

All-Optical Switching Data Center Network Supporting 100Gbps Upgrade and Mixed-Line-Rate Interoperability

Lei Xu, Shaoliang Zhang, Fatih Yaman, Ting Wang, Guangxun Liao*, Kai Chen⁺, Ankit Singla**, Atul Singh, Kishore Ramachandran, Yueping Zhang

NEC Laboratories America, Inc., 4 Independence Way, Princeton, NJ 08540, USA

*CoAdna Photonics, 733 Palomar Ave., Sunnyvale, CA 94085, USA

⁺Department of Electrical Engineering & Computer Science, Northwestern University, Evanston IL 60208, USA

**Department of Computer Science, University of Illinois, Urbana-Champaign, IL 61801, USA

Abstract: Over an 8-rack, 32-server WDM-based all-optical switching data center network, we show the network upgrade using single-wavelength or multi-wavelength 100 Gbps signals, and further show the mixed-line-rate interoperability between 100 Gbps and 10 Gbps signals.

OCIS codes: (060.1155) All-optical networks; (060.4253) Networks, circuit-switched.

1. Introduction

As one of the fast growing sectors in the internet applications, data center networks (DCN) have been generating huge demand on communication bandwidth [1]. For intra DCN communications, through which thousands of computer servers are networked, small form factor optical transceivers and parallel optical transmissions have been adopted to boost the link capacity [1]. However, the current intra DCN communication, which is still based on high-capacity electrical switching fabric, is facing technical challenges of high power consumption and complex manageability during the network migration to higher bit rates. Recently, new DCN architectures utilizing hybrid optical/electrical switching or all-optical switching have attracted a lot of research interests due to their benefits of significantly improved networking performance and lower power consumption [2-4].

Our recent proposal of all-optical switching *Proteus* DCN, which leverages optical wavelength and space switching to achieve network topology malleability, can support higher bandwidth, simplified cabling and network flexibility [4]. Here we use an existing 8-rack, 32-server WDM-based all-optical switching DCN testbed [5] to investigate some critical physical layer issues regarding the future network upgrade to higher bit rates at 100 Gbps, including upgrade using single-wavelength and quad-wavelengths 100 Gbps signals and the mixed-line-rate interoperability between 100 Gbps and 10 Gbps signals.

2. All-optical *Proteus* DCN testbed with 100Gbps links

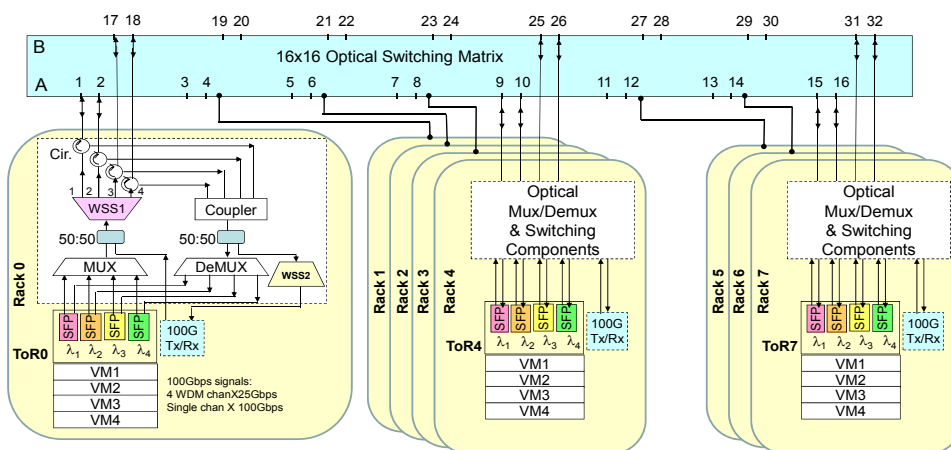


Fig. 1 Testbed for all-optical switching *Proteus* DCN using optical wavelength and space switching and supporting 100 Gbps upgrade. VM: virtual machine; SFP: small form factor pluggable optical transceiver; MUX: optical wavelength multiplexer; DeMUX: optical wavelength demultiplexer; Cir: optical circulator; WSS: wavelength selective switch; ToR: top of rack switch.

Fig.1 shows the testbed of an all-optical switching *Proteus* DCN with the support of 100 Gbps upgrade. There are total 8 server racks (Rack 0-7), and each rack has four virtual machines (VM 1-4). The four 2.5 Gbps SFP WDM transceivers are in 100 GHz channel spacing at 1545.32 nm (Chan 1), 1544.53 nm (Chan 2), 1543.73 nm (Chan 3) and 1542.94 nm (Chan 4). The optical SFP transceivers has ~30 dB optical link power budget. In the transmitting

side, the outputs from the four SFP transceivers are multiplexed, and combined with 100 Gbps optical signals through a 50:50 optical coupler. Please note here that the 100 Gbps optical signal is always set at different wavelengths from the four SFP transceivers. The aggregated WDM signals are grouped into four fiber ports by a CoAdna wavelength selective switch (WSS 1). The optical circulators are adopted to support bi-directional optical communications. The optical switching matrix (OSM) is a 16*16 Polaris optical fiber switching system. Due to the bipartite connectivity of the OSM, the four output ports of the WSS are split into two groups, connecting the ports on both “A” and “B” sides of the OSM. At the receiver side, the 100 Gbps optical signals are selected by a low port count CoAdna WSS (WSS2). The bandwidth of WSS2 can be dynamically adjusted for 100 Gbps in different formats.

