NoSQL 기반 클라우드 사용자 행동 탐지 시스템 설계

안광민 · 이봉환

대전대학교 정보통신공학과

NoSQL-based User Behavior Detection System in Cloud Computing Environment

Kwang-Min Ahn · Bong-Hwan Lee

Daejeon University

E-mail: mangdol@callon.kr · blee@dju.kr

요 약

클라우드 서비스는 모든 자원을 서비스 제공자가 제공하고 다수의 사용자가 공유하기 때문에 서비스 제공자가 사용자의 정보를 더욱 안전하게 보호해야만 한다. 본 논문에서는 모바일 클라우드 서비스의 보안을 강화하기 위해 NoSQL 기반의 비정상 탐지 시스템을 제안한다. 다양한 보안장비에서 발생시키는 보안 로그와 클라우드 노드에서 발생시키는 데이터는 대량의 데이터가 형식이 모두 다른 비정형 데이터이기 때문에 기존의 통합보안 관리 시스템에서 사용하는 관계형 데이터베이스를 사용하여서는 실시간 처리가 어렵다. 제안하는 시스템은 분산처리 환경에서 실시간 처리 및 확장성을 제공하기 때문에 클라우드 환경에서 새롭게 대두되는 보안 문제를 해결할 수 있다.

ABSTRACT

Cloud service provider has to protect client's information securely since all the resources are offered by the service provider, and a large number of users share the resources. In this paper, a NoSQL-based anomaly detection system is proposed in order to enhance the security of mobile cloud services. The existing integrated security management system that uses a relational database can not be used for real-time processing of data since security log from a variety of security equipment and data from cloud node have different data format with unstructured features. The proposed system can resolve the emerging security problem because it provides real time processing and scalability in distributed processing environment.

키워드

NoSQL, 클라우드 컴퓨팅, 보안 로그, 행동 탐지, 비정형 데이터

Ⅰ. 서 론

클라우드 컴퓨팅 서비스는 규모의 경제에 입각한 대규모 분산 컴퓨팅 패러다임으로서 스토리지, 플랫폼, 서비스 등과 같은 거대한 IT 자원들을 가상화와 동적 확장이 가능한 체제로, 사용자가 필요한 만큼을 인터넷을 통하여 사용하는 컴퓨팅 서비스 환경이다. 현재 클라우드 컴퓨팅 시장은 초기 도입기를 거치고 있으며, 웹메일, 블로그, 웹하드 서비스, 웹호스팅 서비스 등이 이미 사용되고 있다. 그러나 본격적인 성장 단계로 진입하기 위해서는 사용자의 요구 수준에 맞는 애플리케이션과 서비스 발굴, 기존 시스템과의 연동성 확대, 보안에 대한 우려 불식 등과 같은 문제들이 선결되어야 한다. 국제 시장 조사 기관인 IDC에서 IT 임원을 대상으로 한 조사 결과 클라우드 컴퓨팅 서비스 사용을 위해 선결되어야 할 과제가 보안이라고 응답하였다[1].

이러한 이유로 현재 클라우드 컴퓨팅 기술 및 보안에 대한 연구가 활발히 진행되고 있으나, 기업별 각기 다른 클라우드 아키텍처로 인해 보안 요소 연구 및 구체적인 대응 방법 연구에 대한 문제가 발생하고 있다. 또한, 상호 호환성 및 연동성 문제와 클라우드 컴퓨팅의 주요 기능인 가상화에 대한 보안 요소를 도출하여 적용하는데 어렵다는 단점이 존재한다. 이러한 문제점에 대한 효과적인 보완을 위해서는 각 벤더 및 기관의 아키텍처의 공통적인 기능이 반영된 아키텍처를 구성하여 가상화 보안 레이어 역할 및 이와 관련된 기능에 대한 분석과 그에 따라 발생 가능한 위협 및 대응방안에 대한 연구가 필요하다[2].

본 논문에서는 NoSQL의 특성, NoSQL과 관계형

데이터베이스를 비교 설명하고 NoSQL을 이용하여 클라우드 사용자 행동 모니터링 시스템을 설계한다.

