Experimental Demonstration of Wavelength-Reconfigurable Optical Packet- and Circuit-Switched Platform for Data Center Networks

Wenjia Zhang*†, Dawei Liu*, Howard Wang* and Keren Bergman*
*Department of Electrical Engineering, Columbia University, 1300 S. W. Mudd, 500 West, 120th St., New York, New York 10027, USA
† State Key Lab of Information Photonics and Optical Commun., Beijing University of Posts and Telecom, Beijing, 100876, China
wz2207@columbia.edu

Abstract: We demonstrate a wavelength-reconfigurable optical packet- and circuit-switched platform for data center networks based on a modular optical switch fabric prototyping system.

1. Introduction

The increasing popularity of cloud computing applications is driving the creation of more powerful data centers (DCs). As these cloud-based applications evolve, conventional over-subscribed hierarchical electronic DC networks, based on commodity Ethernet switches and IP routers, will not be able to satisfy the massively increasing bandwidth requirements of these applications without significant growth in power consumption and cost. As a result, a high-bandwidth and "all-servers-equidistant" network is highly desirable for modern data center designs to overcome these emerging communication bottlenecks and workload placement constraints, which have already restricted data-intensive software models (*e.g.* MapReduce) from being applied across the entire DC. Furthermore, this type of DC network design would effectively improve server utilization and therefore significantly reduce power consumption of the overall system.

Optical interconnects featuring wavelength-division-multiplexing (WDM) and high-radix switching is an attractive candidate technology for high-performance next-generation switch designs. However, a key challenge for photonic interconnects in DCs is in deploying commercially available optical technologies to provide multi-granularity interconnectivity for heterogeneous data applications. Existing research, such as Helios [1] and C-through [2], which have introduced optical switching along with electronic switching into a hybrid DC network design, have achieved high-bandwidth communication using MEMS – a device with relatively slow reconfiguration times (10s of ms) – but have also shown limited benefits when implementing this system into the real DCs [1-2]. A recently proposed architecture combining both wavelength selective and space switching allows the capacity of the each connection to be varied from a few Gb/s to a hundreds of Gb/s on demand [3]. While this architecture is highly flexible in terms of bandwidth allocation, it suffers from slow switching speeds for link establishment.

In this paper, we propose a wavelength-reconfigurable optical packet- and circuit-switched network architecture enabling multi-granular interconnectivity at nanosecond-scale switching speeds. A testbed is constructed on a modular optical switch fabric prototyping platform, supporting technologies enabling diversified switching speeds (from nano-second to second scales) and switching functionalities (packet switching and circuit switching). We verify error-free routing and transmission across a prototype utilizing wavelength selective switches (WSSs) and semiconductor optical amplifiers (SOAs), illustrating the potential to support dynamic wavelength and sub-wavelength bandwidth granularities with different QoS levels in a flexible and programmable fashion.

2. Modular optical switch fabric prototyping platform

In order to facilitate increased functionality and to provide support for future switching technologies in a straight-forward way, we design and develop a modular optical switch fabric prototyping platform (Fig. 1). This platform consists of a 12''×12'' mainboard (Fig. 1(c)) featuring a FPGA; eight SOA-driver daughter boards (Fig. 1(d)), each containing driver circuitry for two SOAs; eight optical receiver daughter boards (Fig. 1(e)), each supporting two optical receivers; and one peripheral Ethernet interface. All the daughter boards are connected to the FPGA (Xilinx Virtex 5), which serves as a central controller providing the desired switching functionality, through sockets on the mainboard. This platform integrates the design of the hybrid packet and circuit switched network [4] to create a versatile testbed that can potentially provide more functionalities. Moreover, the modular design enables easy adaptability to more advanced switching devices, such as PLZT.

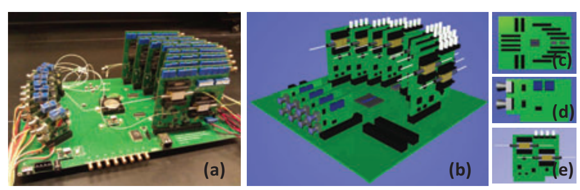

Fig.1 (a) Picture and (b) design schematic of modular optical switch fabric prototyping platform, consisting of (c) mother board; (d) optical receiver card; (e) optical switching card.

3. Optical Switching Architecture

In this work, we propose a wavelength-reconfigurable optical packet- and circuit-switched network to provide bandwidth control at varying levels of granularity. A 2×2 switching module is illustrated in Fig. 2 that primarily consists of two WSSs for wavelength selection, and the aforementioned optical switch platform enabling packet and circuit switching [4]. In this design, wavelength-striped packets with bit-parallel headers are individually routed depending on the encoded address [5]. When a long-lived and high-QoS application occurs, a pre-selected combination of wavelengths, previously utilized by packet payloads, are re-allocated for circuit paths by appropriately configuring the WSS to route these wavelengths to the circuit switching subsystems. The packet- and circuit-switching ports are both managed by the FPGA, in a manner similar to that in [4]. As a result, this architecture can not only route optical packet and circuit traffic simultaneously, but also support wavelength reconfigurability using the WSS.

