

차세대 데이터센터를 위한 광 스위칭 기반 네트워크 기술 동향

Trends in Optical Switching Based Network Technologies for
Next-Generation Data Centers

고재수 [J.S. Ko, jsko@etri.re.kr]
김광준 [K.J. Kim, kjk@etri.re.kr]
이준기 [J.K. Lee, juneki@etri.re.kr]

광네트워크연구그룹 책임연구원
광네트워크연구그룹 책임연구원/PL
광네트워크연구그룹 책임연구원/그룹장

To accommodate the rapid increase in the amount of data traffic, the capacity of datacenters is expanding rapidly. Datacenter networks (DCNs) utilize electronic packet switches. However, the increases in the speed and capacity of electronic devices are slower than that of data expansion. Furthermore, electronics are too well developed to reach very near their physical limits. To achieve a breakthrough under this situation, optical switching schemes have been widely examined to replace or collaborate with incumbent electronic switches in a DCN. This report reviews the current status of such global researches on an optical DCN.

* DOI: 10.22648/ETRI.2018.J.330109

* 본 연구는 2017년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임[2016-0-00573, Photonic Frame 기반 패킷 스위칭 가능한 데이터센터 광 네트워킹 핵심 기술 개발].



본 저작물은 공공누리 제4유형
출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.

2018
Electronics and
Telecommunications
Trends

4차 산업혁명 사회의 초연결
지능과 신뢰 인터넷 기술 특집

- I. 서론
- II. 광 데이터센터 네트워크
기술 동향
- III. 광 패킷 스위칭 기술
- IV. 국내 연구 동향
- V. 결론

I. 서론

전 세계 데이터센터 트래픽은 2015년도에 4.7ZB에 달하였으며, 2020년도에는 약 3배 성장하여 15.3ZB에 이를 것으로 보고 있다. 이 중 데이터센터 내에서 저장, 생성, 처리되는 내부 트래픽이 77%에 달한다[1]. 데이터센터의 ICT 설비는 컴퓨팅을 위한 서버, 데이터 저장을 위한 스토리지, 연결 네트워크로 이루어지는데, 데이터센터의 성능에는 컴퓨팅 용량과 내부 연결 네트워크(DCN: Data Center Network)가 중요한 역할을 한다.

현재 데이터센터 내의 네트워크 구성 방식은 대부분 전자식 패킷 스위칭 방식의 네트워크로 구성되어 있다. 지난 수십 년간 눈부시게 발전해 온 전자 기반 기술은 여전히 성능이 향상되고 있으나 관련 소자 성능이 물리적 한계에 근접하고 있다. 따라서 계속 전자 기반 스위칭 기술에만 의존해 DCN 용량을 확장해 나가는 것은 설치 및 운용 비용 증가, 에너지 소비 증가, 데이터 전달 지연 증가 등의 문제점을 근본적으로 해결하기 어렵다[2]. 이러한 상황에 대한 대안으로 광기반 스위칭 방식을 DCN에 도입하는 방안이 제안되고 있다[3].

광 스위칭 방식에는 광 회선 스위칭(OCS)과 광 패킷 스위칭(OPS) 방식이 있으며, 각각 스위칭 용량 및 스위칭 시간에 있어서 장·단점을 가지고 있다. 광 DCN 기술 발전에 대해서는 초기에는 전자 및 광 스위칭 방식의 네트워크가 혼재된 전자/광 하이브리드 네트워크 구조로 도입이 진행되어 궁극적으로는 전광 스위치 기반의 광 네트워크로 전환될 것으로 보는 시각이 일반적이다[4]. 전자/광 하이브리드 구조에서는 작고 실시간성이 요구되는 패킷 데이터 트래픽은 기존의 이더넷 스위치와 같은 전자식 스위칭 기반 네트워크를 거치고 크고 실시간성이 강조되지 않는 패킷 트래픽은 광 스위칭 기반 네트워크를 거치도록 하는 것이다. 전광 스위칭 기반 네트워크의 실현을 위해서는 데이터센터 패킷 트래픽을 성능

저하 없이 처리할 수 있고 확장 가능한 DCN 구조와 관련 기술 확보가 필요하다. 미래의 데이터센터 네트워크 적용을 목적으로 광 데이터센터 네트워크(광 DCN) 인프라와 핵심 기술을 확보하기 위한 관련 연구가 전 세계적으로 활발하게 이루어지고 있다[5].

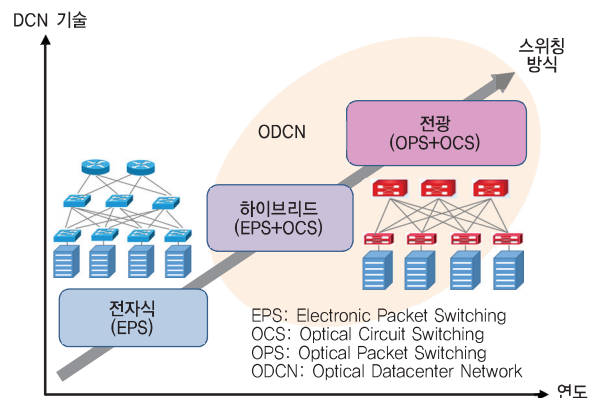
본고에서는 차세대 데이터센터 적용을 위하여 연구되고 있는 광 DCN 기술들과 관련 연구 동향을 소개하고, 이를 통해 개발되고 있는 여러 가지 광 스위칭 기술에 대해 알아본다.

II. 광 데이터센터 네트워크 기술 동향

1. 광 DCN 구조 연구 동향

광 DCN에 대해서는 다양한 구조가 연구되고 있으며 크게 분류하면 전자식 스위칭과 광 스위칭을 혼용하여 사용하도록 하는 전자/광 하이브리드(EPS+OCS) 구조와 광 스위칭만을 사용하는 전광(OPS+OCS) 구조가 있다. 이들은 (그림 1)과 같이 발전되어 갈 것으로 보인다.

