클라우드 컴퓨팅 환경에서 자료은닉기법의 활용에 관한 연구

배수환*, 신용태**
*숭실대학교 융합소프트웨어학과
**숭실대학교 컴퓨터학과
email:shbae0213@naver.com*, shin@ssu.ac.kr**

A Study on Utilization of Data Hiding technique in the Cloud Computing environment

Su-Hwan Bae*, Yong-Tae Shin**
*Dept of Software Convergence, Soongsil University
**Dept of Computing, Soongsil University

요 약

클라우드 시대가 도래함에 따라 클라우드 상에 생성, 저장되는 정보의 보호 또한 중요해지고 있다. 현재 정보보호를 위해 주로 암호화 기법을 사용하고 있지만 정보유출 및 해킹 등 보안사고의 우려로 클라우드 이용률이 저조한 상황이다. 암호화 외에 안전성이 높은 자료은닉을 위해 데이터 마스킹과 블록체인을 활용하는 것이 높은 호환성과 효율성을 가질 수 있는지 연구하고자 한다. 이에 본 논문에서는 암호화를 사용한 이중 사용자 인증, 데이터 마스킹을 사용한 가상 마스킹 데이터베이스, 블록체인을 사용한 데이터베이스 트랜잭션 공유를 제안하였다. 각각의 제안 방법을 통해 비인가된 사용자의 접근을 통제하고, 정보 유출 시 피해 방지와 위변조 방지의 역할을 수행할 수 있게 된다.

1. 서 론

최근 클라우드 컴퓨팅(Cloud Computing) 환경을 이용하여 제공되는 서비스들이 국가와 산업 전반적으로 활발하게 이루어지고 있다. 그러나 최근 빈번하게 공공기관 및 기업들의 개인신상정보나 기밀문서가 유출되면서 사회적인 논란을 일으키고 있다. 또한 현재 사설 클라우드 컴퓨팅을 사용하고 있는 공공기관과 기업들이 공공 클라우드 컴퓨팅 환경으로 이전하고자하는 상황에서 보안 문제에 커다란 우려를 나타내고 있다.[1] 이러한 추세에 따라서 자료은닉에 대한 중요성이 부각되고 있다. 클라우드 컴퓨팅은 시간과 장소의 제약 없이 네트워크를 통하여 자원을 획득하여 사용할 수 있다는 장점이 있지만, 악의적인 사용자에 의한 정보 유출이 취약하다는 문제점을 가지고 있다. 따라서 클라우드 컴퓨팅에서 정보를 저장하고 획득하는 과정에서 자료은닉기법을 활용하여 정보를 더욱 안전하게 보호하는 것이 필요하다. 본 논문에서, 2장에서는 클라우드 컴퓨팅의 요구사항과 자료은닉 기법인 암호화 알고리즘, 데이터 마스킹, 블록체인에 대해서 살펴보고, 3장에서는 각각의 자료은닉기법을 클라우드 컴퓨팅에 적용하여 활용할 수 있는 방법을 제안하였다. 4장에서는 제안한내용을 바탕으로 향후 연구 방향을 제시하였다.*

* 본 연구는 미래창조과학부 및 정보통신기술진흥센터의 대학ICT연구센터육성 지원사업의 연구결로 수행되었음(IITP-2016-H8501-16-1008)

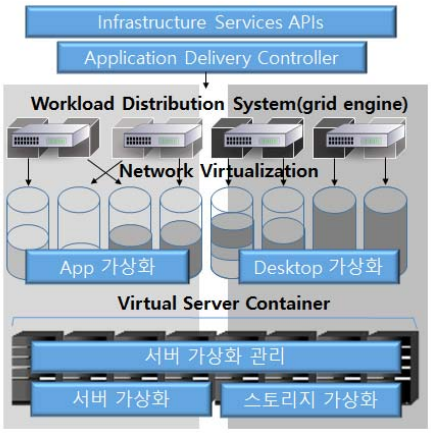
2. 클라우드 컴퓨팅의 요구사항과 자료은닉기법

“클라우드 컴퓨팅은 직접 · 공유된 정보통신기기, 정보통신설비, 소프트웨어 등 정보통신자원을 이용자의 요구나 수요 변화에 따라 정보통신망을 통하여 신축적으로 이용할 수 있도록 하는 정보처리체계를 말한다.”[2] 본 논문에서 살펴볼 자료은닉기법으로는 암호화, 데이터 마스킹, 블록체인 기법이 있다.

