

정리해보면, 클라우드 컴퓨팅 구조는 애플리케이션 서버, 분산 파일 시스템, 분산 데이터베이스, 캐시, 클라우드 스토리지, 대용량 데이터 분석, 클러스터 관리, 서버 가상화 등 많은 요소로 구성되어 있으며 사용자는 웹을 통해 클라우드 서비스에 접속하여 제공하는 애플리케이션을 이용하는 것[9]이다. 이들 요소와 구조의 관계는 다음 [그림 4]와 같다.

<table>
<tr><td rowspan="3">로그
관리</td><td>애플리케이션 서버</td><td>대용량 데이터 분석</td><td rowspan="3">클러스터
관리</td></tr>
<tr><td>SQL(DB)</td><td>캐시</td></tr>
<tr><td>서버 가상화</td><td>클라우드 스토리지</td></tr>
</table>

[그림 4] 클라우드 컴퓨팅의 요소 관계

III. Digital Forensic Software as a Service

본 장에서는 기존 디지털 포렌식 도구의 한계점에 대해 알아보고 그에 따른 클라우드 컴퓨팅 서비스 전환시 고려사항에 대해 살펴본다. 또한 기존 도구를 아무런 수정없이 클라우드 컴퓨팅 서비스에서 동작 시켰을 때 문제점들에 대해 살펴보고, 디지털 포렌식 절차에 따른 기능들을 추려 클라우드 서비스로 만들 수 있는 도구의 구조와 흐름도를 설명한 뒤 프레임워크를 제시하여 절차별 시나리오를 나열한다.

1. 기존 포렌식 도구의 한계와 클라우드 컴퓨팅 서비스 전환시 고려사항

가. 기존 포렌식 도구의 단점

기존 포렌식 도구의 단점은 다음과 같이 4가지로 나눌 수 있다.

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

(2) 기존 도구를 이용하여 디지털 포렌식 수행시 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.

(3) 새로운 형태의 데이터에 대한 도구 기능 확장시에 도구 전용 웹 사이트의 절차에 따라 지속적인 도구의 유지보수 측면에서 일일이 업데이트를 실시해야 한다.

(4) 디지털 포렌식은 즉시 행해 질 수 없고, 도구가 설치된 디지털 포렌식 연구실 등 특정 장소에 이동해야 수행이 가능하다.

위 단점들을 보완하기 위해 새로운 디지털 포렌식 도구가 개발되어야 하며 이에 대해 클라우드 컴퓨팅 기술을 이용하여 보완하고자 한다.

나. 기존 도구의 클라우드 컴퓨팅 서비스 전환

기존의 어플리케이션을 클라우드 컴퓨팅으로 이동시키기 위해서 어플리케이션 자체가 클라우드 플랫폼으로 이동 및 개발 또는 호스팅 하기 적합한 어플리케이션이 있는가 하면, 클라우드를 사용하기에 부적합한 어플리케이션도 있다. 따라서 특정 어플리케이션을 클라우드로 운영하는 것에 대해 실용적인가를 먼저 결정한 후 운영한다. 대부분 EnCase나 FTK와 같은 도구들을 클라우드 컴퓨팅 서비스의 Virtual Machine에 바로 접목 시키면 별다른 수정 없이 클라우드에서 안정적이고, 확장성을 가지며 손쉽게 비용 절감을 이룰 수 있을 것이라 생각하지만, 클라우드에서 어플리케이션을 아무런 수정없이 구동하는 것은 아무런 특징을 가질 수 없다. 일반적으로 단일 컴퓨터에서 동작하도록 디자인된 어플리케이션은 여러 대의 컴퓨터로 된 인프라에서 사용 가능한 확장성이 배제되어 있다[10].

또한 클라우드 컴퓨팅은 데이터에 대해 강한 수준의 정합성이나 견고성보다는 손쉬운 확장성, 시스템 확장 과정이나 장애 상황에서도 서비스를 유지할 수 있는 고가용성, 낮은 비용 등을 요구한다. 따라서 기존

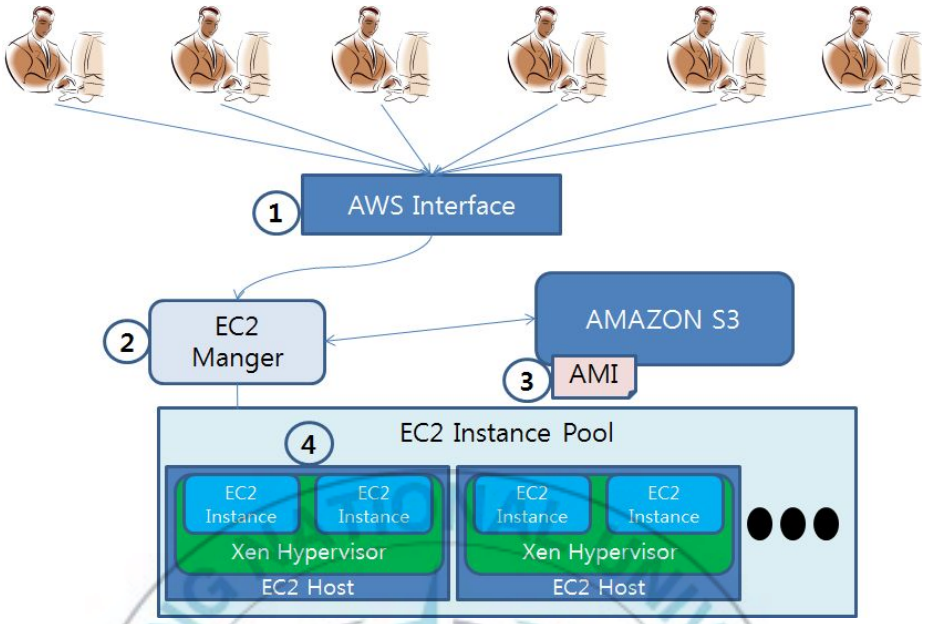
도구가 사용하는 표준 SQL이나 엔티티 간의 관계를 지원하지 않는 데이터 관리 시스템이 등장하게 됐으며, 이런 데이터 관리 시스템을 NoSQL[11]이라고 한다. 때문에 NoSQL은 클라우드 컴퓨팅 서비스의 요구 사항들을 만족시켜주고 있으므로 클라우드 컴퓨팅 플랫폼의 데이터 서비스 계층에 사용되고 있다. NoSQL을 사용할 때에는 구축하려는 시스템에 적합한 솔루션인지 판단하기 위해 다음 <표 1>과 같이 항목들을 비교, 검토한다.

클라우드 컴퓨팅에 맞는 데이터 파일 저장 방식으로는 데이터 파일을 분산 파일 시스템에 저장하고 데이터 관리 시스템에서는 논리적인 관리만을 담당하는 방식이나 관리 시스템 자체적으로 데이터 파일을 저장하는 방식이 있다. NoSQL은 데이터 모델별로 많은 솔루션이 있지만, 이에 대한 솔루션으로 빅테이블[12], 클라우데이터, HBase 등이 해당하며, 클라우데이터나 HBase의 경우 빅테이블을 모방한 솔루션으로 구조가 비슷하다[9].

<표 1> NoSQL 구축시 검토항목

데이터 복제 여부	- 저가의 하드웨어를 사용하기 때문에 데이터는 반드시 복제돼야 한다.
특정 서버 장애시 데이터 연산 수행 가능 여부	- 저가의 하드웨어를 사용하기 때문에 일부 서버에서 장애가 수시로 발생할 수 있다는 가정하에 솔루션을 검토해야 한다. 특정 서버에 장애가 발생한 경우에도 데이터 연산은 수행 가능해야 하며, 장애가 발생한 서버를 새로운 서버로 교체한 후 투입시에도 쉽게 추가 가능해야 한다.
데이터 모델	- 대부분의 NoSQL 솔루션은 개발자에게 익숙하지 않은 데이터 모델을 제시하고 있기 때문에 데이터 모델이 시스템의 요구 사항에 부합해야 한다.
데이터 규모	- NoSQL 솔루션에 저장할 수 있는 데이터 규모가 쉽게 확장 가능한 구조인지 확인한다.
관리의 편의성	- NoSQL은 대부분 분산된 여러 서버에 배치돼 운영된다. 따라서 서버를 구성하거나 추가/삭제 작업이 편리해야 한다.

