클라우드 서비스의 비정상 행동 탐지 시스템

정재윤 [1], 현종환 [1], 홍원기 [1,2]
포항공과대학교 컴퓨터공학과 [1], 포항공과대학교 정보전자융합공학부 [2]

{dejavu94, noraki, jwkhong}@postech.ac.kr

Abnormal Behavior Detection System for Cloud Service

Jae Yoon Chung[1], Jonghwan Hyun[1], James Won-Ki Hong[1,2]
CSE POSTECH[1], ITCE POSTECH[2]

요 약

최근 스토리지, 컴퓨팅 등의 자원을 빌려주고 사용한 만큼 비용을 지불하는 클라우드 서비스가 주목 받고 있다. 이러한 클라우드 서비스는 모든 자원을 서비스 제공자가 사용자에게 필요한 만큼 자원을 할당해주고 사용자는 사용한 만큼 비용을 지불하는 방식이기 때문에 서비스 제공자가 사용자의 정보를 더욱 안전하게 보호해야만 한다. 본 논문에서는 클라우드 서비스의 보안을 강화하기 위한 행동기만의 비정상 탐지 시스템을 제안한다. 클라우드 서비스의 비정상 행동 모니터링을 위하여 본 논문에서는 Open vSwitch 의 트래픽 미러링 기능을 사용하여 하이퍼바이저의 부담을 최소로 하면서 가능한 모든 가상 인스턴스의 행동 정보를 얻을 수 있도록 하였다. 수집된 사용자의 행동 정보는 기계 학습 방법으로 사용자 행동을 구분 함으로서 각 사용자의 정상 행동과 비정상 행동을 구분 할 수 있다. 또한 프로토타입의 구현을 통해 제안하는 비정상 행동 모니터링 및 탐지 방법을 검증하였으며 Random Forest 알고리즘을 사용한 비정상 행동을 탐지방법을 10-fold cross validation 방법으로 검증하였다.

.

Ⅰ. 서 론

최근 스토리지, 컴퓨팅 자원을 빌려주고 사용한 만큼 비용을 지불하는 클라우드 서비스가 주목 받고 있다. 사용자는 thin client 를 사용하여 서버에 집중된 가상화 자원을 사용함으로써 자원 할당의 신축성과 사용 비용 측면에서 상대적으로 유리하다고 알려져 있다. 사용자는 가상화 환경에 접속하여 게임, 동영상 재생, 웹 서핑 등 기존의 모바일 서비스뿐만 아니라 고사양 3D 게임, 복잡한 연산, 대용량 스토리지 사용 등 클라우드의 장점을 활용한 서비스도 제공 받을 수 있다.

클라우드 서비스는 모든 자원을 서비스 제공자가 사용자에게 필요한 만큼 자원을 할당해주고 사용자는 사용한 만큼 비용을 지불하는 방식이기 때문에 서비스 제공자가 사용자의 정보를 더욱 안전하게 보호해야만 한다. 클라우드 서비스의 보안은 단말, 네트워크, 클라우드 인프라를 아우르는 전체 영역에서 이루어져야만 한다.

클라우드에 할당된 가상 인스턴스를 모니터링 하기 위한 방법은 크게 두 가지로 나눌 수 있다. 첫째, 가상 인스턴스에 에이전트 방식의 모니터링 프로그램을 설치하여 해당 프로그램으로부터 인스턴스 상태 정보를 제공 받는 것[1], 둘째, 가상 인스턴스를 관리하는 하이퍼바이저 수준에서 가상 인스턴스의 상태를 모니터링 하는 것[2]이다. 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 모니터링 되고 있음을 인지할 수 없으며 정보를 조작할 수도 없는 방법이다.

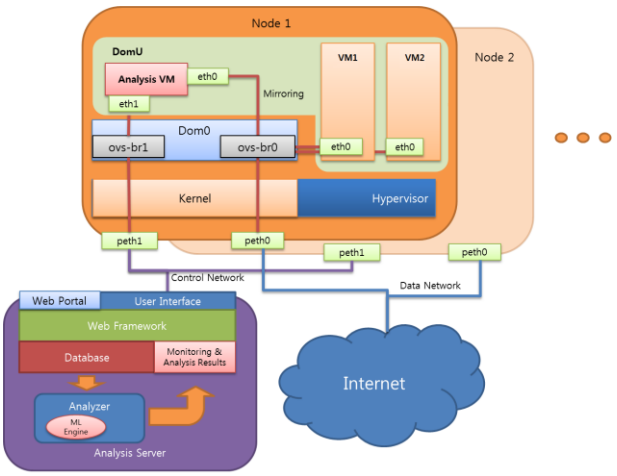

그림 1. 클라우드 서비스 행동 모니터링 시스템 구조

본 연구는 한국연구재단을 통해 교육과학기술부의 세계수준의 연구중심대학육성사업(WCU)(R31-2010-000-10100-0)과, 방송통신위원회의 "HiMang (Highly Manageable Network and Service Architecture for New Generation) 원천 연구" 원천기술개발사업(KCA-2011-10921-05003)의 연구결과로 수행되었음.

본 논문에서는 노드의 가상 스위치에서 트래픽을 미러링 함으로써 하이퍼바이저의 부담을 줄이면서 모든 가상 인스턴스의 네트워크 행동을 모니터링 하고 탐지 할 수 있는 방법을 제시한다.

