

이름 기반 네트워크에서의 캐시 교체 정책 성능 비교

유정민, 최훈규, 권태경*, 최양희*

서울대학교 컴퓨터 공학부

{jmyoo, hgchoi}@mmlab.snu.ac.kr, *{tkkwon, yhchoi}@snu.ac.kr

Performance comparison of Cache replacement policy in Named Data Network

Jungmin Yoo, Hoongyu Choi,

Ted “Taekyong” Kwon, Yanghee Choi

School of Computer Science and Engineering, Seoul National University

요약

인터넷 구조는 50년이 넘도록 크게 변하지 않았지만 사람들이 인터넷을 사용하는 방식에는 큰 변화가 있었다. 사람들이 콘텐츠를 어디서 가져오느냐가 아니라 어떤 콘텐츠를 가져오느냐에 초점을 맞추기 시작하면서, 콘텐츠의 이름으로 통신하는 인터넷 구조인 Content Centric Network (CCN)이 등장하였고 현재 많은 연구가 진행 중이다. NDN에서는 이전에 전달했던 콘텐츠를 캐시에 저장해 놓은 후 이 콘텐츠가 다음에 다시 요청될 경우 해당 콘텐츠의 요청자 쪽으로 바로 전달하기 때문에 캐시의 성능이 중요하다. 이를 위해 지금까지 NDN에서 캐시 성능 향상을 위해 새로운 알고리즘을 제시하는 연구는 활발히 진행되어 왔으나, 네트워크 토폴로지 형태나 콘텐츠 요청 패턴에 따라 기존의 캐시 교체 정책을 분석한 경우는 없다. 따라서 본 논문에서는, 세 가지 캐시 교체 정책인 LRU, LFU, RR의 캐시 적중률을 비교하여 성능을 비교한다.

I. 서론

현재의 인터넷 구조는 IP주소를 기반으로 패킷을 주고받는 방식으로 지난 수십 년간 크게 변하지 않았다. 그러나 사람들이 인터넷을 통해 파일을 공유하거나 비디오 스트리밍을 보는 등, 콘텐츠의 위치가 아닌 콘텐츠의 내용 자체에 관심을 가지게 되면서 NDN이라는 새로운 인터넷 구조가 등장하였고 현재 많은 연구가 진행 중이다[1][4].

NDN에서는 원하는 콘텐츠의 이름으로 Interest packet을 보내어 해당 콘텐츠를 갖고 있는 서버로부터 Data packet을 받아온다. IP주소로 해당 ‘목적지’에서만 콘텐츠를 받아들일 수 있었던 데에 비해, NDN은 콘텐츠의 위치가 어디든 가장 가까운 곳으로부터 원하는 콘텐츠를 받아들일 수 있는 데에 장점이 있다[1][4].

이는 NDN에서 각 노드마다 *In-network caching*을 지원하기 때문에 가능하다. 기존 인터넷에서는 목적지에서 패킷을 받아오는 임무만 수행할 뿐, 어느 경로를 통해 전달되었는지 기록을 남기지 않는다. 그러나 NDN에서는 Pending Information Table (PIT)을 구현하여 Interest packet이 어느 경로를 통해 전달되었는지 기록을 남긴다. 그리고 Data packet을 받아오는 과정에서 PIT의 기록을 통해 요청자 쪽으로 콘텐츠를 전달해준다. 이 과정에서 경로 상의 노드마다 Content Store (CS)에 콘텐츠를 캐시하여 나중에 다른 요청자가 똑같은 콘텐츠를 요청할 때, 더 가까운 노드로부터 콘텐츠를 받아들일 수 있도록 지원한다[1][4].

이 때, 각 노드마다 캐시 교체 정책을 사용하여 전달된 데이터를 저장할 지말지, 캐시 내의 어느 데이터를 제거할지 결정하게 되는데, cache hit ratio를 최대화할 수 있는 정책이 가장 효율적인 정책이라고 말할 수 있다. 일반적으로 LRU를 사용할 때 캐시 성능이 극대화되기 때문에, 지금까지 NDN 캐시 연구에서도 주로 Least Recently Used (LRU) 정책을 고려하였다[1][4]. 여기에 Least Frequently Used (LFU) 와의 성능을 비교한 연구도 진행되어왔지만[2], 네트워크 토폴로지 형태나 콘텐츠 요청 패턴에 따라 기존의 캐시 교체 정책을 분석한 경우는 없다. 따라서 본 논문에서는 LFU와 Random Replacement (RR) 정책을 LRU와 비교하여 네트워크 토폴

로지와 콘텐츠 요청 패턴에 따른 캐시 교체 정책의 성능에 대해 알아보고자 한다.

