

데이터센터 자원 연결 방안 연구

기장근*† , 권기영*

A Study on Connections of Resources in Data Centers

Jang-Geun Ki*† , Kee-Young Kwon*

요 약

최근 인터넷 보급과 함께 폭발적으로 증가하고 있는 클라우드 서비스 등의 데이터 트래픽 급증으로 데이터 센터의 구성 자원들을 효율적으로 연결할 수 있는 초고속 광모듈 네트워크의 필요성이 급증하고 있다. 본 논문에서는 광스위치 연결구조에서 스위칭 셀 동작을 제어하기 위한 알고리즘들을 제안하고, 이에 따른 성능을 시뮬레이션을 통해 비교 분석하였다. 성능 분석결과 본 논문에서 제안하는 알고리즘이 기존 알고리즘에 비해 1:2 이상의 다중연결설정 성공확률이 약 3~7% 정도 향상됨을 보였다.

Abstract

The recent explosion of data traffic, including cloud services, coupled with the Internet penetration has led to a surge in the need for ultra-fast optical networks that can efficiently connect the data center's reconfigurable resources. In this paper, the algorithms for controlling switching cell operation in the optical switch connection structure are proposed, and the resulting performance is compared and analyzed through simulation. Performance analysis results showed that the algorithm proposed in this paper has improved the probability of successful multi-connection setup by about 3 to 7% compared to the existing algorithm.

한글키워드 : 데이터 센터, 자원, 연결 네트워크, 스위치 제어, 성능

keywords : data center, resource, connection network, switch control, performance

1. 서론

서버 컴퓨터와 저장장치 및 이들을 연결하는 네트워크 회선들을 제공하는 데이터 센터는 인터넷의 보급과 함께 폭발적으로 증가하여왔으며, 이들 데이터 센터들은 전통적으로 자신들이 보유한 네트워크 자원들과 기술들을 사용해왔다. 그

러나 시간이 갈수록 클라우드 서비스 등을 위한 데이터 트래픽 급증으로 데이터 센터들의 투자가 대규모로 이루어짐에 따라 초광대역 전송, 저전력 소모 조건 등을 만족시키는 새로운 광 모듈들에 대한 수요가 증가하고 있으며, 데이터 센터 운영자들은 광 네트워크와 같은 새로운 기술에 의존하게 되었다. Technavio의 연구보고서[1]에 따르면 구글, 페이스북, 마이크로소프트 등이 소유하고 있는 데이터 센터들은 에너지 효율적이고 초고성능의 저장장치 인프라를 갖추고 있으며,

* 공주대학교 전기전자제어공학부 교수

† 교신저자: 기장근(email: kjpg@kongju.ac.kr)

접수일자: 2019.11.08. 심사완료: 2019.12.06.

게재확정: 2019.12.20.

광 네트워크 하드웨어가 이러한 데이터 센터들의 핵심요소라고 밝히고 있다. 예를 들어 미국 Iowa 주 Council Bluffs에 위치한 구글의 데이터 센터는 미국 일반 가정 인터넷보다 20만배 빠른 속도로 구글의 웹 사이트들을 연결하는 광 네트워크로 구성되어 있다.

전 세계 광 스위치 시장은 Technavio의 연구 보고서[1]에 따르면 2018-2022년 사이에 연평균 성장률(CAGR: Compound Annual Growth Rate) 11% 정도를 보일 것으로 예상하고 있으며, Mordor Intelligence의 연구보고서[2]에 따르면 2019-2024년 사이의 연평균성장률이 13.41%에 이를 것으로 전망되고 있다.

이와 같이 광통신을 비롯한 스위치 망의 성능이 급속히 향상된 최근에는 새로운 응용 서비스의 출현에 따른 다양한 트래픽 요구조건들을 동적으로 만족시키기 위해 지역적으로 산재해 있는 다양한 컴퓨팅 자원들을 동적으로 할당하고 연결 관리하는 방안에 대한 연구가 활발히 진행되고 있다[3-6].

