

16 개의 호스트를 수용하는 아주 작은 것에서 11,664 개의 호스트를 수용하는 큰 것으로 증가시켜 가면서 생성하였다. 세트 1 과 세트 2 의 차이점은 각 호스트가 생성하는 플로우의 수를 세트 1 은 2 로 설정하고 세트 2 는 4 로 설정한 것이다. 세트 3 은 데이터 센터 네트워크의 규모는 8,192 의 호스트를 수용하는 것으로 고정하고, 각 호스트가 생성하는 플로우의 수를 1 에서 5 로 증가시켜 가면서 생성하였다. 세트 1, 2, 3 은 공통적으로 랙 내부 트래픽 비율을 50%로 고정하였다. 마지막으로 세트 4 는 나머지 인자들은 고정하고 랙 내부 트래픽 비율을 10 에서 90%로 증가시켜 가면서 생성하였다. 모든 데이터 세트들이 포함하고 있는 플로우들은 공통적으로 최대 링크 용량의 10~20% 사이에서 임의로 선정된 트래픽 요구량을 갖도록 생성되었다.

그림 5 는 최적 토폴로지 구성 단계를 통해 얻을 수 있는 전력 절감 비율을 나타낸다. 전력 절감 비율은 링크의 전력 비용은 1 이고 스위치의 전력 비용은 150 이라고 가정하여 계산했다 [7]. 그림 5 에서 “최소” 가 나타내는 것은 그림 2 를 통해 기술한 최소 부분 토폴로지 산출 알고리즘만을 적용하여 얻어진 부분 토폴로지를 그대로 활용했을 때의 전력 절감 비율이고, “A(i, j)” 가 나타내는 것은 산출된 최소 부분 토폴로지에 여유 용량을 제공하기 위해 팻(pod)당 i 개의 집선(aggregate) 스위치와 j 개의 코어(core) 스위치를 추가했을 때의 전력 절감 비율이다. 여기에서 k 값은 Fat-Tree 를 구성할 때 사용되는 인자로, Fat-Tree 는 k 개의 포트를 갖는 동일한 스위치들만으로 구성할 수 있다[1].

