

콘텐츠 중심 네트워크에서 성능 향상을 위한 인기도 기반 캐시 교체 기법

우 태 희*, 박 흥 순°, 김 호 길*

Popularity Based Cache Replacement Scheme to Enhance Performance in Content Centric Networks

Taehee Woo*, Heungsoon Park°, Hogil Kim*

요 약

콘텐츠 중심 네트워크(CCN)는 기존의 IP 방식과는 다른 새로운 네트워킹 패러다임으로써 데이터를 찾는데 필요한 라우팅 정보를 콘텐츠 이름에서 찾는 방식이다. CCN은 노드마다 콘텐츠를 저장할 수 있는 캐시를 가지고 있어서 반복적인 콘텐츠 요청에 효율적으로 처리할 수 있다. 본 논문은 콘텐츠의 인기를 활용한 캐시 교체 기법을 제안하여 기존의 기법보다 캐시의 히트율을 향상시켰고, 이에 따라 서버의 부하 및 Round Trip Time(RTT) 시간이 감소하여 성능이 향상됨을 증명하였다.

Key Words : Content Centric Networking, Cache Replacement Scheme, Popularity

ABSTRACT

Unlike an existing IP routing methods, the Content Centric Network(CCN) is new networking paradigm to find the contents by content name. The CCN can effectively process the content requested repeatedly by the user because of the cache which can be storing a content. This paper proposes a popularity based cache replacement scheme. The proposed scheme improves the hit rate better than a existing scheme. Accordingly reducing the load of server and Round Trip Time(RTT).

I. 서 론

1970년대 초 알파넷의 등장 이후 TCP/IP 기반의 컴퓨터 통신은 오늘날까지 널리 사용되고 있는 가장 기본적인 컴퓨터 네트워크 모델이다. 하지만 단말기의 급격한 증가와 저장매체 비용 감소, 전 세계적 규모의 네트워크 범위 확장 등 초기와 다른 컴퓨터 통신 환경의 변화는 기존 TCP/IP 모델에서 불필요한 위치 탐색 및 반복 전송, 병목 현상 등과 같은 문제점이 대두하였다. 이에 따라 기존 네트워크 모델을 개선한 콘텐츠

중심 네트워크(Content Centric Network, 이하 CCN)와 같은 새로운 정보 중심 네트워크(Information Centric Network) 모델이 등장하게 되었다^[1]. CCN은 노드마다 캐시 기능을 추가하여 수신된 콘텐츠를 저장하고, 이후에 동일한 요청이 올 경우에 콘텐츠의 최초 위치로 요청을 보내지 않고 중간 노드에서 즉시 콘텐츠를 전달하기 때문에 기존의 IP 방식보다 성능이 좋다^[2].

한편, 웹 캐시에 대해 연구한 기존 연구에 따르면 캐시 교체 기법은 네트워크 대역폭과 서버의 부하를

* First Author : Korea Army Academy at Young-Chung(KAAY) Department of Cyber Warfare, seeles0@gmail.com, 학생회원

° Corresponding Author : Joint Forces Military University, heungsoon.park@gmail.com, 종신회원

* Korea Army Academy at Young-Chung(KAAY) Department of Cyber Warfare, khglex@gmail.com, 정회원

논문번호 : KICS2015-08-260, Received August 20, 2015; Revised November 10, 2015; Accepted November 10, 2015

감소시키며, 사용자가 정보를 인식하는데 소요되는 지연시간을 줄일 수 있다^[3]. 그래서 캐시의 히트율(hit ratio)을 높이기 위해서 많은 연구자들이 연구를 해왔고, 캐시의 히트율이 높을수록 성능이 향상됨을 증명하였다^[4]. 노드마다 캐시 기능을 가지는 CCN에서도 기존의 교체 기법보다 더 높은 히트율을 가지는 교체 기법을 적용한다면 더 좋은 성능을 보일 것이다. 우리는 CCN에서 캐시의 히트율을 향상시키기 위해 콘텐츠 인기도를 기반으로한 캐시 교체 기법을 제안한다.

본 논문의 구성은 다음과 같다. 2장은 CCN과 캐시 교체 기법에 대해서 소개하고, 3장에서는 CCN에서 캐시의 히트율 향상을 위한 인기도 기반 캐시 교체 기법을 소개한다. 4장은 제안한 모델의 성능을 평가하고 분석하였으며, 5장에서 결론으로 마무리 지었다.