본 논문에서는 Nirmal Kumbhare의 논문에서 제시된 초고성능 초저지연의 JITA 인프라 구조 [7-8]에서 제안된 광스위치 연결구조의 스위칭 셀 동작을 제어하기 위한 알고리즘들을 제안하고, 이에 따른 성능을 시뮬레이션을 통해 비교 분석하였다.

2. 광 스위치 연결 알고리즘

본 연구의 선행연구에서 개발된 기존 광 스위치 제어 알고리즘(ver.1)[8]에서는 마스터에서 각각의 슬레이브 호스트로 최단 경로를 계산하고 이들을 모아서 마스터와 슬레이브들간 1:n 연결설정을 하였다. 이 방법을 사용할 경우 마스터에서 각 슬레이브까지의 거리는 최단 거리로 연결될 수 있으나 전송 링크와 스위치가 초고속 전송을 지원하

는 광케이블 스위치 시스템으로 구성될 경우 마스터와 슬레이브간 전송지연시간은 큰 의미를 갖지 않으며, 오히려 고정된 크기의 스위칭 시스템에서 좀 더 많은 연결을 지원할 수 있는 효율적인 연결 구조를 갖도록 하는 것이 바람직하다.

본 논문에서는 모든 링크가 양방향으로 동작하고 마스터와 슬레이브 사이의 연결 또한 양방향으로 동작함을 감안하여 연결해야 될 첫번째 슬레이브에서 마스터 방향으로 최단 경로를 먼저 찾고, 두 번째 이후의 슬레이브에서는 앞서 찾았던 경로상의 노드들중 가장 가까운 거리에 위치한 노드를 찾아 그 노드까지의 경로를 추가하는 방법으로 연결설정을 하도록 새로운 알고리즘(ver.2)을 제안하였다. 이와 같은 방법을 사용할 경우 마스터와 각 슬레이브들간의 경로가 중복으로 사용되는 구간이 증가하여 고정된 크기의 스위칭 시스템에서 보다 많은 연결을 수용할 수 있게 된다.

그림 1은 기존의 알고리즘과 본 논문에서 제안하는 알고리즘의 원리 차이를 보여주고 있으며, 그림에서 마스터 M으로부터 슬레이브 S1과 S2로의 경로를 차례로 설정한다고 가정한다. 그림에서 마스터 호스트 M으로부터 슬레이브 호스트 S1으로의 최단 경로가 설정된 후, 추가되는 슬레이브 호스트 S2는 영역1 또는 영역2 또는 영역3에 위치할 수 있다.

기존 알고리즘의 경우 S2 슬레이브 호스트와의 경로를 찾을 때 S2의 위치에 상관없이 무조건 마스터 M으로부터 S2까지의 최단 경로(M-S2)를 설정한다.

그러나 본 논문에서 제안하는 알고리즘에서는 S2 슬레이브가 영역 2에 속할 경우, M-S1 경로와의 최단 거리로 경로를 설정하여 M-X-S2로 마스터에서 S2 슬레이브로의 경로가 설정된다. 그림에서 M, X, S2의 세 지점은 직각 삼각형을 이루며, M-X-S2의 길이는 빗변에 해당하는