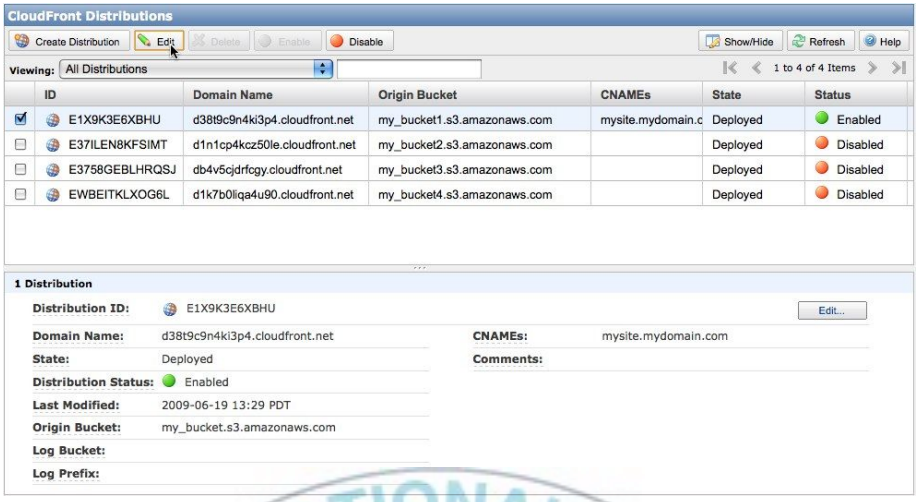
다음 [그림 5]은 아마존 클라우드 서비스(Amazon Web Service 이하 AWS)[13]에 기존의 디지털 포렌식 도구의 이미지를 저장하여 도구가 동작하는 모습을 나타낸다. AWS는 최대의 IaaS 서비스 제공업체로써 클라우드 컴퓨팅 시장의 큰 영향력을 행사하는 기업이다. 처음에는 한두 가지의 작은 서비스로 시작했지만, 이제는 세계의 주요 지역에서 수많은 클라우드 관련 서비스를 제공하며 지속적으로 발전되고 있다.



[그림 5] 아마존 클라우드 서비스(Amazon Web Service)

AWS는 아마존에서 제공하는 이미지들을 통해 기본적인 운영체제나 어플리케이션을 동작 시킬 수 있고, 아마존 머신 이미지(AMI)를 이용하여 사용자가 설정한 운영체제나 설치된 어플리케이션을 AWS에서 사용할 수 있다. 따라서 사용자는 [그림 5]에 기재된 번호에 따라 먼저 AWS에 접속하여 가상 서버 설정을 위한 EC2 매니저를 이용하여 구동시킬 AMI를 검색한다. 이 때, 디지털 포렌식 도구를 미리 설치한 AMI를 아마존 S3에 저장하고, 이 이미지를 선택하여 가상 서버를 구동시킨다. 사용자가 생성한 AMI는 다른 AWS 사용자들에게 제공이 가능하다. 따라서 다중 사용자가 EC2를 사용 할 때 제공 받는 AMI로 인한 도구의 원활한 사용이 가능하다.

다음 [그림 6]은 아마존 웹 서비스 콘솔에서 이미지를 선택하는 것을 나타낸다.

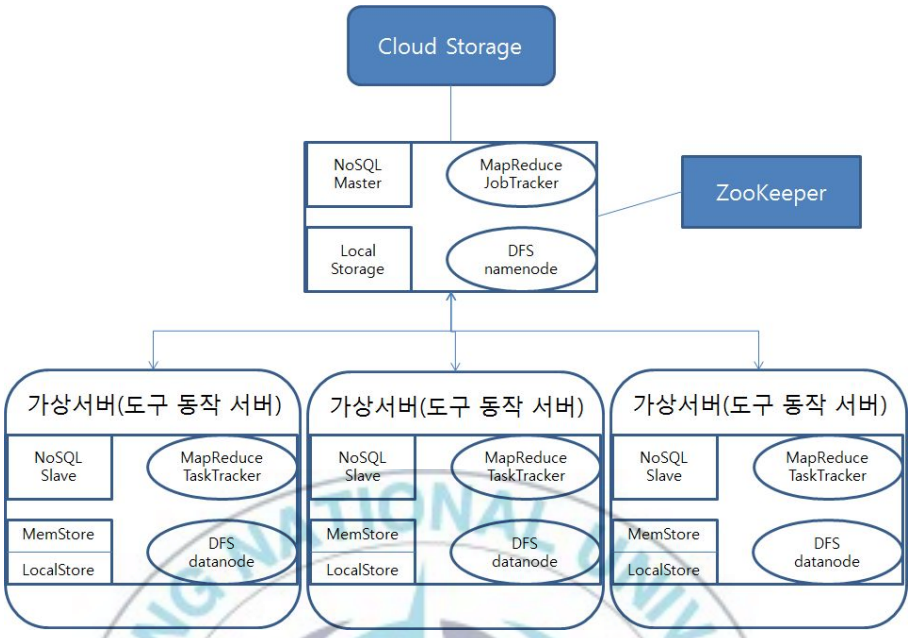


[그림 6] 아마존 웹 서비스 콘솔

AWS를 이용하면 각 EC2의 AMI 이용으로 인한 기존 디지털 포렌식 도구 사용은 가능하지만, 패스워드 해독, 이미지 색인 등 특정 기능들의 빠른 분석을 위한 분산처리를 이용해야 하는 기능들은 따로 찾을 수 없으므로 사용할 수 있도록 설정해야 한다.

이에 대해 AWS는 분산처리를 위해 Amazon Elastic MapReduce를 지원하지만, 이는 EC2와 S3 서비스의 인프라에서 실행되는 호스팅 Hadoop 프레임워크를 이용하여 실행되고 있으며, EC2를 이용하여 서비스를 제공받을 때 이미 사용 되고 있다. 때문에 EC2의 내부 어플리케이션에 있는 특정 기능들을 위하여 분산 처리를 지원하진 않는다.

따라서 Virtual Desktop 서비스에서 동작하는 디지털 포렌식 도구는 고속 분석을 위한 분산 데이터 관리 시스템이 필요하며, [그림 7]과 같이 기존 디지털 포렌식 도구의 단점 보완을 위해 추가되어야 할 필요가 있다.

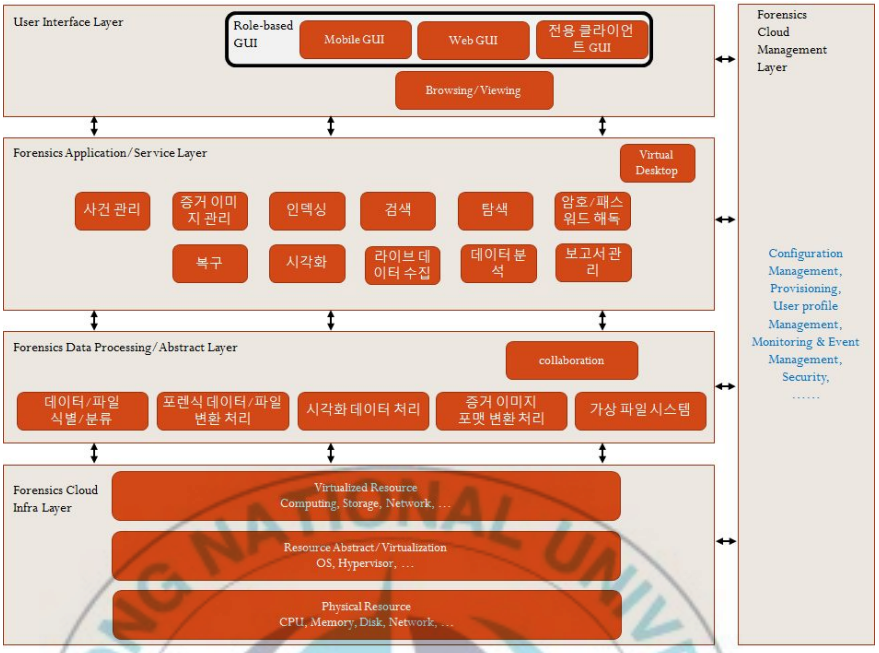


[그림 7] 분산 데이터 관리 시스템의 필요성

또한 위에서 소개했던 AWS와 같은 IaaS 서비스로 디지털 포렌식 도구를 제공하였을 때, 도구의 동작에 있어 해당 도구의 인프라를 사용자가 설정하여 도구를 사용 하는 것 보다, 어플리케이션 도구 사용에 집중할 수 있는 SaaS 형태의 도구가 더 적합할 것이며, 내부적으로 포렌식 도구의 이미지들을 관리하거나 긴급을 요구하는 분석을 위한 가상 서버의 리소스를 관리하는 분산 데이터 관리 시스템, 매니지먼트 시스템들의 역할이 중요하다.

2. DFSaaS(Digital Forensic Software as a Service) 프레임워크

클라우드 컴퓨팅 서비스를 하고 있는 AWS에 대한 기존 포렌식 도구 서비스에 대해 기술했다. 단순히 도구를 사용하는 것에만 집중한다면, 기존의 클라우드 컴퓨팅 서비스를 이용하여 디지털 포렌식 도구의 클라우드 컴퓨팅 화는 간단한 것처럼 보이지만, 도구 기능들의 분석에 대한 분산 처리 기능과 가상 서버의 리소스를 관리하기 위한 기능 등 미흡한 부분이 있다. DFSaaS는 기존 도구에 대한 기능들을 간추려 클라우드 플랫폼 구조로 구성하였고, 내부 기능들은 대부분 클라우드 공급자가 제공하는 클라우드 컴퓨팅 서비스에서 안정적으로 구동 될 수 있다. 현재 공개된 클라우드 컴퓨팅 프레임워크는 NIST의 NIST Cloud Computing Reference Architecture[14]와 IETF의 Cloud Reference Framework[15]가 있으며 이를 참고하여 DFSaaS의 프레임워크를 구축하였고, 아래 [그림 8]과 같이 나타낼 수 있다.