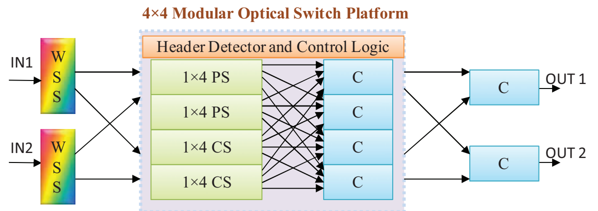

Fig.2 The implementations of a 2×2 wavelength-reconfigurable optical packet- and circuit-switched architecture. (WSS: wavelength selective switch; PS: packet switch; CS: circuit switch; C: coupler)

4. Experiment and Results

In order to validate the feasibility of our proposed design, we construct a testbed comprised of our modular switch platform as an optical space switch and an Optoplex 3-port tunable optical add/drop multiplexer (TOADM) as a WSS. As shown in Fig. 2, the modular switch platform is configured into a 4×4 SOA-based broadcast-and-select optical space switch, in which the top two input ports support optically-addressed packet switching while the remaining two input ports support circuit switching. Eight C-band 100 GHz-spaced continuous-wave (CW) lasers, with wavelengths ranging from 1542.94nm to 1548.51nm, are combined using a coupler and simultaneously modulated with a 10-Gb/s, 2^{31}-1 pseudorandom bit sequence (PRBS) data, which is subsequently decorrelated by 10-km SMF. The control wavelengths – including one frame and two headers – are gated independently into a 89.6-ns long optical packet with a 102.4-ns period using SOAs driven by an Agilent ParBERT and combined with data wavelengths to create wavelength-striped optical packets. The modulated data are injected into the TOADM and then switched using a SOA-based switching fabric. Optical circuits using the same wavelengths are modulated and transmitted in the same way. At the output of the network, one of the 10-Gb/s channels is filtered by a tunable filter, amplified by an erbium-doped fiber amplifier (EDFA) and then followed by a tunable filter. The filtered signal passes through a variable optical attenuator (VOA) and is received by a PIN receiver with transimpedance amplifier (TIA) and limiting amplifier (LA), and subsequently examined on a Bit Error Rate Tester (BERT).

First, we confirm that the 8×10-Gb/s optical payloads are correctly routed through the switch fabric. We then experimentally confirm that all the egressing optical packets and circuits achieve error-free transmission (defined as BERs less than 10^{-12}). Example eye diagrams of a representative channel (C42, λ= 1543.73nm) from both the input and output of the switch are recorded in Fig. 3(a-b). The BER curves in Fig. 4(a) show an approximately 0.8-dB power penalty for the switch fabric. In order to demonstrate correct routing functionality under circuit switching, 10-Gb/s input and output eye diagrams are recorded in Fig. 3(c-d) and sensitivity curves for this channel show a power penalty of 0.6-dB for this switch fabric. Optical spectrum of both the circuit switched channel and the packet switched payload wavelengths after the TOADM are recorded in Fig. 5. The additional power penalty incurred by packet switching is higher than circuit switching at 10-Gb/s because the multi-wavelength packets suffer from waveform distortion due to carrier density variation and saturation effects caused by SOA switches.

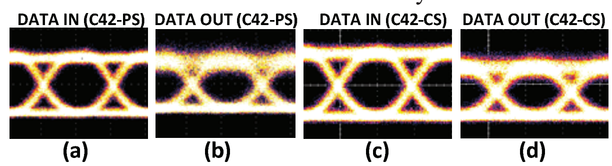

Fig.3 10-Gb/s eye diagrams of input and output of optical packet and circuit of C42 (1543.73nm). PS: packet switch; CS: circuit switch.

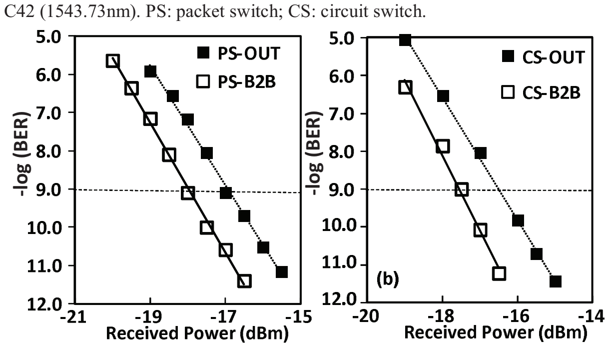

Fig.4 BER curves for the (a) Packet switch (PS); (b) circuit switch (CS)

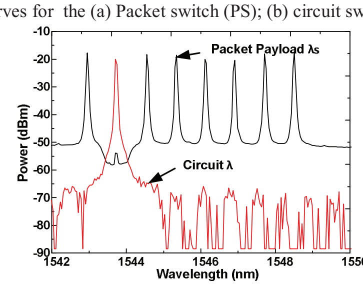

Fig.5 Spectrum of packet payload channels and circuit channel

5. Conclusion

We propose a wavelength-reconfigurable optical packet- and circuit-switch for high-performance computing systems. We experimentally demonstrate a hybrid optical switching testbed that supports error-free routing of 8×10-Gb/s wavelength-striped packets and 10-Gb/s circuits. This demonstration of high-speed optical packet and circuit switching opens the door for deploying commercially available optical technologies to provide optical multi-granularity interconnectivity for heterogeneous data applications in future high-performance computing systems.

The authors gratefully acknowledge support for this work from the National Science Foundation under Award 0811012.

5. References

1. N. Farrington, et al, Proc. of ACM SIGCOMM '10, pp. 339 –350.
2. G. Wang, et al, Proc. of ACM SIGCOMM '10, pp. 327-338.
3. A. Singla et al, Proc. of ACM Hotnets '10, article 8.
4. H. Wang, et al, Proc. of OFC 2010, OTuP3.
5. A. Shacham, et al, IEEE PTL 17 (12) 2742-2744 (Dec 2005).