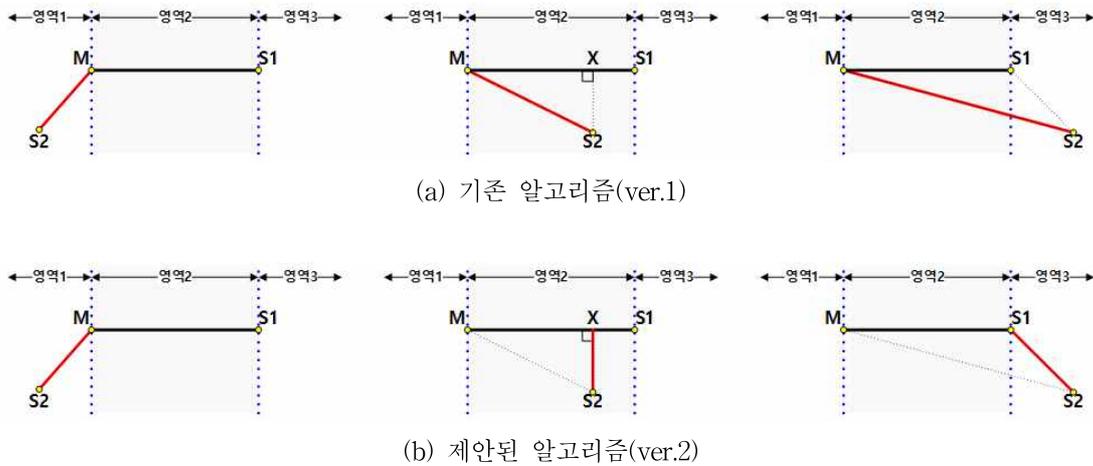


그림 1. 기존 알고리즘(ver.1)과 제안된 알고리즘(ver.2) 동작
 Fig. 1. Operations of ver.1 and ver.2 algorithms

M-S2의 길이보다 길어진다. 반면에 슬레이브 S2를 위해 새로이 추가되는 경로만 생각해 본다면 기존의 알고리즘은 M-S2 경로가 추가되고, 본 논문의 알고리즘에서는 X-S2 경로가 새로 추가되는데, 당연히 X-S2의 길이가 M-S2의 길이보다 짧게 되어 새롭게 사용되어지는 자원(링크와 스위치) 개수가 절약되고, 이 절약된 자원들은 다른 마스터-슬레이브간 연결에 사용될 수 있게 되고, 결과적으로 전체적인 연결 성공 확률이 커지게 된다. S2 슬레이브가 영역 3에 속하는 경우에는 S2에서 S1으로의 최단 경로를 찾아 M-S1-S2로 경로를 설정한다.

3. 알고리즘별 성능 분석

본 논문에서 제안하는 알고리즘(ver.2)과 기존 알고리즘(ver.1)간의 성능향상 정도를 평가하기 위해 먼저 1:n 연결설정 성공확률을 다음과 같이 정의한다. 여기서 1:n은 마스터 호스트가 n개의 슬레이브 호스트들과 연결되어야 함을 의미하고, 마스터로부터 n개 슬레이브 호스트로의 경로가

모두 존재할 경우 연결설정이 성공한 것으로 정의한다.

AxB 2차원 스위치 연결 구조의 경우 최대 수용 가능한 호스트의 수는 2(A+B) 개가 되고, 이 구조에서 가능한 1:n 연결의 최대개수는 이론적으로 아래 식 (1)과 같이 계산된다.

$$N_{max}^{2d} = floor \left[\frac{2(A+B)}{1+n} \right] \quad (1)$$

8x8 2차원 스위치 구조에 대해 1:n 연결 최대개수를 그림 2에 나타내었다.

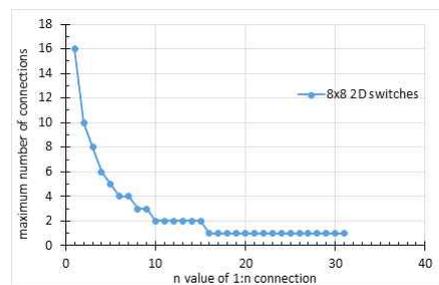


그림 2. 8x8 2차원 스위치 구조의 1:n 연결 최대개수
 Fig. 2. Maximum number of 1:n connections in 8x8 2-dimensional switch architecture

$A \times B \times C$ 3차원 스위치 연결 구조의 경우 최대 수용 가능한 호스트의 수는 $2(AB+BC+CA)$ 개가 되고, 이 구조에서 가능한 1:n 연결의 최대개수는 이론적으로 아래 식 (2)와 같이 계산된다.

$$N_{\max}^{3d} = \text{floor} \left[\frac{2(AB+BC+CA)}{1+n} \right] \quad (2)$$

4x4x4 3차원 스위치 구조에 대해 1:n 연결 최대개수를 그림 3에 나타내었다.