2.1 클라우드 컴퓨팅 요구사항

클라우드 컴퓨팅의 가장 큰 특징은 가상화 기술을 기반으로하며, [그림 1]은 클라우드 컴퓨팅의 구성이다.

[그림 1] 클라우드 컴퓨팅 구성도



클라우드 컴퓨팅의 구성에서 보는 것 같이 서버, 네트워크, 스토리지 등 기반 기술들이 가상화로 구성되어 있다. 가상화된 기반 위에 OS부터 문서를 작업하는 프로그램까지 다양한 응용 프로그램들이 설치되어 사용한다. 이때 사용자는 현재 사용하는 서비스가 가상화되었다는 것을 인지하지 못해도 일반적인 시스템을 사용하는 것과 같이 사용 가능해야한다. 따라서 클라우드 컴퓨팅은 다음 [표 1]과 같은 요구사항과 가상화 기법을 통한 구현이 필요하다.[3]

[표 1] 클라우드 컴퓨팅의 요구사항과 구현 내용

요구사항	구현 내용
효율성 및 자동화	pooling, zero-touch infrastructure
민첩성	automated provisioning, contol
선택의 자유	open, inter-operatability
보안성	data encryption, API firewall

가상화 기술을 사용하여 구현된 클라우드 컴퓨팅 시스템은 크게 5가지의 기능을 수행 할 수 있어야 한다. 첫째, 물리적으로 제한되는 자원의 한계를 높이기 위한 방법(폴링)으로, 여러 개의 물리적 자원을 가상 자원으로 만들어 효율을 올리는 것이다. 둘째, 가상의 자원을 하나의 물리적인 자원처럼 인식하도록 하여 공유 가능하도록 해야 한다. 셋째, 특정 서비스의 기능을 가상 자원에서 실질적으로 존재하는 것처럼 할 수 있는 에뮬레이션을 제공해야 한다. 네 번째, 물리적 자원을 가상화 할 때 상호 매핑을 통하여 서비스 제공 시 영향을 주지 않아야 한다. 마지막으로 보안적인 측면으로 악의적인 사용자의 접근을 통제하고 정보가 유출되어도 정보를 쉽게 사용하지 못하도록 하는 것이 필요하다.

2.2 자료은닉기법

2.2.1 암호화

암호화는 평문을 특정 키값을 사용하여 암호문으로 변경하는 방식이다. 크게 블록 암호화와 스트림 암호화로 구분할 수 있다. 블록 암호화에는 DES, 3-DES, AES, Blowfish, 등이 있고, 스트림 암호화는 RSA, RC4, A5/1, A5/2, 등이 있다. 현재 클라우드 컴퓨팅에서는 암호화를 사용하여 정보를 암호화하거나 사용자의 인증에 사용하고 있다.

2.2.2 데이터 마스킹

데이터 마스킹 기법은 개인의 민감한 정보를 타인이 식별 할 수 없도록 하는 기법이다. 데이터 마스킹 기법에 사용되는 민감한 데이터는 다음의 [표 2]와 같은 텍스트, 이미지, 비디오 등이 될 수 있다.

[표 2] 데이터 마스킹 적용 범위

구분	적용 범위
텍스트	개인신상정보, IP, ID/PW, 등
이미지	지도, 사진, 등
비디오	영화, 드라마, CCTV, 등

2.2.3 블록체인

블록체인 기법은 Satoshi Nakamoto가 고안한 Bitcoin이라는 금융서비스 기법에서 처음으로 소개되었다.[4] 블록체인 기법은 비대칭키 암호화 알고리즘인 RSA를 사용하고 특정 데이터를 모든 사용자에게 공개하고 저장하는 분산기술을 필요로 한다. 모든 사용자에게 동일한 데이터를 저장시키는 방식을 채택함으로서 데이터의 위변조를 원천 봉쇄하는 것이 블록체인 기법의 목적이다. 블록체인 서비스는 'Pro of-of-work', 'Proof-of-stake', 'Consensus-by-bet' 기술을 사용한다.[5]

