클라우드 CDN 환경에서의 효과적인 VM 관리를 위한 네트워크 I/O 성능 분석

현명석, 김희재, 윤찬현
한국과학기술원 전기및전자공학과
e-mail : {tidusqoop,kim881019,chyoun}@kaist.ac.kr

An analysis of Network I/O Performance for Effective VM Management under Cloud CDN Environment

Myeongseok Hyeon, Heejae Kim, and Chan-Hyun Youn
Dept. of Electrical Engineering, KAIST

요 약

최근 콘텐츠 전송 네트워크(content delivery network, CDN)와 클라우드 컴퓨팅(cloud computing)을 결합하여 효과적으로 콘텐츠를 전달하기 위한 방법이 대두되고 있으며 이와 같이 클라우드 컴퓨팅과 결합하여 서비스되는 CDN 을 클라우드 CDN 이라고 칭한다. 본 논문에서는 클라우드 CDN 환경에서의 효과적인 가상 머신(virtual machine, VM) 관리를 위한 캐시서버(caching server)로써의 VM 의 네트워크 I/O 성능 분석을 다룬다. 해당 성능 분석은 엔드 유저(end-user)들과 캐시서버 간 동영상 스트리밍(streaming)을 통하여 이루어졌으며 해당 캐시 서버의 네트워크 I/O 성능에 영향을 주는 다양한 경우에 대하여 진행되었다. 본 논문에서의 성능 분석은 클라우드 CDN 환경에서의 데이터센터(datacenter) 선택 및 요청 라우팅(routing) 등에 적용될 수 있다.

1. 서론

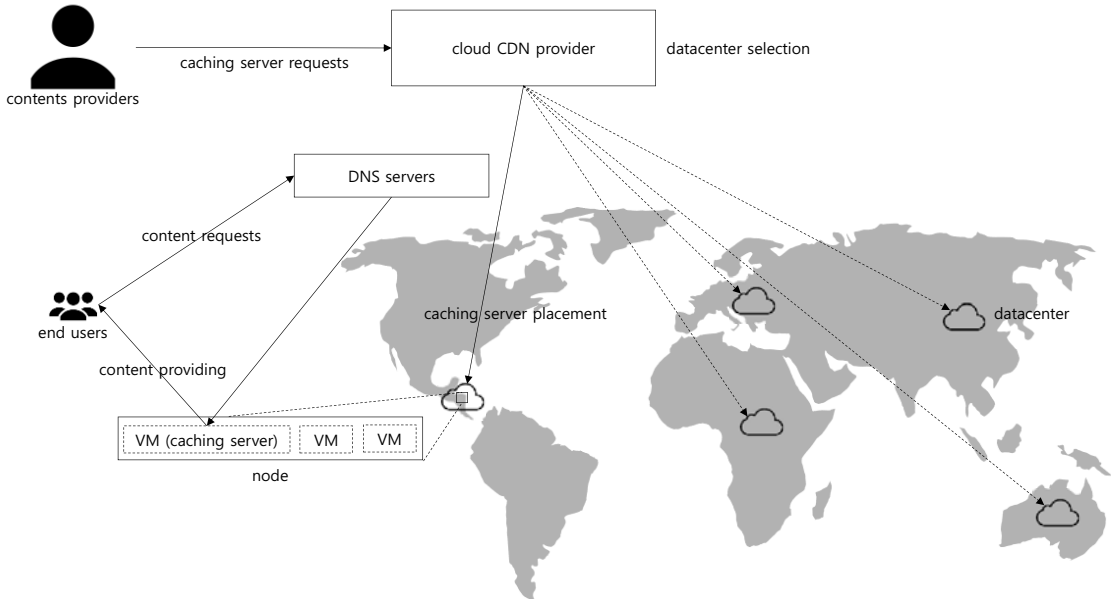
최근 콘텐츠 전송 네트워크(content delivery network, CDN)와 클라우드 컴퓨팅(cloud computing)을 결합하여 효과적으로 콘텐츠를 전달하기 위한 방법이 대두되고 있다. 그 예로 Amazon 과 Rackspace 등의 클라우드 서비스 제공자들에서는 각각 CloudFront [1] 와 Rackspace CDN [2]을 제공하고 있고 Akamai 와 Limelight 등의 CDN 제공자들에서는 그들의 CDN 들에 클라우드 컴퓨팅을 도입했으며 [3, 4] Netflix 와 Youtube 등의 콘텐츠 제공자들 또한 각각 Amazon Web Service 와 Google 클라우드를 사용하여 그들의 서비스를 제공하고 있다 [5, 6]. 이와 같이 클라우드 컴퓨팅과 결합하여 서비스되는 CDN 을 클라우드 CDN 이라고 칭한다 [7].

