

소프트웨어 정의 네트워킹을 이용한 컨텐츠 인기도 기반 캐싱 기법에 관한 연구

곽명철, 장덕현, 권태경*, 최양희*
서울대학교 컴퓨터 공학부

{mckwak, dhchang}@mmlab.snu.ac.kr, *{tk, yhchoi}@snu.ac.kr

A Study on the Content Popularity based Caching in Software Defined Networking

Myungchul Kwak, Dukhyun Chang, Ted " Taekyoung" Kwon, Yanghee Choi
School of Computer Science and Engineering, Seoul National University

요 약

소프트웨어 정의 네트워킹 (Software Defined Networking)은 네트워킹의 제어 기능 (Control plane)과 데이터 전송 기능 (Data plane)을 분리한 후, 제어 기능을 외부 네트워킹 개체에 집중시켜 보다 정밀한 트래픽 제어를 가능케 한다. 이를 통해 SDN은 네트워킹 연구자들이 새로운 네트워킹 기법을 저비용, 고효율로 실험해 볼 수 있게 하고, 또한 중소 규모 네트워킹에 유용한 제어 기능들을 쉽게 구현할 수 있는 인터페이스를 제공함으로써 유망한 미래 인터넷 기술 중 하나로 주목받고 있다. 본 논문에서는 SDN 기술 중 가장 널리 통용되고 있는 표준 인터페이스인 Openflow를 이용해 미래 인터넷의 한 조류인 Content-centric delivery 기법을 구현하고, 해당 구현을 토대로 중계 노드들에 컨텐츠를 캐싱하는 내재적 캐싱 기법을 제안한다. 구체적으로, 네트워킹에 존재하는 컨텐츠 플로우들의 현황을 모니터링 할 수 있는 SDN의 특성을 이용해 각 컨텐츠의 요청 회수에 따라 캐싱 여부를 결정하는 캐싱 기법을 제안하고, 그 유용성에 대해 논하고자 한다.

I. 서론

네트워킹은 산업 및 교육, 그리고 실생활에 이르기까지 사회 여러 분야에서 없어서는 안 될 인프라 기술로 완전히 자리매김 하고있다. 그러나 이러한 성공의 이면에는 이미 세계적으로 널리 정착된 폐쇄적인 네트워킹 프로토콜 및 장비가 새로운 기법의 도입과 이를 통한 혁신을 어렵게 한다는 단점이 존재한다. 따라서 새로운 네트워킹 기법들을 쉽게 적용할 수 있도록 하기 위한 대안으로 개방형 표준 인터페이스를 제안하는 기술이 소프트웨어 정의 네트워킹 (Software Defined Networking, 이하 SDN)이다.[1]

SDN은 네트워킹의 제어 기능 (Control plane)과 데이터 전송 기능 (Data plane)을 분리한 후, 제어 기능을 스위치 외부의 새로운 네트워킹 개체인 컨트롤러에 집중시켜 네트워킹 전체에 대한 보다 개방적이고 유연한 트래픽 제어를 가능케 하는 기술이다. 또한 이 컨트롤러를 간단한 소프트웨어로 정의할 수 있게 하는 인터페이스를 제공함으로써 네트워킹 연구자나 관리자도 하여금 새로운 네트워킹 기법 (라우팅, 보안, 이동성 지원 등)을 쉽게 연구, 개발하고 적용할 수 있도록 돕는다.

본 논문에서는 SDN을 이용해 유망한 미래 인터넷의 한 조류인 컨텐츠 기반 네트워킹 (Content Centric Networking, 이하 CCN)를 구현 하고, 해당 구현을 토대로 내재적 캐싱 기법 중 한 방식을 제안하고자 한다. 구현에는 SDN 기술 중 가장 널리 통용되고 있는 표준 인터페이스인 Openflow[2]를 사용하였다.

II. 본론

A. CCN

CCN[3]의 기본 컨셉은 위치 정보 (Locator)와 식별자 (Identifier)가 혼재된 IP를 이용해 컨텐츠를 전달하는

기존 IP 기반 네트워킹과는 달리 위치 정보가 분리된 컨텐츠 이름만을 이용해 컨텐츠를 전달하는 것이다.

CCN에서는 컨텐츠 전달을 위해 패킷을 두가지 타입으로 구분해 정의한다. 첫번째는 Interest 패킷으로 컨텐츠 이름을 기록하여 컨텐츠를 요청할 때 쓰는 패킷이다. 두번째는 Data 패킷인데, 이 패킷에는 컨텐츠 이름과 더불어 해당하는 컨텐츠가 탑재된다. 따라서 엔드 유저가 Interest 패킷으로 특정 컨텐츠를 요청하면 거기에 해당하는 Data 패킷이 컨텐츠를 보유한 노드로부터 Interest 패킷이 라우팅 됐던 노드들을 따라 유저에게 돌아가는 방식으로 컨텐츠 전달이 이루어진다.

