

서버센트릭 데이터센터를 위한 서버 배치 최적화에 관한 연구

서성원, 박민호
 숭실대학교 정보통신전자공학부
 e-mail : ithinkb@naver.com, mhp@ssu.ac.kr

A Study of Server Deployment for Optimization of Server-Centric Datacenter

Sungwon Seo, Minho Park
 Dept. of Electronic Engineering, Soong-Sil University

요 약

클라우드 컴퓨팅 기술의 발전 및 모바일 디바이스의 확산으로 데이터센터의 수요가 폭발적으로 증가하고 있고, 데이터 센터내 트래픽 또한 급증할 것으로 예측된다. 따라서 데이터센터 내에 서버를 확장 할 수 있고, 고장방지에 뛰어난 서버센트릭 네트워크 구조가 수많은 서버를 효율적으로 관리할 수 있는 대안으로 제시되었다. 본 논문에서는 서버센트릭 데이터센터에서 서버 배치를 통한 트래픽의 변화에 대해 분석하고 효율적으로 트래픽을 감소시키기 위한 휴리스틱(Heuristic) 접근방법을 제안한다.

1. 서론

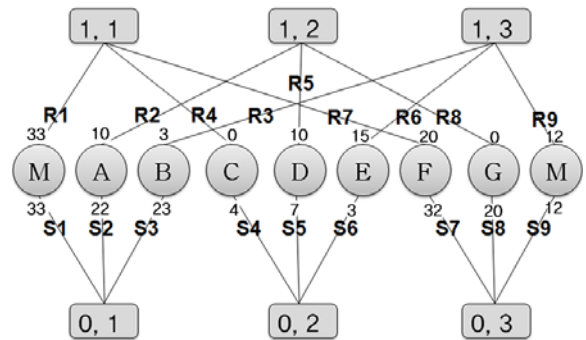
근래 스마트폰, 태블릿의 확산과 클라우드 컴퓨팅이 급부상하면서 데이터센터의 수요가 폭발적으로 증가하고 있다. 시스코는 ‘시스코 글로벌 클라우드 인덱스(CGI) 2012-2017’ 에서 2017년에는 전세계 데이터센터 트래픽이 2012년 대비 3 배 증가한 7.7 Zettabytes (1 Zettabytes =10 억 Terabytes)에 달할 것으로 전망했고 데이터센터 트래픽의 76%가 데이터센터 내부에서 발생할 것으로 예측하였다[1]. 이러한 데이터센터 내부의 급격한 트래픽의 증가를 해결하기 위해 서버가 호스트와 노드로 동작하여 서버의 배치가 자유롭고 서버간의 패킷 전달이 가능한 서버센트릭 구조가 제안되었다. 서버센트릭 구조에서는 서버의 배치에 따라 패킷의 이동경로가 달라지고, 데이터센터 내에서 발생하는 전체 트래픽 또한 달라진다. 트래픽 발생을 최소화하기 위해 서버배치의 모든 경우의 수를 고려하여 최적의 배치를 구할 수 있지만, 다수의 서버를 배치할 경우 모든 경우의 수를 구하는 것은 현실적으로 불가능하다. 본 연구에서는 휴리스틱(Heuristic) 접근법을 이용하여 가장 효율적인 방법으로 데이터센터 내부의 트래픽을 최소화하는 최적의 서버배치 방법을 찾는다.

2. 서버배치 최적화

2.1 서버센트릭 데이터센터

서버센트릭 네트워크는 서버가 호스트와 노드의 역

할을 하는 구조이며[2] 서버가 스위치와 같은 동작을 하기 위해 패킷을 포워딩하는 기능을 포함하는 특징을 가지고 있다.[3] 서버센트릭 데이터센터는 패킷을 포워딩 할 수 있는 서버와 COTS(commodity off-the-shelf) mini-switch 가 연결된 구조로 되어있다.