지금까지 제안된 대표적인 구조들이 <표 1>에 정리되어 있다[2], [3]. 전자/광 하이브리드 구조는 Helios, c-Through, WDM-PON, OSA, HOS, FISSION 등에서 제안하고 있다. Helios에 사용되는 OCS 스위치는 MEMS 기술을 사용하며 WDM을 지원한다. MEMS에서 소비되는 전력은 EPS 스위치에 비해 포트당 약 50배나



(그림 1) DCN 기술의 발전 추세

〈표 1〉 제안된 광 DCN 기술들

제안 기술	구조	주요 특징
Helios[6]	Hybrid	- MEMS 기반의 OCS와 기존 EPS를 혼용한 구조
c-Through[7]	Hybrid	- HyPac 구성에 의한 OCS와 EPS 혼용 구조
WDM-PON[8]	Hybrid	- PON 구조와 이더넷 스위칭을 결합
OSA[9]	Hybrid	- MEMS/WSS 기반의 동적 재구성 가능한 구조
HOS[10]	Hybrid	- SOA와 MEMS를 이용한 하이브리드 구조
FISSION[11]	Hybrid	- 광 버스 백플레인과 EPS 스위치를 결합
Space-WL[12]	전광	- 스케줄러와 SOA 기반 공간-파장 스위치 구조
STIA[13]	전광	- 새로운 공간-시간 상호연결 구조
LIONS[14]	전광	- 수동 AWGR 기반 스위치, 패킷 충돌 회피 AO-NACK 기술 사용
TONAK LION[15]	전광	- 분산 AO-TOKEN과 AO-NACK 구조 결합
MIMO-OFDM [16]	전광	- CAWG 소자, MIMO 및 PSD 기술 기반
Data Vortex [17]	전광	- SOA 기반의 패킷/회선 트래픽 스위칭 - 제어/클럭/페이로드용 WDM 파장 사용
Petabit[18]	전광	- AWG/TWC 기반 3단 Clos 구조, 프레임 기반 스위칭
All-to-All[19]	전광	- WR 기반의 채널 본딩 기술 이용
LIGHTNESS [20]	전광	- 전광 하이브리드 OPS/OCS 구조
COSIGN[21]	전광	- SDN과 전광 기술 기반 TDM/WDM 다중 구조
NEPHELE[22]	전광	- slotted TDMA와 DWDM 스위칭 기술의 결합
HOPR[23]	전광	- 광 라벨 처리기, 광 스위치, 공유 버퍼 사용

절감되는 것으로 확인되었다[6]. c-Through는 HyPac 구성 방법을 도입하여, 하위에 랙 간 고속 통신을 위해 OCS를 구성하고, 상위에 EPS를 구성한다[7]. WDM-PON은 랙 간 통신을 위한 수동 광 네트워크(PON) 구조와 랙 내부 통신을 위한 전통적인 이더넷 스위칭(EPS)을 결합한 하이브리드 솔루션을 이용하였다[8]. OSA는 다른 하이브리드 구조와 달리 ToR 외에 전자식 콤포넌

트를 사용하지 않는 MEMS/WSS 기반의 동적 재구성이 가능한 구조이다[9]. HOS는 기존의 전자식 애그리게이션 스위치와 코어 스위치를 HOS 엣지와 코어 노드로 대체한 것으로 에지 노드는 EPS를 사용하고 코어 노드는 MEMS와 SOA를 사용하여 회선 및 패킷 스위칭이 가능하도록 구성하였다[10]. FISSION은 광 버스 백플레인과 EPS 스위치를 결합하여 수백만 서버 연결을 지원할 수 있도록 하였다[11].

전광 스위칭 기반의 광 DCN 기술의 주요 특징은 다음과 같다. Space-WL은 새로운 공간-파장 스위치 구조를 [12], STIA는 새로운 공간-시간 상호연결 구조를 제안하였으며[13], LIONS는 고성능 데이터센터에 위한 수동 AWGR 기반 저지연 상호연결 광 네트워크 스위치를 제안하여 파장 할당 방식의 광라벨 스위칭을 실현하였다[14]. TONAK LION은 LIONS 스위치의 개선 버전으로 분산 전광 토큰 및 전광 NACK 구조의 AWGR 기반 TONACK 스위치를 제안하여 확장성 한계를 극복하였다[15]. MIMO-OFDM은 CAWG 소자, MIMO 및 PSD 기술에 기반한 새로운 DCN 구조를 도입하였다[16]. Data Vortex는 고성능 컴퓨팅 및 데이터센터 연결을 위한 분산 상호연결 네트워크 구조를 제안하였으며[17], Petabit는 3단 Clos 네트워크 기반의 버퍼 없는 광 스위치 구조를 제안하여 프레임 기반 스위칭을 실현하였다 [18]. All-to-All은 AWGR과 FTL에 의한 채널 본딩 기반 파장 라우팅(WR) 기술을 이용한 구조를 제안하였다[19].

또한, LIGHTNESS는 SDN 제어와 결합되어 데이터 센터내 연결 서비스를 제공하는 전광 하이브리드 OPS/OCS 구조를 제안하였으며[20], COSIGN은 SDN과 전광 기술 기반 TDM/WDM 다중 구조를 제안하였다[21]. NEPHELE는 SDN 기술로 제어되는 이더넷 광 TDMA 데이터센터 네트워크 구조를 채택하였다[22]. 그리고 HOPR은 광 라벨 처리기, 광 스위치, 공유 버퍼를 사용하여 광 DCN을 구성하였다[23].

〈표 2〉 광 DCN을 구성하는 요소 기술

요소 기술	주요 기능/특성	관련 부품/소자
OCS 스위치	크고 긴 패킷 데이터 스위칭	WSS, MEMS, AWG
OPS 스위치	빠르고 짧은 패킷 데이터 스위칭	SOA, TWC, AWGR, AVC, PLZT, FDL
광 ToR	데이터 전송, 광 송수신	FTL, BMR, FPGA
스위칭 제어	경로 설정 및 자원 할당	

광 DCN을 구성하는 요소 기술은 OCS 스위치, OPS 스위치, 광 ToR 스위치, 광 스위칭 제어 기술 등이 있으며, 주요 부품 및 소자로는 TWC, AWG, MEMS, WSS, FTL, BMR 등이 있다. 이를 〈표 2〉에 정리하였다.

2. 주요 광 DCN 기술 연구 프로젝트

본 절에서는 데이터센터 내 적용을 위한 광 네트워크 기술 관련 연구 프로젝트와 관련 기술에 대해 살펴본다. 데이터센터 광 네트워크 관련한 주요 연구 프로젝트로는 유럽의 FP7 및 H2020 프로그램을 통한 몇몇 프로젝트와 일본 NICT 프로그램을 통한 HOPR이 있다. 〈표

