Comparative Studies of All-Optical vs. Electrical vs. Hybrid Switches in Datacom and in Telecom Networks

Xiaohui Ye, Venkatesh Akella, and S. J. B. Yoo
Department of Electrical and Computer Engineering, University of California, Davis, California 95616, USA.
Email: {xye, akella, sbyoo}@ucdavis.edu

Abstract: This paper compares all-optical, electrical, and hybrid switches in the context of datacenter networks and telecommunication networks. Latency and throughput simulations indicate the advantages of hybrid switches over optical and electrical switches.

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

1. Introduction

The Future Internet architecture and technology must support exponentially growing traffic with various quality of service (QoS), class of service (CoS), and type of service (ToS) on a unified platform. The core of today's Internet is supported by relatively static over-provisioned connections of nodes using the wavelength division multiplexing (WDM) technology. As data traffic continues to surge exponentially, the dynamically changing traffic pattern and the exponentially growing traffic capacity are pressing the need for more agile and scalable optical networking to more efficiently accommodating the traffic. In addition to such trends in the 'telecom' networks, the rapid increases in data traffic have also given birth to many large data centers. Today's data centers have grown to many thousand racks consuming nearly 10 MegaWatts of power. Given this power limit, the data centers can achieve much higher productivity if more optical switching with parallelism is exploited. However, today's data centers utilize mainly Ethernet or Infiniband [1] based electronic switches that cost large latency and low throughput. Recently, all-optical and hybrid (optical and electrical) switching technologies have been introduced for both telecom and data center networks to effectively handle data packets. This paper reviews and compares all-optical and hybrid switches in light of state-of-the-art electronic switches.

2. Telecom Switches: optical vs. electrical vs. hybrid

In the national and global telecom backbone network, nodes typically have a degree of connections from 2 to 5. Although the degree of connections is small, the link between two nodes has a high capacity in the range from hundreds of gigabits per second to terabits per second. The high capacity link of the backbone network is realized through DWDM technique, where tens of wavelengths are transmitted in parallel on a fiber link. An end-to-end path usually contains multiple hops and travels through several backbone nodes. The end-to-end latency builds up as the number of hops increases. Unlike in datacom switches to be discussed later, typical telecom switches in the core will be surrounded by the edge routers, where traffic shaping and packet aggregation functions are realized to make the link utilization of the core network more efficient. Hence, even if the core switches are all-optical, they can benefit the electronic processing at the edge.

The electrical core switch uses an electrical switching fabric and requires O/E and E/O conversion at input and output ports. All incoming packets will be first converted from optical signals to electrical signals and stored in the buffer. The control plane then schedules forwarding of the packets from the input queue to the output queue. The aggregate switching capacity of the switching fabric can be extremely high even if most of the packets are through traffic. Such a high capacity electrical switch consists of a combination of many smaller electrical switches and serial-to-parallel/parallel-to-serial converter stages. As each packet must undergo many stages of electrical processing, the latency and power consumption can be high, however the packet loss can be low if buffers are sufficiently large. In electrical switches, contention resolution relies on holding the packets that have no access to the switching resource in the buffer.

In the all-optical switch [2] as shown in Figure 1, only the control plane will be in the electrical domain while the data plane will remain in the optical domain. The all-optical switch can exploit optical parallelism in the wavelength domain to support very high port count beyond thousands. Non-blocking switching and speed up can be easily achieved by the arrayed waveguide grating router (AWGR) as parallel paths use different wavelengths. At the input of the switch fabric, tunable wavelength converters (TWCs) are placed thus enabling wavelength routing from any input to any output. Fixed wavelength converters (FWCs) are also placed at the output of switching fabric to keep the wavelengths used in the output link the same as that of the input link. Although no mature optical buffer is available today, all-optical switches can effectively use all-optical contention resolution in the wavelength, time (the loopback fiber delay lines, FDLs), or space (deflection routing) domain [2]. Since there is no store-and-forward or bit-by-bit processing, the all-optical switch is more power efficient compared with the electrical switch. For wavelength routing, the AWGR switching fabric is passive and consumes no power.

The hybrid optical switch is similar to the all-optical switch except the loopback FDLs are replaced by the loopback shared buffer system, as shown in Figure 2. The loopback shared buffer system only need to occupy one input and one output of the switching fabric instead of *w***m* inputs and outputs in the all-optical switch, where *w* is the total number of wavelengths on each FDL and *m* the number of loopback FDLs. Like the all-optical switch, the control plane first checks the availability of other wavelengths on the same output upon contention happens. If all wavelengths are occupied, the packet will be sent to the loopback shared buffer and wait for retransmission. If the shared buffer does not have space, the packet will be dropped. The loopback shared buffer outperforms loopback FDLs as it occupies fewer ports of the switching fabric and provides arbitrary delay, allowing the delayed packet be sent out immediately after the resource is available. Although the loopback shared electronic buffer consumes power, the size of the loopback shared buffer can be kept small exploiting statistical multiplexing of the packets.