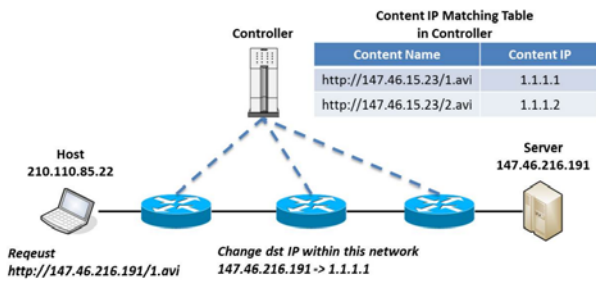
이를 통해 CCN에서는 내재적으로 중계노드에 컨텐츠를 캐싱할 수 있게 되며, 결과적으로 네트워킹 전체 트래픽 양, 컨텐츠 전송 시간 등의 측면에서 좀 더 효율적인 컨텐츠 전달이 가능해진다.

B. Openflow를 이용한 CCN 구현

Openflow 상에 CCN 오퍼레이션을 구현하기 위해 택한 구현 방식은 플로우-동일한 IP를 가진 패킷들의 연속적인 집합-의 IP를 컨텐츠 네임에 매칭되는 컨텐츠 IP로 변환시키는 것이다. <그림 1>의 예시를 이용해 설명하자면, 유저가 컨텐츠를 요청해서 네트워킹 내에 새 플로우가 발생했을 때 컨트롤러는 이 플로우들에 대해 DPI (Deep Packet Inspection)를 수행하여 HTTP 헤더의 컨텐츠 네임 (e.g. http://147.46.15.23/1.avi)을 알아낸다. 그 후 Openflow의 IP 변환 액션을 이용하여 플로우의 IP를 네트워킹 내부에서 독자적으로 사용하는 컨텐츠 IP (e.g. 1.1.1.1)로 변환시킨다. 이 컨텐츠 IP는 각기 컨텐츠 네임과 매칭되며 컨트롤러에 테이블 형태로 저장된다. 물론 요청된 컨텐츠가 유저에게 전달될 때에도 이러한 과정을 거친다.

본 구현에서는 이 컨텐츠 IP를 기반으로 라우팅 및 캐싱을 수행하는데, 이를 통해 원래 호스트나 서버의 IP가 다르더라도 컨텐츠 IP가 같으면 같은 컨텐츠로 판별할 수 있게 되므로 IP 네트워킹 상에서도 CCN과 유사한 방식의 내재적 캐싱을 수행할 수 있게 된다. 이를

위해 실제 구현에서는 콘텐츠 IP 와 매칭되는 패킷들을 해당 스위치 메모리 내에 캐싱하는 Openflow 액션인 Cache 액션을 새로 정의하였다.



<그림 1> HTTP 요청 시 콘텐츠 IP 변환 예시

B. 콘텐츠 인기도 기반 캐싱

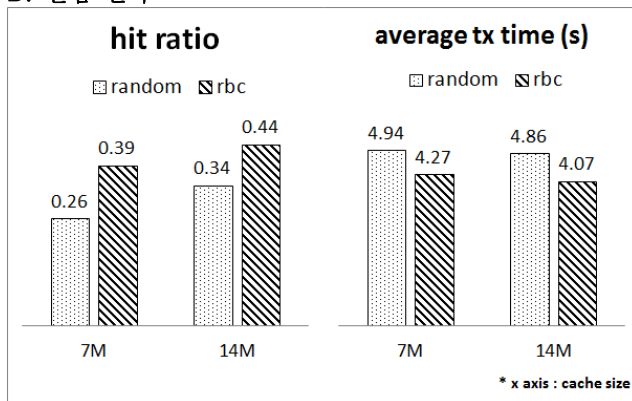
본 논문에서 제안하고자 하는 캐싱 기법은 SDN 의 특성을 이용해 인기도가 높은 콘텐츠를 우선적으로 캐싱하는 기법이다. SDN 의 특성인 네트워크 내 트래픽에 대한 관찰 및 제어 기능을 이용하면 기존 방식(e.g. LRU) 보다 더 명시적으로 인기 있는 콘텐츠를 판별하여 더 효율적인 내재적 캐싱을 수행할 수 있다.