(그림 1) 서버센트릭 구조의 예

2.2 최적화 배치의 필요성

서버센트릭 네트워크에서 서버는 하나의 독립적인 노드로써 동작하므로 서버의 배치 및 이동이 자유롭다. 따라서 네트워크 내에서 발생하는 트래픽이 서버의 배치에 영향을 준다. 서버 배치에 따른 트래픽의 변화를 알아보기 위해 그림 1 과 같이 서버센트릭 구조를 구성하고 서버간에 발생하는 트래픽을 그림 2 와 같이 가정하였다. 이 때 네트워크 내부에 발생한 전

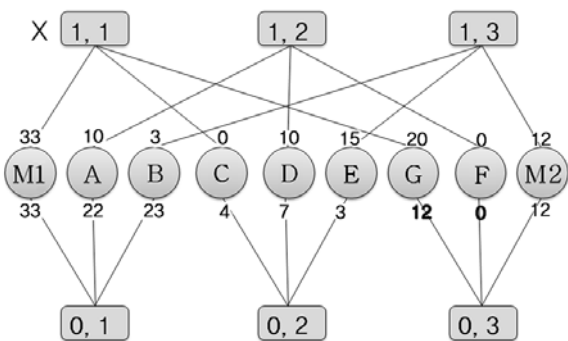
체 트래픽은 259bps 이다.

동일한 조건의 서버센트릭 구조에서 서버배치에 따른 효율의 변화를 확인하기 위하여 그림 1의 서버배치에서 서버 F와 서버 G의 위치를 그림 3과 같이 변경하였다. 이 때 발생한 전체 트래픽은 219bps로 그림 1의 구조보다 15.4% 감소하였다. 동일한 조건의 서버센트릭 구조에서 두 개의 서버만 위치를 변경하였음에도 불구하고 네트워크 효율성이 크게 증가하는 것을 확인할 수 있다. 이는 곧 적절한 서버배치를 통해서 효율성의 극대화, 즉 최적화(optimization)를 달성할 수 있음을 보여준다.

서버센트릭 데이터센터에서 가능한 모든 서버배치의 경우의 수는 전체 서버 개수의 조합으로 구할 수 있다. 그림 1과 같이 9개의 서버가 존재할 경우 $9! (=362880)$ 개의 방법으로 서버를 배치할 수 있고 20개의 서버가 존재할 경우 $20! (=2432902008176640000)$ 번의 서버배치를 통해 최적의 경우를 찾아야 하며 서버의 개수가 하나씩 증가할 때 마다 가능한 경우의 수는 기하급수적으로 증가한다. 따라서 모든 경우의 수를 따져 서버를 배치하는 것은 현실적으로 불가능한 방법이며 서버의 수가 많은 데이터센터 내에서 최적의 서버배치를 효율적으로 찾기 위해서는 휴리스틱(Heuristic)한 접근법이 필요하다[4].

		Destination								
		A	B	C	D	E	F	G	X	
Source	A	■		2						6
	B		■		3					7
	C	2		■						
	D		6		■					
	E					■	4			
	F						8	■		
	G								■	10
	X	10	10						10	■

(그림 2) 트래픽 발생 매트릭스



(그림 3) 서버센트릭 데이터센터의 트래픽 발생 예시

2.3 제안 알고리즘

휴리스틱(Heuristic) 접근방법은 어떠한 문제를 해결할 수 있는 정확한 방법이 존재하지만, 이용 가능한

조건을 가지고 효율적인 해결방법을 구하기 어려운 경우에 최적의 문제해결 방법을 찾기 위해 사용한다 [5]. 휴리스틱 접근법을 이용할 경우 하나의 네트워크 안에서 서버를 배치하는 모든 경우의 수를 고려하지 않고 효율적인 배치가 가능하므로 서버센트릭 데이터센터에서 서버를 효율적으로 배치하는데 효과적이다.

휴리스틱 접근법으로 문제를 해결하기 위해 먼저 한가지 문제를 다루기 쉬운 하위문제로 나누고, 각각의 하위문제를 해결한다. 그 다음 하위문제들의 해결방안들을 하나로 합침으로써 전체적인 문제를 해결한다. 휴리스틱 접근법을 이용하여 서버센트릭 데이터센터에서 서버배치를 최적화하는 방법은 다음과 같다.

서버배치 최적화 알고리즘

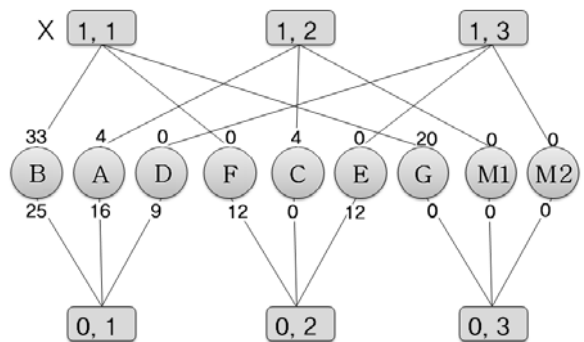
Step1. Source와 Destination 사이의 트래픽이 가장 큰 경우를 찾음