본 논문에서는 캐시 서버(cache server)로써의 가상머신(virtual machine, VM)의 네트워크 I/O 성능 분석을 다룬다. 이는 클라우드 CDN 에서의 캐시 서버 배치를 위한 데이터센터(datacenter) 선택, 요청 라우팅(routing) 등에서 VM 의 네트워크 I/O 성능을 고려하여 수행하도록 함으로써 효과적인 자원 관리를 가능하도록 하는데 목적이 있다. 따라서 해당 성능 분석은 클라우드 CDN 환경을 고려하여 엔드 유저(end-user)들이 캐시 서버에 동영상 스트리밍(streaming) 요청을 하고 캐시 서버에서 해당 요청을 처리하는 환경에서 이루어졌으며 해당 캐시 서버의 네트워크 I/O 성능에 영향을 주는 다양한 경우에 대하여 진행되었다.

2. 클라우드 CDN 환경에서의 VM 관리

그림 1 은 클라우드 CDN 환경을 나타낸다. 해당 클라우드 CDN 환경에서 캐시 서버들은 데이터센터 안에서 VM 으로써 배치된다. 해당 그림에서 콘텐츠 제공자의 캐시 서버 요청을 받은 클라우드 CDN 제공자는 데이터센터 선택을 통하여 캐시 서버들을 데이터센터들에 배치한다. 해당 캐시 서버들은 엔드 유저들의 콘텐츠 요청들을 처리하는 데 사용되며 해당 콘텐츠 요청들은 도메인 네임 서비스(domain name service, DNS) 서버들에서의 요청 라우팅을 통하여 캐시 서버들에 할당된다.

Kim et al. [8]에서는 해당 환경에서의 내쉬 협상 해(Nash bargaining solution, NBS) 기반 데이터센터 선택 알고리즘을 다루었다. 해당 알고리즘의 최적화 문제는 식 1, 2, 3 과 같이 구성되었다. 식 1 에서 u(V)는 유틸리티 함수(utility function)이며 벡터 V 의 원소 v_i 는 결정 변수로써 데이터센터 $i \in I$ 에 배치된 캐시 서버들의 개수를 의미한다. 여기서 I 는 데이터센터들의 집합을 나타낸다. 또한 c, δ_i, n 은 각각 제어 파라메터(parameter), 엔드 유저 수요를 데이터센터의 지리적 위치에 따라 클러스터링(clustering) 하였을 때 데이터센터 i 에서의 엔드 유저 수요 예측 값, 하나의 캐시 서버에서 처리될 수 있는 엔드 유저 요청 수를 의미한다. 식 2 는 v_i 들의 합이 캐시 서버 요청 개수와 예약 VM 개수의 합과 같아야 함을 나타내는 제약 조건으로써 s 와 ρ_i 는 각각 캐시 서버 요청 개수와 데이터



(그림 1) 클라우드 CDN 환경 [8].

센터 i 에서의 예약 VM 개수를 나타낸다. 여기서 예약 VM 은 해당 알고리즘 수행 시 데이터센터에 존재하는 VM 들을 의미한다. 식 3 은 각 데이터센터 i 에서의 v_i 의 최소값을 제한하기 위한 제약조건이다.