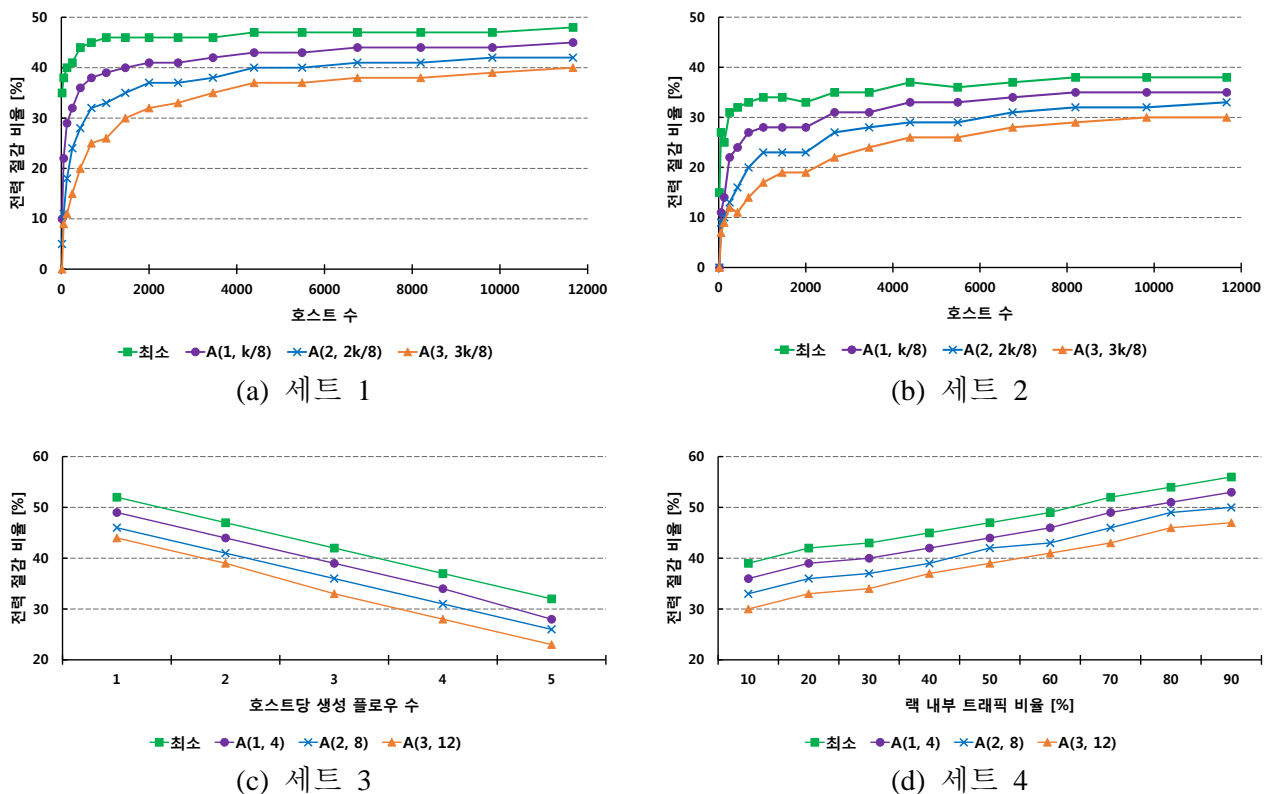


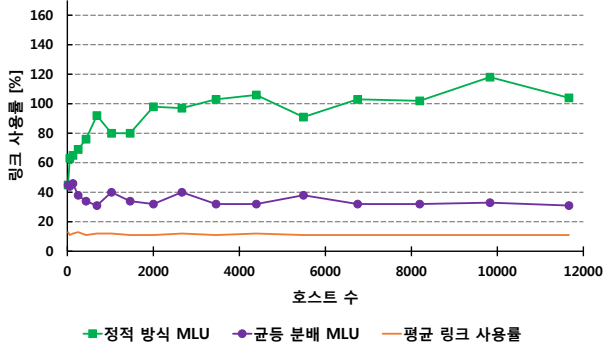
그림 5. 최적 토폴로지 구성 전력 절감 비율

그림 5a 와 b 를 통해 알 수 있는 것은 데이터 센터 네트워크의 규모가 커질수록 절감할 수 있는 전력의 비율이 높아지지만, 어느 정도 수준에 이르면 트래픽 량에 따라서 전력 절감 비율이 특정 값에 수렴한다는 것이다. 이 때, 호스트당 생성 플로우 수가 2 인 세트 1 보다 호스트당 생성 플로우 수가 4 인 세트 2 의 전력 절감 비율이 낮은 것을 알 수 있다. 이러한 특성은 그림 5c 를 통해 보다 명확히 확인할 수 있다. 그림 5c 는 8,192 개의 호스트를 수용하는 데이터 센터 네트워크에서 호스트당 생성 플로우의 수를 1 에서 5 까지 증가시켜가면서 실험한 것으로, 호스트당 생성 플로우의 수가 증가할수록 절감할 수 있는 전력 비율은 낮아지는 것을 알 수 있다.

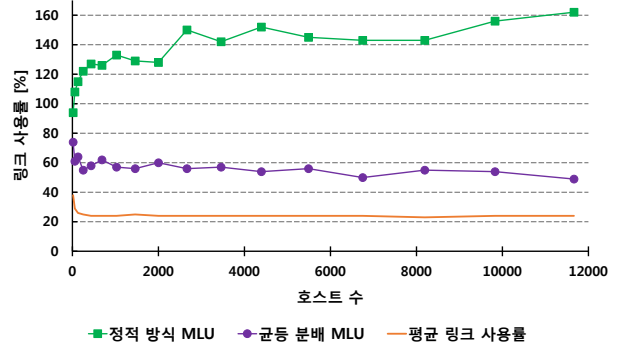
즉, 데이터 센터 네트워크에 존재하는 트래픽 량에 따라 데이터 센터 네트워크가 소모하는 전력을 적응적으로 절감할 수 있었다. 마지막으로, 그림 5d 는 랙 내부 트래픽 비율을 10 에서 90%까지 증가시켜 가면서 실험한 것으로, 랙 내부 트래픽 비율이 높아질수록 전력 절감 비율이 향상되는 것을 알 수 있다. 따라서, 데이터 센터 네트워크의 운용 비용을 절감하기 위해서는 랙 내부 트래픽 비율을 높이도록 데이터 센터 네트워크를 활용하는 어플리케이션들을 배치해야 한다.

네 개의 데이터 세트에 대한 결과를 종합하여 평균을 계산했을 때, 최소 부분 토폴로지를 활용하는 경우 전체 토폴로지를 활용하는 경우에 대비하여 41% 정도 전력을 절감할 수 있었으며, $A(1, k/8)$ 은 35%, $A(2, 2k/8)$ 은 31%, $A(3, 3k/8)$ 은 27%의 전력을 각각 절감할 수 있었다.