Proof-of work

사용자가 변경되는 내용을 하나의 블록으로 생성하고 블록체인에 생성한 블록을 연결시키고 내용을 공유하여 위변조를 방지하는 기법

Proof-of-stake

사용자가 블록 생성, 등록의 독점을 막기 위해 자신이 가지고 있는 정보를 스스로 증명하는 기법

Consensus-by-bet

블록체인의 등록을 위해 네트워크에 참여하는 사람들의 동의가 있어야 등록이 완료되는 기법

3. 자료은닉기법의 활용법 제안

이와 같은 자료은닉기법이 클라우드 컴퓨팅 환경에서 데이터를 저장하거나 검색할 때 자료은닉을 위한 방식으로 사용될 수 있다. 각각의 자료은닉기법은 다음과 같이 사용될 수 있다.

3.1 암호화 알고리즘의 활용법

클라우드 컴퓨팅 환경에서는 스트림 암호화 알고리즘보다는 블록 암호화 알고리즘을 사용하여 데이터의 암호화를

진행한다. 현재 상용화되어있는 클라우드 컴퓨팅 서비스를 살펴보면 상대적으로 암호화 성능이 약한 DES, 3-DES보다는 AES를 사용하고 있다.

본 논문에서 제안하는 방식은 OTP(One Time Password)의 챌린지 · 응답 방식과 AES를 혼합해서 사용하는 사용자 인증 방식이다. 현재 클라우드 서비스를 사용하기 위해서 사용자는 회원가입 절차를 진행해야 한다. 기존에는 사용자가 설정한 패스워드를 통해 인증을 수행했다. 이에 사용자의 패스워드 정보가 탈취되면 추가적인 인증 없이 계정을 사용할 수 있었다. 이러한 문제의 해결을 위해 사용자의 패스워드와 주기적으로 변경되는 AES 암호화 키로 암호화한 패스워드 두 가지를 사용하는 것이다. 두 개의 패스워드를 사용하여 사용자가 클라우드 서비스에 인증을 진행하는 단계는 다음과 같다.

① 클라우드 서버로부터 OTP를 통한 인증으로 자신의 패스워드를 암호화하기 위해 사용된 AES 암호화 키를 획득한다.

② OTP 인증으로 획득한 암호화 키를 사용자 패스워드로 사용하여 로그인을 진행한다.

③ 입력받은 AES 키를 사용하여 암호화되어있는 사용자의 패스워드를 복호화한다.

④ 복호화 결과가 사용자가 회원가입 시 사용한 패스워드와 동일하다면 로그인을 완료한다.

위의 과정에서 볼 수 있듯 OTP와 AES를 통한 이중인증을 사용하여 사용자 인증 과정에서 인가된 사용자와 비인가된 사용자를 판별하여 악의적인 사용자의 접근을 차단할 수 있다.

3.2 데이터 마스킹의 활용법

데이터 마스킹은 정보를 고정 숫자, 문자열, 널값, 특정 배열 등을 사용하여 새로운 내용으로 변경하거나 삭제하는 방식이다. 임의의 숫자, 문자를 더하거나 곱해서 식별정보를 노출시키지 않도록 '임의 잡음 추가' 기법과 소수의 레코드를 선택 한 후, 선택된 항목을 공백으로 변경하고 대체하는 '공백과 대체' 기법을 사용한다.[6]

본 논문에서 제안하는 방식은 클라우드 컴퓨팅 환경에서 개인정보를 데이터베이스에 저장할 때 영구적으로 마스킹하는 가상의 마스킹 데이터베이스를 사용하도록 하는 것이다. 데이터 마스킹 변환 및 제공 절차는 [그림 2]와 같다.