II. 관련연구

2.1 콘텐츠 중심 네트워크^[3]

CCN은 사용자가 정보를 요청하기 위한 Interest 패킷과 Interest 패킷에 대해 응답하고 콘텐츠 청크(chunk)를 수반한 Data 패킷으로 나뉜다. Data 패킷은 Interest 패킷의 역경로로 이동하며 사용자에게 전달된다. CCN의 라우터 구조는 Content Store(CS), Pending Interest Table(PIT), Forwarding Information Base(FIB)로 구성되며 다음과 같이 동작한다.

Interest 패킷이 CCN 라우터에 도착하면, 라우터는 CS에서 도착한 Interest 패킷과 이름이 일치하는 콘텐츠가 있는지 확인을 한다. 만약 있다면, 해당 콘텐츠를 Interest 패킷이 도착한 역경로로 보낸다. 일치하는 콘텐츠가 없다면 PIT에 일치하는 항목이 있는지 확인을 하고, PIT 항목에 있다면 그 항목에 Interest 패킷의 인터페이스를 추가한다. PIT에도 일치하는 항목이 없다면 FIB에 일치하는 항목이 있는지 확인을 하고, 일치하는 항목이 있다면 라우터는 PIT에 새로운 항목을 생성한 후에 FIB의 항목에 따라 Interest 패킷을 전달한다. FIB 항목에도 존재하지 않는다면 Interest 패킷은 폐기된다.

Data 패킷이 CCN 라우터에 도착하면, 라우터는 CS에 도착한 Data 패킷과 일치하는 콘텐츠가 존재하는지 확인을 하고, 존재한다면 폐기한다. CS에 일치하는 콘텐츠가 없다면 PIT에 일치하는 항목이 존재하는지 확인을 하고, 존재한다면 라우터는 도착한 Data 패킷을 CS에 저장하고 일치하는 PIT 항목에 표기된 인터페이스로 Data 패킷을 전송한다. 그리고 Data 패킷을 전송한 후에 해당 PIT 항목은 삭제한다. PIT에도

일치하는 항목이 없으면 Data 패킷은 폐기된다.

2.2 캐시 교체 기법

네트워크에서의 캐시 교체 기법은 데이터를 어디에 위치시킬 것인가에 대한 부분을 주로 다룬다. 노드들의 독립적인 캐시 교체 기법이 아닌 네트워크 전반적인 관점에서 어디에 캐시를 할 것인가에 대한 기법으로 Leave Copy Everywhere(LCE), Leave Copy Down(LCD), Move Copy Down(MCD), Prob 등이 있다^[5,6].

그림 1은 LCE, LCD, MCD, Prob의 캐시 교체 기법 알고리즘에 대해서 묘사하였다. A와 B는 중간노드, C는 콘텐츠를 가지고 있는 서버이며, A의 하위에서 콘텐츠 요청이 이루어지고 있다. LCE부터 알아보면, A가 데이터를 요청받았을 때, 데이터를 가지고 있지 않다면 B와 C를 거치는 일련의 경로를 이동하여 데이터의 요청이 발생한다. C에서 해당 데이터를 보유하고 있어 요청에 대한 데이터가 발송되면 경로에 있는 A와 B 모두 해당 데이터에 대해 캐싱한다. 결과적으로 경로에 있는 모든 중간 노드(A, B)는 데이터를 캐싱한다. 동일한 상황에서 LCD방식은 데이터의 캐싱이 전달경로에서 하위 1홉의 B에게만 발생하며, MCD에서는 하위 1홉의 B에게 데이터가 전달되고 전달된 데이터는 서버 C에서 전달과 동시에 삭제된다. Prob방식은 LCE방식과 유사하나 캐싱 여부를 데이터의 생성시기와 캐시 크기 등을 고려하여 확률적인 방법으로 수행한다.