Table 1(a) DCN testbed routing path

ToR-X	0	0	0	0	0	0	0	1	1	1	1	1	1	2	2	2	2	2	3	3	3	3	4	4	4	4	5	5	6
Hop ToR				3	1	1		0	5			0	3	1	1			4	0			4	0				4		
ToR-Y	1	2	3	4	5	6	7	2	3	4	5	6	7	3	4	5	6	7	4	5	6	7	5	6	7	6	7	6	7

Table 1(b) wavelength assignment of each ToR

ToR-X	0	0	0	0	1	1	1	2	2	3	3	4	4	4	5	6
ToR-Y	1	2	3	7	2	5	6	3	7	4	5	5	6	7	6	7
Wavelength	1	3	4	2	2	4	3	1	4	2	3	1	4	3	2	1

Each server on the DCN testbed can communicate simultaneously with any servers on other racks through the construction of 4-regular ToR graphs [5], and direct connection and hop-by-hop communications are two ways to achieve network connectivity. Based on the traffic demand, the *Proteus* topology manager can reconfigure the OSM to support an optimal network topology, and the WSS to aggregate the wavelengths for dynamical link capacity adjustment. Table 1 (a) and (b) show the network routing path and wavelength assignment of one case of the DCN testbed connectivity.

3. Experimental Results

In our experiments, 100 Gbps optical signals are transmitted through the *Proteus* DCN testbed together with the existing Ethernet traffic carried by the SFP transceivers. The 100 Gbps signals can be generated using a single carrier/wavelength or multiple carrier/wavelengths. The single-wavelength 100 Gbps optical signal can be generated with dual-polarization, quadrature phase shift keying (DP-QPSK) and received with digital coherent receiving technologies. The multi-wavelength 100Gbps optical signals can be 10*10 Gbps or 4*25 Gbps using on-off-keying (OOK) modulation.

(1) Network upgrade with a single-wavelength 100Gbps scheme

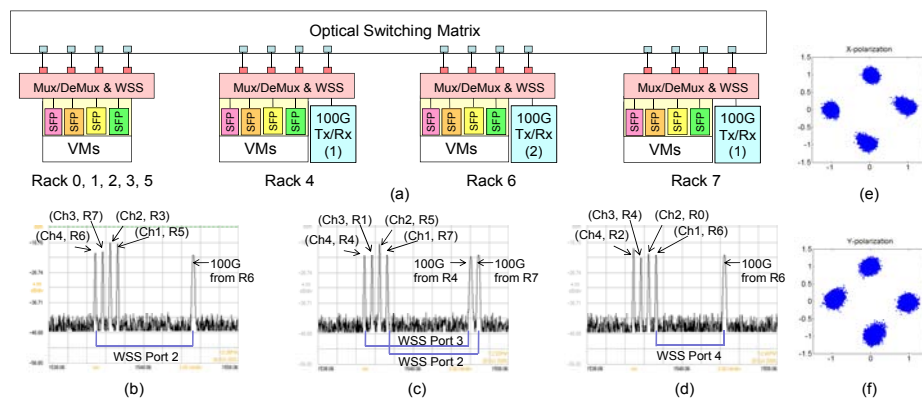


Fig.2 (a) network upgrade with single-wavelength 100Gbps optical signals on Rack 4, 6 and 7. (b) the received optical signal spectrum at Rack 4; (c) the received optical signal spectrum at Rack 6; (d) the received optical signal spectrum at Rack 7. (Ch#, R#) shows the current channel number (Ch#) and its source rack (R#). (e) (f) received 100Gbps signal constellation maps in X and Y polarizations at Rack 7.

Fig. 2 (a) shows the network upgrade using single-wavelength 100 Gbps optical signals, where Rack 4 and 7 each has one 100 Gbps (line rate at 112 Gbps) transceiver and Rack 6 has two 100 Gbps transceivers. Rack 6 is connected with Rack 4 and Rack 7 with 100G links. Fig. 2 (b), (c) and (d) show the received optical spectrum at Rack 4, 6 and 7, respectively. Here the 100Gbps optical signals transmit together with one of the low-speed Ethernet channels (Ch1, 2, 3, 4) and does not cause any interruptions to the existing traffic. Fig. 2(b) shows that the 100 Gbps optical signal from Rack 6 arrives at Rack 4 together with the low-speed signal Ch4 from Rack 6 (Ch4, R6), and both signals arrives at Rack 4 through the WSS port 2.

The single wavelength 100 Gbps optical signal is generated with dual-polarization QPSK modulation format. Fig. 2 (e) and (f) show the received 100 Gbps signal constellation maps in X and Y polarizations at Rack 7. With moderate optical pre-amplification at the receiver side, the 100 Gbps signal can have adequate link power budget.

