Converged Inter/Intra All-Optical DC Network Externally Distributing Optical Carriers to Coherent Transceivers

Ritsuki Hamagami, Masamichi Fujiwara, Naotaka Shibata, Shin Kaneko, Jun-ichi Kani, and Tomoaki Yoshida

NTT Access Network Service Systems Laboratories, NTT Corporation 1-1 Hikari-no-oka, Yokosuka-shi, Kanagawa 239-0847, Japan
ritsuki.hamagami@ntt.com

Abstract: We propose a DCN that directly connects server racks distributed among DCs through ROADM-based nodes. External light sources are introduced to coherent transceivers to avoid laser-diodes being operated under high temperature conditions on top-of-rack switches. © 2024 The Author(s)

1. Introduction

Optical circuit switches (OCS), which have the advantage of low power consumption and bitrate transparent characteristics, are beginning to be applied to higher tier switches for intra data center (DC) communications to meet the rapid traffic growth of DC networks (DCNs) [1]. Several studies have proposed converged inter/intra DCNs that connect optical transceivers (TRxs) mounted on each top of rack (ToR) electrical switch by end-to-end direct optical path through OCS [2,3]. A technical trend in DCNs is that reconfigurable optical add/drop multiplexer (ROADM)-based node configurations are introduced to the DC interconnect (DCI) [4]. Another technical challenge is to connect a ROADM-based metro network to OCS-based DCNs. Here, dense wavelength division multiplexing (DWDM) technology is indispensable for reducing the number of optical fibers while providing direct optical connections between ToRs and is enhanced by being combined with ROADM. The digital coherent technique is desirable for transmitting high-speed DWDM signals through the ROADM-based DCI since digital signal processing (DSP) can adaptively compensate for chromatic dispersion of optical fiber that varies from DWDM signals. One issue is that it is difficult for a wavelength-tunable laser diode (TLD) to meet the strict requirements for coherent transmission under high ambient temperature conditions in a ToR. A similar situation occurs in optical interconnection applications, where the use of external laser sources (ELSs), which are mounted away from electrical switches that emit high heat, has been proposed for intensity modulation and direct-detection (IM-DD)-based TRxs [5].

We propose a converged intra/inter DCN that enables all-optical connections between ToRs distributed among DCs connected to metro-access nodes consisting of ROADM. A TLD is removed from each coherent TRx to ensure thermal reliability, and a continuous wave (CW) light is distributed from the ELS to the TRx. An ELS is composed of distributed-feedback LDs (DFB-LDs) with a fixed wavelength, and an OCS changes the wavelength. We verified the feasibility of the proposed DCN from the viewpoint of transmission performance through demonstration experiments assuming intra/inter DCIs. To the best of our knowledge, this is the first study in which wavelength distributions are applied for both modulation and coherent detection.

2. Proposed Converged Inter/Intra All-Optical Data Center Network

Figures 1 (a) and (b) show the wavelength settings and optical-path routing schemes in OCS-based DCs using the conventional digital coherent TRx and the digital coherent TRx in the proposed DCN, respectively. In Fig. 1 (a), the wavelength setting and optical-path routing are executed by a TLD and the OCS, respectively. The issues are that the TLD is exposed to high temperature in the ToR, and wavelength and optical-path controls are carried out against two types of devices having different response times (TLD and OCS). The former hinders reliable operation of the TLD and the latter makes the control scheme more complicated. The proposed DCN in Fig. 1(b) does not use any light sources, and a CW light is supplied from an ELS through the OCS. The wavelength passes through a compact polarization controller (PC) with an automatic polarization-tracking function [6] and its power is equally divided for modulation and coherent detection. In addition to optical-path routing, wavelength selection is also executed by identical devices (OCS); thus, the control scheme needed in Fig. 1(b) becomes simpler than that in Fig. 1(a). Figure 1(c) shows the proposed DCN with a ring topology. It is assumed that each DC has several hundred server racks, which is typical for edge DCs enabling low latency, and that multiple DCs are connected to a ROADM-based node. The dotted red, dashed blue, and solid green lines show the optical paths of the intra DCI, inter DCI between DCs connected to the same node, and between DCs connected to different nodes, respectively. Each DC consists of a TRx mounted on each ToR (see Fig. 1(b)), an OCS, array waveguide gratings (AWGs), and an ELS for the intra DCI. The ELS for the intra DCI is a single wavelength light source, divided into the same number of TRxs by an