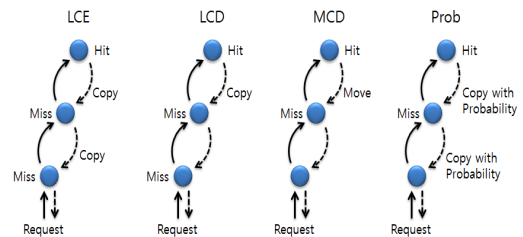


그림 1. LCE, LCD, MCD, Prob 알고리즘 동작
Fig. 1. LCE, LCD, MCD, Prob algorithm operation

2.3 CCN에서의 캐시 교체 기법

CCN노드는 단일 노드 관점에서 LRU 또는 LFU방식의 교체 기법을 기본으로 사용하고 있으며[5], 네트워크적 관점에서는 자신을 경유하는 모든 콘텐츠를 저장하는 LCE방식을 따른다. 하지만 이러한 LCE방식은 그림 2와 같이 다수의 클라이언트들이 인기 있는 콘텐츠를 요청하여 콘텐츠가 제공될시, 중간 지점에 위치한 노드 B~G에 동일한 콘텐츠의 중복 저장

이 발생하여 상대적으로 다른 콘텐츠가 캐싱될 확률을 낮추고 CCN 노드가 가지는 저장 공간을 효율적으로 사용하지 못하는 것을 볼 수 있다.

이러한 네트워크에서의 캐싱 영역에 중복을 줄이기 위한 많은 방안이 연구되고 있다^[7-9]. [7]은 요청 횟수에 따라서 콘텐츠의 청크를 요청자에게 가까운 노드로 순차적으로 전달하는 방안을 제안하였다. 하지만 LCD 방식을 사용하여 하위 노드에 콘텐츠를 전달하고도 상위 노드에서 콘텐츠를 보유하므로 캐싱 영역의 중복성을 야기시킨다. [8]은 중간에 교차되는 노드에 대해서만 캐싱하는 방안을 제안하였으나, 많은 교차점이 발생하는 네트워크에서는 기존의 LCD 방식과 동일하게 동작되므로 많은 중복된 콘텐츠들이 네트워크상에 존재하게 된다. [9]은 요청이 일정한 임계값(Popularity Threshold) 수치에 이르면 연결된 모든 인터페이스에 콘텐츠를 전파를 하는 방안에 대해서 제안하였다. 하지만 임계값에 이르기 전까지 콘텐츠는 일반적인 서버-클라이언트 방식을 따르며, 임계값이 넘어 연결된 모든 인터페이스에 콘텐츠가 전달되더라도 기존의 상위 노드에는 콘텐츠를 보유하므로 중복된 콘텐츠가 발생한다.

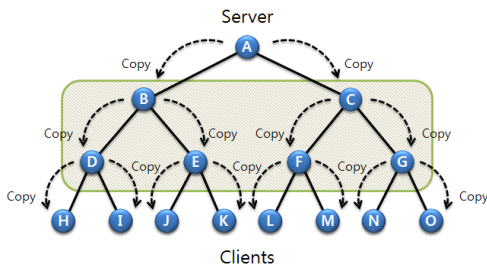


그림 2. 콘텐츠 중복 문제
Fig. 2. Contents Duplication Problem

III. 인기도 기반 캐시 교체 기법 설계

3.1 기본개념

사용자가 많이 찾는 즉, 인기 있는 콘텐츠의 캐싱이 되는 위치는 사용자에게 가까울수록 효과적일 것이다. 따라서 그림 3과 같이 서비스를 제공하고 받는 입장에서 인기가 있는 콘텐츠 일수록 사용자에게 가깝게 위치하고 인기가 낮은 콘텐츠일수록 서버 쪽에 위치하는 것이 이상적인 콘텐츠의 캐싱 위치가 된다.

인기도 기반 캐시 교체 기법에서는 네트워크를 계층적으로 보고 콘텐츠를 제공하는 노드를 상위노드, 콘텐츠를 요청하는 노드를 하위노드로 본다. 하위 노

드에 전송된 콘텐츠들은 소비자에게 전송되었거나 하위 노드에 캐싱이 이루어져 향후 추가적인 요청이 발생하지 않을 것이므로 중간 노드에서 해당 콘텐츠에 대한 캐싱은 불필요하며, 이에 따라 다른 콘텐츠보다 우선적인 교체가 이루어진다. 또한 인기 있는 콘텐츠가 신속하게 전파되기 위해 지역 내에서 인기 있는 콘텐츠를 미리 하위노드에서 캐싱 한다면 중간 노드에서 캐싱이 이루어질 가능성이 더욱 증가한다. 즉, 중간 캐싱 영역에서 콘텐츠의 중복을 줄이고 인기 있는 콘텐츠를 신속히 전파하기 위한 방안으로 인기 있는 콘텐츠를 전달하고 중복되는 콘텐츠는 신속히 제거될 필요가 있다.