(2) Network upgrade with a four-wavelength 4*25 Gbps scheme

While the single-wavelength 100 Gbps scheme can have high spectral efficiency, the four wavelengths 4*25 Gbps (actual line rate is at 4*28 Gbps) scheme can support aggregated 100 Gbps transmission at lower system cost. The 25 Gbps optical signal per wavelength is generally based on OOK modulation. One network benefit from the 4*25 Gbps scheme is the flexibility of allocating the four 25 Gbps channels to different destination racks, as shown in Fig. 3(a) and (b). Fig. 3(c) shows the eye diagram of one of the received 25 Gbps signals.

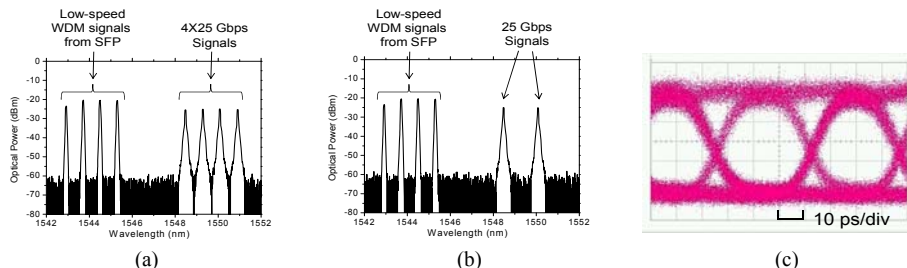


Fig. 3 (a) received optical spectrum with 4*25 Gbps signals (b) received optical spectrum with 2*25 Gbps signals (c) received eye diagram of one 25 Gbps signal

(3) Interoperability between 100G and 10G signals

Since all-optical switching fabric is transparent to the signal bit rates and modulation formats, optical communications at different line rates using different signal formats can co-exist in a DCN. One important issue is the mixed-line-rate interoperability of different optical signals. Table 2 shows the link bit rates between 10 Gbps OOK and different 100 Gbps signal schemes. As for the connection between a 10 Gbps OOK signal and a 100 Gbps DP-QPSK signal, the modulator operation status of the 100G DP-QPSK system needs to be adjusted to generate 10 Gbps optical OOK signal. At the same time, 10 Gbps OOK signal should be detected by the digital coherent receiver which is designed for 100 Gbps DP-QPSK receiving. Fig. 4(a) shows the experimental results of the received 10 Gbps signal generated from a 100 Gbps DP-QPSK transmitter, and Fig. 4 (b) shows the received 10 Gbps signal using the digital coherent receiver. Although it should generally avoid interconnection between a high-speed transmitter and a low-speed transmitter, the interoperability can improve the network flexibility when such scenarios are necessary.

Table 2 Link bit rates between 10 Gbps OOK and different 100 Gbps signal schemes

Signal Format	10G OOK	10X10G OOK	4X25G OOK	100G DP-QPSK
10G OOK	10G	10G	10G	10G
10X10G OOK	10G	100G	40G	10G
4X25G OOK	10G	40G	100G	25G
100G DP-QPSK	10G	10G	25G	100G

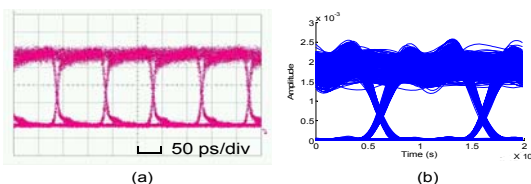


Fig. 4 (a) Received 10 Gbps signal generated from 100 Gbps DP-QPSK transmitter (b) Received 10 Gbps signal using digital coherent receiver used for 100 Gbps DP-QPSK signals.

4. Discussions

The current *Proteus* DCN testbed is based on all-optical switching fabric, which enables the deployment of much higher bit rate, e.g. 100 Gbps, without replacing any of the fiber links and the optical switching elements. The adoption of 100 Gbps signals can bring multiple system benefits and support new network applications. The high-bit-rate optical signals can also be switched “out” of the DCN for inter-DCN communications and/or long-haul transmissions.

Acknowledgement: The authors would like to thank Polatis for the equipment support, and Mr. Don DePaolo and Mr. Yin Shao at NEC Laboratories America for helping with the testbed setup.

References

- [1] C. F. Lam, et al, “Fiber Optic Communication Technologies: What’s Needed for Datacenter Network Operations”, IEEE Comm. Mag, (2010).
- [2] G. Wang, D. G. Andersen, M. Kaminsky, et al, “c-through: Part-time optics in data centers,” in ACM SIGCOMM, 2010.
- [3] N. Farrington, G. Porter, et al, “Helios: a hybrid electrical/optical switch architecture for modular data centers,” in ACM SIGCOMM, 2010.
- [4] A. Singla, A. Singh, K. Ramachandran, et al, “Proteus: a topology malleable data center networks,” in ACM HotNets 2010.
- [5] Details of algorithm design, network management, evaluation and prototype implementation of Proteus proposal are presented in another paper (currently under review). Here, we focus on upgradability to 100G.