

효율적인 콘텐츠 배포를 위한 전달 계수 기반의 전송 기법

하태훈, 권태경, 최양희
서울대학교 전기컴퓨터공학부

thha@mmlab.snu.ac.kr, tkkwon@snu.ac.kr, yhchoi@snu.ac.kr

Propagation Coefficient of Swarm for Efficient Content Distribution

Taehoon Ha, Ted “Taekyoung” Kwon, Yanghee Choi
School of Computer Science and Engineering, Seoul National University

요약

Peer-to-peer 프로토콜은 사용자 간에 데이터를 전달하는 효율적인 방법 중의 하나이다. Peer-to-Peer 프로토콜을 효율적으로 운영하기 위한 다양한 방법이 연구되고 있다. 그 중 Content Propagation Metric 을 이용한 방법은 유틸 대역폭을 활용하여 전체 Peer 간의 대역폭을 최대한으로 활용할 목적을 가지고 디자인되었다[1]. 그러나 CPM 을 이용한 방법은 특정 peer 로 데이터가 독식될 우려가 있다. 본 논문에서 이 문제점을 논의하고 그 부분을 보완한 새로운 방법을 제안한다. 본 기법을 이용하면 전체적인 peer 간에 적절한 균형을 유지한 상태에서 전체 네트워크 자원을 활용할 수 있을 것이다.

I. 서론

멀티미디어 콘텐츠의 배포가 전체 인터넷 트래픽의 많은 부분을 차지하고 있다[2]. 점점 대용량의 콘텐츠를 배포하면서 네트워크의 사용량은 지속적으로 증가하고 있다. 특히 Youtube 등을 통한 영상 자료의 전달이나, Peer-to-Peer 네트워크를 통한 대용량 영상 자료의 전송이 일상적이 되어가고 있다. 이렇게 전송되는 데이터의 양이 증가하면서 전체 네트워크를 효율적으로 운영하는 방법이 중요하게 되었다.

Peer-to-Peer 프로토콜에서는 각 Peer 들이 정보의 요청자인 동시에 정보의 공급자로서의 역할도 수행하면서 콘텐츠의 빠른 전파와 네트워크 대역폭의 높은 활용도를 이루고 있다. 그러나 peer-to-peer 프로토콜에서는 각 peer 가 어느 상태 peer 로 전달하느냐에 따라 전체적인 데이터 전파효율이 달라질 수 있다. 이를 수치화하여 peer 별로 전달되는 데이터량을 조절하는 기법이 Content Propagation Metric(이하 CPM)을 이용한 관리기법이다.

CPM이란 하나의 Peer가 자신이 속한 각 Swarm으로 자신의 데이터를 보낼 때, 일정 시간 동안 자신이 그 시간 동안 보낸 데이터가 각 Swarm 내부의 Peer들로 전달된 총량을 뜻한다. Swarm의 CPM이 높다는 것은 해당 Peer가 Swarm으로 보낸 데이터가 Swarm 전체적으로 많이 복제되어 전파되었다는 것을 의미하며, 이는 해당 Swarm으로 데이터를 보낸 행위의 효과가 크다는 것을 뜻한다. CPM을 이용한 방식은 일정 시간 후 각 Swarm별로 CPM을 비교한 후 낮은 CPM을 가진 Swarm으로 보내는 대역폭을 줄이고 보다 높은 CPM을 가진 Swarm 쪽으로 보내는 대역폭을 증가시킴으로써, 전체적으로 좀 더 많은 데이터 전파량을 발생하는 방식이다

하지만 이렇게 보다 효율적인 Swarm 쪽으로 대역폭을 몰아주게 되면, 반대로 비효율적인 Swarm에 속한

peer들은 대역폭을 배정받지 못하여 전혀 통신을 할 수 없는 가능성이 있다. 이 논문에서는 과도하게 대역폭이 감소 현상에 대한 원인과 개선 아이디어를 제시하고자 한다. 또한 개선 아이디어를 시뮬레이션한 결과를 이전 결과와 비교하여 개선 효과를 보이하고자 한다. 이 아이디어를 사용한다면 전체적인 전달의 효율을 꺾임과 동시에 Peer 간에 스마트하고 Fair한 결과도 얻을 수 있을 것으로 예상된다.