$$\text{maximize } u(V)=\sum_{i\in I}\log\left(v_i-\frac{c\cdot\delta_i}{n}\right) \quad (1)$$

$$\text{subject to } \sum_{i\in I}v_i-\left(s+\sum_i\rho_i\right)=0 \quad (2)$$

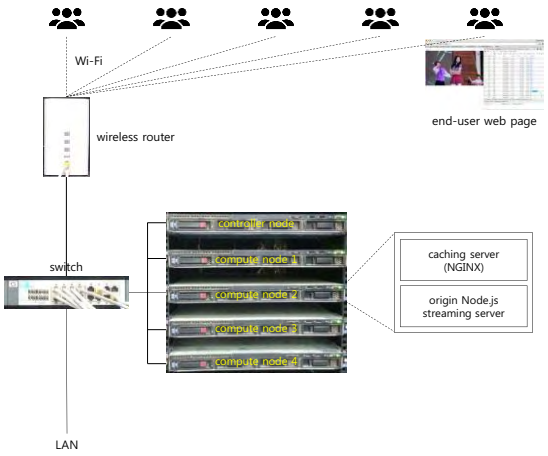
$$v_i\geq\max\left(\frac{c\cdot\delta_i}{n},\rho_i\right),\forall i\in I. \quad (3)$$

해당 알고리즘에서 n 은 캐시 서버 안의 프로세스(process) 및 쓰레드(thread) 간 자원 경쟁과 동일 노드(node) 내 VM 간 자원 경쟁 등으로 인하여 캐시 서버의 네트워크 I/O 성능에 영향을 미칠 수 있다. 따라서 n 은 노드 및 VM 들의 성능을 고려하여 결정되어야 하며 3 장에서는 이를 위한 VM 의 네트워크 I/O 성능 분석을 다룬다.

3. 캐시 서버로써의 VM 의 네트워크 I/O 성능 분석

3.1 실험 환경 및 시나리오

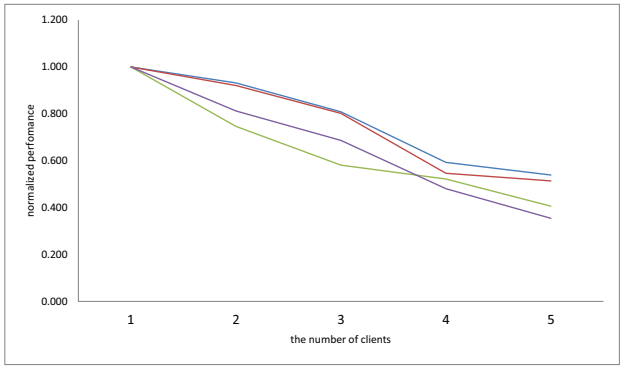
캐시 서버로써의 VM 의 네트워크 I/O 성능 분석을 위하여 그림 2 과 같은 실험 환경을 구성하였다. 본 실험에서는 오픈스택(OpenStack) [9]을 사용하여 클라우드 환경을 구축하였고 해당 노드들은 각각 두 개의 쿼드코어 프로세서(quad-core processor) Intel Xeon Processor E5620, 14 GB 의 메모리(memory), 1 TB 의 디스크(disk)로 이루어져 있으며, 하나의 스위치에 연결되어 있다. 앤드 유저들은 노드와의 통신을 위해 이 스위치에 연결된 무선공유기에 접속한다.



(그림 2) 실험 환경.

본 실험을 위하여 compute node 2 에 두 개의 VM 을 생성하였고, VM 은 각각 1 VCPU, 2 GB 메모리, 80 GB 의 디스크 사양을 가진다. VM 1 에는 NGINX [10]를 사용한 캐시 서버와 Node.js [11]기반의 스트리밍 서버를 구성하였다.

엔드 유저들은 캐시 서버에 동영상 스트리밍을 요청하고, 캐시 서버는 요청된 콘텐츠의 캐싱 여부에 따라 응답하거나 요청을 스트리밍 서버에 라우팅한다. 본 실험에서는 이 과정에서의 스트리밍 완료 시간 및 대역폭을 크롬 개발자 도구(Chrome developer tools) [12]을 통하여 측정하였다. 더불어 해당 실험 과정에서 추가적으로 VM 2 에서 CPU, 메모리, 디스크 I/O 집약적 작업으로써 각각 Compress-7zip[13], Stress[14], Bonnie++[15]를 실행했을 때의 네트워크 I/O 성능 분석 또한 진행하였다.