〈표 3〉 광 DCN 기술 연구 프로젝트

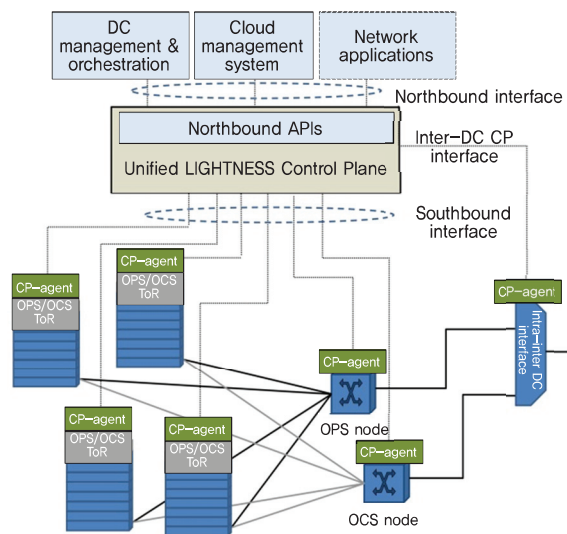
항 목	LIGHTNESS	COSIGN	NEPHELE	HOPR
프로그램	F7	F7	H2020	NICT
연구목표	고성능 DCN 인프라 개발	새로운 DCN 구조 정의/구현	미래 DCN 인프라 개발	HOPR 기반 기술 개발
네트워크 구조	Flattened DCN	Flattened/WDM 링	WDM 링 (Multi-plane)	Flat Torus DCN
스위칭 구성	Hybrid OCS/OPS	Multi SDM/TDM	Hybrid OCS/OPS	Single Platform
제어평면	SDN 기반	SDN 기반	SDN 기반	SDN 기반
용량 (패킷/회선)	4×4/ 100×100	4×4/ 192×192	4×4/ 80×80	8×8
주요기술	광 ToR, FPGA-NIC, OPS 스위치	광 ToR, FPGA-SIC, OXS 스위치	POD/ToR 스위치, TDMA 스케줄러	라벨 처리, B&S 스위치, 공유버퍼
연구기간	'12.11~ '15.10	'14.1~ '16.12	'15.2~ '18.1	'11.2~ '16.3
참여기관	UoB 등 6개	DTU 등 13개	NTUA 등 7개	NTT 등 4개

3)은 광 DCN 관련 연구 프로젝트들을 요약하였다.

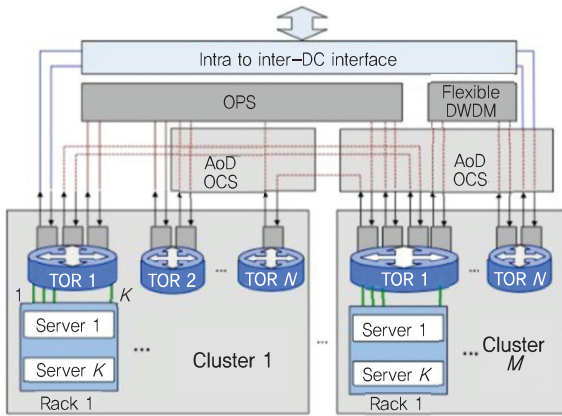
유럽에서는 LIGHTNESS[20]와 COSIGN[21] 프로젝트가 F7 프로그램에서 수행되었고, 현재 H2020 프로그램에서 NEPHELE[22] 프로젝트가 진행되고 있다. LIGHTNESS는 데이터센터용 인프라를 제공하는 것을 목표로 Hybrid OPS/OCS DCN 구조를 제안하였다[20]. COSIGN은 최신 광 기술을 최대한 사용하여 미래 광 DCN에 적용할 수 있는 솔루션을 제안하였다[24]. NEPHELE는 slotted TDMA와 WDM 기술을 이용하여 확장성이 큰 데이터센터 내 광 네트워크 구조를 제안하여 검증 중에 있다[22], [25].

가. LIGHTNESS

LIGHTNESS의 DCN 구조는 (그림 2)와 같이 기존 이더넷 스위치 기반 DCN의 평탄(Flattened) 구조에 이더넷 스위치 대신에 OPS와 OCS를 적용하여 구성한다. OPS의 경우 SOA와 라벨 처리기를 사용하여 4×4 패킷 스위칭을 실현하였으며, OCS의 경우 Polatis 파이버 스위치를 사용하여 100×100회선 스위칭 용량을 실현하였다. 100G 용량을 처리하는 FPGA 기반 광 NIC는 랙 간 통신용으로 10×10G DWDM CFP와 랙내 통신용으로



(그림 2) LIGHTNESS DCN 구조[20]



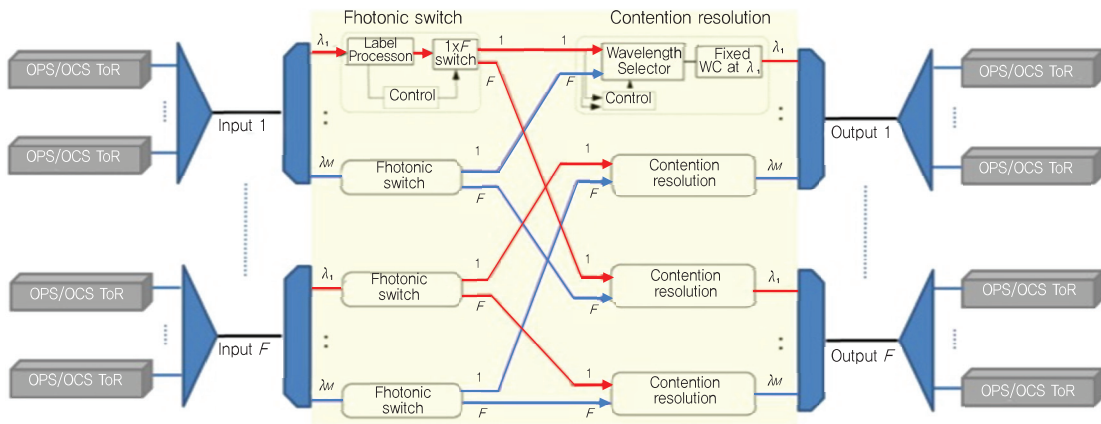
(그림 3) AoD OCS를 사용한 연결 구조[20]

24x10G VCSEL MiniPod 인터페이스를 구비하고 있으며, 트래픽 전달은 데이터의 특성에 따라 구분하여 다른 스위칭 경로로 전달하도록 한다. 즉, 용량이 크고 실시간성을 요구하지 않는 데이터 트래픽은 OCS로 전달하고 용량이 작으며 실시간성이 중요한 트래픽은 OPS로 전달한다. 랙 간 연결을 위한 광 ToR 스위치는 WSS 또는 AWG를 사용하며, 포트당 100Gbps DP-QPSK 데이터 25개 파장을 지원할 수 있다. WSS 광 ToR의 경우 가변 파장, AWG 광 ToR의 경우 고정 파장을 사용한다. 사용된 NIC, OPS, OCS, ToR 스위치는 중앙 집중형의 SDN 기반 통합 제어 평면에 의해 제어된다[26].

