

머신러닝 기반의 NFV 관리를 위한 모니터링 프레임워크 구조

정세연¹, 이도영¹, 김가연¹, 유재형², 홍원기¹¹ 포항공과대학교 컴퓨터공학과² 포항공과대학교 정보통신대학원

{jsy0906, dylee90, gayen0n, styoo, jwkhong}@postech.ac.kr

Design of Monitoring Framework
for NFV Management based on Machine LearningSeyeon Jeong¹, Doyoung Lee¹, Ga-Yeon Kim¹, Jae-Hyoung Yoo², James Won-Ki Hong¹¹Department of Computer Science and Engineering, POSTECH²Graduate School of Information Technology, POSTECH

요약

네트워크 기능 가상화(Network Function Virtualization, NFV)는 기존 미들박스 중심의 네트워크 대비 보다 유연하고 기민한 환경을 제공하지만, 가상화 환경의 도입에 따른 다양하고 새로운 네트워크 관리 이슈가 발생하고 있다. 이를 효과적으로 대처하기 위해서, NFV 네트워크 관리에 자율 네트워킹(Autonomous Networking) 개념을 적용하려는 연구가 활발히 진행되고 있는데, 이는 NFV 환경에서 수집되는 모니터링 데이터를 바탕으로 다양한 네트워크 관리 이슈에 대한 머신러닝 분석을 전제한다. 따라서 머신러닝 기반의 NFV 관리를 위한 효과적인 모니터링 환경의 필요성이 제기되고 있다. 본 논문에서는 NFV의 가상화 환경에서 발생하는 광범위한 모니터링 데이터를 수집하며, 수집된 데이터의 전처리 과정을 통해 NFV 관리 기능에 대한 효과적인 머신러닝 분석을 가능케하는 NFV 모니터링 프레임워크의 구조를 제안한다.

I. 서론

네트워크 기능 가상화(Network Function Virtualization, NFV)는 기존 전통적인 네트워크 환경에서 핵심적인 기능을 수행하는 물리 네트워크 장치(Middle-box)의 주요 기능을 소프트웨어로 구현하여, 이를 가상 머신, 가상 스토리지 및 가상 네트워크 등의 가상 자원을 이용하는 형태로 실행하려는 새로운 네트워킹 패러다임을 말한다. 이러한 소프트웨어 기반의 NFV 환경을 통해, 네트워크 사업자 또는 운영자는 기존 네트워크 대비 장비 투자비(CAPEX)와 운용 비용(OPEX)을 절감하고, 서비스 대응 및 트래픽 변화에 신속하게 대처할 수 있을 것으로 전망되고 있다.

NFV 도입에 따른 대표적인 네트워크 관리 이슈로는 가상 네트워크 기능(Virtualized Network Function, 이하 VNF)의 라이프사이클(Lifecycle) 관리가 있다. 기존의 미들박스 기반 운용 환경과 달리, 통신사업자의 NFV 환경에서는 수천~수만 개 이상의 가상화된 서버로 대표되는 가상 자원 위에서 (가상) 네트워크 기능이 수행되므로, 이에 대한 관리 복잡도가 매우 증가한다. 즉, 기존 네트워크 대비 보다 향상된 유연성, 확장성 등을 제공하지만 네트워크 관리자의 업무가 가중되며 잘못된 구성에 따른 에러에 취약하다는 문제점이 제기되고 있다.

따라서, 이에 대한 해결책으로 국내외 주요 통신사업자로부터 NFV 환경에 자율 네트워킹(Autonomous Networking) 및 인지 네트워킹(Cognitive Networking) 개념을 적용하려는 연구가 활발히 진행되고 있으며, 이는 오늘날 컴퓨터 비전 및 언어 처리 등의 분야에서 혁신적인 발전을 이룬 머신러닝(Machine Learning) 및 딥러닝(Deep Learning) 기술을 네트워크 관리에 효과적으로 접

목하는 것을 전제로 하고 있다.

본 논문에서는 주요 NFV 관리 기능을 머신러닝 기반으로 처리하기 위한 전제 조건인 NFV 모니터링 프레임워크를 제안한다. 이는 NFV 환경에 대한 전반적인 데이터 수집, 관리 및 머신러닝을 위한 데이터 전처리를 효과적으로 수행하는 구조를 포함하며, 기존 방식과의 비교를 통해 향후 연구 방향을 제시한다.