[그림 8] DFSaaS Framework

프레임워크를 보면 맨 위 User Interface Layer는 접속하는 도구에 대한 GUI와 브라우저를 담당한다. 아래 <표 2>와 같이 User Interface Layer는 도구에 접속하는 사용자들에 대한 도구 화면을 담당한다고 생각하면 된다.

<표 2> User Interface Layer

User Interface Layer	
Role-Based GUI	접속하는 도구에 대한 GUI를 담당하여 도구에 접속한 기기가 모바일인지, 데스크탑인지, 전용 클라이언트인지를 식별해 해당 GUI에 접근한다.
Browsing/ Viewing	도구를 사용할 때 새로운 창을 띄우거나 현재 사용하는 GUI에서 새로운 화면이 갱신될 때 마다 동작한다.

다음 두 번째 Layer인 Forensics Application/Service는 사건 보고서 생성, 관리, 수집, 분석 등 디지털 포렌식 도구의 전반적인 기능들을 사용할 수 있도록 서비스 환경을 제공한다.

<표 3> Forensics Application/Service Layer

Forensics Application/Service Layer	
사건 관리, 보고서 관리, 증거 이미지 관리 인덱싱, 검색, 탐색, 암호/패스워드 해독, 복구, 시각화, 라이브 데이터 수집, 데이터 분석	분석관, 수집관, 사건관리관이 수사 및 분석에 관한 모든 기능들을 이용한다.
Virtual Desktop	필요에 따라 기존의 디지털 포렌식 수행 도구들을 사용 할 수 있게 가상 데스크탑 서비스를 제공한다.

Forensics Data Processing/Abstract Layer는 Application/Service Layer에서 제공하는 기능들에 대한 실질적인 데이터 처리가 되는 Layer로 백그라운드에서 동작한다.

<표 4> Forensics Data Processing/Abstract Layer

Forensics Data Processing/Abstract Layer	
데이터/파일 식별/분류, 포렌식 데이터/파일 변환 처리, 시각화 데이터 처리, 증거 이미지 포맷 변환 처리 가상 파일 시스템	대용량 데이터 처리가 이루어지는 부분으로 Service Layer에서 사용자가 색인이나 검색, 이미지 포맷 등 대용량 데이터 처리를 요청하면 Processing Layer에서 백그라운드로 처리가 된다.
Collaboration	디지털 포렌식이 적용되는 경찰이나 검찰, 법원 등의 지사 협업 업무를 웹 컨퍼런스와 온라인 협업 서비스를 확대하여 보다 신속한 의사 결정이 가능하다.

Forensics Cloud Infra Layer는 DFSaaS의 클라우드 컴퓨팅 서비스의 근간이 되는 물리적 하드웨어가 위치한 곳이라고 생각하면 된다. 클라우드 컴퓨팅을 실제 동작시키고 있는 네트워크, 스토리지, 가상 서버 등이 위치하며, 이들을 이용해 클라우드 컴퓨팅 서비스를 운용한다.

<표 5> Forensics Cloud Infra Layer

Forensics Cloud Infra Layer	
가상 리소스	사용자에게 OS, Storage, Compute 등 가상 환경을 제공하여 클라우드 컴퓨팅 서비스가 동작할 수 있도록 한다.
물리 리소스	실제 클라우드 컴퓨팅 서비스가 동작하는 CPU, Memory, disk, Network 등이 위치한다.

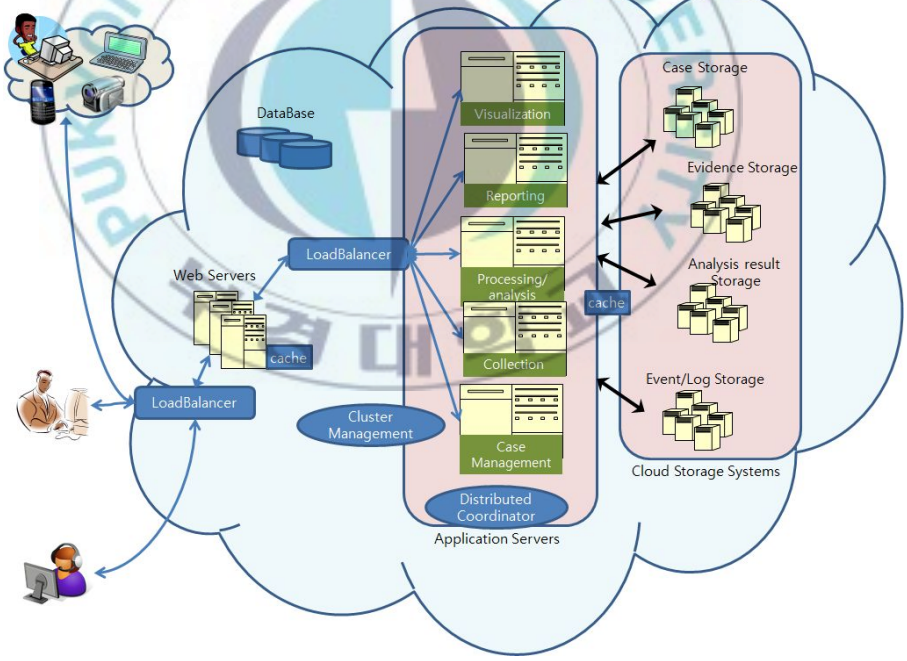
마지막으로 Forensics Cloud Management Layer는 DFSaaS 클라우드 컴퓨팅 서비스에 대한 모든 Layer를 관리하며 등록된 사용자, Layer 설정 관리, 각 사용자 모니터링 및 과금 정책 및 가상 서버 리소스 관리, 분산 데이터 시스템 관리 등 DFSaaS의 전반적인 관리를 목적으로 한다.

<표 6> Forensics Cloud Management Layer

Forensics Cloud Infra Layer	
Management, Monitering, Security	각 가상 서버 설정, 사용자 정보 관리, 사용자 권한 설정, 사용자 모니터링 및 과금, 가상 서버 리소스 관리, 데이터 관리 및 각 클라우드 컴퓨팅 서비스 보안 설정 등 DFSaaS 에 대한 관리를 담당한다.

3. DFSaaS(Digital Forensic Software as a Service) 구조

기존의 통합 도구는 단일 플랫폼에 소프트웨어 형태로 설치되어 있다. 디지털 포렌식을 수행하려면 도구가 설치된 포렌식 연구실로 이동하여 수행하여야 하며, 이는 최근 늘어난 디지털 증거의 분석 처리 속도를 저하하게 만드는 요인이 될 수 있다. 따라서 웹 상이나 전용 클라이언트로 접속하여 어디서든 포렌식을 수행할 수 있어야 한다. DFSaaS 프레임워크에 대한 도구 구조를 다음 [그림 9]와 같이 제시 한다.



[그림 9] Digital Forensic Software as a Service 구조

도구 사용은 웹 서버군에서 인증 방식을 선택하여 사용자 인증을 거친 뒤에 어플리케이션 서버군에 있는 기능들을 이용한다. 어플리케이션 서버 기능들은 사건 관리, 증거 수집, 증거 분석, 사건 보고가 존재 하며, 각 서버군은 캐시로 인한 프로세싱 속도가 증가한다. 사용자 그룹에서 로드밸런서를 통해 웹 서버 군과 어플리케이션 서버 군에 접속하여 서버의 자원에 대해 관리하며 다중 사용자를 위한 분산 코디네이터로 서버를 관리한다. DFSaaS의 각 구성 요소는 다음과 같이 나눌 수 있다.

가. 웹 서버군

데이터베이스에 저장된 사용자 정보로 인증을 하거나 공인인증서를 이용한 사용자 인증 기능, 어플리케이션 서버에 있는 도구 기능을 요청하는 작업, 백그라운드로 캐시를 이용하여 서버의 처리 속도를 높이는 기능 등이 포함된다.

나. 어플리케이션 서버군

- 사건 관리 : 포렌식 절차의 조사 준비 및 현장 대응 단계에 해당하는 기능으로 사건 Case 생성, 수사 담당자 지정, 사용자 별 업무 내용 기록, 사건 기록 저장소에 저장 기능 등이 포함된다.
- 증거 수집 : 포렌식 절차의 증거 확보 및 수집, 증거 이송 및 확인에 해당하여 도구가 수행할 수 있는 기능은 휘발성 데이터 수집 기능과 현장에서 이루어 질 수 있는 사진 촬영, 증거물 위치 촬영 등 현장 인증 데이터 전송 기능, 이송을 위한 Chain of custody 기능이 포함된다.
- 증거 분석 : 조사 및 분석에 해당하여 증거 이미지 사본 생성,

저장된 이미지를 읽어서 인덱스 데이터베이스를 생성하는 기능, 비트 단위 검색, 인덱스 기반 검색, 정규표현식 검색, 패턴 검색 등 검색 기능이 존재하며, 삭제된 데이터를 복구하거나 비할당 영역에 있는 데이터를 복구하는 데이터 복구 기능, 패스워드 해독 기능, 은닉 데이터 탐지인 스테가노그라피 등 안티포렌식 대응 기능, 분석자의 보고서 작성을 쉽게 도와주기 위한 연관성 데이터 분석 등이 포함되어야 한다.