II. 본론

2.1 NDN의 캐시 교체 정책

지금까지의 NDN 연구에서는 캐시 교체 정책으로 주로 LRU를 고려하고 있다. 그 이유는 가장 최근에 사용하지 않은 데이터를 캐시에서 제거하는 방식으로, 구현이 단순하기 때문이다. 또한 최근에 요청된 콘텐츠는 가까운 미래에 또 요청된다는 것을 고려했기 때문에 성능 또한 우수하여 캐시 교체 정책으로 주로 사용된다.

본 논문에서는 LRU뿐만 아니라 LFU와 RR의 성능 또한 cache hit ratio를 비교하여 검증한다. LFU는 사용되는 빈도수가 가장 적은 데이터를 캐시에서 먼저 제거하는 정책으로, 자주 요청되는 콘텐츠는 곧바로 다시 요청될 확률이 높기 때문에 cache hit ratio가 높을 것으로 예상된다. RR은 말 그대로 캐시에서 임의의 데이터를 선택하여 제거하는 방식이다. 이 방식을 사용하면, cache hit이 발생하는 데이터는 요청자 쪽으로 전달해주고 cache miss가 발생한 요청은 상위 노드로 전달하는 cache filtering을 완화시켜주어 LRU보다 캐시 내에 더 다양한 콘텐츠를 저장하게 되고 따라서 cache hit ratio가 높을 것으로 예상된다[5]. 이 세 가지 캐시 교체 정책의 cache hit ratio를 비교·분석하기 위해 아래와 같이 실험을 구현하였다.

2.2 실험

2.2.1 실험환경

실험은 두 가지의 대표적인 네트워크 토폴로지(그림 1)에서 진행되었다. 두 토폴로지 모두 end node(주황색 노드)에서 콘텐츠를 요청하도록 하였고, (a)Tree topology에서는 노드마다 600개, (b)Mesh topology에서는 노드마다 800개씩 총 2400개를 요청하였다. 콘텐츠 요청 패턴에 따라서 분석하기 위해 Zipf distribution의 parameter를 0.8/1.0/1.2로 다르게 설정해 주었다. 이 parameter값이 클수록 콘텐츠 요청 패턴이 편향되어

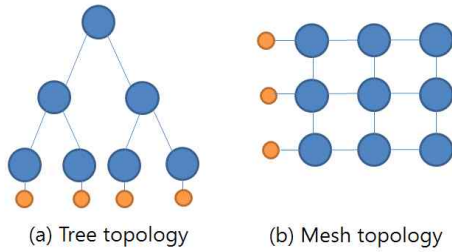


그림 1. 토폴로지

있음을 말한다. Busy한 상태의 Web server의 Zipf parameter가 일반적으로 1.2정도이다[3]. End node를 제외한 나머지 노드들은 각각 100개의 다른 콘텐츠를 가지고 있다. NDN에서는 storage가 제한적이기 때문에 모든 콘텐츠를 저장할 수 없다. 따라서 총 cache size는 총 콘텐츠 개수의 5%로 설정하였다[3]. 이 밖의 실험환경은 표 1과 같다.

total # of request	2400
cache size	5%
Poisson distribution	0.1
Content request pattern	Zipf distribution (parameter = 0.8/1.0/1.2)
packet size	10M

표 1. 실험환경

2.2.2 캐시 교체 정책의 성능 비교

그림 2, 그림 3은 각각 tree topology와 mesh topology에서 콘텐츠 요청 패턴에 따라 세 가지 캐시 교체 정책인 LRU, RR, LFU의 cache hit ratio를 비교한 결과 그래프이다.

두 그래프 모두 Zipf parameter가 클수록, 즉 특정 콘텐츠가 요청되는 빈도수가 높을수록 cache hit ratio가 커지는 것을 볼 수 있다. 이는 캐시 교체 알고리즘과는 상관없이, 자주 요청되는 특정 콘텐츠들이 캐시에 주로 남아있게 되어 또 다시 요청될 경우 cache hit이 나기 쉽기 때문이다. RR을 사용하여 임의의 콘텐츠를 교체한다 하더라도 인기 있는 콘텐츠는 곧 다시 캐시에 남아있게 된다.