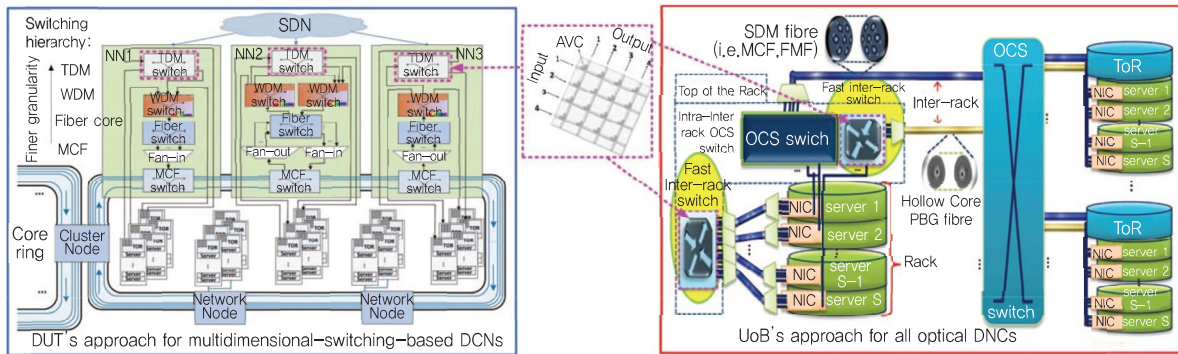
LIGHTNESS에서 ToR과 연결되어 DC 내 및 DC 간

스위칭을 수행하는 OCS 노드는 (그림 3)과 같이 AoD 구조를 적용하였다. AoD 구조는 3D MEMS 스위치와 여러 광 신호처리 모듈들(SSS, OPS 스위치, EDFA 등)이 연결되는 광 백플레인과 노드 입출력으로 구성된다. AoD 구조에서는 여러 광 모듈들 간 연결이 고정되어 있지 않아 수요에 따라 유연하게 조정할 수 있는 장점이 있다[20].

LIGHTNESS에서 OPS 스위치는 (그림 4)와 같은 구조를 갖는다[20]. 입출력 포트는 N 개($N=FXM$, 즉 F 개 파이버가 각각 M 개의 파장을 가짐)의 포트 수를 가진다. 각 파장은 각 ToR 스위치에 할당된다. 포토닉 스위치는 라벨 처리기, SOA 기반의 B&S $1x F$ 스위치, 스위치 컨트롤러로 구성된다. 충돌회피 블록은 $Fx1$ 파장선택기(WS)와 파장 변환기(WC)로 구성되어 목적지 ToR 스위치에 도달하는 출력 파이버에 동일한 파장이 두 개 이상 할당되지 않도록 한다. 스위칭 정보는 인밴드 RF 톤 라벨링 기술을 사용하여 처리시간을 줄이고 확장성을 개선하였다. 클라이언트 측의 라벨 발생부에서는 M 개의 이진 부호화된 RF 톤을 페이로드의 스펙트럼 인밴드에 삽입하여 M 개의 파장에 실어 보낸다. 이 방식은 MxN 라벨 비트와 2^{MxN} 포트로 확장시킬 수 있고, 라벨이 병렬로 처리되어 톤의 수에 관계없이 처리 시간이 결정되는 장점이 있다[26].



(그림 4) LIGHTNESS의 OPS 스위치 구조[20]



(그림 5) COSIGN DCN 구조[24]

나. COSIGN

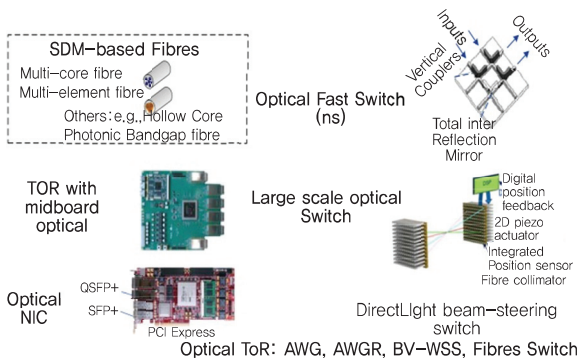
COSIGN 프로젝트를 통해 덴마크 공대(Technical University of Denmark)에서 제안한 WDM 링 구조와 영국 브리스톨대학교(UoB)에서 제안한 평탄 구조가 (그림 5)에 나와 있다. WDM 링 구조의 경우, 스위칭 방식은 다차원 SDM/TDM 스위칭으로 스위칭 계위별로 이중의 MCF/Fiber/WDM/TDM 스위치를 적용하여 적절한 단위로 스위칭을 하도록 하였다. SDM 기반 파이버는 MCF, MEF, HCPBF 등이 있다. 특히 NIC 대신에 FPGA 기반 SIC를 사용하여 기능별(CPU, 메모리, 스토리지)로 모아 놓은 컴퓨터 블레이드(Disaggregated computer blade)를 광 스위치와 효율적으로 연결할 수 있도록 하였다. 평탄 구조에서는 풀 메시로 WSS/AWG 기반의 광 ToR 스위치를 사용하여 다른 ToR과 직접 연

결하였다. OCS 용 스위치는 MEMS 스위치에 비해 손실이 적은 192x192 Polatis 파이버 스위치[27]를 사용하였고, OPS용 고속 스위치는 InP 기반 칩 기술을 사용한 4x4 AVC OXS 스위치를 사용하여 구성하였다. COSIGN에 사용된 주요 모듈/부품이 (그림 6)에 나와 있다. COSIGN은 광 데이터 평면 자원을 점진적으로 도입함으로써 전광 DCN을 향하여 광 데이터 평면 기술 발전 및 성능 개선을 도모하고 있다[24].

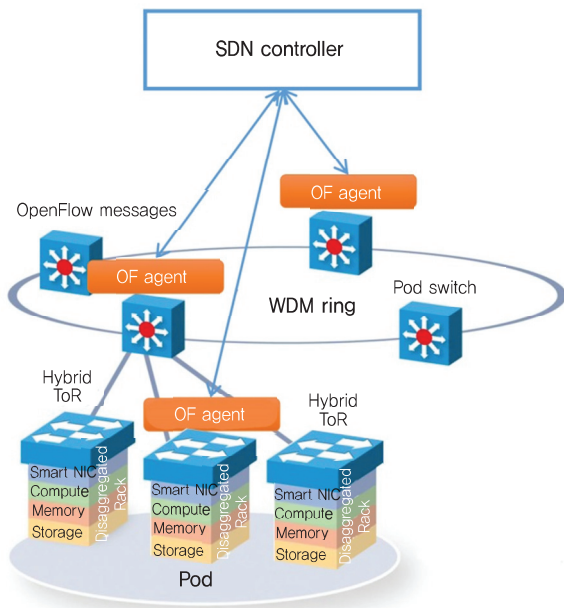
4x4 AVC 기반 OXS는 (그림 5)와 같이 스위칭 그리드 매트릭스의 크로스 포인트에 위치한 16개의 독립된 스위칭 셀로 구성된다. 각각의 AVC 셀 토폴로지는 On/Off 스위칭을 수행한다. 현재 개발 중인 2세대 4x4 OXS 모듈은 AVC와 유사한 모듈리식 InP 기반 칩 구조로 PDL 특성이 우수하다.