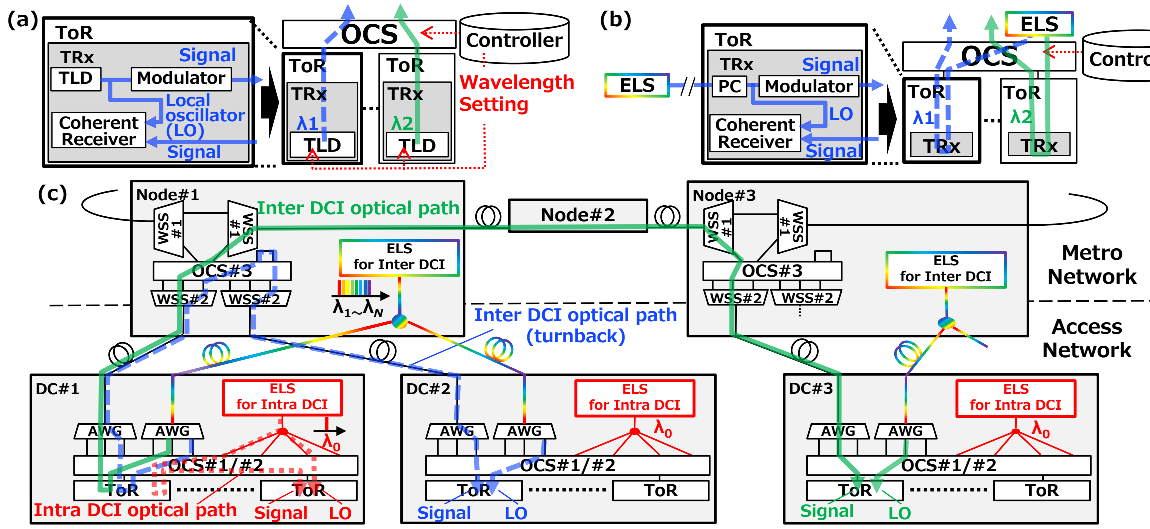

Fig. 1: Wavelength settings and optical-path routing schemes in OCS-based DCs. (a) With conventional digital coherent TRx. (b) With digital coherent TRx used in proposed DCN. (c) Proposed converged all-optical intra/inter DCN.

optical splitter (OS) then connected to OCS#1. An ELS for the inter DCI, however, is a DWDM light source. The ELS for the inter DCI is located in each node and divided into the same number of DCs connected to the same node for cost reduction, and launched in an access fiber. After transmission, it is demultiplexed by the AWG then connected to OCS#1. All the distributed wavelengths are connected to OCS#1, thus the wavelength supplied to each TRx can be changed by reconfiguring the connecting ports of OCS#1, and both intra and inter DCIs are executed by an identical colorless TRx. The supplied wavelength is modulated at the Tx of each TRx and is input to OCS#2. The intra DCI path is directly made through OCS#2, and the signal is received at the Rx of a different TRx (see dotted red line). The DWDM signals for the inter DCI pass through OCS#2, multiplexed by the AWG, then transmitted over an access fiber to Node#1. Node#1 implements extended-ROADMs, where some modifications are made to typical configurations. They consist of wave selective switches (WSS#1 and #2) and OCS#3. WSS#1 is used in the metro ring network to select wavelength groups for pass through and add/drop. The number of add/drop wavelength groups is equal to that of DCs connected to Node#1 and each wavelength group is connected to the target DC. WSS#2 is connected to WSS#1, and an access fiber through OCS#3 is used to multi/demutiplex the wavelength groups described above. OCS#3 is used to execute turnback connections to the TRx in different DCs connected to Node#1. Thus, WSS#2 also muti/demultiplexes the wavelength groups passing through these turnback paths (see dashed blue line). The turnback signals are transmitted over an access fiber, demultiplexed by the AWG in DC#2, and are received. The wavelength groups added to the metro ring network through WSS#1 are transmitted to Node#3 via Node#2 and dropped (see solid green line). Each WDM signal is then received at DC#3.

