

논문 2017-1-6

데이터센터 구성요소 동적 연결 구조 성능 분석

기장근*

Performance Analysis on the Dynamic Connection Structure in Data Centers

Jang-Geun Ki*

요 약

미래의 데이터 인프라 구조는 아주 복잡한 응용 서비스 프로그램들이 수행되는 수많은 컴퓨터들과 저장 장치들로 구성된다. 미래의 응용 서비스들은 비선형적 확장성, 상대적으로 예측하기 힘든 성장률 등의 특징을 가지며, 이러한 요구조건들의 동적인 변화는 결국 컴퓨터, 저장장치 같은 하드웨어 자원들의 효율적인 배치 및 융통성 있는 사용 방안 마련을 요구하게 된다. 본 연구에서는 광 크로스바 스위치 구조를 이용한 재구성 가능 데이터 센터 인프라 구조에 대한 방안에 대해 광스위치 구조를 제어하는 알고리즘을 개발하고 성능 분석을 수행 하였다.

Abstract

The data communication infrastructure in the future consists of hundreds of thousands of computers and storages running complex application service programs. Future applications are characterized by non-linear scalability and relatively unpredictable rapid growth. In order to meet the dynamic requirements of application services, the method for effective allocation and flexible usage of hardware resources such as computers and storages is needed. In this paper, for the reconfigurable flexible data center infrastructure, the optical switch control algorithm is developed and simulated for performance evaluation.

한글키워드 : 데이터센터, 광 크로스바 스위치, 제어 알고리즘, 성능분석

keywords : data center, optical crossbar switch, control algorithm, performance evaluation

1. 서 론

일반적으로 컴퓨터 제조업자, 소프트웨어 개발자 그리고 통신 서비스 사업자들은 특정한 한 종류의 응용 서비스를 제공하는 최적화된 방안을

지원하기 위해 많은 시간과 자원을 투자하고 있다. 그러나 이러한 기존의 방식은 동적으로 끊임 없이 변화되는 미래 응용 서비스들의 요구조건을 충족시키기 어렵고, 안정성 및 서비스 품질 저하 등의 문제점을 야기한다. 또한 고정된 종류의 부하를 지원하도록 최적화된 데이터 센터에 새로운 종류의 부하 요구가 추가될 경우 데이터 센터를 확장하는데 매우 많은 비용이 요구된다[1,2]. 따라서 차세대 데이터 센터의 경우 다양한 종류의

* 공주대학교 전기전자제어공학부
(email: kjg@kongju.ac.kr)

접수일자: 2016.5.22. 심사완료: 2016.6.7.

게재확정: 2016.6.18.

응용 서비스 타입과 자원들을 원하는 방식으로 동적으로 변경 또는 구성 가능한 구조를 갖도록 설계하는 것이 매우 중요하다.

본 논문에서는 고정 구조를 갖는 기존의 데이터 센터 단점을 극복하기 위해 동적으로 구성 요소들을 연결할 수 있도록 제안된 광스위치 구조 [3]를 제어하는 알고리즘을 개발하고 성능 분석을 수행 하였다.

2. 광스위치 연결 구조

(1) 1:1 연결 지원 구조

데이터 센터에서 특정 응용 서비스를 지원하기 위해 필요한 성능 요구조건(대역폭, 지연시간 등)을 만족시키기 위해 자원들의 동적 할당이 필요하며, Nirmal Kumbhare 등에 의해 제안된 광스위치 구조[3]는 그림 1과 같다.

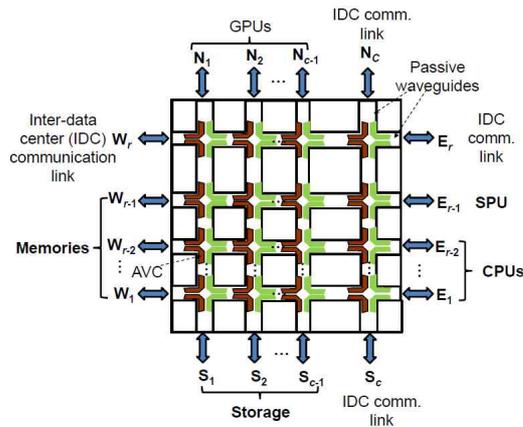


그림 1. Nirmal Kumbhare의 광스위치 구조[3]
Fig. 1. Optical switch of Nirmal Kumbhare

그림 1의 광스위치 구조에서 각 스위칭 셀의 동작 원리는 그림 2와 같다. 각 스위칭 셀은 4개의 포트를 가지며, 스위칭 셀 상태 0은 포트들이 연결되지 않은 상태를 의미하고, 상태 1은 포트1

과 2, 포트 3과 4가 연결된 상태를 의미한다. 상태 2에서는 그림에 나타난 것과 같이 포트 1-3, 포트 2-4가 각각 연결되고, 상태 3에서는 포트 1-4, 포트 2-3이 각각 연결된다.