다. NEPHELE

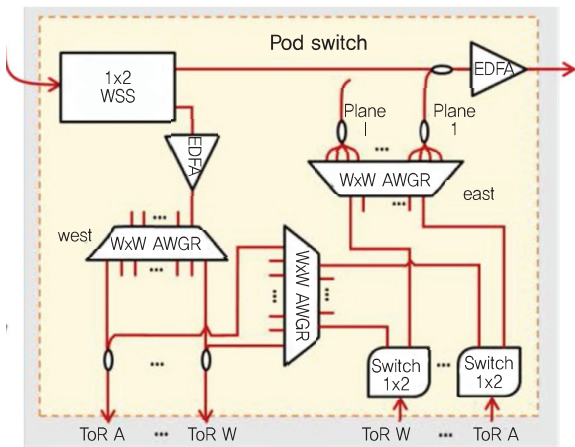
NEPHELE는 (그림 7)과 같이 WDM 링 구조를 기본으로 하고 있으며, 이를 중첩하여 다중 평면 구조로 확장할 수 있다. 광 스위치는 ToR 스위치와 POD 스위치가 있다. POD 스위치는 WDM 링 내에 하나의 노드를 구성하며 몇 개의 랙을 한 그룹으로 묶는다. POD 스위치에서는 POD 내 연결 및 POD 간 연결을 위해 OPS와 OCS 기능을 모두 수행할 수 있다. 이를 위해 POD 스위치는 (그림 8)과 같이 3개의 1x80 AWG와 80개의 1x2



(그림 6) COSIGN DCN에 사용된 디바이스들[21]



(그림 7) NEPHELE DCN 구조[25]



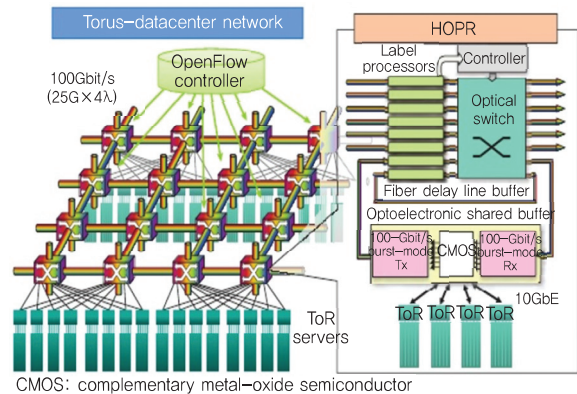
(그림 8) NEPHELE POD 스위치 구성도[22]

WSS 스위치를 사용하여 80x80 용량을 구현하였다. ToR 스위치와의 연결은 1x2 PLZT 스위치와 커플러로 이루어진다. ToR 스위치는 고속 파장가변 송신기(FTL)와 버스트 모드 수신기(BMR), 데이터 패킷 처리용 FPGA, 4x4 PLZT 스위치를 구비하고 있다. NEPHELE SDN 컨트롤러를 통해 POD 및 ToR 스위치를 제어할 수 있다. DCN의 제어는 TDMA 스케줄러에 의해 수행되는데, TDMA 동작은 타임 슬롯 기반으로 이루어진다.

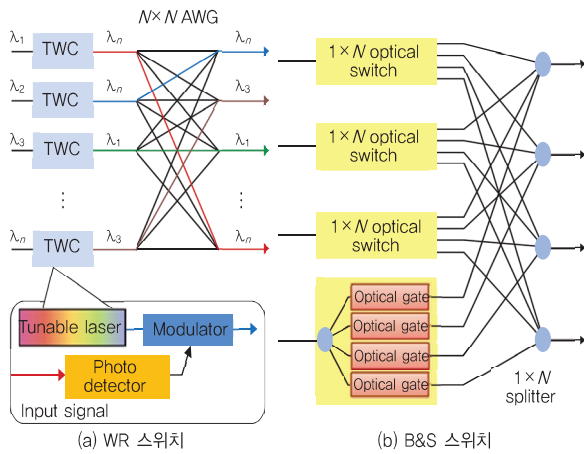
한 스케줄의 길이는 16.8ms이며, 패킷 길이 200 μ s와 가드 타임 10 μ s인 80개의 슬롯을 갖는다. 스케줄의 길이는 요구된 트래픽 패턴에 따라 변할 수 있다.

라. HOPR

일본 NTT에서는 (그림 9)와 같이 HOPR과 OpenFlow 컨트롤러를 결합한 다차원의 토러스(Torus) 토폴로지 DCN 구조를 제안하였다[23]. 토러스 토폴로지는 간단하고 신속한 라우팅 알고리즘이 가능하여 충돌 회피용 에너지 소모형 버퍼의 사용을 줄일 수 있다. 이 구조에서 OpenFlow 기반의 중앙집중형 제어와 라벨 처리 기반의 분산 제어의 결합으로 단일 하드웨어 플랫폼상에서 OPS와 OCS를 모두 실현할 수 있다. 토러스 DCN은 확장성이 뛰어나고 목적지 루트가 다양하여 노드 장애 시 경로보호 능력이 우수하다는 장점이 있으나, 멀티 홉 전송이 필요하다는 단점을 가지고 있다[28]. HOPR은 100Gb/s 링크 속도를 지원하며 라벨 정보를 이용한 패킷 포워딩으로 광 스위칭을 수행한다. 구성을 보면 크게 광 패킷 스위치와 공유 버퍼 두 부분으로 나누며, 광 패킷 스위치는 라벨 처리기, 8x8 광 스위치, FDL로 구성된다. 공유 버퍼는 CMOS와 100G (4x25G) 파장가변/BM 광 송/수신기로 이루어지며, 40G (4x10GbE) ToR 스위치와 광 스위치 간 브릿지 역할을 수행한다. FDL과 공유 버퍼는 충돌 회피용 대체 경로로 사용되며, 또한,



(그림 9) HOPR과 Torus-DCN 구조[23]



(그림 10) WR 스위치와 B&S 스위치 구조[23]

FDL은 OCS 패킷 경로용으로도 사용된다[29].

HOPR에서 사용된 고속 광 스위치로는 WR 스위치와 B&S 스위치가 있다. 이 두 스위치의 구조가 (그림 10)에 나와 있다. WR 스위치는 파장 가변 변환기(TWC)와 $N \times N$ AWG로 구성된다. AWG의 출력 포트는 입력 파장에 의해 결정되기 때문에 TWC에서 패킷 파장을 원하는 파장으로 변환함으로써 입력 패킷을 원하는 출력 포트에 포워드시킬 수 있다. WR 광 스위치는 AWG 손실 증가 없이 파장-채널 간격을 좁혀서 포트 수 N 을 증가시킬 수 있으므로 확장성이 우수하나, WDM 패킷 처리에 적합하지 않다. B&S 스위치는 $1 \times N$ 광 스플리터와 광 게이트로 구성되며, 원하는 출력 포트에 연결된 광 게이트를 사용하여 빛을 통과시키거나 차단시킴으로써 광 스위칭이 수행된다. B&S 광 스위치는 스위치 포트 수가 증가하면 분할 손실이 증가하므로 확장성이 제한되나, 광 게이트의 간단한 제어로 빛을 차단 또는 통과시킬 수

<표 4> 스위치 구조에 따른 HOPR 비교

구분	WR 스위치	B&S 스위치
스위칭 용량	8×8 (10Gbps)	8×8 (100Gbps)
스위칭 시간	10ns	10ns
광패킷 속도	10Gbps	4×25 Gbps
소비 전력	2.25W/Gbps	0.09W/Gbps
노드 지연	380ns	100ns

있는 장점이 있다. 광 게이트는 EAM 게이트를 사용하였다[23]. <표 4>는 스위치 구조에 따른 HOPR 시작품을 비교한 것이다. B&S 스위치 방식이 소비 전력과 노드 지연에서 우수한 특성을 가지고 있음을 알 수 있다.