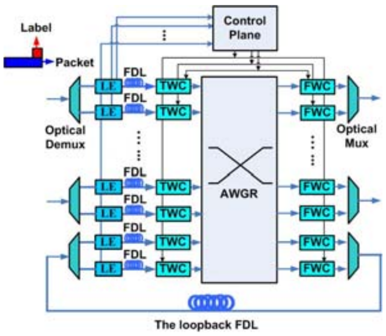

Figure 1 The all-optical switch for telecom backbone network

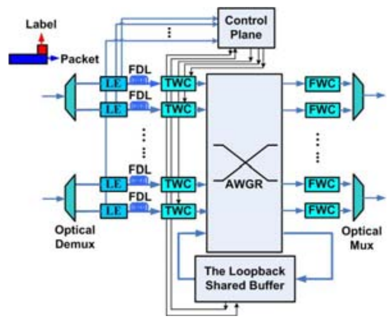

Figure 2 The optical hybrid switch for telecom network

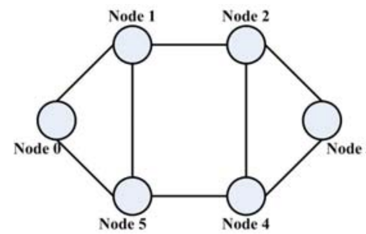

Figure 3 Six-node topology for simulation

Figure 3 shows the 6-node simulation topology for the performance study. Figure 4 and 5 show the latency and packet drop rate comparison for the all-optical switch versus the optical hybrid switch. Compared to the all-optical switch, the optical hybrid switch can achieve slightly lower end-to-end network latency since the loopback shared buffer can have shorter delay than FDL when the resource becomes available. Moreover, the optical hybrid switch has a lower packet drop rate since the electrical buffer can hold the packet until the resource becomes available.

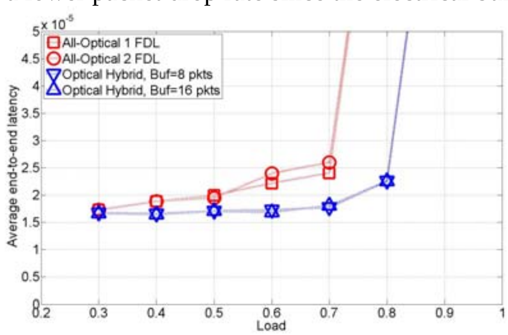

Figure 4 The end-to-end latency comparison for different switches in telecom network

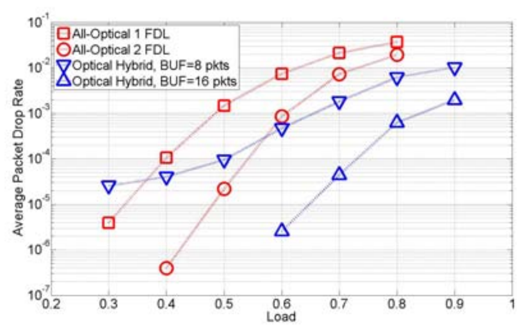

Figure 5 The packet drop rate comparison for different switches in telecom network

3. Data Center Switches: optical vs. electrical vs. hybrid

In contrast to telecom backbone network switches, data center switches usually need to connect hundreds or even more end nodes. In addition, data center networks have a strict requirement on latency. As more and more parallel computing based applications used in the data center network, several hundred nanoseconds or even tens of nanoseconds latency difference will critically affect the performance of application.

A single electrical switch cannot easily scale to hundreds ports, the current design trend uses many cheap electrical switches with small port count to build a switching network [3] that can connect to hundreds or even more nodes. Although latency experienced in each switch is small, the accumulated end-to-end latency across multiple hops may still become significantly large. In the following performance study, we will use the electrical flatten butterfly network [4] as a representative for electrical switches. In comparison, optical switching fabric, like AWGR, can support much higher port count than an electrical switching fabric, so that a single stage optical switch can connect to more end nodes than a single stage electrical switch. Although multi-stage switching system is still needed when we want to connect thousands or even more nodes, the number of switches required will be much smaller and the network topology will be much simpler. As the optical switch does not use store and forward mechanism, packets that do not experience contention will have very small switching latency. The latency

OThQ4.pdf

experience by packets that face contention depends on how fast the requested resource is available.

Figure 6 shows the architecture for an optical hybrid switch designed for data center network [5]. Compared with the optical hybrid switch for telecom network discussed above, the switch used for data center network does not need FWCs on the switching fabric output side. More importantly, the switch will have asymmetric link capacity for the incoming and outgoing link, thus further exploiting benefits from wavelength parallelism. Figure 7 shows an all-optical switch that replaces the shared loopback buffer in Figure 6 with multiple FDLs.