- 사건 보고 : 보고 및 제출 단계에 해당하여 각 결과물들을 종합해 보고서를 만들고, 제출해야 할 문서 포맷에 맞게 보고서를 변환한다. 이때 원활한 보고서 작성을 위해 분석 완료된 결과물들을 시각화하여 보여주는 분석결과 뷰어 기능 등이 포함되어야 한다.

다. 분산 데이터 시스템

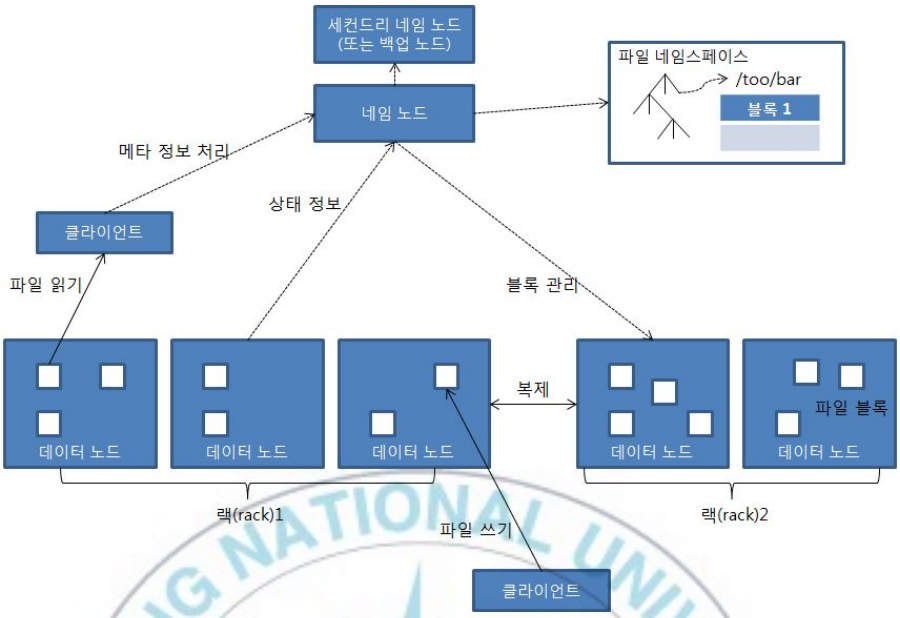
- 사건 기록 저장소 : 사건 case, 수사관 업무 내용 저장 및 최종 분석 결과 보고서를 저장한다.
- 증거 저장소 : 현장에서 수집하는 증거물 저장, 증거 이미지 원본 저장, 수집된 활성데이터 이미지 사본 저장, 현장 인증 데이터 저장 및 chain of custody가 저장된다.
- 증거 분석 결과 저장소 : 분석된 이미지 사본, 고속 처리를 위한 인덱싱 결과물 저장, 검색 결과, 데이터 복구 결과, 안티포렌식 대응을 위한 패스워드 해독 결과 등 분석된 결과가 저장된다.
- 이벤트/로그 저장소 : 도구 사용에 발생된 이벤트를 저장하는 사용자 이벤트 저장과 사용자의 사용 행위 로그가 저장된다.

4. DFSaaS(Digital Forensic Software as a Service) 흐름도

DFSaaS 구조[그림 5]는 크게 웹 서버군, 어플리케이션 서버군, 클라우드 스토리지 시스템으로 나뉜다. 전체적인 흐름은 Hadoop 파일 시스템으로 동작하며 각 구조들은 개별적인 흐름을 갖고 있다. 고속 처리를 위한 부분은 Hadoop MapReduce를 이용한 분산처리가 적용된다[16].

가. Hadoop 파일 시스템

DFSaaS는 Hadoop 파일 시스템[16][17]으로 구성되어 운영된다고 가정 할 수 있다. Hadoop 파일 시스템은 구글 파일 시스템의 설계와 아이디어를 이용해 개발 됐으며, 많은 부분에서 구글 파일 시스템과 유사하다. 다음 [그림 10]은 Hadoop 파일 시스템 구성도를 나타낸 것이다.

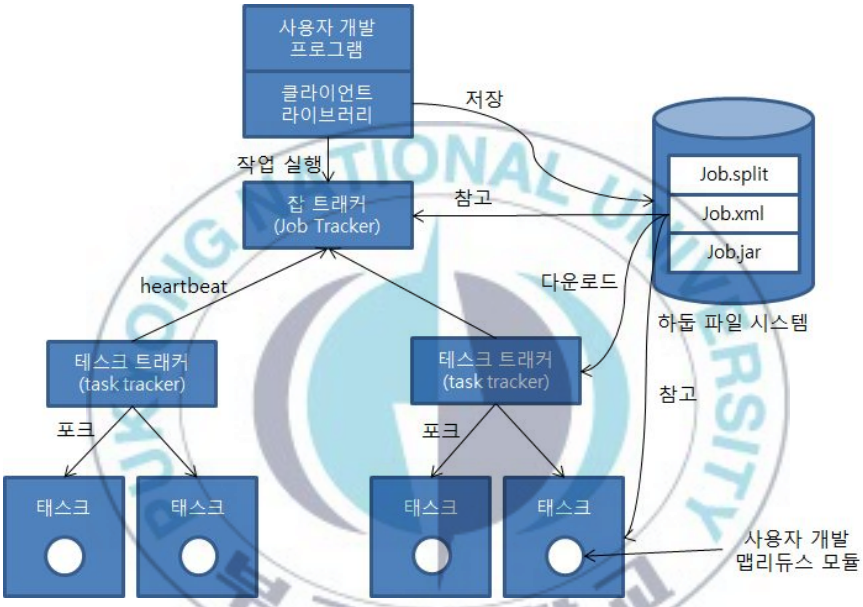


[그림 10] Hadoop 파일 시스템

Hadoop 파일 시스템은 하나의 네임 노드 서버, 세컨드리 네임 노드와 다수의 데이터 노드 서버들로 구성된다. 네임 노드는 파일 시스템의 네임스페이스(디렉토리, 파일명, 파일블럭 등)를 관리하면서 클라이언트의 파일 접근 요청을 처리한다. Hadoop 파일 시스템에서 파일데이터는 블록 단위로 나뉘어 여러 데이터 노드에 분산돼 저장된다. 블록은 구글 파일 시스템의 청크와 동일한 개념으로 가용성을 보장하기 위해 복제돼 다른 데이터 노드에 저장된다. 특정 서버의 장애 발생 시 자동으로 이를 감지해 장애가 발생하지 않은 서버의 복사본을 이용해 다른 노드에 복제본을 추가로 생성한다.

나. Hadoop MapReduce 시스템

Hadoop MapReduce도 Hadoop 파일 시스템과 유사한 분산 시스템을 구성한다. 다음 [그림 11]은 Hadoop MapReduce의 시스템 구성을 나타낸 그림이다.



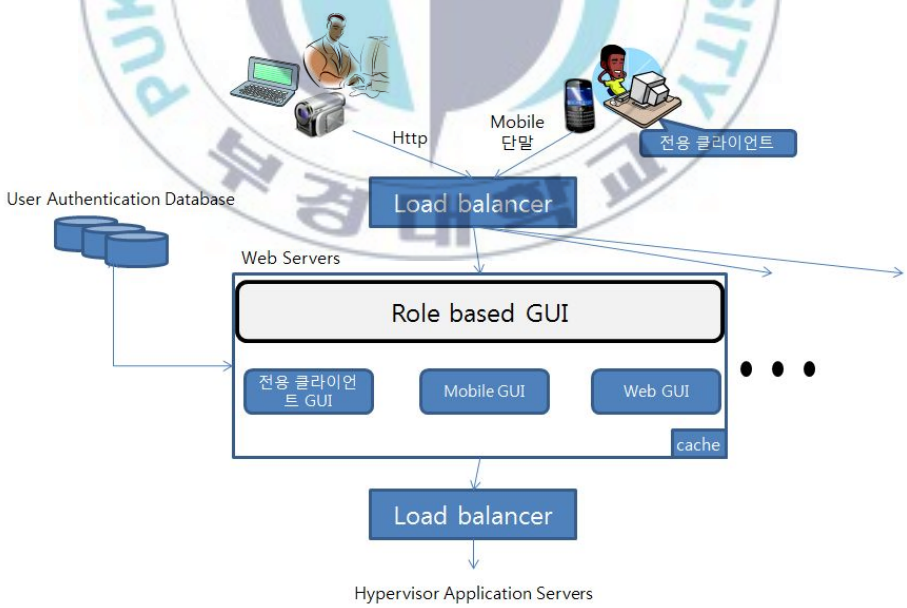
[그림 11] Hadoop MapReduce 시스템

Hadoop MapReduce 시스템은 잡 트래커, 태스크 트래커, 클라이언트 라이브러리로 구성된다. 잡 트래커는 전체 작업을 관리하는 기능을 수행하는 마스터 역할을 수행하는 서버이고, 태스크 트래커는 사용자가 요청한 작업을 실행하며 하나의 MapReduce 클러스터는 한 개의 잡 트래커와 여러 개의 태스크 트래커로 구성된다. 잡 트래커는 주로 Hadoop 파일 시스템의 네임 노드가 실행되는 서버에서 실행되고, 태스크 트래커는 데이터 노드가 실행되는 서버에 실행한다. 클라

이언트 라이브러리는 사용자가 다양한 입력 형태와 분산 처리를 사용할 수 있는 메커니즘을 제공하며, 사용자가 개발한 프로그램을 잡 트래커로 작업하도록 요청하고 작업 결과를 모니터링 할 수 있는 API를 제공한다.