III. 광 패킷 스위칭 기술

광 DCN에 적용되는 광 스위칭 기술로는 OCS와 OBS/OPS가 있으며, OCS의 경우는 스위칭 시간이 수~수십 ms가 소요되므로 기존의 제어 기술로 문제가 없다. OBS/OPS의 경우는 수 μ s/ns 이내를 만족하여야 하므로 고속 스위칭을 위한 새로운 데이터 전송과 제어 기술이 필요하다. 광 패킷 스위칭을 위한 데이터 전송과 스위칭 제어 방식을 <표 5>에 비교하였다. 데이터 전송 방식으로는 타임슬롯(TS) 전송 방식과 패킷/라벨 전송 방식이 있다. 타임슬롯 방식은 타임슬라이스 또는 TDM 기반 방식이라고도 하며, 패킷 데이터를 시분할된 슬롯에 할당하여 전송하는 방식이다[30]. 슬롯과 여러 슬롯으로 이루어진 프레임의 타이밍을 유지하기 위해 동기화가 필요하며, 외부 동기 방식과 내부 동기 방식이 있다. 외부 동기는 OTSS[31]의 GPS와 IEEE 1588을 이용하는 시간 동기 방식과 TSON[32]의 E1 신호를 이용하는 프레임 동기 방식이 있다[33]. 내부 동기는 NEPHELE

<표 5> 광 패킷 스위칭 기술 비교

전송 방식	스위칭 방식	제어 방식	적용 예
타임슬롯	비동기 TS	RTSA, 시그널링	OTSS[31]
	동기 TS	RWTA, 시그널링	TSON[32]
	동기 TS	RTSA, 스케줄링	NEPHELE
	비동기 TS	스케줄링, OMAC	OPST[34]
패킷/라벨	B&S	인밴드 광라벨, ACK	LIGHTNESS
	WR	인밴드 광라벨, 버퍼/AO-NACK	LIONS
	WR	AO-NACK, AO-TOKEN	TONACK LION
	WR, B&S	인밴드 광라벨, 버퍼	HOPR
혼용	Hybrid	스케줄링, 아웃밴드 제어패킷, ACK, 버퍼	HOS

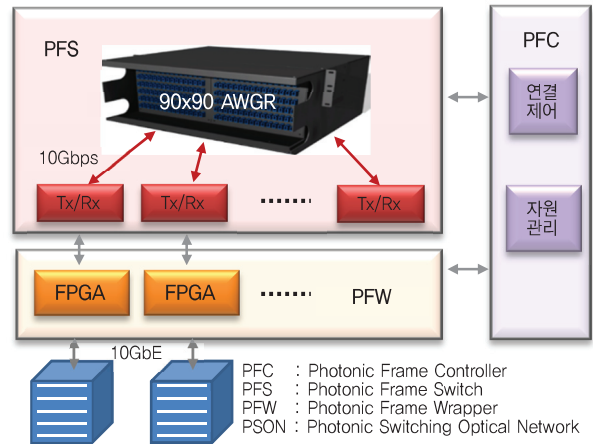
에서 적용하는 바와 같이 프레임내 동기 워드(패턴)을 보내어 유지하는 방식이다. 스위칭 방식에 있어서는 스위칭 시 TS의 타이밍 동기가 필요한 동기 TS 방식과 타이밍 동기가 필요가 없는 비동기 TS 방식이 있다. 타임슬롯 방식에서는 스위칭 경로 설정과 타임슬롯을 할당하기 위하여 RTSA/RWTA와 같은 대역폭 할당 알고리즘이 필요하며, 자원 충돌 회피를 위한 시그널링/스케줄링 절차가 필요하다. OPST[34]에서 적용한 OMAC 메커니즘은 CSMA-CA MAC 프로토콜을 이용하여 광 채널을 감시하여 충돌을 회피하는 방식으로 링 구조의 광 버스트 전송에 적합하다.

패킷/라벨 방식은 라벨을 사용하여 스위칭 정보를 인밴드 또는 아웃밴드에 실어 보낸다. 두 가지 충돌 회피 방식이 있는데, 버퍼(FDL, 공유 버퍼 등)를 사용하여 목적지가 중복되게 할당된 패킷을 버퍼에 저장하였다가 보내거나, 버퍼를 사용하지 않고 특정 신호(ACK 등)를 수신 측으로 보내어 패킷을 재전송하게 한다. 스위칭 방식으로는 파장변환기와 AWGR을 사용한 WR 방식을 사용하는 경우와 광 게이트를 사용한 B&S 방식을 사용하는 경우가 있다.

혼용 방식은 HOS에서 사용한 것으로, 아웃 밴드 제어용 패킷을 통해 정보를 전달하며, TDM 슬롯을 통해 회선 및 패킷 데이터를 전달할 수 있도록 함으로써 자원 활용도를 높일 수 있다.

VI. 국내 연구 동향

국내에서도 ETRI의 PSON 프로젝트에서 광 DCN 기술 관련 연구가 진행되고 있다[5]. (그림 11)에서와 같이 PSON은 전광 방식 구조로 3개의 기능, 즉 PFS, PFW, PFC로 구성된다. PFS는 스위칭 패브릭으로서 FTL, BMR, 90x90 AWGR로 구성되며, 수 μ s 이내의 스위칭 시간과 360x360 포트의 스위칭 용량을 만족한다. PFW는 서버로부터의 10GbE 데이터를 타임슬롯 방식인



(그림 11) PSON의 구조

PSON 고유의 PF 프레임에 할당하여 PFS로 보내는 PF 생성 기능과 PFS로부터 오는 PF 프레임을 중단하여 10GbE 데이터를 추출하여 서버로 보내주는 PF 중단 기능을 수행한다. PF 중단 및 생성 기능은 고성능 FPGA를 이용하여 구현한다. PFC는 데이터센터내 네트워크의 연결 제어 및 자원 관리 기능을 통합적으로 수행한다. 신속한 경로 설정과 타임슬롯 할당을 위해 새로운 스케줄링 알고리즘을 제안하여 적용하였다[35]. PSON은 1홉 방식의 단순한 구조를 가지므로 DCN의 전광화에 유리한 장점을 갖고 있다.