Ⅱ. 관련 연구

2.1 NoSQL

관계형 데이터베이스는 구조적 데이터 저장소의 솔루션으로서 가장 많이 사용되어 왔다 하지만 관계형 데이터베이스는 대규모의 데이터를 저장하기 위한 스케일 아웃 방식으로 확장하기 어렵고 데이터의 복잡한 조인 과정을 거쳐 형성되어있어 대용량 데이터 처리 시 병목현상이 발생할 수 있으며, 수십억 건 이상의 데이터를 저장하고 있는 테이블에 대해 인덱스의 재구성 칼럼의 추가 등과 같은 작업은 시스템 운영 상황에서 불가능하다는 단점이 있다.

이에 기존의 관계형 데이터베이스와는 다른 형태의 데이터 저장소에 대한 요구 사항이 제시되었고 여기에 부합한 데이터 저장소를 별도로 개발해 사용하기 시작했다. 구글의 Bigtable, 아마존의 Dynamo 등과 같은 저장소가 발표되면서 NoSQL에 대한 관심이 증가했고 그동안 특별한 용도로만 사용되던 일부 오픈 소스들도NoSQL로 분류되면서 NoSQL이 새로운 트렌드가 되었다[3.]

NoSQL은 관계형 데이터베이스의 한계를 극복하기 위한 데이터 저장소의 새로운 형태로써 오픈소스 중심으로 발전하고 있으며, 인덱스와 데이터가 분리돼 별도로 운용되며 스키마에 대한 정의가 자유롭다. 또 저렴한 PC 서버로 수평적 확장이 가능한 분산 아키텍처로서 초대용량 비정형 데이터 처리에 적합한 경량화된 분산 데이터 저장소라 할 수 있다.

또 관계형 데이터 모델이 아닌 키-값 또는 키-값을 응용한 데이터모델 안정적이고 고가의 하드웨어가 아닌 다수의 값싼 하드웨어 이용 데이터는 분산된 노드에 복제되어 저장되는 특성을 가지고 있을 때 NoSQL로 분류하며, 데이터모델에 따라 키-값, 칼럼, 문서, 그래프로 구분할 수 있다.

표 1. NoSQL 종류

구 분	데이터베이스
Key-Value	memcached, Dynamo. Volemort, Tokyo Cabinet, Redis
Column	Bigtable, Cloudata, Hbase, Hypertable, Cassandra
Document	MongoDB, CouchDB
Graph	Neo4j, FlockDB, InfiniteGraph

2.2 클라우드 모니터링 시스템

클라우드에 할당된 가상 인스턴스를 모니터링하기 위한 방법은 크게 두 가지로 나눌 수 있다 첫째는 가상 인스턴스에 에이전트 방식의 모니터링 프로그램을 설치하여 해당 프로그램으로부터 인스턴스 상태 정보를 제공 받는 방법[3]이고, 둘째는 가상 인스턴스를 관리하는 하이퍼바이저 수준에서 가상 인스턴스의 상태를 모니터링 하는 방법이다[4]. 에이전트를 사용한 모니터링은 가상 모바일 인스턴스의 상태를 자세하게 모니터링 할 수 있다는 장점이 있으나 공격자가 에이전트를 변조하여 에이전트를 무력화 시킬 수 있다는 단점이 있다. 하이퍼바이저 영역에서의 모니터링은 얻을 수 있는 정보가 제한적이고 방법이 복잡할 뿐만 아니라 복잡한 하이퍼바이저가 가져오는 보안 취약성도 고려해야 한다. 하지만 하이퍼바이저가 공격 당하지 않는다면 가상 모바일 인스턴스에서는 자신이 모니터링되고 있음을 인지할 수 없으며 정보를 조작할 수도 없는 장점이 있다[5].

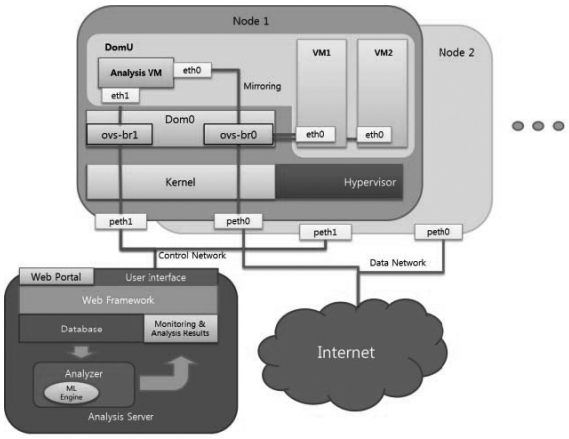

그림 1. 클라우드 행동 모니터링 시스템 구조

2.3 통합보안관리 시스템(ESM)

