

미래 콘텐츠 네트워크에서의 규모성 문제와 해결 방안

이문영, 조기덕, 박건우, 권태경, 최양희

서울대학교 전기컴퓨터공학부

{mylee, kdcho, kwpark}@mmlab.snu.ac.kr, {tkkwon, yhchoi}@snu.ac.kr

A Scalability Issue and a Solution in Future Content Networks

Munyoung Lee, Kideok Cho, Kunwoo Park, Ted "Taekyoung" Kwon and Yanghee Choi

School of Computer Science and Engineering, Seoul National University.

{mylee, kdcho, kwpark}@mmlab.snu.ac.kr, {tkkwon, yhchoi}@snu.ac.kr

요 약

현재의 인터넷은 일대일 통신을 지원하기 위해 설계되어 콘텐츠 다운로드 시 중복 다운로드 및 원거리 다운로드 등의 문제가 발생한다. 이러한 문제를 해결하기 위해 네트워크 내의 저장공간을 활용하여 콘텐츠를 전송할 수 있도록 콘텐츠를 중심으로 네트워크를 새롭게 설계하려는 연구가 세계적으로 활발히 진행되고 있다. 콘텐츠 네트워크에 저장된 콘텐츠를 찾기 위해서는 라우터의 라우팅 테이블에 콘텐츠에 대한 메타 정보가 저장되어야만 한다. 하지만 미래인터넷에서는 네트워크 내에 저장되는 콘텐츠 양이 상당하여 규모성(scalability) 문제가 필연적으로 발생한다. 본 논문에서는 계층적 이름 구조를 사용하는 경우에도 라우터의 규모성 문제가 발생함을 보이고, 이러한 문제를 해결하기 위한 미래 콘텐츠 네트워크의 요구사항을 도출한다.

I. 서론

현재 전 세계 수 많은 사람들이 이용하고 있는 인터넷은 본래 일대일 통신을 지원하기 위해 설계되었다. 하지만 최근의 연구결과에 따르면 현재 인터넷 트래픽은 일대일 통신 트래픽이 아닌 콘텐츠 획득을 주 목적으로 하는 트래픽(e.g. p2p, VoD)이 대부분을 차지하고 있음을 알 수 있다 [1]. 이는 실제 사용자의 대부분이 콘텐츠를 다운로드 하는데 인터넷을 사용하고 있음을 의미한다. 그러나 현재의 인터넷은 content-aware 하지 않아서 콘텐츠 다운로드 트래픽을 효율적으로 처리하지 못하는 문제가 있다. 그 예로 일대일 통신 기반의 인터넷에서는 여러 사용자가 동일한 콘텐츠를 요청하는 경우에도 서버로부터 동일한 콘텐츠를 중복해서 받아오는 중복 다운로드 문제가 발생한다. 또한 콘텐츠가 가까운 곳에 저장되어 있어도 원거리에 있는 서버로부터 받아오는 원거리 다운로드 문제도 발생한다. 위와 같은 문제를 해결하기 위해 콘텐츠 중심의 네트워크를 새롭게 설계하려는 연구가 세계적으로

활발히 진행되고 있다 [2,3]. 이러한 아키텍처들은 content-aware 한 라우팅 방식으로 동작하기 때문에 가까운 곳에 있는 콘텐츠를 효율적으로 찾을 수 있고, 이를 통해 중복 다운로드 문제 및 원거리 다운로드 문제를 해결하게 된다.

Content-aware 라우팅을 통해 네트워크 내에 저장된 콘텐츠를 찾기 위해서는 라우터의 라우팅 테이블에 콘텐츠에 대한 식별자(CID)나 메타 정보가 저장되어 라우팅에 사용되어야 한다. 하지만 미래 인터넷에서는 네트워크 내에 저장되는 콘텐츠 양이 상당하여 필연적으로 규모성(scalability) 문제가 발생할 것으로 기대된다. 특히 기존의 계층적인 구조를 가진 IP 주소를 쓰는 경우에도 코어 라우터에서 라우팅 규모성(routing scalability) 문제가 발생하는 것을 고려할 때[6], 편평한(flat)한 콘텐츠 식별자(CID)를 라우팅 엔트리로 사용하는 라우팅 방식에서는 규모성 문제가 더욱 커질 것으로 예상할 수 있다. 이를 막기 위해 URL 과 유사한 계층적 이름 구조를 사용할 수 있다. 하지만 이러한

경우에도 라우터의 규모성 문제가 발생하게 된다. 본 논문에서는 URL 과 같은 계층적 이름 구조를 사용하는 경우에도 라우터의 규모성 문제가 발생함을 보이고, 이러한 문제를 해결하기 위한 미래 콘텐츠 네트워크의 요구사항을 도출한다.