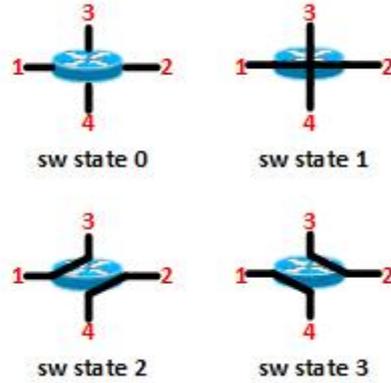


그림 2. 스위칭 셀 동작 모드
Fig. 2. Operation mode of switch cell

(2) 1:2 연결 지원 구조

그림 2에 나타난 동작 모드를 지원하는 스위칭 셀은 1:2 연결을 지원할 수 없다. 따라서 본 논문에서는 그림 3에 나타난 것과 같은 동작 모드를 지원하는 확장 스위칭 셀을 사용하는 구조를 제안하고, 이에 따른 스위치 제어 알고리즘 개발 및 성능 분석을 수행하였다.

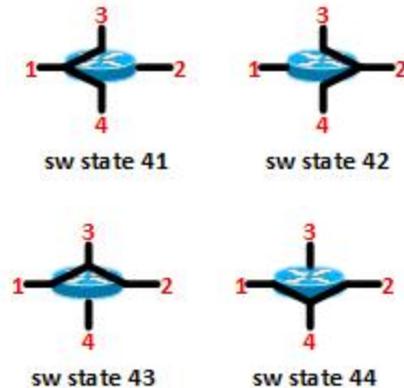


그림 3. 확장된 스위칭 셀 동작 모드
Fig. 3. Operation mode of extended switch cell

3. 스위치 제어 알고리즘 개발 및 성능 분석

(1) 1:1 연결 지원 알고리즘 및 성능분석

데이터 센터 구성 자원간의 1:1 연결을 효율적으로 지원하기 위해 자원간 최단경로를 계산하는 알고리즘을 개발하고 이에 따라 스위칭 셀들의 상태를 자동으로 제어하는 소프트웨어 프로그램을 개발하였다. 개발된 프로그램의 기능 검증 및 성능분석을 위해 최근 SDN(Software Defined Network) 연구에 많이 사용되고 있는 Mininet을 이용하였다.

Mininet은 가상 호스트들과, 스위치들 및 컨트롤러 그리고 이들을 연결하는 링크들을 제공하는 네트워크 에뮬레이터이다. Mininet의 스위치를 제어하는 컨트롤러 프로그램을 개발하기 위해 컴포넌트 기반의 SDN 프레임워크를 제공하는 Ryu 오픈 플로우 컨트롤러를 사용하였으며, Python 언어를 사용해 스위치 제어 프로그램을 개발하였다.

개발된 스위치 제어 프로그램의 기능을 검증하기 위해 Mininet에서 NxN 스위치들과 호스트들의 연결 구조를 자동으로 생성할 수 있는 프로그램을 개발하였다. 그림 4는 구성 예로 14x14 스위치들과 가장자리 스위치들에 연결된 호스트들의 연결 구조를 보여주고 있다.

그림 4의 14x14 스위치들은 이들의 동작을 제어하는 Ryu 컨트롤러와 연결되어 있으며, 본 논문에서 개발된 Ryu 컨트롤러 프로그램 코드에서 내보내는 명령어에 따라 동작모드가 결정된다.