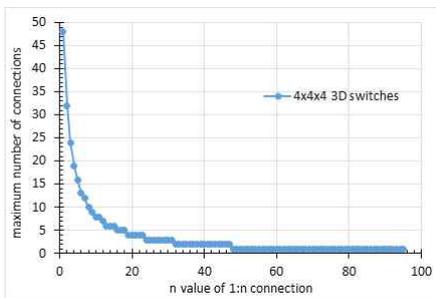


그림 3. 4x4x4 3차원 스위치 구조의 1:n 연결 최대개수

Fig. 3. Maximum number of 1:n connections in 4x4x4 3-dimensional switch architecture

1:n 연결설정 성공확률은 실제 실험에서 성공한 1:n 연결 설정 개수를 이론적으로 가능한 최대연결개수로 나누어 계산한다.

$$P^{2d} = \frac{N_{\text{experiment}}}{N_{\max}^{2d}}, \quad P^{3d} = \frac{N_{\text{experiment}}}{N_{\max}^{3d}} \quad (3)$$

그림 4에 8x8 2차원 스위치 구조에서 기존 알고리즘(ver.1)과 본 논문에서 제안된 새로운 알고리즘(ver.2)의 연결성공 확률을 비교하여 나타내었다. 그림에서 확인할 수 있듯이 1:1 연결의 경우에는 기존 알고리즘(ver.1)과 새로운 알고리즘(ver.2)의 차이가 없으나, 1:2 이상의 연결에 대한 연결성공확률은 최소 3%에서 최대 약 7%까지 새로운 알고리즘(ver.2)이 더 큼을 볼 수 있다.

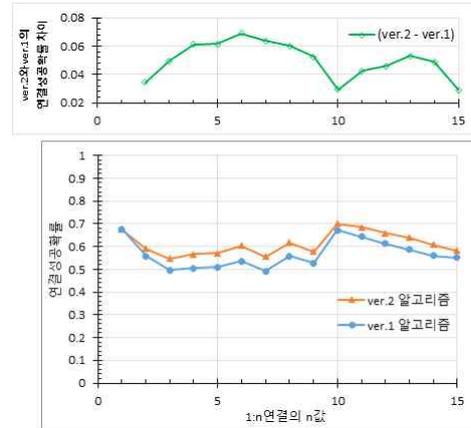


그림 4. 8x8 2차원 스위치 구조의 1:n 연결성공확률
Fig. 4. 1:n connection success probability in 8x8 2-dimensional switch architecture

그림 5는 8x8 2차원 스위치 구조에서 성공한 1:n 연결들의 마스터부터 슬레이브까지의 평균 경로 길이를 나타내고 있다. 그림에서 새로운 알고리즘의 평균 연결 길이가 기존 알고리즘에 비해 좀 더 긴 것을 볼 수 있는데 이는 새로운 알고리즘의 경우 마스터에서 각 슬레이브까지의 거리를 최소화하는 것이 목적이 아니라 가능한 한 각 슬레이브로 향하는 공통 경로를 최대한 공유함으로써 보다 많은 연결설정이 가능하도록 설계되었기 때문이다.