II. 본론

A. 계층적 이름 구조

계층적 이름 구조는 IP 주소와 유사하게 콘텐츠 식별자(CID)를 계층으로 나눠서 관리하는 이름 구조를 지칭한다. 그림 1 은 계층적 이름 구조를 사용하는 대표적인 예로 Content Centric Network (CCN) 아키텍처[3]에서의 콘텐츠의 이름 URL 을 보여준다. 이 계층적 이름 구조는 ABC.com 의 media 란 디렉토리에 존재하는 video.mpg 란 이름의 콘텐츠를 나타내고, 또한 해당 콘텐츠는 3 개의 다른 버전으로 구성되어 있고, 버전 2 는 두 개의 파일 덩어리(chunk)로 구성되어 있음을 의미한다. 이러한 계층적 이름 구조는 집합(aggregation)이 가능하기 때문에 편평한(flat)한 구조의 콘텐츠 식별자(CID)를 라우팅 테이블의 엔트리로 사용하는 것에 비해 규모성 문제를 다소 완화할 수 있다. 하지만 계층적 이름 구조는 완벽한 집합(aggregation)이 되지 않기 때문에 규모성의 문제가 여전히 남아있다.

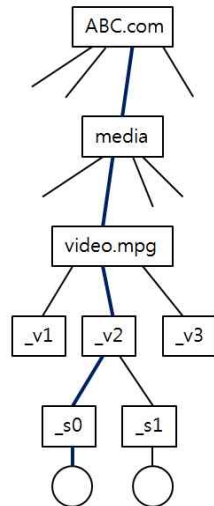


그림 1. 계층적 이름 구조의 형태

B. 계층적 이름 구조의 규모성 문제

계층적 이름 구조를 라우팅 테이블의 엔트리로 사용하여도 크게 두 가지 이유 (집합의 한계, 라우터에 캐쉬된 콘텐츠를 고려)로 규모성의 문제가 존재함을 다음의 시나리오를 통해 알 수 있다. 그림 2 는 사용자가

각각 ABC.com 과 CNN.com 에 존재하는 content A 와 content B 를 요청하는 경우를 보여준다. ABC.com 과 CNN.com 는 47.46.2.55 와 47.46.1.51 의 IP 주소를 가지고 있고, 라우터 A-E 와 함께 그림과 같은 토폴로지를 구성하고 있다. 또한, 설명을 위해 라우터 D 가 가지고 있는 라우팅 테이블을 계층적 이름 구조를 사용하는 경우(왼쪽 테이블)와 IP 주소를 사용하는 경우(오른쪽 테이블)로 나누어서 나타내었다. 먼저 IP 주소를 사용하는 경우에는 유사한 IP 주소를 가지고 있는 ABC.com 과 CNN.com 이 47.46 으로 집합이 되기 때문에 라우팅 테이블에 하나의 엔트리로 유지하면 된다. 반면 계층적 이름 구조를 사용하는 경우에는 ABC.com, CNN.com, BBC.com 을 “.com” 이라는 하나의 엔트리로 집합할 수 없다. 따라서 계층적 이름 구조의 장점을 활용하지 못하고 ABC.com 과 CNN.com 을 라우팅 테이블에 각각의 엔트리로 유지해야 하는 경우가 발생한다. 즉, 계층적 구조로 표현되는 최상위 루트는 콘텐츠가 존재하는 서버의 위치가 되어야 하므로 계층적 이름 구조는 IP 주소에 비해 집합이 제한적이다. 또한 콘텐츠 도달(content reachability)을 보장하기 위해서는 라우팅 테이블은 모든 서버의 위치에 대한 라우팅 엔트리를 유지해야 한다. 하지만 네트워크에는 수 많은 서버가 존재하므로 이를 라우팅 테이블에 모두 유지하는 것은 다소 한계가 있다. 따라서 계층적 이름 구조를 사용해도 규모성의 문제가 존재함을 쉽게 예측할 수 있다.