통합보안관리 시스템(ESM, Enterprise Security Management)은 일반적으로 이벤트를 발생시키는 침입차단시스템이나 침입방지시스템, 게이트웨이, 라우터 등의 장비에서 정보(패킷의 정보, 트래픽 양 등)를 취합하여 실시간으로 수집 서버에 전송하는 에이전트(Agent)①, 각 장비에 탑재되어 있는 에이전트를 통하여 들어온 정보들을 수집 또는 시간 별, 종류별로 정리하고 실시간 또는 주기적으로 데이터베이스 서버 및 분석 서버에 정보를 전송하는 수집 서버②, 수집 서버에서 정리된 정보를 데이터베이스에 저장하는 시스템인 데이터베이스 서버③, 수집 서버에서 들어온 데이터 및 데이터베이스에 저장된 정보를 바탕으로 현 네트워크의 상태 및 위기 상황을 시간별 종류별

로 분석하고 정리하여 그 결과를 사용자에게 알려주고 현상 및 결과 등의 정보를 데이터베이스에 다시 저장하는 시스템인 분석서버④로 이루어져있다. 각각의 구성요소는 구성에 따라 단/복수로 확장 또는 축소될 수 있는 구조를 가지고 있다.

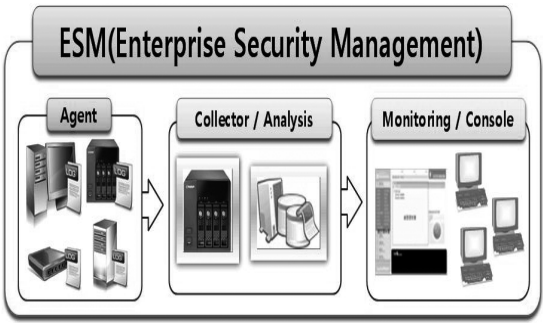

그림 2. 통합보안관리 시스템 구성도

현재 대부분의 통합보안관리 시스템에서는 관계형 데이터베이스를 사용하고 있는데, 이는 실시간 처리에 한계점을 드러내고 있다. 다시 말해,여러 가지 네트워킹 공격 등으로 트래픽이 발생하면 시스템에서는 보안로그를 실시간으로 수집하여 이를 토대로 분석, 처리하고 각각의 공격 형태에 맞는 효율적인 방어를 취하게 되는데, DDoS 등 대용량 트래픽을 발생시키는 공격이 발생하게 되면 네트워크 정보가 저장되는 데이터베이스에 보안로그 수집 및 저장 시 병목현상이 발생하게 되고 이로 인해 시스템의 성능이 저하되어 데이터를 효율적으로 수집 및 분석을 할 수 없게 되거나 혹은, 과도한 부하로 저장된 DB에 문제가 발생하게 되면 모든 보안 로그를 손실하게 될 수도 있다. 이와 같은 문제점들을 해결하는 방법으로 여러 가지 고성능 H/W 장비 도입을 제시하고 있지만 이는 성능에 대비해 고비용을 요구하므로 쉽게 수용할 수 없는 단점이 있다[6].

Ⅲ. 클라우드 사용자 행동탐지 시스템

그림 3은 본 논문에서 제안하는 클라우드 사용자 행동 탐지 시스템에 NoSQL를 적용한 확장성 있는 사용자 행동탐지 시스템의 구성도이다. 외부에서 보안 이벤트 정보가 발생하면 LogCollecter (LoadBalancer)는 보안 로그로 분류된 대량의 정보들을 수집하여 보안로그들을 적절히 시스템에 분배하는 역할을 한다. mongod는 데이터를 저장, 관리하는 서버로 분산 복제 정책을 적용한다. 또한, mongod(configsvr)은 데이터 분산 저장을 위한 샤딩을 적용하고, 파티셔닝에 대한 정보를 관리한다. mongos는 클라이언트의 요청을 받아 환경설정 서버의 파티셔닝 정보를 참고해 보안 로그를 적절히 데이터 서버로 포워딩하는 라우팅 역할을 수행한다[7]. MongoDB는 서버의 사용불가 유무에 따라 동적으로 서버의 마스터-슬레이브 연결을 조정하며, 이에 맞추어 자동으로 샤딩 정보가 변경되고, 이를 통해 대용량의 정보를 성능저하 없이 실시간으로 저장한다[8].