다. 도구 접속 흐름도

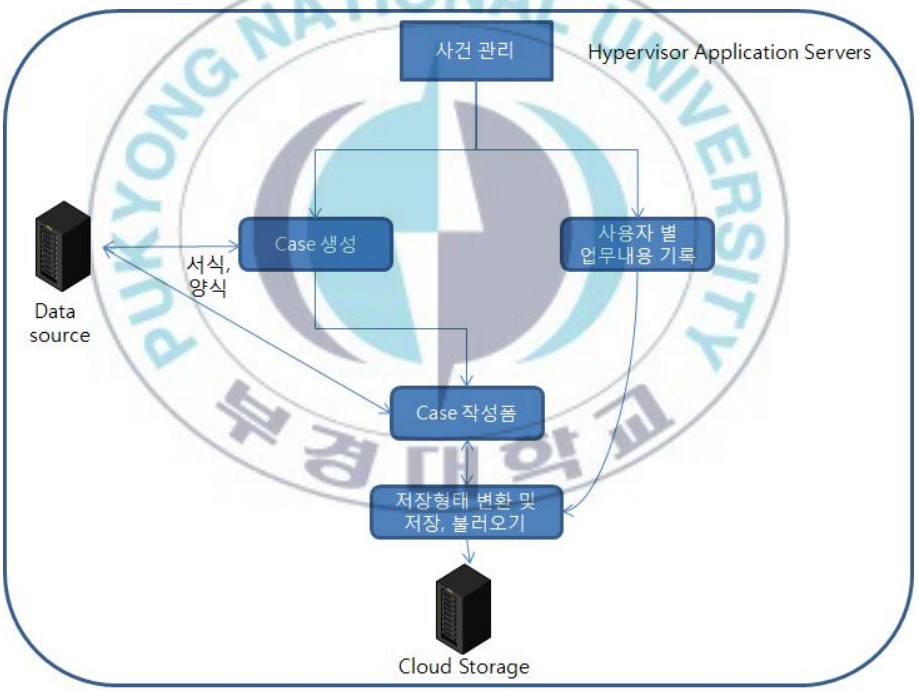
웹 또는 모바일 단말, 전용 클라이언트로 도구에 접속하며 접속시 사용자 인증을 거치게 된다. 사용자 인증은 인증 데이터베이스에 있는 사용자 정보를 이용할 수 있으며, 공인인증서를 이용한 인증 방법이 있을 수 있다. 인증을 거친 사용자는 도구에 접속하며 로드밸런서는 사용되지 않는 웹 서버군을 선택해 사용자에게 리소스를 할당한다. 다음 [그림 12]은 도구 접속 흐름도를 나타낸다.



[그림 12] 도구 접속 흐름도

라. 사건 관리 흐름도

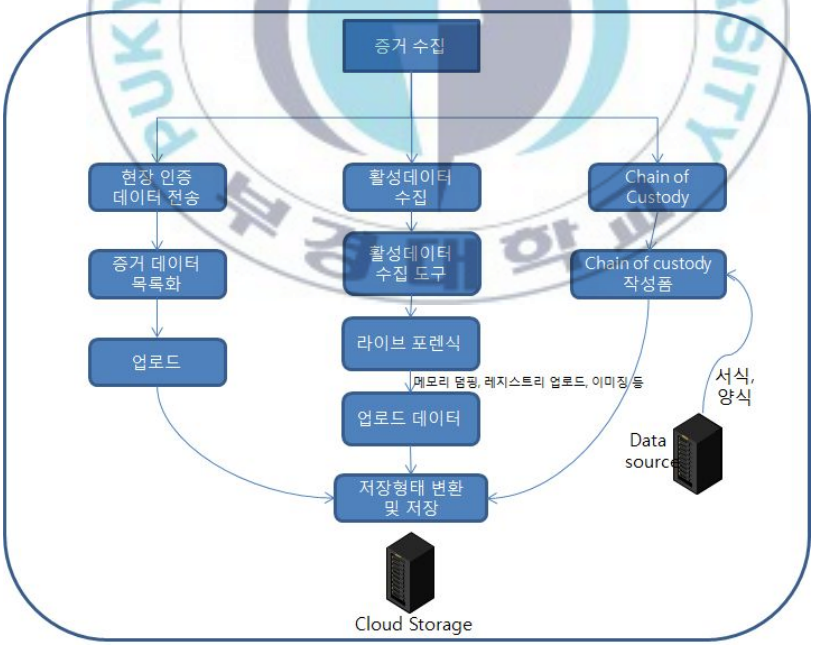
Case를 생성하는 기능과 사용자 별 업무내용을 기록하는 기능으로 나눌 수 있다. Case는 데이터 소스 저장소에서 서식이나 양식을 불러와 생성하며, Case 작성 폼에서 작성한다. 스토리지에 저장될 때, 저장 형태에 맞게 변환이 되어 저장되고, 사용자가 저장된 Case를 불러 업무내용을 기록하고 싶을 때 다시 Case를 불러와서 Case 작성폼에서 작업한다. 다음 [그림 13]는 사건 관리 흐름도를 나타낸다.



[그림 13] 사건 관리 흐름도

마. 증거 수집 흐름도

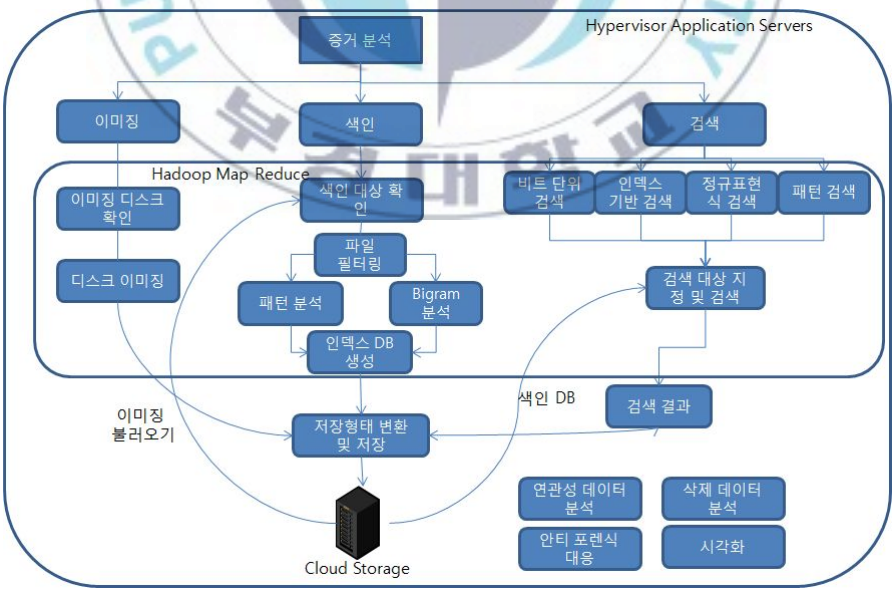
도구를 이용한 수집 기능으로 현장 인증 데이터 전송 기능, 활성 데이터 수집 기능, 전송을 위한 Chain of custody 작성 기능 등이 있을 수 있다. 현장에서 촬영한 사진이나 비디오 파일 등을 도구에 바로 업로드 해 데이터들을 목록화 하고 스토리지에 저장한다. 활성데이터를 수집해야할 경우에는 활성데이터 수집도구를 이용하여 라이브 포렌식을 수행, 메모리 덤프나 레지스트리를 수집하고 하드웨어를 압수하지 못할 경우에는 디스크 이미징을 통해 증거를 수집한 후 도구에 업로드 하여 저장한다. Chain of custody는 데이터 소스 저장소에서 역시 서식과 양식을 받아 작성 폼에서 작성 후 저장된다.
다음 [그림 14]은 증거 수집 흐름도를 나타낸다.