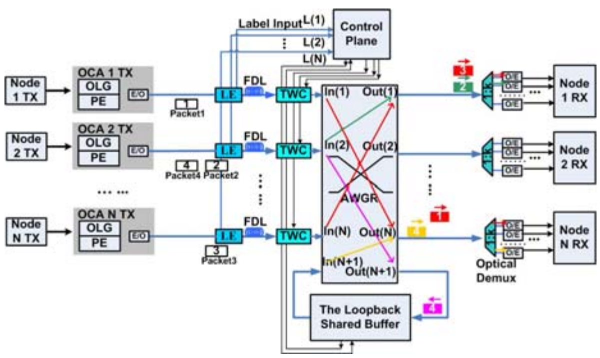

Figure 6 The optical hybrid switch (DOS) for the data center network

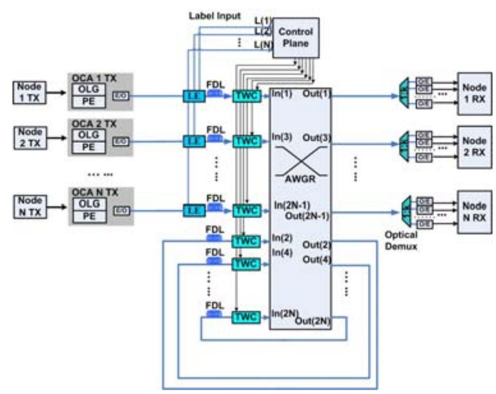

Figure 7 The all-optical switch for the data center network

Figures 8 and 9 show the latency and the effective bandwidth comparison for different switches. The simulations focused on a 128-node network. The electrical flatten butterfly network has the worst performance compared to the other two optical switches. The latency increases faster and the effective bandwidth becomes saturated even the load is not heavy. On the other side, both all-optical switch and optical hybrid switch have very low latency for most of configurations. And the effective bandwidth increases as the load increases except for all-optical switch with k=1 (no redundancy at the output receiver) at high load range. The hybrid optical switches can achieve the best performance in terms of latency for the same reason as described for the telecom switches.

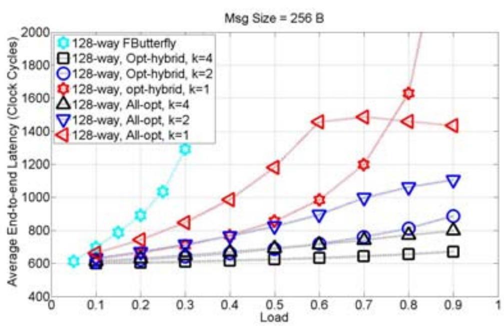

Figure 8 The end-to-end latency comparison of different switches for the data center network

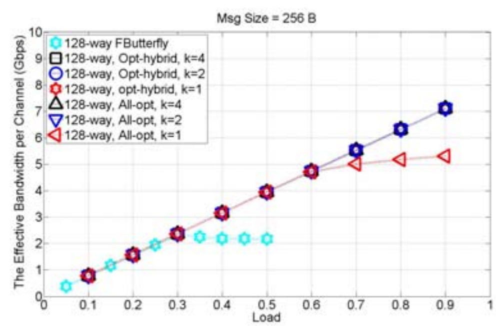

Figure 9 The throughput comparison of different switches for the data center network

4. Conclusion:

This paper compared all-optical, electrical, and optical hybrid switches in both telecommunication and data center networks. Optical switches outperform electrical switches in terms of latency, power, and throughput. The optical hybrid switch can achieve the lowest latency as arbitrary delay is available but it may consume more power than the all-optical switch.

References :

[1] InfiniBand Architecture Specification, Volume 1, Release 1.0. http://www.infinibandta.org/specs.

[2] Haijun Yang and S. J. Ben Yoo, "All-Optical Variable Buffering Strategies and Switch Fabric Architectures for Future All-Optical Data Routers", in Special Issue on Optical Networks, IEEE/OSA Journal of Lightwave Technology, Vol. 23, No. 10, pp. 3321-3330, October, 2005.

[3] M. Al-Fares, A. Loukissas, and A. Vahdat, "A Scalable, Commodity Data Center Network Architecture," in SIGCOMM'08, August 2008.

[4] J. Kim, W. J. Dally, and D. Abts, "Flattened butterfly: a cost-efficient topology for high-radix networks," in the 34th annual international symposium on Computer architecture (ISCA), 2007, pp. 126 - 137.

[5] Xiaohui Ye, Paul Mejia, Yawei Yin, Roberto Proietti, S. J. B. Yoo, and Venkatesh Akella, "DOS - A scalable Optical Switch for Datacenters", in ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS) 2010, October, 2010.

This work was supported in part by DoD contract #H88230-08-C-0202.