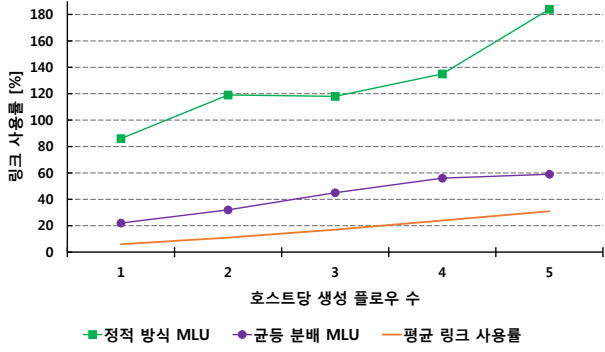
그림 6 은 제안한 트래픽 균등 분배 알고리즘을 데이터 센터 네트워크 전체 토폴로지에 적용했을 때의 최대 링크 사용률(MLU)을 정적인 경로 선정 방식을 적용했을 때의 최대 링크 사용률과 비교하고 있다. 실험 결과 네 개의 데이터 세트가 모두 비슷한 양상을 보였다. 즉, 정적인 경로 선정 방식을 적용했을 때는 최대 링크 사용률이 매우 높았으며 평균 링크 사용률과 매우 큰 편차를 보인 반면, 제안한 트래픽 균등 분배 방식을 적용한 경우, 최대 링크 사용률이 평균 링크 사용률과 비교하여 안정적인 편차를 보이며 낮게 유지되었다. 평균적으로 보았을 때 정적인 경로 선정 방식을 적용한 경우 최대 링크 사용률이 111%였고, 트래픽 균등 분배를 적용한 경우 44%로, 제안한 트래픽 균등 분배 방법을 적용하면 정적인 방식에 비하여 최대 링크 사용률을 60%까지 낮출 수 있었다.



