Optical Burst-over-Circuit Switching for Multi-Granularity Traffic in Data Centers

Qirui Huang, Yong-Kee Yeo, and Luying Zhou
Institute for Infocomm Research, A*STAR, 1 Fusionopolis Way, #21-01 Connexis, Singapore
ghuang@i2r.a-star.edu.sg

Abstract: We propose a composite optical crossconnect by adding a fast wavelength routing module to a MEMS-based switch. This crossconnect efficiently addresses the requirements of burst and circuit switching in data center optical networks.

OCIS codes: (060.4250) Networks; (060.4510) Optical communications.

1. Introduction

Driven by the deployment of many cloud-based applications and services, the exponential growth of data traffic demands powerful data centers that can support hundreds or even thousands of server racks with huge bandwidth and low latency [1]. Although optical fiber links have been widely deployed in modern data centers to enhance the transmission capacity between racks, the switching tasks are performed by electronic switches in which a large number of O/E/O (optical/electrical/optical) converters are used. The large power consumption and heat dissipation caused by electronic switches is one of the factors that limit the evolution of future data centers. Thus optical switching has been proposed to take over some of switching tasks originally performed by electronic switches [2]. Among various rack-to-rack switching solutions, the hybrid electronic/optical switch architecture relies on MEMS (Micro-Electro-Mechanical Systems)-based optical switches for high-throughput connections and electronic switches for low-latency connections [3, 4]. Although this scheme can provide high bandwidth to data center traffic, due to the slow reconfiguration time (typically in 10 to 20 milliseconds) of MEMS-based optical switch, it needs to perform complex traffic monitoring and estimation to batch the traffic for filling the optical links effectively. Our recent study in [5] shows that the switching time of an optical switch can play an important role in the network bandwidth utilization. In order to provide sufficient switching time for multi-granularity traffic in data centers, we propose to implement optical burst switching over circuit switching. In particular, we present a new optical crossconnect design based on a wavelength-assisted switch and a MEMS-based optical switch. Experiments are carried out to investigate the feasibility and dynamic switching performance of this crossconnect.

2. Principle and Architecture

Fast tunable transceivers/lasers combined with arrayed waveguide grating routers (AWGR) is an attractive solution for small granularity switching (burst or packet level) as they can provide reconfiguration time of a few microseconds or faster [6, 7]. However, from implementation point of view, it is costly to deploy fast tunable transceivers/lasers in all the top of rack (ToR) switches compared to the use of fixed-wavelength transceivers/lasers, especially for a data center that contains thousand of ToR switches. From another perspective, a large volume of bursty traffic may not exist in most data centers. A study on data center traffic in [8] found that bursty traffic pattern exists only in a few subsets of the switches while most of the switches have relatively static traffic patterns. Therefore a small number of fast-reconfigurable optical switches could sufficiently support the bursty traffic in most data centers.

Fig. 1. Optical switching scheme in data center.

Fig. 2. Architecture of the proposed BoCS Crossconnect.
Motivated by this, we propose a new optical switching scheme for data centers. As illustrated in Fig. 1, in the lower logical layer, the ToR switches are interconnected through a non-blocking optical switch that is used for the circuit-switched traffic that does not have requirements for fast switching speed; the upper logical layer, which consists of fast-reconfigurable optical switches, will handle the burst-switched traffic from ToR switches. When burst-switched traffic is generated among a subset of ToR switches, it will be routed to one fast-reconfigurable optical switch in the upper layer. Meanwhile, this scheme can be easily upgraded to support more ToR switches for burst switching by adding the fast-reconfigurable optical switches as needed. Therefore it can satisfy the burst switching requirement for the ToR switches and avoid a full adoption of tunable transceivers in a data center.

With this design, we propose an optical burst-over-circuit switching (BoCS) crossconnect. It is a hybrid switch architecture comprising an \(N \times N\) MEMS switch and a \(K \times K\) fast-reconfigurable wavelength-assisted switch. As shown in Fig. 2, the wavelength-assisted switch, which is attached to a subset of ports of the MEMS switch, is composed of \(K\) fast tunable wavelength converters (TWC) and a \(K \times K\) AWGR. The ToR switches are connected to the MEMS switch and each of them has a set of fixed-wavelength optical transceivers (typically 4 \(\times\) 10G SFP+/1 \(\times\) 40G QSFP). A central controller is used to manage the MEMS switch and the wavelength-assisted switch for circuit-switched traffic and burst-switched traffic, respectively. When there is a circuit-switched request between two ToR switches, the controller will simply configure the corresponding mirrors in the MEMS switch to set up a connection between them. When there is a burst-switched request among a group of ToR switches, the MEMS switch just set up connections for these ToR switches to the input and output ports of the wavelength-assisted switch. Then the controller will configure the TWCs in the wavelength-assisted switch to route the incoming optical bursts to appropriate output ports through the AWGR. It should be noted that the wavelength-assisted switch can be shared among all the ToR switches and thus it can support any-to-any switching without needing a port count as large as that of the MEMS switch, i.e., the number \(K\) can be smaller than \(N\). Hence this scheme avoids the requirement for a large number of TWCs. Detailed analysis on the optimal ratio of \(K\) to \(N\) and its influence on network layer performance will be discussed in another publication. This paper primarily focuses on the feasibility and physical layer performance of this crossconnect.