II. 본론

1. Content Propagation Metric

CPM이란 단위 시간 당 해당 Swarm으로 데이터가 전파된 총량을 뜻한다. Figure 1 상에서 만약 p 가 s_1 에 속한 peer로 b_1 의 데이터를 전송하였을 경우에 s_1 에 속한 Peer 중 해당 데이터를 원하는 Peer들 사이에는 그 데이터가 상호전달이 될 것이다. 단위 시간 동안 해당 Swarm 내에서 그 데이터가 전달되는 총량을 측정하고, b_1 뿐 아니라 전달되는 다른 모든 종류의 데이터에 대해 해당 측정을 수행하여 계산하면 P 에서 s_1 에 대한 CPM을 측정할 수 있다.

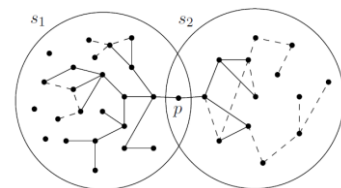


Figure 1 두 swarm 간의 데이터 확산

CPM이 측정된 후에는 이 결과값을 이용하여 해당 Swarm에 대한 전송대역폭을 정할 수 있다. CPM이 Swarm 내에 확산되는 데이터의 총량을 의미하므로, CPM이 높은 쪽은 낮은 쪽에 비해 Peer 간에 데이터

전달이 활발하며 즉 서로 간에 존재하는 대역폭을 보다 많이 사용한다고 볼 수 있다. 그러므로 CPM 이 높은 쪽에 좀 더 많은 데이터를 전송하여 같은 양의 데이터를 보내더라도 시스템 전체적인 측면에서의 대역폭 활용률과 데이터 전달율을 높일 수 있다.

그렇지만 위와 같이 대역폭을 조절할 경우 한쪽으로 대역폭이 몰려 버릴 가능성이 있다. 위의 경우에서 s1 으로 가는 대역폭을 증가할 경우 거기에 비례하여 해당 swarm 으로 전송되는 데이터양도 증가하게 되고 증가된 데이터 양이 해당 swarm 의 CPM 으로 나타나게 된다. 이렇게 지속적으로 한쪽의 대역폭이 증가하게 되면 반대로 다른 swarm 의 peer 들은 아무런 데이터를 받게 되지 못할 수도 있다.

2. Propagation Coefficient

CPM 을 이용한 방식에서 위의 문제는 대역폭 배정의 기준으로 사용하는 CPM 이 CPM 으로 인한 전송량 변화 결과에 영향을 받는, 출력이 입력으로 재이용되는 일종의 순환 구조에 있다. 이로 인해 다른 변동사항이 없음에도 불구하고 한쪽으로 대역폭 배정이 자연스럽게 치우치게 되는 것이다. 이를 해결하기 위해 본 논문에서는 새로운 아이디어를 소개한다.

$$Bandwidth_{peer_i} = Total\ Bandwidth \times \frac{CPM_i}{\sum_i CPM_i}$$

새로운 식을 보면 기존의 CPM 대신, 주어진 대역폭 대비 CPM 으로 계산되는 Propagation Coefficient 을 새로운 대역폭 배정 기준으로 사용하게 된다. CPM 을 Swarm 에 배정되는 대역폭에 기준함으로써 진정으로 해당 Swarm 의 전달 계수를 알 수 있게 된다. 이 수식으로 대역폭을 배정하게 되면 전달 계수가 큰 쪽으로 대역폭을 배정하다가도 추가되는 대역폭 대비 증가되는 CPM 이 감소되게 되면 점점 대역폭 배정 증가폭도 줄어들게 되는 것이다. 이렇게 Propagation Coefficient 을 기준으로 수신 peer 간 대역폭을 배정하게 됨으로써 큰 대역폭이 주어진 쪽에서 대용량의 데이터 전송이 일어나는 것과 작은 대역폭이 주어진 쪽에서 작은 용량의 데이터 전송이 일어나는 것이 경쟁이 가능하게 되며, 약한 Peer 쪽에서도 일정 수준 이상의 대역폭을 확보할 수 있게 된다.