1:1 연결에 대한 기능 검증 및 성능 분석을 위해 시뮬레이션한 결과의 한 예를 그림 5에 나타내었다. 그림 5는 그림 4의 스위치 네트워크 구조에서 임의로 2개의 호스트들을 차례로 선택한 후 이들을 최단경로로 연결하도록 스위치의 상태

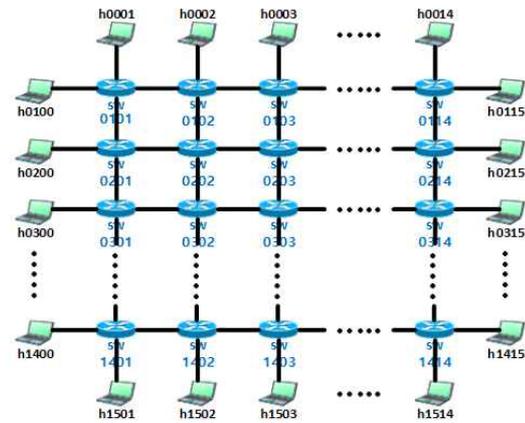
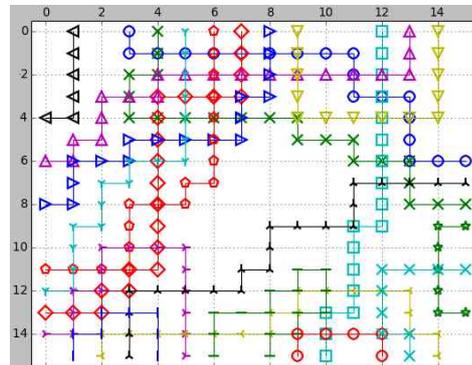


그림 4. 스위치 네트워크 구조
Fig. 4. Switch network architecture



연결성공율 =성공연결수/총연결수	0.643 =18/28
평균 연결 길이	15.167
스위치 미사용율 =스위치 미사용수/총수	0.148 =29/196

호스트 연결 리스트	
h0003 -- h0615	h1100 -- h0006
h0004 -- h0815	h1200 -- h0005
h0007 -- h1300	h1400 -- h1505
h0012 -- h1510	h1502 -- h1514
h0013 -- h0600	h1503 -- h0715
h0014 -- h0009	h1504 -- h1501
h0400 -- h0001	h1508 -- h1506
h0800 -- h0008	h1509 -- h1512
h0915 -- h1315	h1513 -- h1115

그림 5. 스위치 네트워크 1:1 연결 예
Fig. 5. Switch network 1:1 connection

를 조정하는 과정을 거쳐 얻어진 결과의 한 예를 보여준다.

표 1에는 14x14 스위치 구조에서 총 56개의 호스트들 중 2개씩 쌍을 지어 최대 몇 개까지의 경로를 설정할 수 있는지와 각 경로의 평균 길이, 사용되지 않은 스위치의 개수 등에 대한 시뮬레이션 평균 결과를 나타내었다. 1:1 연결의 경우 총 호스트 수의 절반인 28개까지의 연결이 성립될 수 있으며, 본 논문에서 개발된 최단거리 경로 설정 스위치 제어 알고리즘을 사용할 경우 평균 18,536개의 연결이 성공적으로 설정되고, 성공된 1:1 연결의 평균 길이는 호스트에서 호스트까지 16,215 홉을 가진다. 전체 14x14=196개의 스위치들중 1:1 연결에 사용되지 않은 스위치의 평균 개수는 약 17.724개로 스위치 미사용율은 $17.724/196 = 0.090$ 으로 나타났다.

표 1. 1:1 연결 시뮬레이션 결과
Table 1. Simulation result of 1:1 connection

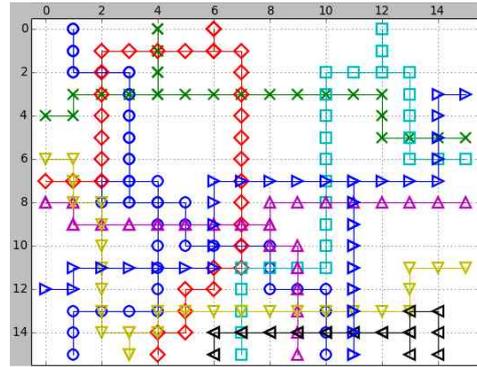
항목	결과 값
연결 성공 확률	0.662
연결 성공 개수	18,536
평균 연결 길이	16,215
미사용 스위치 평균 개수	17.724
스위치 미사용율	0.090

(2) 1:2 연결 지원 알고리즘 및 성능분석

1:2 연결 지원을 위해 그림 3에 나타내었던 확장된 스위칭 셀 동작모드를 지원하도록 프로그램된 시뮬레이션 결과의 한 예를 그림 6에 나타내었다. 1:2 연결의 경우 14x14 스위치 구조의 총 56개 호스트들간 성립될 수 있는 연결은 최대 $\text{int}(56/3)=18$ 개다.