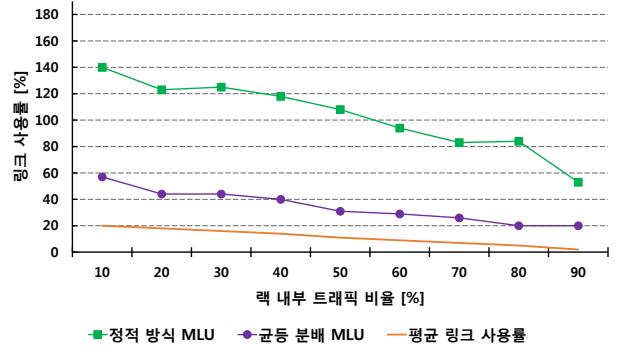
(a) 세트 1



(b) 세트 2

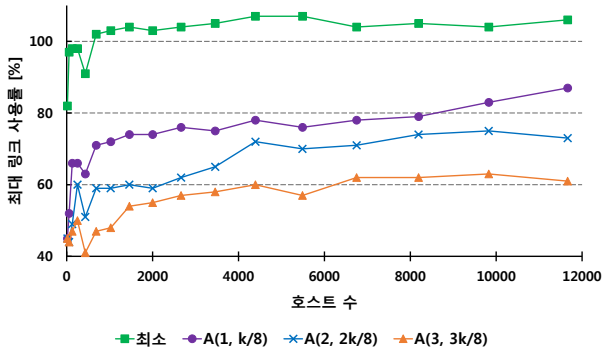


(c) 세트 3

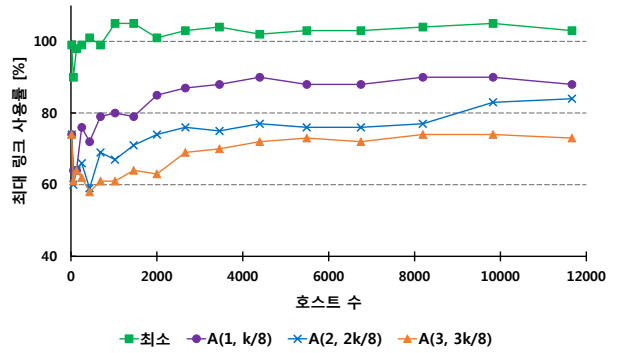


(d) 세트 4

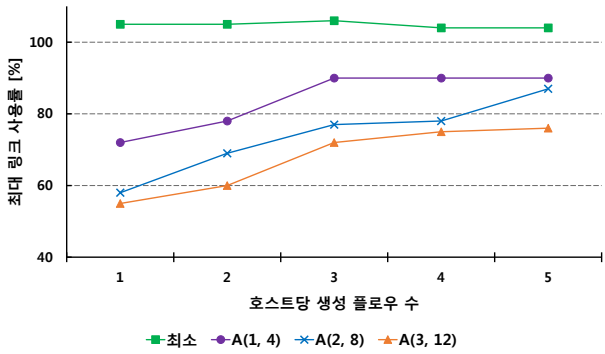
그림 6. 전체 토폴로지에 트래픽 균등 분배 적용



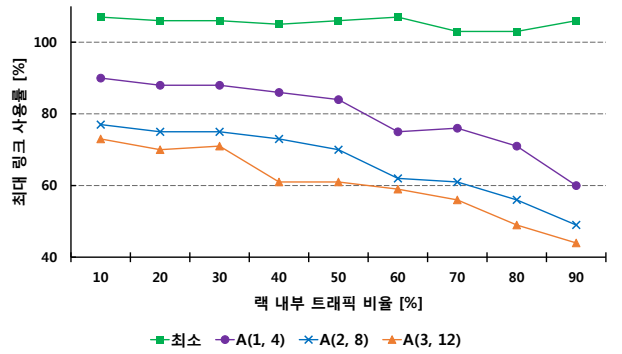
(a) 세트 1



(b) 세트 2



(c) 세트 3



(d) 세트 4

그림 7. 최적 토폴로지에 트래픽 균등 분배 적용

그림 7 은 제안한 트래픽 균등 분배 방식을 최적 토폴로지 단계를 통해 산출된 데이터 센터 네트워크 부분 토폴로지에 적용했을 때의 최대 링크 사용률을 나타내고 있다. 최소 부분 토폴로지 산출 알고리즘을 통해 얻어진 “최소”의 부분 토폴로지에 트래픽 균등 분배를 적용한 경우에는 대부분의 경우 최대 링크 사용률이 100%를 넘는 반면, 여분의 스위치와 링크를 추가한 경우 이러한 최대 링크 사용률을 혼잡이 발생하지 않는 수준으로 낮출 수 있었다. 최소의 부분 토폴로지에 트래픽 균등 분배를 적용했을 때 최대 링크 사용률이 100%를 넘는 것은 최소 토폴로지 산출(그림 2)과 트래픽 균등 분배(그림 3) 알고리즘의 플로우 할당 순서 및 할당 기준이 다르기 때문이다. 평균적으로 보았을 때, 최소의 부분 토폴로지를 활용하는 경우 전력 절감 비율은 41%였고, 최대 링크 사용률은 102%였다. 여기에 여분의 스위치를 추가한 A(1, k/8)의 전력 절감 비율은 35%, 최대 링크 사용률은 78%이며, A(2, 2k/8)의 전력 절감 비율은 31%, 최대 링크 사용률은 68%, A(3, 3k/8)의 전력 절감 비율은 27%, 최대 링크 사용률은 61%였다. 즉, 여분의 스위치를 추가할수록 전력 절감 비율과 최대 링크 사용률이 낮아졌다. 이러한 결과를 활용하여 어느 정도 수준으로 여분의 스위치와 링크를 추가할 것인지는 네트워크 운영자가 데이터 센터 네트워크의 운영상 특성을 고려하여 적절히 설정하여야 한다.

5. 결론 및 향후 과제

현재의 데이터 센터 네트워크는 비효율적인 전력 소모에 따른 높은 운영 비용과 정적인 경로 선정 방식에 의한 혼잡 발생 가능성이라는 문제를 가지고 있다. 본 논문에서는 이러한 데이터 센터 네트워크의 문제점을 해결하기 위하여 SDN 기술을 활용한 동적인 트래픽 엔지니어링 방법을 제안했다. 제안한 방법은 최적 토폴로지 구성과 트래픽 균등 분배의 두 단계로 구성되어 있다. 최적 토폴로지 구성 단계를 통해서 주어진 트래픽 요구량을 만족시킬 수 있는 최적의 부분 토폴로지를 찾아내고, 트래픽 균등 분배 단계를 통해서 찾아진 부분 토폴로지 내에서 트래픽을 균등하게 분배함으로써 데이터 센터 네트워크가 소모하는 전력을 절감하는 동시에 최대 링크 사용률도 낮추어 혼잡 발생 가능성도 줄일 수 있었다. 제안한 방식을 검증하기 위하여 프로토타입을 구현하였고, 시뮬레이션 실험 결과도 제시하였다.