[그림 14] 증거 수집 흐름도

바. 증거 분석 흐름도

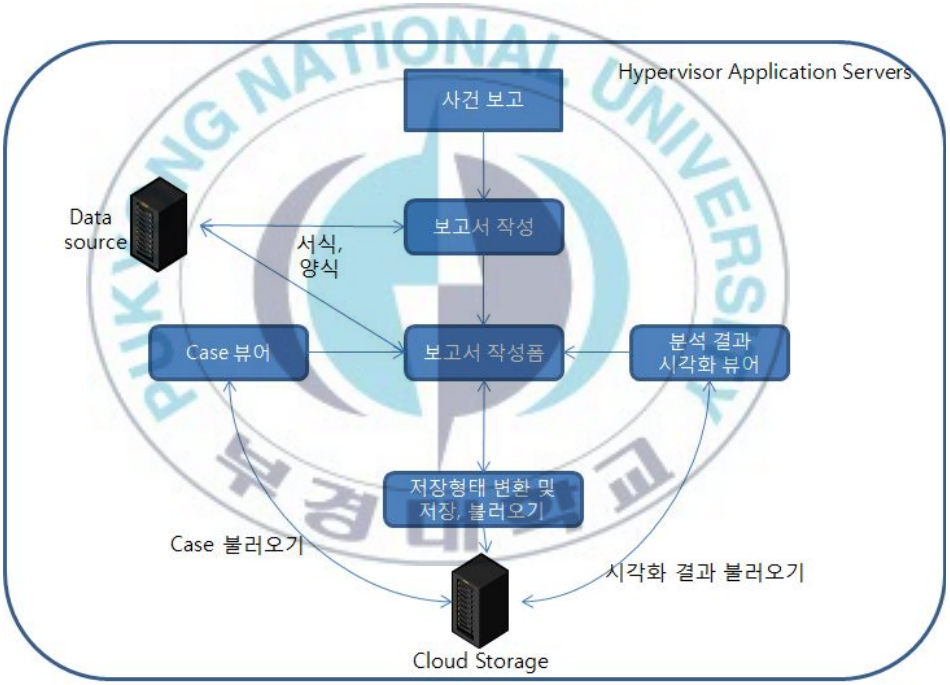
증거 분석에 따른 기능들은 이미징, 색인, 검색, 삭제 데이터 분석, 연관성 데이터 분석, 안티포렌식 대응, 시각화가 있을 수 있다. 논문에서는 이미징, 색인, 검색에 관한 흐름도를 제시하며 차후 지속적인 업데이트가 이루어져야 할 것이다. 데이터 이미징을 위해서 이미징 디스크를 확인하는 작업을 거친 후 디스크 이미징 사본 생성을 한다. 색인은 대상을 확인 한 후 파일 필터링을 통해 각 분석이 이루어지고 인덱스 데이터베이스를 생성하여 저장한다. 검색 방법은 비트 단위 검색, 인덱스 기반 검색, 정규표현식 검색, 패턴 검색 방법 등이 있을 수 있다. 역시 검색 대상을 지정하여 검색하고 검색된 결과는 스토리지에 저장된다. 다음 [그림 15]은 증거 분석 흐름도를 나타낸다.



[그림 15] 증거 분석 흐름도

사. 사건 보고 흐름도

분석된 데이터를 참고로 사건에 대한 보고서를 작성하여 저장한다. 사건 관리 흐름도와 비슷하게 보고서 작성폼에서 작성하며 데이터 소스 저장소로부터 서식이나 양식 등을 제공받는다. 이때, 보고서 작성을 위해 사건 Case를 참고하기 위한 Case 뷰어 기능, 분석 결과를 보다 쉽게 알게 하도록 시각화 기능을 통해 보고서를 작성한다. 다음 [그림 16]는 사건 보고 흐름도를 나타낸다.



[그림 16] 사건 보고 흐름도

5. DFSaaS(Digital Forensic Software as a Service) 특징

기존의 통합 도구는 디지털 포렌식 수행에 대한 많은 제약이 따른다 상황에 따라서 다른 도구들을 이용해야 하고, 도구가 설치된 장소에 가지 않으면 포렌식을 수행 할 수 없다. 앞서 나열한 기존 디지털 포렌식 도구의 단점과 DFSaaS 도구의 특징을 비교하여 본다.

다음 <표 7>은 기존 디지털 포렌식 도구의 단점과 DFSaaS의 도입으로 얻을 수 있는 특징들을 나열한 것이다.

<표 7> DFSaaS 도입으로 인한 장점

기존 포렌식 도구의 단점	DFSaaS 도입으로 인한 장점
단일 머신에서 동작하는 디지털 포렌식 도구의 단점	클라우드 컴퓨팅 기반 도구 서비스로 인해 하드웨어적 한계가 사라짐. 필요한 경우, 필요한만큼 즉시 컴퓨팅 자원들이 제공됨
단일 도구 사용으로 인한 디지털 포렌식 수행 제약	필요 기능의 추가 구입 없이, 필요할 때 필요한 기능을 서비스 받아 사용 가능. 또한 클라우드 서비스의 분산/병렬 처리 그리고 협업 지원을 통해 고속의 분석 작업 가능
도구의 유지보수를 위한 기능 확장시 전용 웹사이트의 절차에 따라 일일이 업데이트 실시	필요한 기능을 바로 업데이트하여 따로 설치할 필요가 없음
도구가 설치된 장소에서만 디지털 포렌식 수행 가능	네트워크가 설치되어 있는 곳이라면 어디서든 DFSaaS 도구 사용 가능

따라서 DFSaaS 도구의 장점들은 아래와 같이 설명 할 수 있다.

(1) 온라인 협업

온라인 도구로써 사내 협업이 가능하다. 또한 사용자 인증을 거치게 되면 어디서든 디지털 포렌식 수행이 가능하다.

(2) 하드웨어적 한계점 극복

클라우드 컴퓨팅 기반 도구로써 기존 단일 머신에서 가졌던 스토리지, 메모리 용량 등 하드웨어적 한계가 사라지고, 내부 분산처리로 인한 데이터 분석 속도 또한 증가 하여 기존의 한계점이 사라진다.

(3) 도구 유지보수

도구에 필요한 기능들을 업데이트하여 따로 다운로드나 설치할 필요 없이 바로 사용 가능하다.

(4) 단일 도구의 포렌식 수행 제약 극복

기존 도구의 이미지들을 저장하여 필요시 Virtual Desktop 을 이용하여 사용 할 수 있다. 이는 기존의 많은 도구들을 하나의 도구로 묶어 사용 가능하게 하며, 많은 업무들을 한번에 이룰 수 있게 한다.

현재 클라우드 컴퓨팅 서비스를 진행하고 있는 AWS에 DFSaaS를 적용시킬 경우 Interface Layer와 Infra Layer를 제외한 Application, Processing/Abstract Layer 등은 제공되지 않으므로 앞서 3절에서 설명한 도구의 특징들을 파악하여 분산 데이터 관리 시스템 및 포렌식 도구를 위한 각 기능들이 추가되어야할 필요가 있다.

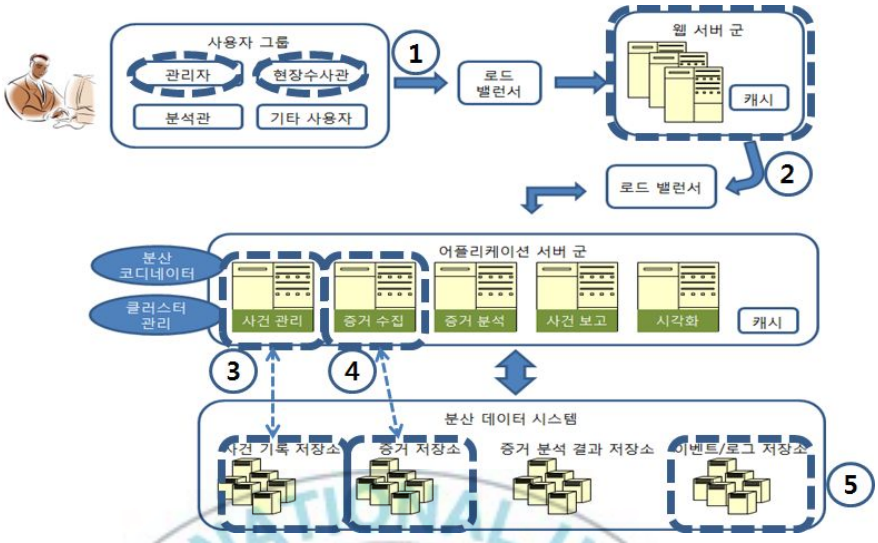
또한 사용자의 편의를 위하여 도구 사용자와 관리자를 분리하여 사용자의 경우 도구의 사용에만 집중 할 수 있도록 Infra가 설정된 상황에서 서비스를 이용 할 수 있어야 할 것이며 관리자는 가상 서버 관리, 분산 데이터 시스템 관리 및 긴급 상황시 가상 서버들의 리소스를 조절하여 빠른 업무 처리를 가능하게 하는 리소스 매니지먼트 시스템이 필요하다.

6. DFSaaS(Digital Forensic Software as a Service) 시나리오

앞서 설명했던 디지털 포렌식의 6가지 절차별로 나누고 컴퓨터, 이동전화 포렌식 가이드라인[7][8]에 따라 DFSaaS 도구의 사용 시나리오를 다음과 같이 제시한다.