3. Experimental Setup and Results

To validate the feasibility and physical layer performance of the proposed BoCS crossconnect, we conduct a proof-of-concept experiment. The experimental setup is shown in Fig. 3(a). A pulse pattern generator (PPG 1) is used to provide 250-KHz trigger to another PPG (PPG 2) for generating 10-Gb/s burst pattern signals with pseudorandom binary sequence length of 2\(^{31}\)-1. Each burst has fixed temporal length of 3.2 \(\mu\)s and the guard time between two adjacent bursts is set to 0.8 \(\mu\)s. Then an optical transmitter (Tx) (based on DFB laser and LiNbO\(_3\) intensity modulator) converts the electrical bursts into optical bursts and send them to an input of the BoCS crossconnect. As shown in Fig. 3(b), the BoCS crossconnect is composed of an \(8 \times 8\) MEMS optical crossconnect (OMM8x8x-2) and a wavelength-assisted switch formed by a TWC and a \(32 \times 32\) AWGR. (Due to availability of components in our lab, an \(8 \times 8\) MEMS switch has been used, but larger optical MEMS switches are commercially available). After passing through the MEMS switch, the optical bursts go to the TWC where they will be converted to appropriate wavelength so that they can be routed by the AWGR according to their destined output ports. The AWGR has 100 GHz channel spacing and 6 dB insertion loss. The TWC is realized based on cross gain modulation (XGM) in counter-propagating configuration using semiconductor optical amplifier (SOA) (CIP SOA-XN-OEC-1550). A fast tunable laser module (Syntune S3500) reported in [6] is used to provide the probe. To perform burst switching function, the switching time of the tunable laser is set to 0.8 \(\mu\)s and a personal computer (PC) is employed as central controller synchronized with PPG1 to configure the MEMS switch, tunable laser module and the SOA.

![Fig. 3. (a) Experimental setup; (b) picture of the BoCS crossconnect](Image)
First of all, we examine the burst switching function of the BoCS crossconnect. The wavelength of the input optical burst signals is 1550.12 nm with 0 dBm average power. To simplify the operation, the controller will periodically tune the wavelength of the tunable laser among four output channels of the AWGR (Ch 1=1535.04 nm, Ch 9=1541.35 nm, Ch 24=1553.33 nm and Ch 32=1559.79 nm) so that individual burst is converted to these channels sequentially and then routed to corresponding output ports of the AWGR. From the temporal waveform of these signals shown in Fig. 4(b), it is clearly seen that the input burst signals are successfully switched to corresponding output ports with the same repetition period of 16 µs. This confirms correct burst switching function in the BoCS crossconnect.

Then the physical layer performance of the BoCS crossconnect is examined. Fig. 5(a) compares the eye diagrams between the input and output signals. There is no obvious distortion in eye diagram at all of the four channels. The bit error rate (BER) curves of the input and output signals are plotted in Fig. 5(b). It is shown that there is a discrepancy in power penalty among the output channels. This is due to the different wavelength conversion efficiencies in the TWC. The largest power penalty for 10^-9 BER is 2.5 dB at Ch 32 while the smallest is 1.8 dB at Ch 1. Hence, there is no significant degradation on physical layer performance in the BoCS crossconnect.

Fig. 4. (a) Spectrum of input and output signals; (b) traced waveforms of input and output signals.

Fig. 5. (a) Eye diagrams of input and output signals; (b) BER performance of the crossconnect.

4. Conclusions

We propose a solution to enhance the performance of MEMS-based optical crossconnect to allow it to support burst-switched traffic. By adding a wavelength-assisted switch to the ports of a MEMS-based switch, we show that the composite crossconnect can support a switching speed of 0.8 microseconds. Using this solution, it is now possible to provide optical burst switching services to all ToR switches without requiring a very large optical switch fabric that has fast switching time.

This work is supported by Singapore-Japan Joint research grant (A*STAR-JST) under Contract No. 1021630070.

References