V. 결론

데이터센터 내 트래픽 증가와 새로운 서비스 요구에 따른 네트워크 인프라의 확장과 네트워크 저지연 및 저전력이 요구되고 있으나, 기존 전자식 패킷 스위칭 기술 기반의 네트워크로는 한계가 있다. 데이터센터 네트워크의 확장성이 자유롭고 저지연을 만족하며 소비전력이 증가되지 않는 광 스위칭 기술의 도입으로 이를 극복할 수 있다. 세계적으로 광 데이터센터 네트워크 구조와 광 스위칭 방식에 대해 여러 가지 연구가 진행되어 왔다.

전자식 패킷 스위칭(EPS) 기술을 대체하는 광 스위칭 기술로 광 회선 스위칭(OCS)과 광 패킷 스위칭(OPS) 기

술이 있다. OCS는 일부 기술이 상용화되어 있어 도입이 가능하나, OPS는 아직 상용화하기에는 기술 완성도가 떨어져 연구가 더 필요하다. 따라서 OCS가 우선 적용되어 EPS와 혼용한 전자/광 하이브리드 방식이 점차 도입될 것으로 보이며, 궁극적으로 광 회선 및 광 패킷 스위칭을 모두 지원하는 전광 스위칭 기반의 광 DCN으로 발전할 것으로 예상된다.

용어해설

데이터센터 네트워크(DCN) 데이터센터 내에서 정보 유동을 위해 전송 및 스위칭 설비로 구성되는 네트워크이다. 현재 DCN은 전자식 패킷 스위칭 기반으로 구성되어 있으나, 향후 트래픽 증가와 새로운 서비스 요구에 따른 확장성, 비용, 전력소비, 지연 등의 문제로 한계에 직면할 것으로 예상됨에 따라 광 스위칭 기반의 광 DCN으로 진화될 전망이다.

약어 정리

ACK	Acknowledgement	FPGA	Field Programmable Gate Array
AO	All-Optical	FTL	Fast Tunable Laser
AoD	Architecture-on-Demand	FISSION	Flexible Interconnection of Scalable Systems Integrated using Optical Networks
AVC	Active Vertical Coupling	GPS	Global Positioning System
AWG	Arrayed Waveguide Grating	HCPBF	Hollow Core Photonic Bandgap Fibre
AWGR	Arrayed Waveguide Grating Router	HOPR	Hybrid Optoelectronic Router
BM	Burst-Mode	HyPac	Hybrid Packet and Circuit Switched Data Center Network Architecture
BMR	Burst-Mode Receiver	ICT	Information & Communication Technology
B&S	Broadcast & Select	InP	Indium Phosphide
CAWG	cyclic Arrayed Waveguide Grating	LIGHTNESS	Low Latency and High Throughput Dynamic Network Infrastructures for High Performance Data Centre Interconnects
CFP	Centum Form Factor Pluggable	LIONS	Low-latency Interconnect Optical Network Switch
COSIGN	Combining Optics and SDN In next Generation data centre Networks	MAC	Media Access Control
CSMA-CA	Carrier Sense Multiple Access/with Collision Avoidance	MCF	Multi-Core Fibre
DC	Data Center	MEF	Multi-Element Fibre
DCN	Data Center Network	MEMS	Micro Electro Mechanical Systems
DP-QPSK	Dual-Polarization Quadrature Phase Shift Keying	MIMO	Multiple-Input Multiple-Output
DWDM	Dense Wavelength Division Multiplexing	NACK	Negative Acknowledgement
EAM	Electro-Absorption Modulator	NEPHELE	eNd to End scalable and dynamically reconfigurable oPtical arcHitecture for application-awarE SDN cLoud datacentErs
EDFA	Erbium Doped Fiber Amplifier	NIC	Network Interface Card
FDL	Fiber Delay Line	NTUA	National Technical University of Athens
		OBS	Optical Burst Switching
		OCS	Optical Circuit Switching
		ODCN	Optical Data Center Network
		OFDM	Orthogonal Frequency Division Multiplexing
		OMAC	Optical Media Access Control
		OPS	Optical Packet Switching
		OPST	Optical Packet Switch and Transport
		OSA	Optical Switching Architecture
		PFC	Photonic Frame Controller
		PFS	Photonic Frame Switch
		PFC	Photonic Frame Wrapper
		PLZT	Perovskite Lead Lanthanum Zirconate Titanate

PDL	Polarization Dependent Loss
PON	Passive Optical Network
PSD	Parallel Signal Detection
PSON	Photonic Switching Optical Network
RTSA	Routing and Time Slot Allocation
RWTA	Routing, Wavelength and Time-slice Assignment
SDM	Space Division Multiplexing
SDN	Software Defined Networking
SIC	Switch & Interface Card
SOA	Semiconductor Optical Amplifier
SSS	Spectrum Selective Switches
STIA	Space-Time Optical Interconnection Architecture
TDM	Time Division (or Domain) Multiplexing
TDMA	Time Division Multiple Access
ToR	Top of the Rack (switch)
TS	Time Slot or Time Slice
TSON	Time Shared Optical Network
TWC	Tunable Wavelength Converter
UoB	University of Bristol
VCSEL	Vertical-Cavity Surface-Emitting Laser
WC	Wavelength Converter
WDM	Wavelength Division Multiplexing
WR	Wavelength Routing
WS	Wavelength Selector
WSS	Wavelength Selective Switch