본 논문에서 제안한 알고리즘들은 입력으로 예측된 트래픽 매트릭스를 요구한다. 이러한 트래픽 매트릭스를 예측하는 연구들은 기존 인터넷 서비스 제공자들을 위한 백본 네트워크에 대해서는 많이 이루어졌지만, 데이터 센터 네트워크의 트래픽 매트릭스를 예측하는 것은 초기 단계이다. 따라서, 데이터 센터 네트워크의 트래픽 매트릭스를 정확히 예측하는 것은 중요한 향후 과제 중 하나이다. 또한, 제안한 방식은 플로우들의 경로를 동적으로 변경하게 되는데, 이러한 동적인 경로 변경이 플로우들의 성능에 미치는 영향을 측정하고 평가하는 과정도 중요한 향후 연구 주제이다.

6. 참고 문헌

- [1] M. Al-Fares, A. Loukissas, and A. Vahdat, “A scalable, commodity data center network architecture,” in *Proc. ACM SIGCOMM '08*, (Seattle, USA), pp. 63-74, Aug. 17-22, 2008.
- [2] A. Greenberg, J. R. Hamilton, N. Jain, S. Kandular, C. Kim, P. Lahiri, D. A. Maltz, P. Patel, and S. Sengupta, “VL2: A scalable and flexible data center network,” in *Proc. ACM SIGCOMM '09*, (Barcelona, Spain), pp. 51-62, Aug. 17-21, 2009.
- [3] R. N. Mysore, A. Pamboris, N. Farrington, N. Huang, P. Miri, S. Radhakrishnan, V. Subramanya, and A. Vahdat, “PortLand: A scalable fault-tolerant layer 2 data center network fabric,” in *Proc. ACM SIGCOMM '09*, (Barcelona, Spain), pp. 39-50, Aug. 17-21, 2009.
- [4] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu, “DCell: A scalable and fault-tolerant network structure for data centers,” in *Proc. ACM SIGCOMM '08*, (Seattle, USA), pp. 75-86, Aug. 17-22, 2008.
- [5] C. Guo, G. Lu, D. Li, H. Wu, X. Zhang, Y. Shi, C. Tian, Y. Zhang, and S. Lu, “BCube: A high performance, server-centric network architecture for modular data centers,” in *Proc. ACM SIGCOMM '09*, (Barcelona, Spain), pp. 63-74, Aug. 17-21, 2009.
- [6] A. Singla, C.-Y. Hong, L. Popa, and P. B. Godfrey, “Jellyfish: Networking data centers randomly,” in *Proc.*

- 9th *USENIX Symposium on Networked Systems Design and Implementation (NSDI '12)*, (San Jose, USA), pp. 1-14, Apr. 25-27, 2012.
- [7] B. Heller, S. Seetharaman, P. Mahadevan, Y. Yiakoumis, P. Sharma, S. Banerjee, and N. McKeown, "ElasticTree: Saving energy in data center networks," in *Proc. 7th USENIX Symposium on Networked Systems Design and Implementation (NSDI '10)*, (San Jose, USA), pp. 1-16, Apr. 28-30, 2010.
- [8] T. Benson, A. Akella, and D. A. Maltz, "Network traffic characteristics of data centers in the wild," in *Proc. ACM Internet Measurement Conference 2010 (IMC '10)*, (Melbourne, Australia), pp. 267-280, Nov. 1-3, 2010.
- [9] "Software-defined networking (SDN) definition." <https://www.opennetworking.org/sdn-resources/sdn-definition>. [Online; accessed Oct. 21, 2013].
- [10] N. Wang, K. H. Ho, G. Pavlou, and M. Howarth, "An overview of routing optimization for Internet traffic engineering," *IEEE Communications Surveys and Tutorials*, vol. 10, no. 1, pp. 36-56, 2008.
- [11] "Mininet: An instant virtual network on your laptop (or other pc)." <http://mininet.org/>. [Online; accessed Nov. 27, 2013].
- [12] "Open vSwitch: An open virtual switch." <http://openvswitch.org/>. [Online; accessed Dec. 2, 2013].
- [13] "Floodlight OpenFlow controller." <http://www.projectfloodlight.org/floodlight/>. [Online; accessed Oct. 23, 2013].
- [14] M. Al-Fares, S. Radhakrishnan, B. Raghavan, N. Huang, and A. Vahdat, "Hedera: Dynamic flow scheduling for data center networks," in *Proc. 7th USENIX Symposium on Networked Systems Design and Implementation (NSDI '10)*, (San Jose, USA), pp. 1-15, Apr. 28-30, 2010.
- [15] T. Benson, A. Anand, A. Akella, and M. Zhang, "MicroTE: Fine grained traffic engineering for data centers," in *Proc. 7th International Conference on emerging Networking EXperiments and Technologies (CoNEXT '11)*, (Tokyo, Japan), pp. 1-12, Dec. 6-9, 2011.
- [16] F. P. Tso and D. P. Pezaros, "Improving data center network utilization using near-optimal traffic engineering," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, pp. 1139-1148, June 2013.
- [17] C. Hopps, "Analysis of an Equal-Cost Multi-Path algorithm." RFC 2992, Nov. 2000.
- [18] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "OpenFlow: Enabling innovation in campus networks," *ACM SIGCOMM Computer Communications Review*, vol. 38, no. 2, pp. 69-119, 2008.