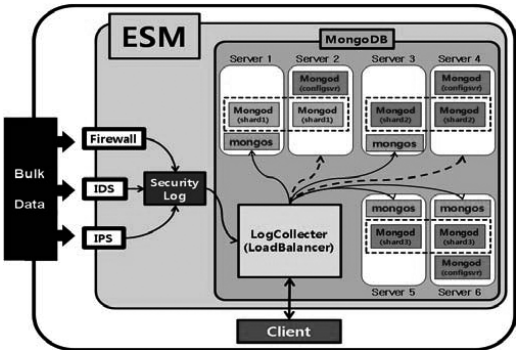

그림 3. NoSQL 기반 행동 탐지 개념도

본 논문에서는 보안 정보를 수집하고 분석하기 위한 가상 인스턴스를 추가로 할당한다. 각 노드의 가상 스위치는 가상 인스턴스로 오가는 모든 트래픽을 미러링하여 할당된 분석용 수집기로 포워딩 한다. 네트워크 정보 수집 및 분석을 위한 가상 머신은 패킷의 패이로드를 포함한 모든 트래픽 정보를 포워딩 받으며 하이퍼바이저를 통해 자원을 할당 받게 된다. 이러한 트래픽 분석용 가상 머신은 각 노드에 하나씩 할당하며 트래픽 정보를 분석 서버로 전송하게 된다.

그림 4. 기존 시스템과 제안 시스템의 비교

Ⅴ. 결 론

향후 클라우드 컴퓨팅 수요 증가와 동시에 클라우드 컴퓨팅에 대한 새로운 취약점과 다양한 공격 방법이 발생할 것이다. 또한 다양화되는 클라우드 서비스별 다른 특징을 가지는 가상화 및 관련 아키텍처 단계에 대한 보안성 문제도 예상된다. 따라서 지속적인 클라우드 컴퓨팅 환경의 취약점 및 악성코드에 대한 대응 방법 및 보호 대책 연구가 진행되어야 할 것이다.

이에 본 논문에서는 클라우드 컴퓨팅의 보안을 강화하기 위해 NoSQL 기반의 비정상 행위 탐지 시스템을 제안하였다. 본 논문에서 제안하는 시스템을 적용할 경우 클라우드의 규모가 확장되는

것과 상관없이 다수의 서버 간에 유동적으로 역할을 변경하여 수집 시스템을 관리해 나가기 때문에 대용량 보안 정보 수집 시 시스템 부하에 따른 성능저하를 예방할 수 있다.

Acknowledgement

본 연구는 교육과학기술부와 한국연구재단의 지역혁신인력양성사업으로 수행된 연구결과임.

참고문헌

[1] Asia Pacific End-User Cloud Computing Servey, IDC, Sep. 2009.
[2] 정순기, 정만현, 손태식, 문종섭, “클라우드 컴퓨팅 가상화 보안을 위한 아키텍처 구성 및 기능 분석 연구”, 보안공학연구논문지, 제8권 제5호, 2011년 10월.
[3] F. Baiardi and D. Sgandurra, “Building trustworthy intrusion detection through vm introspection”, Proceedings of the third international symposium on information assurance and security, pp. 209-214. 2007.
[4] B.D. Payne, M.D.P. de Carbone, and W. Lee, “Secure and Flexible Monitoring of Virtual Machines”, Computer Security Applications Conference, Florida, USA, Dec. 10-14, 2007.
[5] 정재윤, 현종환, 홍원기, “클라우드 서비스의 비정상 행동 탐지 시스템”, 한국통신학회 동계종합학술논문집, 2012.
[6] 차지훈, 이승하, 김양우, “MongoDB 기반 보안로그 수집 시스템 설계”, 한국인터넷정보학회 논문지, 제12권 2호, 2011.
[7] MongoDB, http:ww.mongodb.org, MongoDB Working Foundation
[8] Michael Dirolf and Kristina Chodorow, "MongoDB The Definitive Guide", O'Reilly Media, 2010.