(그림 3) 엔드 유저 수에 따른 정규 성능. 여기서 파란색, 빨간색, 녹색, 보라색은 각각 y = 0, 1, 2, 3 을 나타냄.

<표 1> 엔드 유저 수에 따른 정규 성능.

x \ y	0	1	2	3
1	1.000	1.000	1.000	1.000
2	0.932	0.921	0.747	0.812
3	0.808	0.802	0.581	0.687
4	0.593	0.547	0.522	0.481
5	0.539	0.514	0.406	0.354

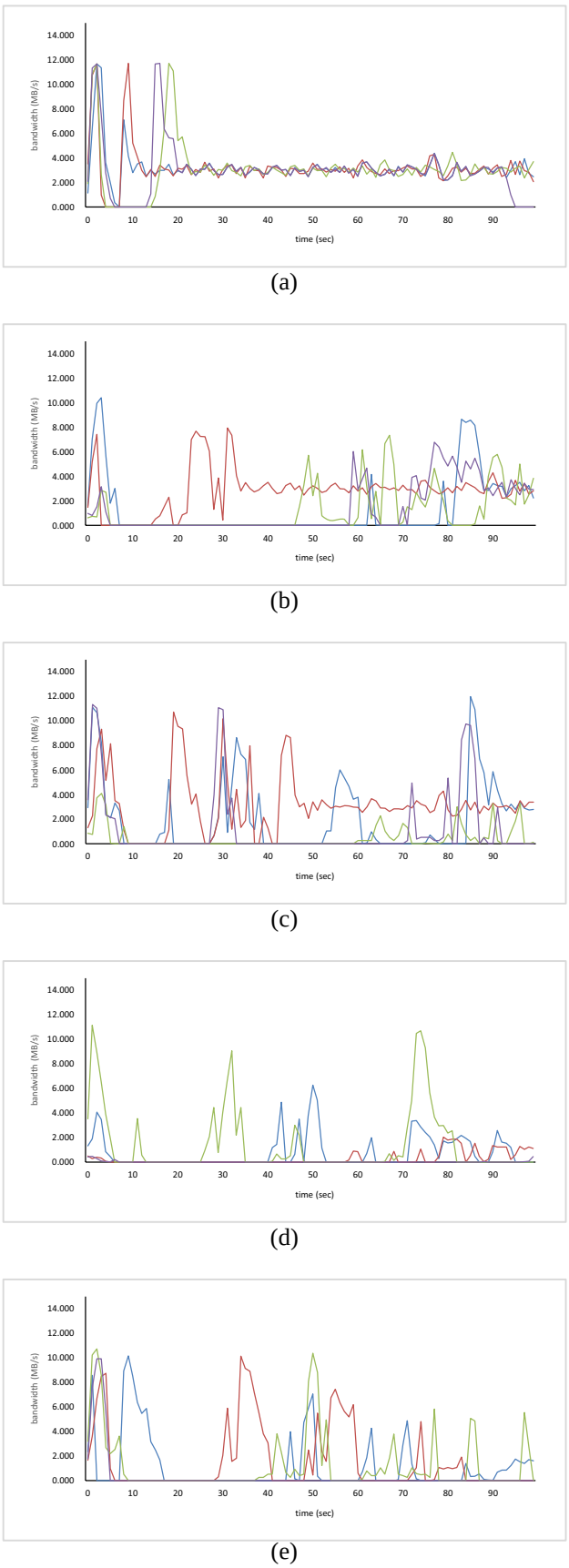
3.2 실험 결과 및 고찰

그림 3 은 엔드 유저 수에 따른 캐시 서버의 정규 성능을 나타낸다. 정규 성능 NP_{xy} 는 식 (4)와 같이 나타나며 여기서 P_{xy} 는 엔드 유저 수가 x 이고 VM 2 에서 수행되고 있는 작업의 종류가 y 일 때의 성능으로써 본 실험에서는 스트리밍 완료 시간의 역수를 의미한다. 여기서 y = 0, 1, 2, 3 은 각각 VM 2 에서 오직 백그라운드(background) 프로세스들만 실행되고 있을 때, 추가적으로 Compress-7zip 이 실행될 때, 추가적으로 Stress 가 실행될 때, 추가적으로 Bonnie++이 실행될 때를 의미한다.

$$NP_{xy} = \frac{P_{10}}{P_{xy}}. \qquad (4)$$

실험 결과를 통하여 본 실험 환경에서 네트워크 I/O 성능은 엔드 유저 수에 따라 지수적 감소를 보이며 CPU 집약적 작업보다 메모리 및 디스크 I/O 집약적 작업에 더 민감함을 알 수 있다.