서 신 석

2008 인하대학교, 컴퓨터공학과 학사

2014 포항공과대학교, 컴퓨터공학과 박사

<관심분야> SDN, 데이터 센터 네트워킹, 상황정보 관리, 자율 컴퓨팅



리 건

2007 연변과학기술대학교, 전자통신공학과 학사

2012 포항공과대학교, 컴퓨터공학과 석사

2012 ~ 현재 포항공과대학교, 정보전자융합공학부 박사과정

<관심분야> SDN, OpenFlow, 클라우드 컴퓨팅, 모바일 장치 관리, 스마트 그리드



현 중 환

2011 포항공과대학교, 컴퓨터공학과 학사

2011 ~ 현재 포항공과대학교, 컴퓨터공학과 통합과정

<관심분야> 네트워크 트래픽 모니터링, SDN



유 재 형

1983 연세대학교 전자공학과 학사

1985 연세대학교 전자공학과 석사

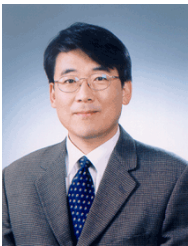
1999 연세대학교 컴퓨터공학과 박사

1986 ~ 2012: KT 네트워크 연구소

2012 ~ 2013: KAIST 전기 및 전자공학과 연구부교수

2013 ~ 현재: POSTECH 컴퓨터공학과 연구부교수

<관심분야> 네트워크 관리 및 보안, SDN, OpenFlow



홍 원 기

1983 Univ. of Western Ontario, BSc in Computer Science

1985 Univ. of Western Ontario, MS in Computer Science

1985 ~ 1986 Univ. of Western Ontario, Lecturer

1986 ~ 1991 Univ. of Waterloo, PhD in Computer Science

1991 ~ 1992 Univ. of Waterloo, Post-Doc Fellow

1992 ~ 1995 Univ. of Western Ontario, 연구교수

1995 ~ 현재 포항공과대학교, 컴퓨터공학과 교수

2007 ~ 2011 포항공과대학교, 정보통신대학원장

2007 ~ 2010 포항공과대학교, 정보통신연구소 연구소장

2008 ~ 2010 포항공과대학교, 컴퓨터공학과 주임교수

2008 ~ 2012 포항공과대학교, 정보전자융합공학부장

2012 ~ 2014 KT 종합기술원, 원장

2008 ~ 현재 포항공과대학교, 컴퓨터공학과 교수

<관심분야> 네트워크 트래픽 모니터링, 네트워크 및 시스템 관리



이 영 우

1984 숭실대학교, 전자공학과 학사
1990 숭실대학교, 전자공학과 석사
2005 충남대학교, 컴퓨터공학과 박사
1990~현재 kt 종합기술원 Infra연구소 상무

<관심분야> Transport SDN, All IP Networking, Network Service Management



황 찬 규

1984 경북대학교 전자공학과 학사
1986 경북대학교 전자공학과 석사
1986~현재 kt 종합기술원 Infra 연구소 팀장

<관심분야> Transport SDN, All IP Networking, Network Service Management, Network Traffic Engineering



백 성 복

1990 강원대학교 전자계산학과 학사
1992 숭실대학교 전자계산학과 석사
2004 Illinois Institute of Technology 전자계산학과 박사
1995~현재 kt 종합기술원 Infra 연구소

<관심분야> SDN, Network Service Management, Network Traffic Engineering