II. 관련 연구

NFV 환경에서는 일반적으로 범용 물리 머신(서버)을 하이퍼바이저(Hypervisor)를 통해 가상 머신으로 가상화시킨 후, 그 위에 VNF를 동작시킨다. 따라서 VNF가 동작하는 가상 머신 및 해당 가상 머신을 수용하고 있는 물리 머신의 리소스 상태와 사용량 현황을 모니터링하는 것이 일반적이다. ETSI NFV 워킹그룹에서는 NFV 환경에서 수집 및 활용 가능한 대표적인 모니터링 데이터를 제시하고 있다 [1].

Sandpiper[2]는 하이퍼바이저(Xen) 수준에서 수집되는 정보만을 이용한 모니터링(Black-box)과 개별 가상 머신에 위치한 별도의 모니터링 에이전트로부터 수집되는 Guest OS 수준 모니터링(Gray-box)에 대한 성능 및 정확도 관점에서의 비교 분석을 통해 Gray-box 모니터링의 실효성을 제시한다. 그러나 NFV와 같이 대량의 가상 머신을 활용하는 환경에서의 확장성은 검증되지 않았으며 VNF의 QoS 수준 데이터 수집이 불가능하다.

대표적인 클라우드 운영 플랫폼인 OpenStack은 구축된 가상화 환경에 대한 자체적인 Telemetry 서비스(Ceilometer)를 제공하고 있다 [3]. 다수의 NFV 환경에서 기반 기술로 OpenStack을 사용한다는 점에서 내장된 모니터링 기능을 이용하여 호환성과 확장성을 확보할 수

있지만, 대량의 (중복) 데이터를 저장 및 조회하는데 있어서 성능 저하 문제가 알려져 있다.

T-NOVA[4]는 다수의 오픈소스를 활용하여 NFV 환경에 대한 MANO (Management and Network Orchestration) 시스템을 구축한 연구로서, NFV 모니터링 기능을 포함한다. 해당 연구에서는 물리/가상 리소스 모니터링을 위해 OpenStack Ceilometer 를 활용하고, SDN 컨트롤러를 통해 네트워크 상태를 파악하며 개별 가상 머신에 모니터링 에이전트를 배치하여 VNF 수준의 모니터링 정보를 수집하지만, 머신러닝 분석에 적합한 별도의 수집 데이터 전처리 과정을 포함하지 않는다.

III. NFV 모니터링 프레임워크 구조 제안

본 논문에서는 관련 연구에서 제시된 확장성, 성능, 데이터 전처리 이슈를 고려한 NFV 모니터링 프레임워크의 구조를 제안한다 (그림 1). 본 구조에서 수집되는 모니터링 데이터는 VNF Deployment, Service Chaining, Auto-Scaling, VNF Migration 등 대표적인 NFV 관리 기능[5]에 대한 머신러닝 분석 기반 자동화된 정책 (Policy) 제어에 활용될 수 있다.

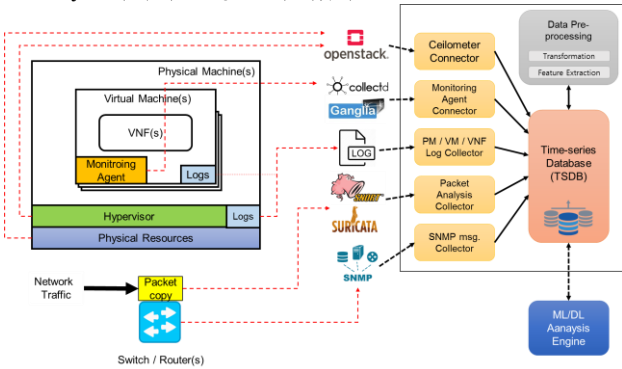


그림 1 제안하는 NFV 모니터링 프레임워크 구조

제안하는 NFV 모니터링 프레임워크는 크게 네 가지 특징을 가진다.

1. 확장된 물리/가상 머신 리소스 상태 및 사용량 정보 수집
2. 시계열 데이터베이스(Time-Series Database) 활용
3. 가상 머신 간 패킷 수준 모니터링 기능
4. 머신러닝 데이터 전처리 기능

VNF 가 동작하는 물리 머신(서버) 및 가상 머신의 실시간 리소스 상태 및 VNF 의 리소스 사용량 정보를 수집하기 위해서 기본적으로 OpenStack Ceilometer 를 활용한다. Ceilometer 를 통해 VNF 가 동작하는 (가상) 머신의 CPU 사용량, 메모리 사용량, 디스크 및 네트워크 I/O 횟수 등의 정보를 수집한다. 이에 더해, 개별 가상 머신에 모니터링 에이전트(collectd, Ganglia 등)를 배치하여 Guest OS 수준 정보(TCP 세션 개수 등) 및 VNF 의 QoS 수준 정보(서비스 지연시간 등)를 수집해서 머신러닝 분석에 활용될 데이터의 다양성을 확보하며, NFV 각 구성요소에서 발생하는 로그 데이터를 수집한다. 이는 더 많고 다양한 데이터를 활용하여 데이터 간에 숨은(Hidden) 상관관계 및 패턴 파악을 통해 [5]의 각 관리 기능별 머신러닝 분석 결과의 정확도를 높이는 것을 목적으로 한다.