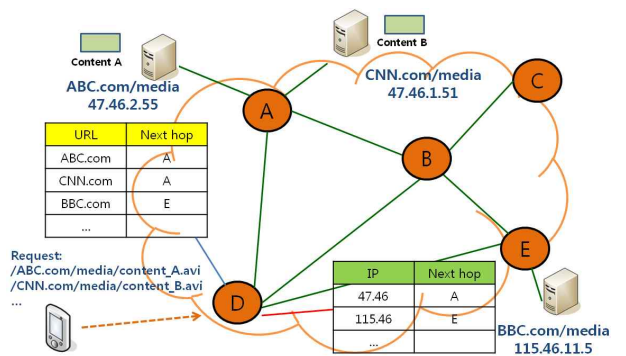


그림 2. 계층적 이름 구조 사용시 집합의 한계

콘텐츠 네트워크에서는 저장공간을 가지고 있는 라우터가 적절한 정책에 의해 콘텐츠를 캐쉬(cache)하고, 이렇게 캐쉬된 콘텐츠를 다음 번 요청에 효과적으로 활용하는 것을 고려하고 있다. 그림 3 은 라우터 B 와 E 에 content A 와 content B 가 저장되어 있는 상황에서의

라우팅 테이블의 모습을 보여준다. content B 를 캐쉬하고 있는 라우터 B 와 E 의 위치를 원래 서버인 CNN.com 과 구별하기 위해서는 캐쉬된 콘텐츠마다 라우팅 엔트리를 유지해야한다. 즉, 라우터에 캐쉬된 콘텐츠를 라우팅 테이블에 유지하기 위해서는 네트워크 곳곳에 캐쉬된 콘텐츠마다 엔트리를 유지해야 하므로 라우팅 테이블의 사이즈가 급격히 증가하게 된다. 따라서 라우터에 캐쉬된 콘텐츠 정보를 라우팅 테이블에 포함하는 경우 기존에 존재하던 규모성 문제가 더욱 심각해지게 된다. 또한, 라우터에 캐쉬된 콘텐츠는 제한된 라우터의 저장 공간 때문에 일정 시간이 지난 후 삭제 될 수도 있다. 이처럼 캐쉬된 콘텐츠의 정보는 실시간으로 변하기 때문에 이를 반영하면서 규모가 큰 라우팅 테이블 유지하는 것은 쉽지 않다.

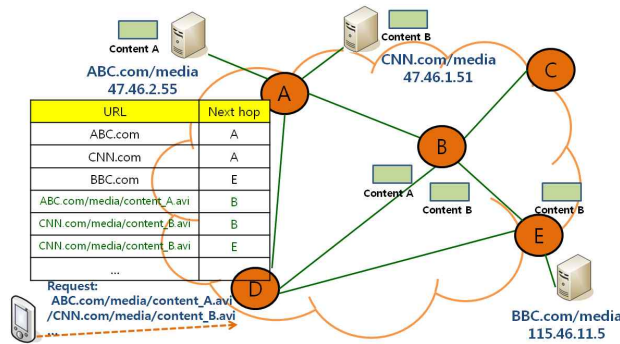


그림 3. 라우터에 캐쉬된 콘텐츠를 고려하는 경우의 라우팅 테이블

C. 미래 콘텐츠 네트워크의 요구사항

앞 절에서는 계층적 이름의 구조를 사용하는 경우에도 여전히 규모성의 문제가 존재함을 살펴보았다. 이러한 규모성의 문제를 해결하기 위해서 다음과 같은 콘텐츠 네트워크의 요구사항이 필요하다.