3. Experiments

Figure 3 shows the experimental setup and Table I summarizes the insertion losses of passive devises used. Note that the OS and OCS#1–#3 are substituted with optical attenuators since they are wavelength independent. The DCN assumed is a ring topology with 3 nodes and each node is connected to 4 DCs, where each span length is 20 km. The number of racks in each DC is about 300, 15% of which communicate with those in other DCs. The dotted red, dashed blue, and solid green lines show the optical paths of the (i) intra DCI in DC#1 (TRx#1–TRx#2; 0 km), (ii) inter DCI between DC#1 and DC#2 (TRx#1–TRx#3; 40 km), and (iii) that between DC#1 and DC#3 (TRx#1–TRx#4; 80 km), respectively. The ELS for (i) in DC#1 consists of a DFB-LD, 1×8 OS, erbium doped fiber amplifier (EDFA#1), and 1×32 OS and supplies CW light with the same wavelength to 256 racks. The wavelength, output power, and linewidth of the DFB-LD are 1550.92 nm, 10.0 dBm, and 1.5 MHz, respectively. The CW light is amplified to 10.1 dBm/ch by EDFA#1, assuming that the saturation output power of EDFA#1 is 28.0 dBm. The ELS for (ii) and (iii) in Node#1 consists of a DFB-LD, 48-channel AWG with 100–GHz spacing, and 1×4 OS and supplies different wavelengths to 48 racks. The wavelength from the ELS for (ii) and (iii) is amplified to 10.1 dBm/ch by EDFA#1. In the Tx, a CW light is modulated by a LN-IQ modulator at 28-Gbaud QPSK signals (for 50 Gbps) with a pseudo random bit sequence pattern of 2^{15}-1. The modulated signal for (i) is received at TRx#2 in DC#1. Those of (ii) and (iii) are amplified to -3.0 dBm/ch by EDFA#2. In the Rx, received signals are captured using a real-time oscilloscope at the sampling rate of 100 GS/s, down-sampled to 56 GS/s, and demodulated by offline DSP. The DSP blocks consist of a 25-tap finite impulse response filter operated using a constant modulus algorithm and followed by carrier frequency/phase offset compensations.

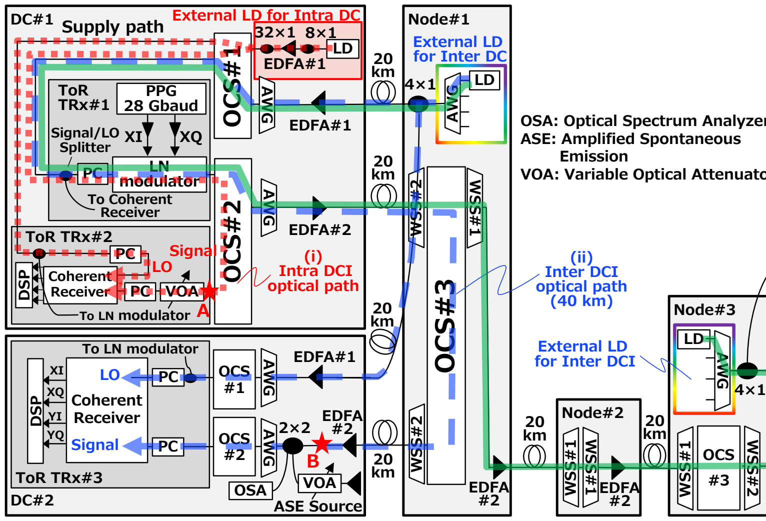

Table I: Insertion loss of passive devices.

Parameters	Values
PC	2.0 dB
AWG	3.4 - 3.9 dB
WSS	3.9 – 4.2 dB
1×*N* OS* (including signal/LO splitter)	3.5log$_2$(*N*) dB
OCS#1/#2 (DC)*	3.0 dB
OCS#3 (Node)*	2.0 dB
Single Mode Fiber (20 km)	4.0 - 4.3 dB

*These are substituted by optical attenuators.

Fig. 2: Experimental setup.