이를 위해 각 CCN 노드들은 데이터의 전송이 이루어진 후 모든 페이스로 해당 데이터가 전송되었는지를 확인하여 이에 해당할 경우 해당 데이터를 최우선 교체 우선순위로 할당한다. 이를 통해 CCN 노드는 인기 있는 콘텐츠를 신속하게 하위 CCN노드에게 배포하고 자신의 CS에서 우선적으로 삭제하여 하위 순위의 인기 있는 콘텐츠를 받아들일 수 있다.

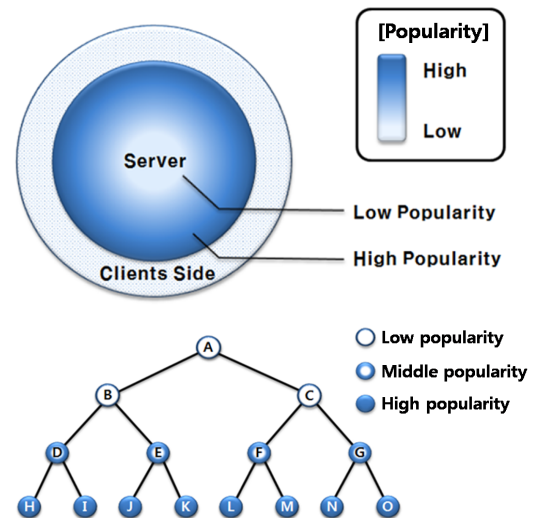


그림 3. 최적화된 콘텐츠 분배
Fig. 3. Optimized contents distribution

3.2 시스템 설계

인기도가 높은 콘텐츠는 높은 확률로 자신의 하위 노드들에게 해당 콘텐츠를 전파한다. 이를 위해 시스템은 자신의 인접 노드들이 얼마나 해당 콘텐츠를 요청하였는지를 기록하며 이를 위해 CS에 콘텐츠별로 요청한 노드들을 기록하기 위한 공간을 설계하였다. 인기도에 따른 확률적 전파 시에는 브로드캐스트로 전파가 되며 이후 해당 콘텐츠는 우선적으로 교체가

된다.

3.2.1 인기도 판단

네트워크에서 인기도는 일반적으로 알려진 Zipf 분포^[10]를 따른다고 가정한다. 만약 각각의 CCN 노드가 콘텐츠별 인기도를 미리 알고 있을 경우 Zipf 분포에 따라 인기도 높은 콘텐츠를 미리 캐싱하는 것이 적절하지만, 전체 네트워크에서 각 콘텐츠별 인기도를 알 수 있는 방법은 제한될 수밖에 없으며 특정 지역에서의 콘텐츠 인기도가 전체 네트워크에서의 콘텐츠 인기도를 항상 따른다고 보기는 어렵다. 또한 CCN에서는 콘텐츠가 매번 서버에서 제공되는 것이 아닌 중간 캐시에서 제공될 수 있으며, 노드 간 콘텐츠의 인기도에 대한 정보가 교환되지 않기 때문에 노드는 자신이 수신하고 전송하는 콘텐츠가 전체 네트워크에서 어느 정도의 인기도를 갖고 있는지 더더욱 알 수 없다.