기존의 Ceilometer 기반 모니터링에서 대량의 수집 데이터를 처리할 때 성능이 저하 되므로 OpenStack Gnocchi 프로젝트에서는 시계열 데이터베이스 및 고속 데이터 인덱싱 기술을 활용하여 이를 해결하고자 한다. 제안하는 구조에서는 Gnocchi 프로젝트의 아키텍처를 참고로 데이터 처리 파이프라인을 구성한다.

서버 가상화 환경에서 가상 스위치로 연결되는 가상 머신 간에 트래픽은 일반적으로 네트워크 플로우(Flow) 수준에서 모니터링 된다(NetFlow, OpenFlow 등). 그러나 네트워크 상태에 대한 보다 정확한 분석과 침입 트래픽 탐지 등을 위해서는 패킷 수준 모니터링이 요구되며, 이를 위해 사전 연구로 진행된 Virtual TAP 기능을 활용한다 (그림 2). Virtual TAP 은 가상 스위치에서 OpenFlow Group Table 기능을 이용, 패킷 복제 기능을 수행하며 SDN 컨트롤러를 통해서 일괄적인 복제 정책 제어가 가능하다.

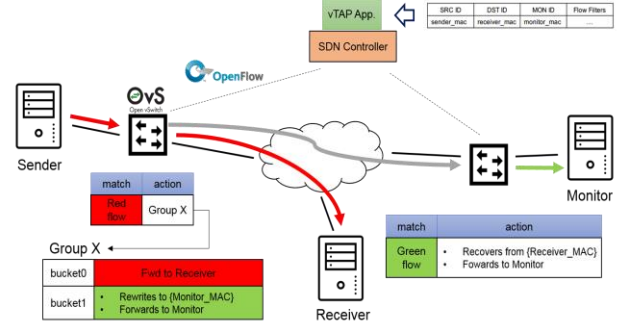


그림 2 패킷 복제를 위한 Virtual TAP 기능 구현

수집된 모니터링 데이터를 머신러닝 분석에 적합한 형태로 가공하기 위해서 별도의 전처리 과정이 필요하다. 이를 위해 모니터링(수집) 시점의 Timestamp 를 기준으로 (key, value) 형태로 정형화하며, 데이터가 입력되는 분석 모델(인공신경망, 강화학습 등)의 정확도에 따라 높은 상관관계를 보이는 데이터 특징(Feature)만을 추출하여 Noise 를 제거한다.

V. 결론

본 논문에서는 자율 네트워크에서 비롯된 머신러닝 기반 NFV 관리의 필요성을 제기하며 이를 실현하기 위한 핵심 기능인 NFV 모니터링 프레임워크의 구조를 제안한다. 향후 연구를 위해 OpenStack 기반의 NFV 테스트베드를 구성하고 모니터링 프레임워크를 구현, 머신러닝 기반의 NFV 관리 기능의 실효성을 검증할 예정이다.

ACKNOWLEDGMENT

이 논문은 2017 년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No. 2017-0-00195, 멀티 서비스를 지원하는 프로그래머블 스위치 제어 기술 개발)

참고 문헌

- [1] ETSI GS NFV IFA027, "Performance Measurement Specifications," 2018.
- [2] Wood, Timothy, et al. "Sandpiper: Black-box and gray-box resource management for virtual machines." *Computer Networks* 53.17 (2009): 2923-2938.
- [3] Sharma, M. Abhishek, and Monica O. Joshi. "Openstack Ceilometer Data Analytics & Predictions." *Cloud Computing in Emerging Markets (CCEM), 2016 IEEE International Conference on*. IEEE, 2016.
- [4] Xilouris, Georgios, et al. "T-NOVA: A marketplace for virtualized network functions." *Networks and Communications (EuCNC), 2014 European Conference on*. IEEE, 2014.
- [5] IRTF Network Function Virtualization Research Group, "Areas of Interest," 2018, (<https://irtf.org/nfvrg>).