Figures 3(a)–(c) show the experimental results. The round, square, and triangular plots show the data for power-split ratios between orthogonal polarizations (a) of 1.0, 0.5, and 0.0, where a = 1.0 denotes that the polarization of incoming signal is aligned to the X-axis of the Rx. We assume the use of staircase feedforward error correction (SC-FEC) with a pre-FEC bit error rate (BER) threshold of 4.5e-3. Figure 3(a) shows BERs against received power for (i). The measured receiver sensitivity and received signal power at point A in Fig. 2 were -33.6 and -10.7 dBm, respectively; thus, the receiver-sensitivity margin of 22.9 dB was obtained. Figures 3(b) and (c) show BERs against the optical signal-to-noise ratio (OSNR) for (ii) and (iii), respectively. Path (ii) showed negligible OSNR penalty after 40-km transmission. The required OSNR and OSNR measured at point B were 10.1 and 28.8 dB, respectively; thus, the OSNR margin of 18.7 dB was obtained. Path (iii) showed the required OSNR penalty after 80 -km transmission of 0.8 dB. The required OSNR and measured OSNR at point C were 11.2 and 29.9 dB, respectively. In spite of the OSNR penalty induced by chromatic dispersion and spectrum narrowing by six WSS, the OSNR margin of 18.7 dB, equivalent to that of (ii), was still obtained.

4. Conclusion

We experimentally evaluated the feasibility of the proposed DCN assuming a ring topology with 3 nodes, 4 DCs per node, and about 300 server racks per DC. The transmission experiments for 50-Gbps signals, which meets bandwidth requirements on ToRs, showed sufficient margin for a received power of 22.7 dB against intra DCI and those for the required OSNR of 18.7 dB against inter DCIs (40 km turnback and 80 km passing through 3 nodes). These margins indicate bitrate scalability to more than 100 Gbps.

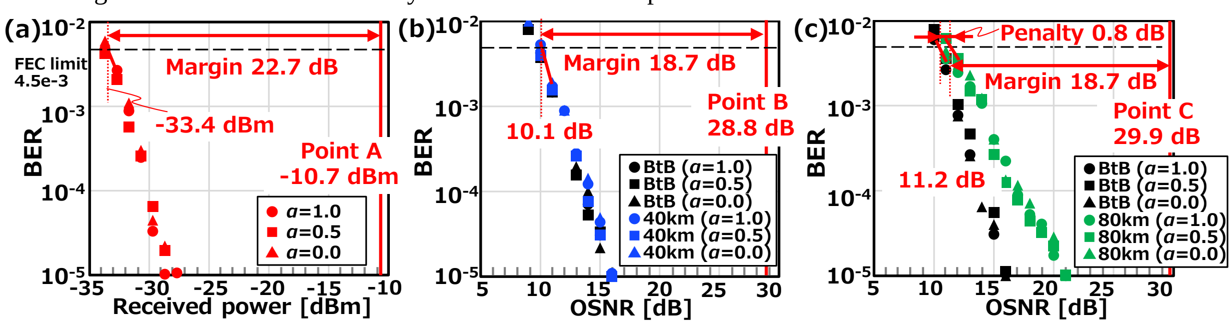

Fig. 3: Experimental results. (a) BERs against received power for path (i). (b)–(c) BERs against OSNR for paths (ii) and (iii).

5. References

[1] L. Poutievski, et al., "Jupiter Evolving: Transforming Google's Datacenter Network via Optical Circuit Switches and Software-Defined Networking," ACM SIGCOMM 2022, 66–85.

[2] S. Yan, et al., "Archon: A Function Programmable Optical Interconnect Architecture for Transparent Intra and Inter Data Center SDM/TDM/WDM Networking," JLT, vol. 33, no. 8, pp. 1586-1595 2015.

[3] J. M. D. Mendinueta et al., "Time-Division Packet Spatial Super-Channel Switching System with 53.3 Tb/s/Port for Converged Inter/Intra data Center Optical Networks," JLT, vol. 37, no. 3, pp. 677-687, 2019.

[4] C. Xie et al., "Open and Disaggregated Optical Transport Networks for Data Center Interconnects [Invited]," JOCN, vol. 12, no. 6, pp. C12-C22, 2020.

[5] T. Sawamura, et al., "8-Channel CWDM TOSA for CPO External Laser Sources Employing a Blind Mate Optical Connector," OFC2023, W4B.1.

[6] T. Gui, et al., "Real-Time Demonstration of Homodyne Coherent Bidirectional Transmission for Next-Generation Data Center Interconnects," JLT, vol. 39, no. 4, pp. 1231-1238, 2021.