표 2에는 1:2 연결 시뮬레이션 결과들의 평균을 정리하여 나타내었다. 표에서 알 수 있듯이

총 18개의 가능한 1:2 연결중 평균 8.18개의 연결이 성공(평균 성공확률 0.454)하였으며, 1:2 연결의 송신 호스트에서 각 수신 호스트로의 평균 경로 길이는 17 홉(hop)으로 나타났다.



연결성공율 =성공연결수/총연결수	0.444 =8/18
평균 연결 길이	17
스위치 미사용율 =스위치 미사용수/총수	0.230 =45/196

호스트 연결 리스트	
h0001 -- h1501	h0800 -- h0815
h0001 -- h1510	h0800 -- h1509
h0004 -- h0400	h1503 -- h0600
h0004 -- h0515	h1503 -- h1115
h0006 -- h0700	h1506 -- h1513
h0006 -- h1504	h1506 -- h1514
h0012 -- h0615	h1511 -- h0315
h0012 -- h1507	h1511 -- h1200

그림 6. 스위치 네트워크 1:2 연결 예
Fig. 6. Switch network of 1:2 connection

표 2. 1:2 연결 시뮬레이션 결과
Table 2. Simulation result of 1:2 connection

항목	결과 값
연결 성공 확률	0.454
연결 성공 개수	8.18
평균 연결 길이	16.762
미사용 스위치 평균 개수	39.472
스위치 미사용율	0.201

4. 결론

구글, 아마존, 이베이 등과 같은 데이터 인프라 구조는 아주 복잡한 소프트웨어 응용 프로그램들이 수행되는 수많은 컴퓨터들과 저장 장치들로 구성된다. 2013년부터 2020년 사이에 모바일, 소셜 네트워크, 빅 데이터 기술 분야에 대한 투자는 클라이언트/서버 기술에 대한 과거의 투자에 비해 20배 이상 성장할 것으로 예측되고 있다[5].

미래의 응용 서비스들을 살펴보면 비선형적 확장성, 상대적으로 예측하기 힘든 성장률 등의 특징을 가지며, 이러한 요구조건들의 동적인 변화는 결국 자원들의 효율적인 배치 및 융통성 있는 사용 방안 마련을 요구하게 된다.

따라서 융통성있는 빌딩 블록들을 부하 요구 조건에 따라 자동적이고 동적으로 조합해 사용 가능토록 하는 물리적인 데이터 센터 인프라 구조에 대한 연구가 필요하다.

본 연구에서는 광 크로스바 스위치 구조를 이용한 재구성 가능 데이터 센터 인프라 구조에 대한 방안에 대해 광스위치 구조를 제어하는 알고리즘을 개발하고 성능 분석을 수행 하였다.

앞으로 보다 융통성 있는 포트간 연결을 지원하는 확장된 광 스위치 셀 구조를 설계하고, 이에 따른 스위치 제어 알고리즘 개발 및 보다 다양한 성능분석을 수행할 예정이다.

참고 문헌

[1] M. Al-Fares, A. Loukissas, A. Vahdat, "A scalable, commodity data center network architecture", ACM Special Interest Group on Data Communication (SIGCOMM) Review, Vol.38, Issue 4, pp.63-74, Aug. 2008.
 [2] F. Fargo, C. Tunc, Y. Al-Nashif, A.

Akoglu, S. Hariri, "Autonomic workload and resources management of cloud computing services", International Conference on Cloud and Autonomic Computing (ICCAC 2014), pp.101-110, Sep. 2014.

[3] 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), Nov. 29, 2016.
 [4] Siamak Azodolmolky, Martin Nordal Peterseny, Anna Manolova Fagertuny, Philipp Wieder, Sarah Ren'ee Rueppy, Ramin Yahyapour, "SONEP: A Software-Defined Optical Network Emulation Platform", International Conference on ONDM (Optical Network Design and Modeling), Stockholm, Sweden, May 19-22, 2014.
 [5] IDC, "The 3rd Platform: Enabling Digital Transformation", Nov. 2013.

저 자 소 개



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
 2016.9-2017.8 Univ. of Arizona 방문교수
 1992.3-현재 : 공주대학교 공과대학 전기 전자제어공학부 교수
 <주관심분야>통신프로토콜, 이동통신시스템