Ⅱ. 본론

그림 1은 본 논문에서 제안하는 클라우드 서비스의 네트워크 행동 모니터링 시스템의 구조이다. 모든 가상 자원을 관리하는 Dom0 영역에서 가상 브릿지를 모니터링 할 경우 패킷의 페이로드를 포함한 모든 정보를 수집 할 수 있다. 하지만 하이퍼바이저나 Dom0 가 많은 I/O 처리나 복잡한 분석을 수행 할 경우 할당된 가상 인스턴스의 성능을 저하 시킨다. 또한 하이퍼바이저는 가능한 최소의 기능만 하도록 유지하는 것이 추가적인 보안 취약점을 만들지 않는다는 점에서 유리하다. 본 논문에서는 이러한 문제점을 해결하기 위해 네트워크 정보를 수집하고 분석하기 위한 가상 인스턴스를 추가로 할당하였다. 각 노드의 가상 스위치는 가상 인스턴스로 오가는 모든 트래픽을 미러링 하여 할당된 분석용 가상 머신으로 포워딩 한다. 네트워크 정보 수집 및 분석을 위한 가상 머신은 패킷의 페이로드를 포함한 모든 트래픽 정보를 포워딩 받으며 하이퍼바이저를 통해 자원을 할당 받게 된다. 이러한 트래픽 분석용 가상 머신은 각 노드에 하나씩 할당 하며 트래픽 정보를 분석 서버로 전송하게 된다.

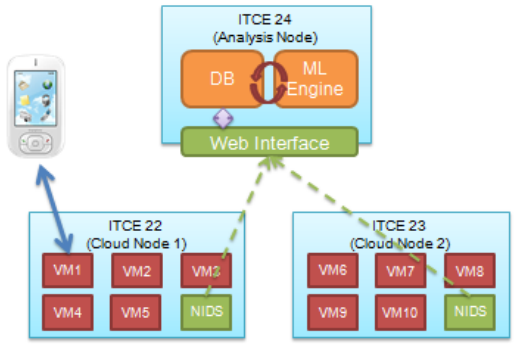

그림 2. 클라우드 서비스의 비정상 행동 모니터링 시스템 프로토타입

본 논문에서는 Open vSwitch[3]의 트래픽 미러링 기능을 사용하여 POSTECH ITCE 클러스터에 비정상 행동 탐지 시스템을 구현하였다 (그림 2). POSTECH ITCE 클러스터는 150 개 노드로 구성된 고성능 컴퓨팅을 위한 클러스터 시스템이다. POSTECH ITCE 클러스터의 22 번, 23 번 머신에 각 5 개의 가상 머신을 할당하고 각 단말의 네트워크 정보를 수집하였다. 수집된 네트워크 정보는 24 번 머신에서 수집하여 Random Forest 알고리즘을 사용해 행동을 탐지하였다. 사용자의 행동은 1) 사용자가 VM 을 사용하고 있지 않은 Inactive 상태; 2) 사용자가 VM 을 사용하고 있는 Active 상태; 3) 비정상 응용프로그램이 실행되고 있는 Abnormal 상태로 구분하였다. 프로토타입 구현에서는 패킷 헤더 정보만을 사용해 네트워크 행동 정보를 매 1 분마다 접속중인 원격 호스트의 수, 패킷, 플로우, 트래픽 양, 주요 포트 별 트래픽 양을 추출하여 분석 서버로 전송하였다. 수집된 행동 정보는 Random Forest 알고리즘[4]을 사용하여 행동을 탐지 했으며 10-fold cross validation 방법을 사용하여 정확도를 검증하였다.

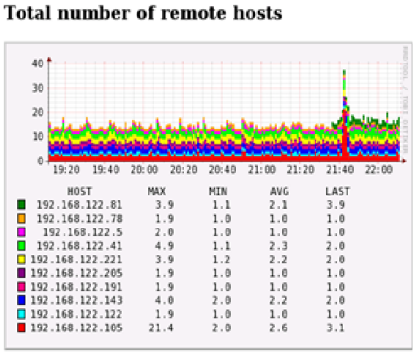

그림 3. 가상 머신 10 개의 모니터링 예시 (원격 호스트 수)

표 1. Random Forest 알고리즘을 사용한 비정상 행동 탐지 방법의 10-fold cross validation 결과

	TP Rate	FP Rate	Class
Accuracy	0.991	0	Inactive
	0.944	0	Active
	1	0.02	Abnormal
Weighted Avg.	0.988	0.008	

Ⅲ. 결론

클라우드 서비스의 보안을 강화하기 위해 행동 기반의 비정상 행위 탐지 시스템을 구현하였다. 가상 스위치의 트래픽 미러링 기능을 사용하여 호스트의 모든 트래픽을 수집하는 가상 머신으로 포워딩 하여 하이퍼바이저의 부담 없이 네트워크 정보를 수집하였다. 실제 POSTECH 클러스터 시스템에 총 10 개의 가상 머신을 설치하고 수집된 행동 정보는 Random Forest 알고리즘을 사용하여 비정상 행동 여부를 탐지하는데 사용되었다.

추후 연구로서 트레이닝을 위한 데이터 수집 방법과 SVM 등 다양한 기계 학습 알고리즘의 비교 분석 연구를 수행하여 실제 운용되고 있는 클라우드 데이터를 대상으로 정확하게 비정상 행동을 탐지하는 연구를 수행할 계획이다.

참 고 문 헌

[1] F. Baiardi and D. Sgandurra, “Building trustworthy intrusion detection through vm introspection”, Proceedings of the third international symposium on information assurance and security, 2007, pp. 209- 214.

[2] 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.

[3] Open vSwitch, http://openvswitch.org/.

[4] L. Breiman, “Random Forests”, Machine Learning, vol. 45, No. 1, 2001, pp. 5-32.