구체적으로는 앞서 정의한 콘텐츠 IP 테이블에 요청 횟수 항목을 추가하여 콘텐츠 IP 별로 요청된 횟수를 기록하고, 특정 횟수 이상 요청받은 콘텐츠를 인기있는 콘텐츠로 판별하는 방식을 사용하였다. 콘트롤러는 인기있는 콘텐츠를 판별한 후 루트 상의 랜덤한 위치에 있는 한 스위치에 Cache 엔트리를 삽입함으로써 해당 콘텐츠를 캐싱한다. 효율성 비교를 위해 랜덤한 확률로 캐싱 여부를 결정하는 단순한 기법을 비교군으로 삼았고, 비교 메트릭으로는 캐시 히트율과 평균 콘텐츠 전송 시간, 그리고 히트된 콘텐츠 용량을 스위치에 캐싱된 콘텐츠 용량으로 나눈 값인 캐시 활용도를 차용하였다.

C. 실험 환경

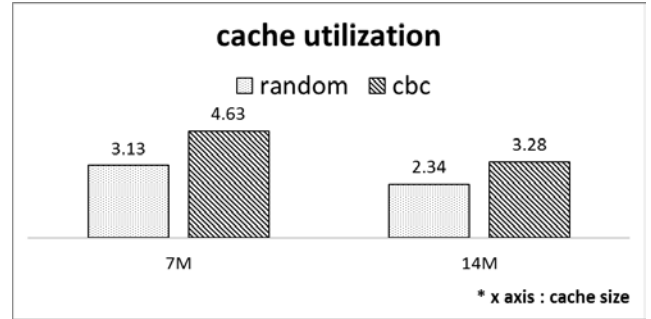
캐싱 여부를 결정하는 조건만을 고려한 실험이므로 간단히 <그림 1>과 동일한 토폴로지를 사용하였다. 콘텐츠는 $\alpha = 1.2$, $N = 2000$ 인 Zipf 분포를 따르도록 하여 각 실험마다 총 100 번 요청하였고, 구현에는 C++ 기반 콘트롤러인 NOX[4]와, Openflow 네트워크 환경을 구성하기 위한 에뮬레이터인 Mininet-HiFi[5], 그리고 C 기반 가상 Openflow 스위치인 Open vSwitch[6]를 이용하였다.

D. 실험 결과



<그림 2> 캐시 히트율 및 전송 시간 비교 그래프

<그림 2>는 직관적인 결과로서 캐시 사이즈가 동일할 때 인기도 기반 캐싱 (pbc)이 랜덤하게 캐싱한 경우 (random)보다 캐시 히트율이 높으며 따라서 평균 전송시간도 더 짧은 것을 확인할 수 있다. <그림 3>은 캐시 활용도 그래프이며 앞 실험과 마찬가지로 캐시 사이즈가 동일할 때 인기도 기반 캐싱 기법의 캐시 활용도가 더 높다. 이는 인기도 기반 캐싱 기법이 히트될 가능성이 높은 콘텐츠를 더 많이 캐싱하였음을 의미하며, 본 기법을 통해 네트워크 내의 한정된 자원인 캐시 공간을 더 효율적으로 활용할 수 있음을 시사한다.



<그림 3> 캐시 활용도 비교 그래프

III. 결론

본 논문에서는 최근 주목 받고 있는 SDN 기술을 활용한 콘텐츠 인기도 기반의 내재적 캐싱 기법을 제안하였다. 본 기법을 통해 보다 더 명시적으로 인기도가 높은 콘텐츠를 캐싱했을 때 전송 효율이 높을 뿐만 아니라 네트워크 내의 한정된 스토리지 자원을 더 효율적으로 활용할 수 있었다. 후속 연구로는 SDN 의 네트워크 관찰 기능을 좀 더 정밀하게 적용해 보다 더 효율적으로 캐싱을 수행할 수 있는 기법을 연구하고 있다.

ACKNOWLEDGMENT

본 연구는 한국방송통신전파진흥원의 차세대 통신 네트워크 원천 기술 개발사업 (10913-05004: 미래인터넷에서의 이동환경 및 네트워크 다양성 지원구조 연구)의 일환으로 수행되었음. 이 연구를 위해 연구장비를 지원하고 공간을 제공한 서울대학교 컴퓨터연구소에 감사 드립니다.

참 고 문 헌

- [1] 윤빈영, 이범철, Dan Pitt, “ 미래 네트워킹 기술 SDN ”, 전자통신동향분석, 27 권, 2 호, 129-136 쪽, 2012 년 4 월.
- [2] N. McKeown, et al, “ Openflow: Enabling Innovation in Campus Networks ”, ACM SIGCOMM Computer Communication Review, vol. 38, iss. 2, pp 69-74, April 2008
- [3] V. Jacobson, et al, “ Networking named content ”, in proceedings of CoNEXT ' 09, December 2009.
- [4] NOX, <http://www.noxrepo.org/>
- [5] Mininet-HiFi, <http://mininet.org/>
- [6] Open vSwitch, <http://openvswitch.org/>