Step2. Source와 Destination 사이의 이동경로가 최소가 되도록 Source와 Destination을 재배치

Step3. Step1부터 반복

종료조건. 전체 트래픽 변화량을 재배치전과 비교하여 변화량이 개선되지 않는 결과가 반복될 때 종료

네트워크 내에서 Source 서버와 Destination 서버 사이에 전송하는 패킷의 크기가 큰 경우, 패킷이 경유하게 되는 노드가 많을수록 네트워크 전체에 걸리는 트래픽의 부하가 커진다. 그러므로 전송하는 패킷의 크기가 큰 경우 Source 서버와 Destination 서버 사이의 거리가 감소하면 전체 네트워크에 걸리는 트래픽 또한 감소하게 된다. 따라서 서버배치 최적화 알고리즘은 데이터센터내의 트래픽 증가를 유발하는 요인을 서버가 전송하는 패킷의 크기가 큰 경우와 생성된 패킷이 이동하는 거리가 긴 경우로 나누고 두 문제를 각각 해결함으로써 전체 트래픽 발생을 감소시켰다.



(그림 4) 서버배치에 따른 트래픽 변화

그림 1의 구조에 서버배치 최적화 알고리즘을 적용

할 경우 그림 4 와 같은 결과를 얻을 수 있다. 그림 4 의 전체 트래픽은 135bps 로 그림 1 의 전체 트래픽에 비해 47.9% 감소한 것을 확인할 수 있고, 표 1 의 구간별 변화량을 분석하면 부분적으로 트래픽이 증가한 구간이 있지만 전체적으로 트래픽이 감소함을 알 수 있다. 따라서 서버배치 최적화 알고리즘의 Step1, Step2 과정을 반복함으로써 구간별 지역 최소점을 구하고 전체 네트워크 트래픽의 전역 최소점을 구하여 전체 트래픽을 감소시킴으로써 서버센터릭 데이터 센터에서 효율적인 서버배치를 할 수 있다.

(그림 5) Simulated Annealing 예시

구간	트래픽 1	트래픽 2	변화량
R1	33	16	-17
R2	10	4	-6
R3	3	0	-3
R4	0	17	17
R5	10	4	-6
R6	15	0	-15
R7	20	20	0
R8	0	0	0
R9	12	0	-12
S1	33	25	-8
S2	22	16	-6
S3	23	9	-14
S4	4	12	8
S5	7	0	-7
S6	3	12	9
S7	32	0	-32
S8	20	0	-20
S9	12	0	-12
총 트래픽	259	135	-124

(표 1) 구간별 트래픽 변화량

3. 결론

본 연구에서는 서버센터릭 데이터센터에서 서버배치의 최적화를 통한 트래픽의 감소에 대해 알아보았다. 서버를 배치하는 방법에는 많은 경우의 수가 있고, 배치되는 서버의 위치에 따라 데이터센터 내부에서 발생하는 전체 트래픽의 크기가 다양하다. 따라서 효율적인 데이터센터의 운영을 위해 서버배치의 최적화가 필요하며 휴리스틱(Heuristic) 접근법을 이용한 서버배치 최적화 알고리즘이 서버센터릭 데이터센터 내 트래픽을 감소시키는데 효과적일 것으로 기대된다.

참고문헌

[1] Cisco. Global Cloud Index(CGI). <http://www.cisco.com/c/en/us/solutions/service-provider/global-cloud-index-gci/index.html#~Overview>.
 [2] Chuanxiang Guo, Guohan lu, Dan Li, Haitao Wu, Xuan

Zhang, Yunfeng Shi, Chen Tian, Yongguang Zhang, Songwu Lu. (2009). BCube: A High Performance, Server-centric Network Architecture for Modular Data Centers.
 [3] Kai Peng, Rongheng Lin, Binbin Huang, Hua Zou, and Fangchun Yang. (2013). Node Importance of Data Center Network Based on Contribution Matrix of Information Entropy. JOURNAL OF NETWORKS, 8(6), 1248-1254.
 [4] S. Kirkpatrick, C. D. Gelatt, Jr., M. P. Vecchi. (1983). Optimization by Simulated Annealing. SCIENCE, 220(4598), 671-680.
 [5] Rafael Marti, Gerhard Reinelt. (2011). The Linear Ordering Problem Exact and Heuristic Methods in Combinatorial Optimization (Chapter 2). Springer