참고문헌

- [1] Cisco Public, Cisco Global Cloud Index: Forecast and Methodology 2015-2020, White Paper, 2016.
- [2] A. Hammadi and L. Mhamdi, "A Survey on Architectures and Energy Efficiency in Data Center Networks," *Comput. Commun.*, vol. 40, Mar. 2014, pp. 1-21.
- [3] L. Dittmann et al., "A Roadmap for Evolving Towards Optical Intra-Data-Center Networks," In *Proc. Eur. Conf. Opt. Commun.*, Dusseldorf, Germany, Sept. 2016, pp. 1-3.
- [4] C. Kachris, K. Kanonakis, and I. Tomkos, "Optical Interconnection Networks in Data Centers: Recent Trends and Future Challenges," *IEEE Commun. Mag.*, vol. 51, no. 9, Sept. 2013, pp. 39-45.
- [5] 유연철 외, "초저지연 인프라 기술," 전자통신동향분석, 제32권 제1호, 2017. 2, pp. 13-23,
- [6] N. Farrington et al., "Helios: a Hybrid Electrical/Optical Switch Architecture for Modular Data Centers," *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 4, 2010, pp. 339-350.
- [7] G. Wang et al., "C-Through: Part-Time Optics in Data Centers," *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 4, 2010, pp. 327-338.
- [8] C. Kachris and I. Tomkos, "Power Consumption Evaluation of Hybrid WDM PON Networks for Data Centers," In *Eur. Conf. Netw. Opt. Commun.*, Newcastle, UK, July 2011, pp. 118-121.
- [9] K. Chen et al., "OSA: An Optical Switching Architecture for Data Center Networks with Unprecedented Flexibility," *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, Apr. 2014, pp. 498-511.
- [10] M. Fiorani, S. Aleksic, and M. Casoni, "Hybrid Optical Switching for Data Center Networks," *J. Electr. Comput. Eng.*, vol. 2014, 2014, pp. 139213:1-139213:13.
- [11] A. Kushwaha et al., "Flexible Interconnection of Scalable Systems Integrated Using Optical Networks (FISSION) Data-Center-Concepts and Demonstration," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 9, no. 7, July 2017, pp. 585-600.
- [12] O. Liboiron-Ladouceur, I. Cerutti, P.G. Raponi, N.Andriolli, and P. Castoldi, "Energy-Efficient Design of a Scalable Optical Multiplane Interconnection Architecture," *IEEE J. Sel. Topics Quantum Electron.*, vol. 17, no. 2, 2011, pp. 1-7.
- [13] I. Cerutti, P.G. Raponi, N. Andriolli, P. Castoldi, and O. Liboiron-Ladouceur, "Designing Energy-Efficient Data Center Networks Using Space-Time Optical Interconnection Architectures," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 2, Mar./Apr. 2013.
- [14] Y. Yin, R. Proietti, X. Ye, C.J. Nitta, V. Akella, and S.J.B. Yoo, "LIONS: An AWGR-Based Low-Latency Optical Switch for High-Performance Computing and Data Centers," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 2, Mar./Apr. 2013.
- [15] R. Proietti et al., "Scalable Optical Interconnect Architecture Using AWGR-Based TONAK LION Switch With Limited Number of Wavelengths," *J. Lightwave Technol.*, vol. 31, no. 24, Dec. 2013, pp. 4087-4097.
- [16] P.N. Ji, D. Qian, K. Kanonakis, C. Kachris, and I. Tomkos, "Design and Evaluation of a Flexible-Bandwidth OFDM-Based Intra Data Center Interconnect," *IEEE J. Sel. Topics*

- Quantum Electron.*, vol. 19, Mar/Apr. 2013.
- [17] O. Liboiron-Ladouceur et al., "The Data Vortex Optical Packet Switched Interconnection Network," *J. Lightwave Technol.*, vol. 26, no. 13, July 2008, pp. 1777-1789.
- [18] K. Xia, Y. Kao, M. Yang, and H.H. Chao, "Petabit Optical Switch for Data Center Networks," Technical report, Polytechnic Institute of NYU, 2010.
- [19] Z. Cao, R. Proietti, M. Clements, and S.B.J. Yoo, "Experimental Demonstration of Flexible Bandwidth Optical Data Center Core Network with All-to-All Interconnectivity," *J. Lightwave Technol.*, vol. 33, no. 8, Apr. 2015, pp. 1578-1585.
- [20] LIGHTNESS D2.2, "Design document for the proposed network architecture," 2013, Accessed 2017. <http://www.ict-lightness.eu/>
- [21] COSIGN Deliverable D1.4, "Architecture design," 2015, Accessed 2017. <http://www.fp7-cosign.eu/>
- [22] K. Tokas et al., "Slotted TDMA and Optically Switched Network for Disaggregated Datacenters," *Int. Conf. Trans. Opt. Netw.*, Girona, Spain, July 2017, pp. 1-5.
- [23] T. Segawa, Y. muranaka, and R. Takahashi, "High-speed Optical Packet Switching for Photonic Datacenter Networks," *NTT Techn. Rev.*, vol. 14, no. 1, Jan. 2016, pp. 1-7.
- [24] Xi Chen et al., "Monolithic InP-Based Fast Optical Switch Module for Optical Networks of the Future," *Int. Conf. Photon. Switching*, Florence, Italy, Sept. 2015, pp. 294-296.
- [25] P. Bakopoulos, "Optical Switching for Scalable and Programmable Data Center Networks," *Symp. Opt. Interconnect Data Centres*, Dusseldorf, Germany, June 2016.
- [26] LIGHTNESS D3.2, "Implementation Results of the OPS Switch, the OCS Switch, and the TOR Switch," 2014.
- [27] N. Parsons, "High radix All-Optical Switches for Software-Defined Datacentre Networks," *Proc. Eur. Conf. Opt. Commun.*, Dusseldorf, Germany, Sept. 2016, pp. 1-3.
- [28] K. Kitayama et al., "Torus-Topology Data Center Network Based on Optical Packet/Agile Circuit Switching with Intelligent Flow Management," *J. Lightwave Technol.*, vol. 33, no. 5, Mar., 2015, pp. 1063-1071.
- [29] Y. Huang, Y. Yoshida, S. Ibrahim, R. Takahashi, and K. Kitayama, "Load Balancing in Switch-Fabric Type of Torus OPS Data Center Networks With Hybrid Optoelectronic Routers," *OptoElectron. Commun. Conf. Held Jointly Int. Conf. Photon. Switching*, Niigata, Japan, July 2016, pp. 1-3.
- [30] 고체수 외, "데이터센터 내 광 네트워크를 위한 시간 동기 기반의 포토닉 프레임 스위칭 방식," *Conf. Electron. Inform. Commun.*, Daejeon, Rep. of Korea, Dec. 2016, pp. 243-245.
- [31] N. Hua and X. Zheng, "Optical Time Slice Switching (OTSS): An All-Optical Sub-wavelength Solution Based on Time Synchronization," *Proc. Asia Commun. Photon. Conf.*, Beijing, China, Nov. 2013.
- [32] G.S. Zervas et al., "Time Shared Optical Network (TSON): A Novel Metro Architecture for Flexible Multi-Granular Services," *Opt. Exp.*, vol. 19, no. 26, 2001, pp. B509-B514.
- [33] 고체수, 김광준, 이준기, "패킷 기반 전광 네트워크를 위한 시간 동기 기술," 2017년 대한전자공학회 하계종합학술대회, 2017. 8, pp. 1633-1636.
- [34] J. Dunne, T. Farrell, and J. Shields, "Optical Packet Switch and Transport: A New Metro Platform to Reduce Costs and Power by 50% to 75% While Simultaneously Increasing Deterministic Performance Levels," *Int. Conf. Trans. Opt. Netw.*, Azores, Portugal, 2009, pp. 1-5.
- [35] J. Song, K. Han, D. Kim, C. Park, and K. Kim, "Low Delay Switch Scheduling for Data Center Optical Switch," *Int. Conf. Inform. Commun. Technol. Convergence.*, Jeju, Rep. of Korea, Oct. 2017, pp. 1250-1252.