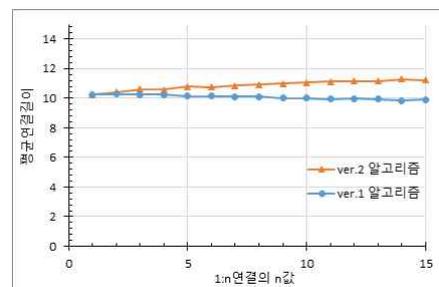


그림 5. 8x8 2차원 스위치 구조에서 1:n 연결의 평균경로길이
Fig. 5. Average path length of 1:n connection in 8x8 2-dimensional switch architecture

그림 6은 8x8 2차원 스위치 구조에서 1:n 연결을 위해 사용된 스위치의 비율을 보여주고 있다. 그림 6의 스위치 사용률과 앞에 나타내었던 그림 4의 연결 성공률 그림에서 확인할 수 있듯이 본 논문에서 제안된 알고리즘(ver.2)이 기존 알고리즘(ver.1)에 비해 연결성공확률은 높으면서도 실제 사용된 스위치의 비율은 더 작아 리소스 사용 효율이 더 높음을 알 수 있다.

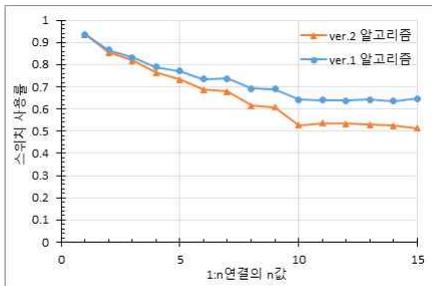


그림 6. 8x8 2차원 스위치 구조에서 1:n 연결의 스위치 사용률

Fig. 6. Used switch ratio of 1:n connection in 8x8 2-dimensional switch architecture

결론적으로 기존 알고리즘은 마스터 호스트에서 각 슬레이브 호스트까지의 거리를 최단 거리로 연결하도록 알고리즘이 동작하며, 본 논문에서 제안하는 새로운 알고리즘은 마스터 호스트에서 각 슬레이브 호스트까지의 연결경로가 가능한 한 중복되면서 최단 거리로 연결되도록 동작한다. 따라서 실험 결과에서도 확인할 수 있듯이 새로운 알고리즘은 기존 알고리즘에 비해 평균연결길이는 증가하되 연결성공확률은 커짐을 알 수 있다. 일반적으로 광케이블로 연결되어 초고속 전송대역을 제공하는 광 스위치 시스템의 경우 연결길이에 따른 전송지연은 상대적으로 다른 요인들에 비해 무시할 수 있으므로, 유한한 고정크기의 스위칭 시스템에서 설정 가능한 연결개수를 증가시켜 연결성공 확률을 크게 하는 본 논문의 알고리즘이 기존의 알고리즘보다 바람직한 것으로 판단된다.

4. 결론

인터넷의 발달과 함께 인터넷을 통해 서버 컴퓨터와 저장장치 및 이들을 연결하는 클라우드 서비스를 제공하는 데이터 센터가 폭발적으로 증가해 왔다. 구글, 페이스북, 마이크로소프트와 같은 글로벌 대기업 들이 소유하고 있는 데이터 센터들은 초절전 에너지 시스템, 초고성능 컴퓨팅 및 저장장치, 그리고 이들을 초광대역 전송속도로 연결하는 광스위칭 네트워크 인프라를 갖추고 있으며, 미래의 다양한 응용 서비스들의 특성에 따른 트래픽 요구조건들을 동적으로 만족시키기 위해 지역적으로 산재해 있는 다양한 컴퓨팅 자원들을 동적으로 할당하고 연결 관리하는 방안에 대한 연구가 활발히 진행되고 있다.

본 논문에서는 이러한 데이터 센터 인프라의 핵심 역할을 수행하는 광스위칭 연결구조에서 스위칭 셀 동작을 제어하기 위한 알고리즘들을 제안하고, 이에 따른 성능을 시뮬레이션을 통해 비교 분석하였다. 성능 분석결과 본 논문에서 제안하는 알고리즘이 기존 알고리즘에 비해 연결설정 성공확률이 약 3~7% 정도 향상됨을 보였다.