그림 4 는 시간에 따른 동영상 스트리밍의 대역폭을 나타낸다. 실험 결과를 통하여 대역폭은 시간에 따라 점차 수렴함을 보이며 이는 엔드 유저 수가 작고 네트워크 I/O 성능에 덜 민감한 자원일 수록 두드러진다. 따라서 본 실험 환경에서 스트리밍 서버는 네트워크 I/O 성능 감소가 적을수록 끊김 없는 실행이 보장되는 특정 속도를 유지하려고 함을 알 수 있다.



(그림 4) 시간에 따른 동영상 스트리밍 대역폭 변화. 여기서 파란색, 빨간색, 녹색, 보라색은 각각 y = 0, 1, 2, 3 을 나타냄. (a) x = 1, (b) x = 2, (c) x = 3, (d) x = 4, (e) x = 5.

4. 결론

본 논문에서는 클라우드 CDN 환경에서의 효과적인 VM 관리를 위한 캐시 서버로써의 VM 의 네트워크 I/O 성능 분석을 다루었다. 실험 결과로써 동영상 스트리밍 시 네트워크 I/O 성능은 엔드 유저 수에 따라 지수적 감소를 보였으며 CPU 집약적 작업보다 메모리 및 디스크 I/O 집약적 작업에 더 민감함을 알 수 있었다. 또한 해당 대역폭은 시간에 따라 점차 수렴함을 보였으며 이는 엔드 유저 수가 작고 네트워크 성능에 덜 민감한 자원일수록 두드려졌다. 해당 실험 결과는 클라우드 CDN 에서의 데이터센터 선택 및 요청 라우팅에 적용되어 효과적인 자원 관리를 가능하게 하는 데 도움을 줄 수 있을 것이다.

Acknowledgement

본 연구는 미래부가 지원한 2012 년도 정부(교육과학기술부)의 재원으로 한국연구재단-클라우드 Collaboration 기술 사업과 BK21 사업(No. 2012-0006425), 미래창조과학부 및 정보통신산업진흥원의 대학 IT 연구센터 육성지원 사업, 2014 년 정보통신·방송(ICT) 연구개발사업의 연구결과로 수행되었음 (NIPA-2014(H0301-14-1020)).

참고문헌

[1] CloudFront, available: http://aws.amazon.com/ko/cloudfront/.
[2] Rackspace CDN, available: http://www.rackspace.com/cloud/cdn-content-delivery-network/.
[3] Akamai, available: http://www.akamai.co.kr/.
[4] Limelight, available: http://www.limelight.com/.
[5] AWS Case Study: Netflix, http://aws.amazon.com/ko/solutions/case-studies/netflix/.
[6] Google Cloud Boosting YouTube Upload Speeds, available: http://blogs.wsj.com/digits/2011/02/14/google-cloud-boosting-youtube-upload-speeds/.
[7] F. Chen, K. Guo, J. Lin, and T. L. Porta, "Intra-cloud Lightning: Building CDNs in the Cloud," in INFOCOM, 2012.
[8] H. Kim, Y. K. Ha, and C. H. Youn, “Nash Bargaining Solution-based Datacenter Selection under Cloud Content Delivery Network Environment,” submitted to IEEE COMSOC MMTC E-Letter, 2015.
[9] OpenStack, available: http://www.openstack.org/.
[10] NGINX, available: http://www.nginx.org/.
[11] Node.js available: https://nodejs.org/.
[12] Chrome developer tools, available: https://developer.chrome.com/devtools/.
[13] Phoronix Test Suite, available: http://www.phoronix-test-suite.com/.
[14] Stress, available: http://pkgs.repoforge.org/stress/
[15] Bonnie++. Available: http://www.coker.com.au/bonnie++/