(1) 조사 준비, 현장 대응

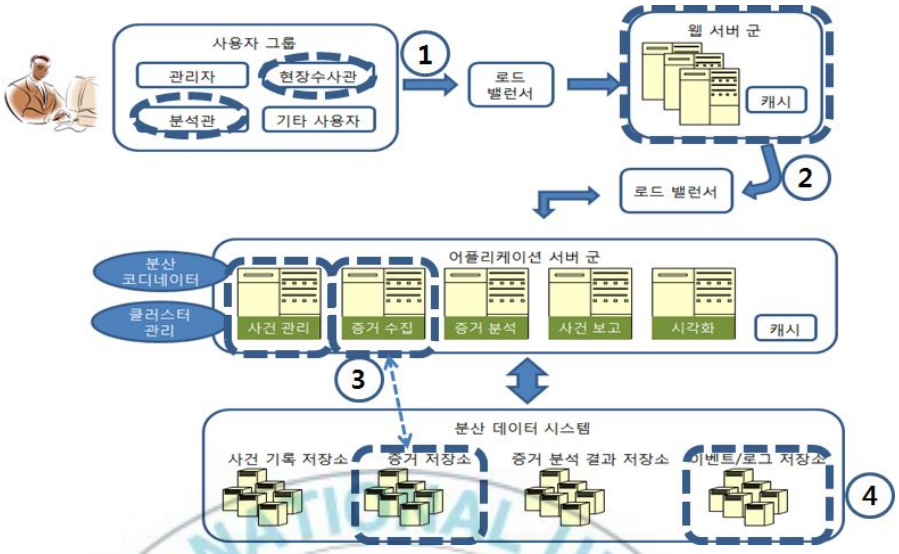
사건 현장에서 디지털 증거를 포함하고 있는 매체를 수집하기 위한 조사 준비 과정이다. 디지털 증거물 수집 계획의 수립, 디지털 증거 수집 팀 구성, 수집 및 분석에 필요한 장비 확보 등이 이루어 진다. 이에 대해 수사관은 도구에 접속하여 사건 관리 기능에서 사건 Case를 생성, 어떤 사건인지 업무 내용을 기록하고 저장한다. 차후 생성한 사건 Case를 지속적으로 불러서 사건에 대한 업무를 기록한다. 지난 사건 기록들을 열람할 때에는 지난 Case를 불러와서 참고한다. 현장에 출동한 증거 수집가는 디지털 포렌식을 수행하면서 발생되는 기기 촬영, 관계자 협조 데이터, 물리적 기기 위치 촬영 등 현장 인증 데이터들을 수집하여 발생한 데이터 파일들을 도구 접속 전용 클라이언트를 이용해 증거 수집에 있는 업로드 기능을 이용해 증거 저장소에 저장한다. 단계에 상관없이 도구 사용자들이 도구를 사용할 때 발생한 이벤트나 로그 역시 로그 저장소에 저장된다. 다음 [그림 17]은 조사 준비, 현장 대응에 대한 시나리오를 나타낸다.



[그림 17] 조사 준비, 현장 대응 시나리오

(2) 증거 확보 및 수집, 이송 및 확인

수사관이 현장에 출동하여 수집 대상에 있는 증거 데이터 들을 모두 수집한다. 앞서 현장 대응에서 수집했던 현장 인증 데이터 뿐만 아니라 도구를 이용해 수집할 수 있는 활성데이터, 디스크 이미징 등이 있을 수 있으며 DFSaaS 도구의 특성상 웹상에서 하드웨어 내부적인 데이터를 수집하기엔 한계가 있으므로 라이브 포렌식 도구(USB 클라이언트 등)를 이용하여 수집한 후 저장소에 저장한다. 각 증거데이터 들을 이송 및 저장 시 Chain of custody에 따라 증거에 대한 목록을 첨부한다. 분석관은 Chain of custody를 통해 수집된 증거 목록을 확인하고 증거가 잘 보관되었다고 인증한다.



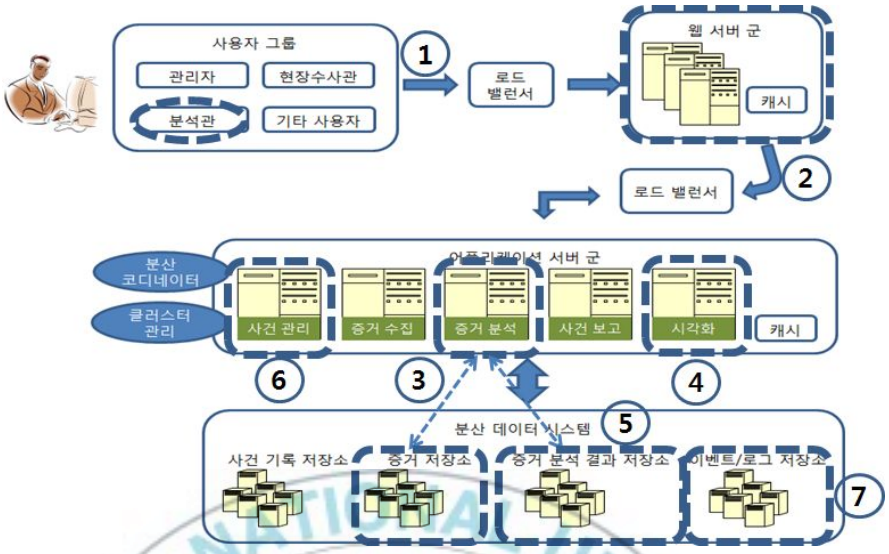
[그림 18] 증거 확보 및 수집, 이송 및 확인 시나리오

(3) 조사 및 분석

분석관은 현장 수사관이 수집하고 저장한 저장소를 통해 증거들을 불러와 증거 분석에 있는 기능들을 이용해 데이터를 분석한다. 디지털 증거 분석은 디스크 이미지를 이용하는 것이 일반적이나 여의치 않을 경우 원본 이미지를 직접 이용 할 수 있다. 하지만 이런 경우 원본 이미지의 데이터 변경을 최소화하고 변경 사항에 대한 이해를 하고 있어야 하며 이러한 사실을 최종 보고서에 포함하여야 한다. 먼저 증거 저장소에 저장된 원본 이미지를 불러와 증거 분석에 있는 이미징 기능을 이용하여 사본을 생성, 증거 저장소에 저장한다. 이후 내부 데이터를 해칠만한 악성 코드의 감염 여부를 확인하여 고속 데이터 검색을 위해 증거 분석의 기능 중 인덱싱 기능을 이용하여 증거 저장소에 저장된 사본 이미지에 대한 인덱싱을 실시한다. 이 후 증거 분석의 검색 기능을 이용하여 사본 이미지에 대한 전체 검색을 실시

하고, 삭제된 데이터나 은닉, 암호화된 데이터들을 추출한다. 중요한 디지털 증거는 저장된 데이터 뿐 아니라 이미 삭제된 데이터에서도 매우 중요한 역할을 할 수 있다. 따라서 증거 저장소의 사본 이미지 내부의 섹터 갭이나 슬랙 스페이스 등에 위치한 삭제된 데이터를 복구하기 위해 증거 분석 기능에 있는 데이터 복구를 이용하여 데이터를 복구해야 한다. 암호화된 파일이 있을 경우 증거 분석의 암호 해독기능을 이용하여 파일 암호 분석을 실시한다. 파일의 타임라인이나 은닉 파일을 분석 할 때 시각화 기능을 이용하여 파일간의 유사도, 분포도, 사용자의 흔적, 파일이 갖고 있는 내부 정보 등을 시각화 하여 분석관이 보다 쉬운 분석을 할 수 있도록 해야 할 것이며, 원활한 보고서 작성을 위해 각 분석 결과가 나올 때 마다 보고서를 작성해, 분석 결과 저장소에 저장한다. 수집관이 현장에서 증거 수집 기능을 이용해 라이브 포렌식 데이터를 전송하여 증거 저장소에 저장하면 분석관은 저장된 데이터를 불러와 데이터 분석을 실시하면서 수집되지 않은 분석에 필요한 데이터를 요구할 수 있다. 예로 인터넷 사용 흔적을 분석 중 네트워크 증거가 필요하다고 판단되면 통신망에 대한 증거물을 추가로 더 수집한다.

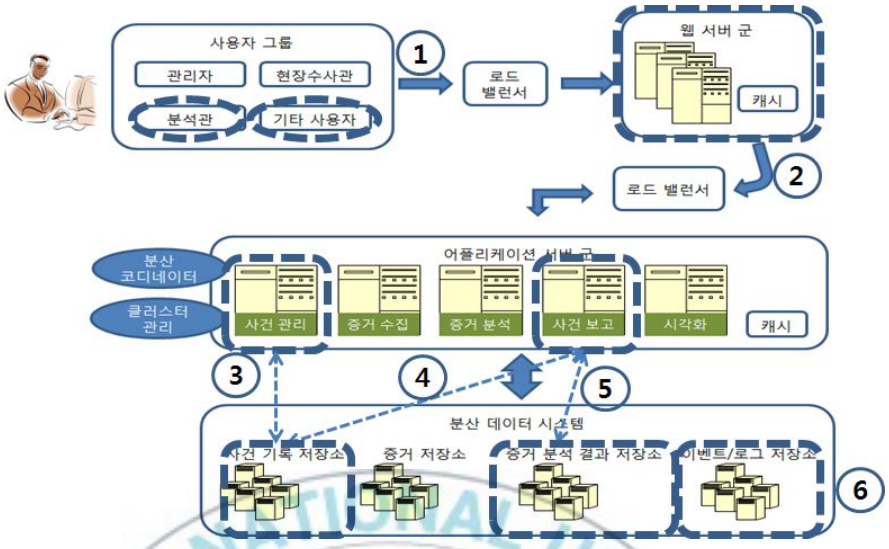
증거를 분석하는 다수의 분석관이 도구를 사용하여 분석 할 수 있다. 이 때 긴급을 요하는 암호 해독이나, 데이터 검색이 있을 수 있는데, 리소스가 사용되지 않는 서버군이나 작업 중요도가 낮은 서버군의 리소스들을 해당 어플리케이션 서버군에 할당하여 타 서버군보다 빠른 분석이 이루어 질 수 있도록 한다.