따라서 콘텐츠별 전체 인기도에 대한 정보를 별도로 수신하지 않는 한, 노드의 콘텐츠 인기도 판단은 지역적인 방법을 사용할 수밖에 없다. 노드가 인기도를 판단할 수 있는 기준으로는 콘텐츠의 참조횟수와 연결되어 있는 Face별 요청여부를 삼을 수 있다. 우리가 제안한 기법에서는 자신과 인접한 노드들 중에서 얼마나 많은 노드들이 해당 콘텐츠를 요청하는가를 통해 판단하고, 이를 전송확률(P)이라고 정의하며 다음과 같이 수식으로 표현한다.

$$P = \frac{\# \text{ of Req.Face} + \# \text{ of Rcv.Face}}{\# \text{ of Total Face}} \times 100 \quad (1)$$

수식 (1)에서 노드의 전체 Face는 자신과 연결된 인접 노드들의 수이며, 이는 자신과 통신이 이루어진 Face의 수를 의미한다. 콘텐츠의 인기도 판단은 각 콘텐츠별로 계산되며 전체 Face 대비 얼마나 많은 Face에서 해당 콘텐츠를 요청했는지를 통합하여 계산한다. 계산된 전송확률은 해당 데이터를 요청하지 않은 하위 다른 노드들로 전송할 확률이 되며, 전송확률이 높은 콘텐츠는 더 높은 확률로 자신의 하위 노드들에게 전파된다.



[CS table]

[Modified CS table]

그림 4. 수정된 CS 테이블
Fig. 4. Modified CS Table

3.2.2 CS 구조 변경

전송확률을 판단하기 위해서는 전송확률을 판단하고 기록할 수 있는 공간이 필요하다. 이 공간을 위해서 그림 4와 같이 CCN 노드 모델에서 Interest 패킷과 Data 패킷이 출입되는 Face를 저장할 수 있는 공간인 “Forwarded Face” 열을 추가하였다. Forwarded Face에는 콘텐츠를 수신한 Face ID와 PIT 항목에 있는 수신된 Interest 패킷의 Face를 기록한다. 요청한 콘텐츠가 도착하여 CS에 등록될 때 콘텐츠를 수신한 Face 번호가 최초 Forwarded Face에 기록되고 PIT 항목에 기록되어 있던 Face 번호들이 순차적으로 기록된다. CS에 존재하는 콘텐츠를 다른 Face에서 요청할 경우 해당 Face가 해당 콘텐츠의 Forwarded Face에 추가적으로 기록된다. Forwarded Face에 이미 기록된 Face에서 재요청이 발생하는 경우 Face 번호는 신규로 추가되지 않으며 콘텐츠를 인접 노드로 전파하는 것에 대한 판단만 이루어진다.

3.2.3 콘텐츠의 전파

콘텐츠는 수식 (1)의 전송확률에 의해 확률적으로 전파되며, 인기도 판단과 전파여부 결정은 해당 콘텐츠가 참조될 때마다 발생한다. 전송확률이 결정되면 CCN 노드는 최초 콘텐츠를 수신한 Face를 제외한 자신 주변의 모든 Face로 콘텐츠를 전파한다. 콘텐츠의 전송이 이루어진 후 노드는 해당 콘텐츠를 더 이상 참조될 가치가 없는 자료로 판단하고 우선적으로 교체한다. 상위 노드에서 데이터가 전파된 동작을 푸싱(Pushing)이라고 정의하며, 이렇게 전파된 콘텐츠는 CS에 신규로 저장한다. 이렇게 푸싱에 의해 캐싱된 콘텐츠는 CCN 노드에서 Interest 패킷에 의해 수신된 다른 콘텐츠와 동일한 캐싱 기법을 적용받는다.

그림 5는 CCN 노드에서 콘텐츠 요청에 따른 Forwarded Face의 변화와 콘텐츠의 전파 이후 캐싱의 우선 교체과정을 나타낸다. (a) ‘Face 0’에서 콘텐츠 요청 발생시 PIT에 해당 Face번호가 기록되며, (b) 콘텐츠가 도착하면 콘텐츠가 도착한 ‘Face 2’와 PIT에 있던 ‘Face 0’가 Forwarded Face에 기록되고 인기도 판단을 시행한다. (c) 이후 ‘Face 1’에서 추가적인 콘텐츠 요청이 발생하면 노드A는 Forwarded Face에 해당 Face를 기록하고 전송확률을 판단한다. (d) 인기도 판단결과 전파확률은 100%라고 가정했을 때, 노드는 해당 콘텐츠를 보내준 최초 Face를 제외한 모든 Face로 해당 콘텐츠를 전파하고, 추가적인 캐싱 상황 발생시 모든 Face로 보낸 콘텐츠를 우선적으로 교체한다.