또한, 전반적으로 LFU가 LRU보다 1.5배 이상 성능이 좋았으며 RR은 두 토폴로지 형태 모두에서 LRU에 떨어지지 않는 성능을 보였다. LFU의 성능이 LRU보다 좋은 것은 콘텐츠의 인기도로 설명할 수 있다. 앞에서 말한 대로, LFU는 가장 적은 빈도로 요청되는 콘텐츠를 교체하고 LRU는 가장 오래된 콘텐츠를 교체하는 방식이다. LRU는 최근에 요청된 콘텐츠는 가까운 미래에 또 요청된다는 것을 고려한 반면, LFU는 콘텐츠의 인기도를 반영하기 때문에 앞으로 자주 요청될 콘텐츠가 캐시 내에 저장되어 있기 쉽다. 따라서 LFU의 cache hit ratio가 LRU보다 높게 나타났다. RR은 LRU와 비슷한 성능을 보였는데, 이는 캐시로부터 임의의 콘텐츠를 교체하여 캐시 내에 다양한 콘텐츠를 저장하기 때문에 cache miss가 날 확률이 작기 때문이다.

토폴로지 형태는 크게 영향을 미치지 않았다. Tree topology에서는 콘텐츠 요청 패턴이 편향되어 있을 때(parameter=1.2) LFU와 LRU의 성능 차이(1.60배)가 mesh topology(1.22배)보다 더 컸다. 그 이유는 Tree topology에서는 hierarchical하게 저장하기 때문에 편향된 콘텐츠 요청에 더 영향을 받기 때문이다.

III. 결론

지금까지, NDN에서 필수적인 요소인 캐시의 성능에 대해 분석하기 위

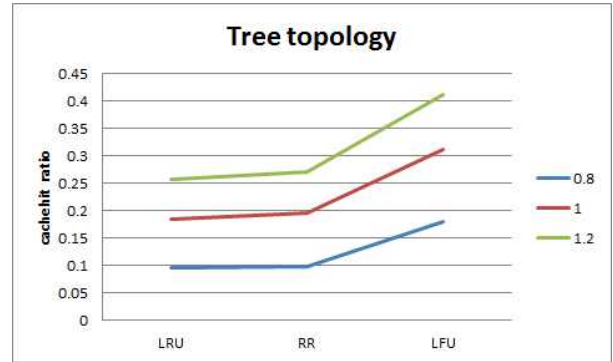


그림 2. Tree topology에서 cache hit ratio 비교

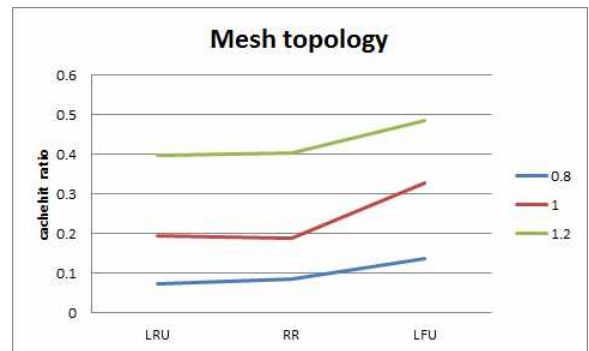


그림 3. Mesh topology에서 cache hit ratio 비교

해, 네트워크 토폴로지 형태와 콘텐츠 요청 패턴에 따라서 기존의 캐시 교체 정책들의 성능을 비교하였다. 그 결과, 전반적으로 LFU가 LRU보다 1.5배 이상 성능이 좋았으며 RR은 예상과 달리 LRU에 떨어지지 않는 성능을 보였다. 이 결과에 착안하여, 향후 연구로 캐시 내 콘텐츠의 다양성을 높이는 방안에 대해 연구해 볼 계획이다.

ACKNOWLEDGMENT

본 연구는 미래창조과학부가 지원한 2013년 정보통신·방송(ICT) 연구개발사업의 연구결과로 수행되었음. 본 연구는 한국정보화진흥원(NIA)의 미래네트워크연구시험망(KOREN) 사업 지원과제의 연구결과로 수행되었음 (13-951-00-001).

참고 문헌

- [1] V. Jacobson, D. Smetters, J. Thornton, M. Plass, N. Briggs, R. Braynard, "Networking named content," Proc. ACM CoNEXT, 2009.
- [2] Zhe Li, Gwendal Simon, Annie Gravey, "Caching Policies for In-Networking Caching", IEEE, 2012.
- [3] Dario Rossi, Giuseppe Rossini, "Caching Performance of content centric networks under multi-path routing", Technical report, 2011.
- [4] Named Data Networking, <http://named-data.net>
- [5] Guangwei Bai, Carey Williamson, "Time-Domain Analysis of Web Cache Filter Effects", Performance Evaluation - Special issue: Distributed systems performance