[그림 19] 조사 및 분석 시나리오

(4) 보고 및 제출

분석관은 분석이 끝난 결과물들을 이용해 보고서를 작성한다. 증거 분석 기능들을 이용하여 증거 저장소에 저장된 디지털 증거 데이터들의 분석 완료된 데이터들은 증거 분석 결과 저장소에 저장되어 있으므로, 시각화 기능을 이용하여 분석이 완료된 데이터 들을 증거 분석 결과 저장소에서 불러와 분석관이 보고서를 보다 쉽게 적을 수 있도록 한다. 최종 보고서는 상호 데이터를 이용하여 최대한 쉽고 명확하게 작성하여야 한다. 작성이 끝난 보고서는 사건 기록 저장소에 저장하여 변호사나 법정 관련자가 사건 관리 기능을 이용하여 보고서를 열람 할 수 있어야 한다.



[그림 20] 보고 및 제출 시나리오

IV. 결론

본 논문에서는 늘어나는 디지털 증거와 온라인 연동 데이터에 대비해 기존의 단일 플랫폼에서 동작하는 디지털 포렌식 도구가 아닌 클라우드 컴퓨팅 환경을 접목시켜 DFSaaS 프레임워크를 제시하고 구조, 내부 흐름도와 도구 사용 시나리오를 제시하였다. 따라서 이를 이용해 도구를 구현 할 수 있으며, 웹 또는 전용 클라이언트를 이용하여 도구를 이용하면 언제 어디서든 디지털 포렌식 업무를 수행 할 수 있으며, 저장된 최종 보고서를 따로 첨부하여 제출하지 않아도, 법정 관련자들은 도구에 접속해 저장된 보고서를 검토하고나, 지난 사건 Case를 살펴 볼 수도 있다.

하지만 논문에서 다루고 있는 구조나 흐름도는 기본적인 것에 불과하며, 각각의 수사 단계와 필요한 상세 기능들은 지속적으로 연구하여 디지털 포렌식 수사 단계가 완벽히 클라우드 컴퓨팅 도구로써 적용될 수 있어야 할 것이다. 추가 고려사항으로 기존 클라우드 서비스에 DFSaaS가 적용 될 경우, 3절에서 설명하였던 도구에 대한 특징들을 파악하여 Application Layer, Processing/Abstract Layer 등 주요 기능들과 분산 데이터 관리 시스템 들이 고려되어야 할 것이다.

또한 제시한 프레임워크와 구조를 기반으로 실제 도구가 구현되어 수집과 분석에 관해 디지털 포렌식 수행 뿐 아니라 법원과 연계한 클라우드 서비스 역시 연구되어야 하며 실사용자들에 대한 요구 사항 등 많은 부분에 있어서 연구가 필요하고 이를 개선해 나가야 할 것이다.

참고 문헌

[1] 이태림, 신상욱, “디지털 포렌식을 위한 증거 분석 도구의 신뢰성 검증”, 정보보호학회논문지 제21권 제3호, 2011.6

[2] 이주영, 은성경, 홍도원, “디지털 포렌식 패러다임 전환 : 포렌식 클라우드”, 디지털 포렌식 기술 워크샵 , pp 101 ~ 107, 2010

[3] Peter Mell, Timothy Grance, The NIST Definition of Cloud Computing(Draft), NIST, Special Publication 800-145, 2011

[4] 이상진, “디지털 포렌식 개론”, 이룬출판사, 2010. 8

[5] Forensic-Proof(Digital Forensic Community), http://forensic-proof.com

[6] 정익래, 홍도원, 정교일 “디지털 포렌식 기술 및 동향”, 전자통신동향분석 제 22권 제 1호, 2007년 2월

[7] 정보통신단체표준 “컴퓨터 포렌식 가이드라인”, TTAS.KO-12.0058, 한국정보통신기술협회, 2007, 12

[8] 정보통신단체표준 “이동 전화 포렌식 가이드라인”, TTAS.KO-12.0059, 한국정보통신기술협회, 2007, 12

[9] 김형준, 조준호, 안성화, 김병준, “클라우드 컴퓨팅 구현 기술”, 에이콘출판, 2011. 1

[10] Darryl Chantry “애플리케이션을 클라우드에 매핑”, May 27, 2010

[11] Wikipedia-NoSQL, http://en.wikipedia.org/wiki/NoSQL

[12] Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach, Mike Burrows, Tushar Chandra, Andrew Fikes, Robert E. Gruber, “Bigtable : A Distributed

Storage System for Structured Data", ACM Transactions on computer Systems, Volume 26, Issue 2, June 2008

[13] AWS(Amazon Web Service), http://aws.amazon/com/ec2/

[14] B.Khasnabish, et al., "Cloud Reference Framework", IETF Internet-Draft, draft-khansnabish-cloud-reference-framework-00.tx, 2010, 12, 31

[15] NIST SP 500-292, "NIST Cloud Computing Reference Architecture", Sept. 8, 2011

[16] hadoop, http://hadoop.apache.org

[17] 톰 화이트, "Hadoop 완벽 가이드", 한빛미디어, 2010. 5

[18] 한국인터넷진흥원 "클라우드 서비스 융·복합 모델 발굴 및 사업화 추진 연구", KISA-WP-2010-0044, 한국과학기술원, 2010. 11

[19] 크리스포터 M. 모이어 "개념, 패턴, 그리고 프로젝트 클라우드 기반 애플리케이션 개발", 제이펍, 2011. 10

[20] 조지 리스, "클라우드 컴퓨팅 애플리케이션 아키텍쳐", 지앤선, 2010. 2

감사의 글

많은 사람들의 논문 후기를 읽어보게 되면 정말 번갯불에 콩 구워먹듯이 시간이 흘러간다고 합니다. 에이 설마 그렇게 빨리 갈까 라는 생각을 했지만 저 역시 똑같은 말을 쓰고 있군요. 부경대학교 LACUC 연구실에서 석사 과정을 무사히 마치면서 감사의 글을 올립니다.

열심히 하겠습니다. 라는 말만 가지고 연구실에 찾아와 대학원 진학을 도와주신 지도교수님 신상욱 교수님 정말 감사드립니다. 지나고 보면, 항상 부족하고 서툴렀던 절 묵묵히 지도해주시고 배려해주셔서 무사히 마칠 수 있게 되었습니다. 또한 바쁘신 와중에도 제 논문을 평가 해주시고 검토해주셨던 송하주 교수님, 논문의 세세한 부분까지 신경써주셨던 신원 교수님께도 감사의 말씀을 드립니다.

부경대학교라는 새로운 출발점에서 LACUC 연구실 멤버들은 가족과 같았습니다. 연구실에서 식사를 같이하며 항상 저에게 조언을 해주시고 나아갈 길 및 인생을 상담해주신 이혜주 사모님, 반할 것 같은 수완 선배, 웃음이 매력적이며 연구실을 나서는 순간 옆집 형님 같은, 하지만 일할땐 무시무시한 태림 선배, 연구실에서 유일하게 알고 있었던, 그리고 연구실 생활에 수많은 도움이 되었던 주영이, 연구실에 처음 들어와 친해졌던 김별명 헌이, 많은 대화의 끝에 서로를 이해했던? 중원이, 동갑내기 훤칠하고 키크고 잘생기며 소심한 준용이, 성훈이, 수학과 수재 헌민이 그리고 우리 연구실에 고생이 많은 남자 막내 김종특 인혁이, 그리고 연구실 레이디 수빈이, 형주, 혜주 석사 생활에 참많은 도움을 주신 영신이형, 태웅이형, 원준이, 중기, 은혜, 다니엘, 성지, 혜진이 모두모두

너무 고맙고 또 너무 고맙습니다. 그리고 인생의 동반자 윤제, 석주, 대성이(태성이), 만제형, 아마 우린 죽을 때 까지 볼 거야

마지막으로 석사 생활을 끝마칠 수 있게 뒷바라지 해주신 가족, 특히 우리 부모님, 형님, 형수님, 조카 소연이, 그리고 2년 동안 함께 학교 생활을 해오며 하나하나 다 챙겨주던 해란이에게 너무너무 고맙고 사랑합니다.

전 정말 복이 많은 사람입니다. 위 감사의 글은 정말 축약하고 축약해서 적은 감사의 글입니다. 말로 하면 다 못할 정도로 많은 고마움을 느끼고 있습니다. 길다면 길고, 짧다면 짧은 석사 2년 생활을 끝마칠 수 있게 도와주셔서 너무 감사합니다. 미처 언급을 하지 못했던 고마운 분들에게도 감사의 마음을 전합니다.