3. 시뮬레이션 환경

두 기법을 실험하고 비교하기 위한 시뮬레이션을 수행하기 위해 Python 언어로 간단한 시뮬레이터를 직접 구현하였다. 리눅스 머신 한대에서 시뮬레이션 하였으며, Figure2 과 같은 Topology 를 이용하였으며 10 초마다 대역폭을 재배정하도록 하였다. 영향요인을 줄이고 비교 차이를 확실하게 나타내기 위해 전송되는 데이터는 시작 시점에 하나의 콘텐츠만 전송하도록 하였다. 시험 Topology 내의 모든 Peer 는 해당 데이터를 요청하는 것으로 가정하였다. 공급자 Peer 의 대역폭만 최대량이 제한되어 있고, 다른 모든 Peer 간의 통신 대역폭은 제한이 없는 것으로 가정하였다.

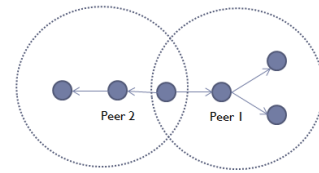


Figure 2 시뮬레이션 토폴로지

4. 시뮬레이션 결과

시뮬레이션 결과 대역폭의 치우침이 개선된 것을 볼 수 있다. Figure3 그래프의 경우 적용 전의 대역폭 변화 추이를 표현하고 있다. 이 경우 peer1 쪽으로는 가능한 최대의 대역폭을 점유하고 있고, peer2 는 전혀 대역폭을 배정받지 못하고 있는 것을 볼 수 있다. 그리고 peer1 이 완료된 후에야 전체 대역폭을 점유하게 된다. 반면에 적용 후의 대역폭 변화 추이를 표현하고 있는 Figure4 그래프에서는 peer1 이 peer2 에 비해 좀 더 많은 대역폭을 배정받긴 하지만 비교적 치우침 없이 데이터를 전달받고 있음을 볼 수 있다. 부가적인 효과로 대역폭의 변화폭이 줄어들어 전체 완료 시간도 115 에서 102 로 줄어들게 되었다

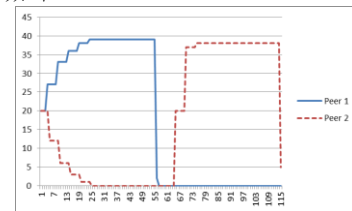


Figure 3 적용 전 peer 별 전송 속도

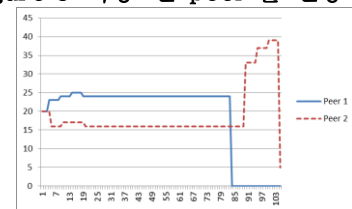


Figure 4 적용 후 peer 별 전송 속도

III. 결론

전송할 데이터가 많아지고 있는 현재, 회선의 대역폭을 최대한 효율적으로 활용하는 것이 중요하다. 본 논문에서 제시하는 Propagation Coefficient 를 이용하면 기존에 가능했던 전체 시스템 대역폭의 효율적인 사용뿐 만 아니라 peer 들 간의 심각한 치우침 없는 균형잡힌 데이터 전송을 가능하게 할 수 있다.

ACKNOWLEDGMENT

본 연구는 미래창조과학부가 지원한 2013 년 정보통신방송 (ICT) 연구개발사업의 연구결과로 수행되었음.

참 고 문 헌

[1] Ryan S. Peterson, Bernard Wong, Emin GnSirer: A Content Propagation Metric for Efficient Content Distribution. SIGCOMM Conference, August 2011.
 [2] Ipoque - Internet Studies
<http://www.ipoque.com/resources/internet-studies>