1) 일부 콘텐츠 정보만 저장

네트워크에 존재하는 콘텐츠는 그 숫자가 매우 많고, 빠른 속도로 증가하고 있기 때문에 모든 콘텐츠의 정보를 유지하는 것은 한계가 있다. 따라서 콘텐츠 네트워크에서의 라우팅 규모성 문제를 해결하기 위해서는 필연적으로 제한된 숫자의 콘텐츠의 정보만을 라우팅 테이블에 유지해야 한다. 예를 들어 각각의 라우터가 캐쉬하고 있는 콘텐츠에 대한 정보만을 이웃 라우터간에 교환하여 테이블을 유지한다면 규모성 문제를 완화할 수 있을 것이다.

2) 콘텐츠 도달(content reachability)을 보장

하지만 일부 콘텐츠의 정보만 저장하는 경우, 찾으려는 콘텐츠의 정보가 라우팅 테이블에 없다면 해당 콘텐츠가 네트워크 내에 존재해도 이를 찾지 못하는 문제가 발생할 수 있다. 따라서 라우팅 테이블에 정보가 없어도 콘텐츠 도달(content reachability)을 보장할 수 있도록 보완적인 기법이 필요하다. 이를 위해 라우팅 엔트리에 없는 콘텐츠를 찾을 때는 CCN 과 유사하게 플러딩(flooding) 방식을 사용하여 찾을 수도 있다 [3,4]. 혹은 플러딩으로 인한 라우팅 오버헤드를 줄이기 위해 기존의 IP 라우팅 기법 등과 결합하여 이용하는 방법을 고려할 수 있다.

3) 기타 요구 사항

콘텐츠 네트워크의 성능을 높이기 위해서는 네트워크 내에 존재하는 복수의 콘텐츠(multiple copy)를 활용해야한다. 따라서 BitTorrent [5]와 같이 여러 라우터로부터 콘텐츠를 동시에 받아서 다운로드 속도를 향상시키는 방법이 필요하다. 또한 하위 호환성(backward compatibility)을 위해 기존의 라우터와 content-aware한 라우터(C-router)가 네트워크에 공존하는 경우에는 C-router 간에 오버레이를 구성한 후, 정보를 교환하도록 하여 문제 없이 동작할 수 있도록 해야 한다.

III. 결론

본 논문에서는 계층적 이름 구조를 사용하는 경우에도 라우터의 규모성 문제가 발생함을 보이고, 이러한 문제를 해결하기 위한 미래 콘텐츠 네트워크의 요구사항을 도출하였다. 미래 콘텐츠 네트워크에서는 비약적으로 증가하는 콘텐츠의 숫자를 고려하여 모든 콘텐츠가 아닌 일부 콘텐츠의 정보만 라우팅 정보로 활용해야 하며, 보완적인 기법을 통해 라우팅 테이블에 없는 콘텐츠의 도달(content reachability)을 보장할 수 있어야 함을 요구사항으로 도출하였다. 향후 우리는 본 연구에서 도출된 요구사항을 만족하는 라우팅 기법을 제안하여 다양한 실험 및 후속 연구를 진행할 계획이다.

감사의 글

본 연구는 기초기술연구회의 NAP 과제 지원으로 수행되었습니다. 이 연구를 위해 연구장비를 지원하고 공간을 제공한 서울대학교 컴퓨터연구소에 감사드립니다.

참고 문헌

- [1] ipoque. Internet Study 2008/2009.
<http://www.ipoque.com/resources/internetstudies/internet-study-2008-2009>.
- [2] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica, "A data-oriented (and beyond) network architecture," Proc. ACM SIGCOMM, Aug. 2007.
- [3] V. Jacobson, D. Smetters, J. Thornton, M. Plass, N. Briggs, and R. Braynard, "Networking named content," Proc. ACM CoNEXT, Dec. 2009.
- [4] A. Carzaniga, M.J. Rutherford, and A.L. Wolf, "A Routing Scheme for Content-Based Networking". Proceedings of IEEE INFOCOM 2004. Hong Kong, China. March, 2004
- [5] B. Cohen, "Incentives build robustness in BitTorrent," Proc. Workshop on Economics of Peer-to-Peer Systems, May 2003.
- [6] Hitesh Ballani, Paul Francis, Tuan Cao, Jia Wang, "Making Routers Last Longer with ViAggre," ACM NSDI 2009, April 2009, Boston