참고 문헌

[1] Technavio, "Global Optical Switches Market 2018-2022", Technavio's ICT reports, Jan. 2018. <https://www.businesswire.com/news/home/20180129005685/en/Global-Optical-Switches-Market---Increasing-Data>

[2] Mordor Intelligence, "OPTICAL SWITCHES MARKET - GROWTH, TRENDS AND FORECAST (2019 - 2024)", 2018. <https://www.mordorintelligence.com/industry-reports/optical-switches-market>

- [3] Klaus-Tycho Foerster, Manya Ghobadi, Stefan Schmid, "Characterizing the algorithmic complexity of reconfigurable data center architectures", Proceedings of the 2018 Symposium on Architectures for Networking and Communications Systems, pp.89-96, July 2018. <https://doi.org/10.1145/3230718.3230722>,
- [4] Chong Liu, Maotong Xu, Suresh Subramaniam, "A Reconfigurable High-Performance Optical Data Center Architecture", GLOBECOM 2016 (IEEE Global Communications Conference), Washington DC, USA, Dec., 2016. <https://doi.org/10.1109/GLOCOM.2016.7841539>,
- [5] Andromachi Chatzieftheriou, Sergey Legtchenko, Hugh Williams, Antony Rowstron, "Larry: practical network reconfigurability in the data center", Proceedings of the 15th USENIX Conference on Networked Systems Design and Implementation, pp.141-156, April 2018. ISBN: 978-1-931971-43-0
- [6] Monia Ghobadi, Ratul Mahajan, Amar Phanishayee, Nikhil Devanur, Janardhan Kulkarni, Gireeja Ranade, Pierre-Alexandre Blanche, Houman Rastegarfar, Madeleine Glick, Daniel Kilper, "ProjecToR: Agile Reconfigurable Data Center Interconnect", Proceedings of the 2016 ACM SIGCOMM Conference, pp.216-229, Florianopolis, Brazil, Aug., 2016. <https://doi.org/10.1145/2934872.2934911>,
- [7] Nirmal Kumbhare, Cihan Tunc, Salim Hariri, Ivan Djordjevic, Ali Akoglu, Howard Jay Siegel, "Just In Time Architecture (JITA) for Dynamically Composable Data Centers", IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), Agadir, Morocco, Nov.29-Dec.2, 2016. <https://doi.org/10.1109/AICCSA.2016.7945778>
- [8] 기장근, 권기영, "스위치 연결구조 시뮬레이션 소프트웨어 개발", 한국소프트웨어감정평가학회 논문지, 14권 1호, pp.41-46, ISSN 2092-8114, Jun.. 2018. [http://www.i3.or.kr/html/paper/2018-1/\(6\)2018-1.pdf](http://www.i3.or.kr/html/paper/2018-1/(6)2018-1.pdf)
- [9] Mininet, <http://mininet.org/>, 2018.
- [10] Ryu, "Component-based software defined networking framework", <https://osrg.github.io/ryu/>, 2018.

저자 소개



기장근(Jang-Geun Ki)

1986.2 고려대학교 전자공학과 졸업
 1988.2 고려대학교 전자공학과 석사
 1992.2 고려대학교 전자공학과 박사
 2002.6-2003.6 Univ. of Arizona 방문교수
 2010.6-2011.8 Univ. of Arizona 방문교수
 2016.8-2017.8 Univ. of Arizona 방문교수
 1992.3-현재 : 공주대학교 공과대학 전기전자 제어공학부 교수
 <주관심분야> 통신프로토콜, 이동통신시스템



권기영(Kee-Young Kwon)

1981.2. 고려대 전자공학과 졸업
 1983.2. KAIST 전기및전자공학과 석사
 1988.2. KAIST 전기및전자공학과 박사
 1988.3.-1991.2 (주)삼성전자 기흥 반도체연구소 선임연구원
 1991.3.-현재 공주대학교 공과대학 전기전자 제어공학부 교수
 2000.3.-2001.2. Southern Methodist University 방문교수
 <주관심분야> 반도체, 광통신