[그림 2] 데이터 마스킹 변환 및 제공 과정

우선 클라우드 컴퓨팅 서비스에 사용되는 메인 데이터베이스와 가상의 마스킹 데이터베이스를 구성한다. 사용자 A가 자신의 개인정보를 데이터베이스에 저장할 때, 우선 메인 데이터베이스에 저장한다. 이후 메인 데이터베이스에 저장되어있는 개인정보를 데이터 마스킹 처리하여 임의의 데이터로 변경시킨다. 변경된 개인정보들은 가상의 마스킹 데이터베이스에 별도로 저장된다. 이후 사용자 A가 자신의 개인정보를 요청하면 메인 데이터베이스에 있는 실제 데이터를 전송받는 것이 아니라 마스킹 데이터베이스에 있는 임의의 데이터를 전송받도록 한다. 개인정보를 수정해야하는 경우를 제외하고는 모든 서비스에서 개인정보를 요청하는 경우에는 마스킹 데이터베이스와 연결하도록 설정하여 변조된 정보만 받아볼 수 있도록 하는 것이다. 해당 방법을 사용했을 때 실제 개인정보의 소유자라면 자신의 개인정보는 알고 있기 때문에 개인정보를 요청해야 할 필요가 없다. 하지만 타인이 자신의 계정을 사용하여 개인정보를 요청했을 때 개인정보가 유출되지 않는다는 장점이 있다.

3.3 블록체인의 활용법

초기에 고안된 블록체인은 금융서비스를 위하여 고안되었기 때문에 블록의 내용이 금융거래 내역이었다. 클라우드 컴퓨팅에서 블록체인을 적용하기 위해서는 어떤 정보를 사용자들에게 공유할 것인지가 중요하다.

본 논문에서 제안하는 방식은 클라우드 데이터베이스의 트랜잭션을 블록체인으로 보호하는 것이다. 클라우드 환경에서 데이터베이스에 저장되는 내용은 사용자, 구성 서버, 저장된 자료들의 위치 등의 정보이다. 사용자가 제공되는 서비스를 사용할 때 클라우드 서버로 요청을 보내면 데이터베이스에서 요청받은 내용을 확인하고 요청한 기능을 제공한다. 이때 요청받은 정보를 확인하고 제공하는 기능은 트랜잭션으로 미리 구성되어 데이터베이스에 저장되어있다. 트랜잭션이 위변조 되거나 문제가 발생하는 경우 서비스에 직접적인 영향을 주거나 정보의 유출로 이어질 수 있

다는 것이다. 따라서 블록체인 기법에서 사용자들에게 공유해야하는 자원을 트랜잭션으로 설정하여 모든 사용자들에게 공유하는 방법을 사용한다면, 악의적인 사용자에 트랜잭션이 변경되어 위변조 되는 것을 봉쇄할 수 있다.

4. 결 론

최근 클라우드 컴퓨팅을 활용한 서비스들이 많아지면서 네트워크를 통한 자료은닉에 대한 중요성이 커지고 있다. 따라서 앞에서 살펴본바와 같이 암호화 알고리즘을 통한 데이터 암호화, 데이터 마스킹을 통한 가상 마스킹 데이터베이스의 활용, 블록체인 기법을 통한 데이터베이스 트랜잭션 보호 기법을 제안하였다. 제안한 방식을 사용했을 때 악의적인 사용자의 접근을 차단하고, 정보의 유출시 보호장치로의 역할을 수행할 수 있다. 또한 위변조를 방지함으로서 정보유출과 서비스의 영향을 줄 수 있는 행위를 사전에 차단할 수 있는 역할을 수행할 수 있다. 향후 본 논문에서 제안한 기술들을 적용한 연구를 진행하여 보안 기능을 강화하고 취약점을 보완하고자 한다.

참고문헌

[1] 한국정보통신기술협회, “공공 클라우드 컴퓨팅의 보안 및 프라이버시 보호 지침”, 2011.12.21

[2] 클라우드 컴퓨팅 발전 및 이용자 보호에 관한 법률 제2조

[3]정현준, 가상화 기술의 동향 및 주요 이슈(I), 2013.2.16.

[4] Satoshi Nakamoto, “Bitcoin:A Peer-to-Peer Electronic Cash System”, 2008.10.31.

[5] 김진화, 정명호, 김재모, 유영석, “Block Chain Primer 블록체인의 기술적 이해 및 도입을 위한 첫걸음”, 2016.03.24.

[6] 미래창조과학부, 빅데이터 활용을 위한 개인정보 비식별화 사례집, 2014.05.01