All-Optical Token Technique for Distributed Contention Resolution in AWGR-based Optical Interconnects

Roberto Proietti, Christopher Nitta, Yawei Yin, Venkatesh Akella, and S.J.B. Yoo
Department of Electrical and Computer Engineering, University of California, Davis, California 95616
Author e-mail address: rproietti@ucdavis.edu

Abstract: This paper studies the networking performance of a 10Gb/s 64-port AWGR-based optical switch with distributed all-optical contention resolution based on saturation effect in R-SOAs. Significant reduction in latency compared to FBF architecture is observed.

OCIS codes: (200.4650) Optical Interconnects; (200.6715) Switching.

1. Introduction

Optical interconnects have emerged as a promising method to realize high-port count, low-latency, and high-throughput networks in high-performance computing systems and data centers. Several research projects have already proposed architectures for optical interconnects [1-3]. A bottleneck to the scalability of these architectures can arise from the centralized control plane, which can limit the optical switch port count due to the maximum number of I/O resources of currently available ASICs. In [4] we reported a first-time demonstration of an all-optical token technique for distributed contention resolution in AWGR-based interconnects. This paper analyzes the networking performance of a 10 Gb/s 64-node interconnect architecture based on the technique in [4] and compare with the performance of the optical switch in [5] and an electrical flattened butterfly (FBF) interconnect network.

2. Working principle of the all-optical token technique and architecture

Fig 1a shows the interconnect architecture. An NxN AWGR is at the core of the system. Each input port is connected to a line-card (Line-card_i TX). Each line-card receives packets from a host (H_i), which buffers them in an input queue (I-Q). The packets are then transmitted in the optical domain by means of a transmitter equipped with a fast tunable laser (TL) [6]. A polarization-diversified (PD) scheme can be used to avoid interference between the token-based control plane messages and data packets. An optical circulator (C) is used to extract the counterpropagating token. Each AWGR output is then connected to a polarization beam splitter (PBS) to separate the token and data path. One PBS output connects to a reflective SOA (R-SOA). The other PBS output connects to a linecard (Line-card_i RX), which buffers the packets in an egress queue (E-Q) and transmits them to the destination. Fig.1b explains the working principle of the token technique. At t_2, Node1 needs to send a packet to output N.

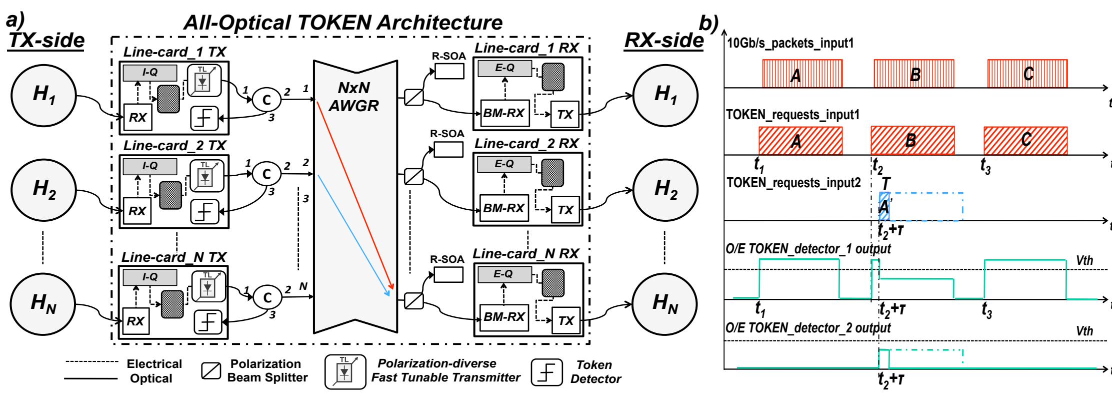

Figure 1. An All-Optical Token Technique for Contention Resolution in AWGR-based Optical Interconnects; An All-Optical Token

Then it tunes its TL to λ_{1N} to generate token request B. The RSOA at output N reflects the signal extracted by the PBS, which reaches the linecard1 TD with an optical power $P_{TO1.}$ The TD generates an electrical signal with V_p= V_{TO1}. At $t_2 + T$, Node2 has also a packet for output N and it tunes its TL to λ_{2N} to generate the token request A′. RSOA reflects the token request extracted by the PBS, which reaches the linecard2 TD with an optical power $P_{TO2.}$ The TD generates an electrical signal with V_p= V_{TO2}. Let us assume that the R-SOA works in strong saturation regime. At $t = t_2$, P_{TO1}= X dBm and V_{TO1} > Vth. A comparator in the TD generates a logic "1" and Node1 gains the token and starts the transmission of packet B on a polarization orthogonal to the token request. At $t = t_2 + T$, due to the saturation effect in the R-SOA, P_{TO2}= (X-3) dBm. Then $V_{TO2} = V_{TO1}/2 < V_{th}$. The token is not available until Node1 finishes to transmit (linecard2 needs to retry with a random backup time). Then Node2 has to stop the token request A′. Note that the voltage drop in TD1 O/E output at $t = t_2 + T$ can be used to sense whether or not other nodes need to access the same output. This information could be used to implement a method of a fairness protocol.

3. Performance Analysis

We developed a cycle-accurate architecture level simulator with user selectable parameters such as: the number of nodes in the network (N), the number of wavelengths per output port (k), the data rate, and distance between the switch and nodes. The parameters assumed for the simulations were *N*=64, *k*=4, 10Gbps data rate, 10m switch distance. Note that, in order to implement the proposed architecture with *k*=4, it is necessary to use 4 RSOAs at each AWGR output port. The 4 RSOAs connects to each AWGR output port through a 1:4 optical demux, as explained in [3]. The packet size was assumed to be 256B (corresponding to a cache line), and a Bernoulli distribution was used to determine when a packet should be injected. The baseline architecture against which we compare is the AWGR-based switch with Distributed Loopback Buffer (DLB) described in [5]. We also compared the performance of RSOA-based scheme to a state-of-the-art electrical switch based on the Flattened-ButterFly (FBF). Uniform Random and Hot-Spot (all packets are destined to the *N*/2 node) traffic patters were used. The throughput and latency results of the synthetic traffic patterns are shown in Fig.2 Fig.2a shows the throughput as a function of offered load (in GB/s) for the Uniform Random traffic. The RSOA (AO-TOKEN) performs quite well compared to the FBF, saturating at ~60GB/s or capable of ~150% of the load of the FBF. When compared to the DLB, the RSOA-based architecture (AO-TOKEN) comes within 17% of the DLB. Note that, despite DLB performs better, it relies on a centralized control plane that strongly limits the switch port count scalability [5]. Fig. 2b shows the average packet latency in nanoseconds (ns) as a function of offered load. Again the performance of RSOA is between that of the FBF and the DLB. The Hot-Spot throughput results are shown in Fig. 2c and 2d. Note that the offered load is limited to 5GB/s (four wavelengths at 10Gbps). The AO-TOKEN achieves ~90% of the throughput of the DLB on Hot-Spot traffic, which impressive result considering that the Hot-Spot pattern represents the worst-case situation (maximum number of nodes attempting to acquire a single token). Notice that the AWGR-based switches outperforms the FBF due to the multiple receiver groups of *k*

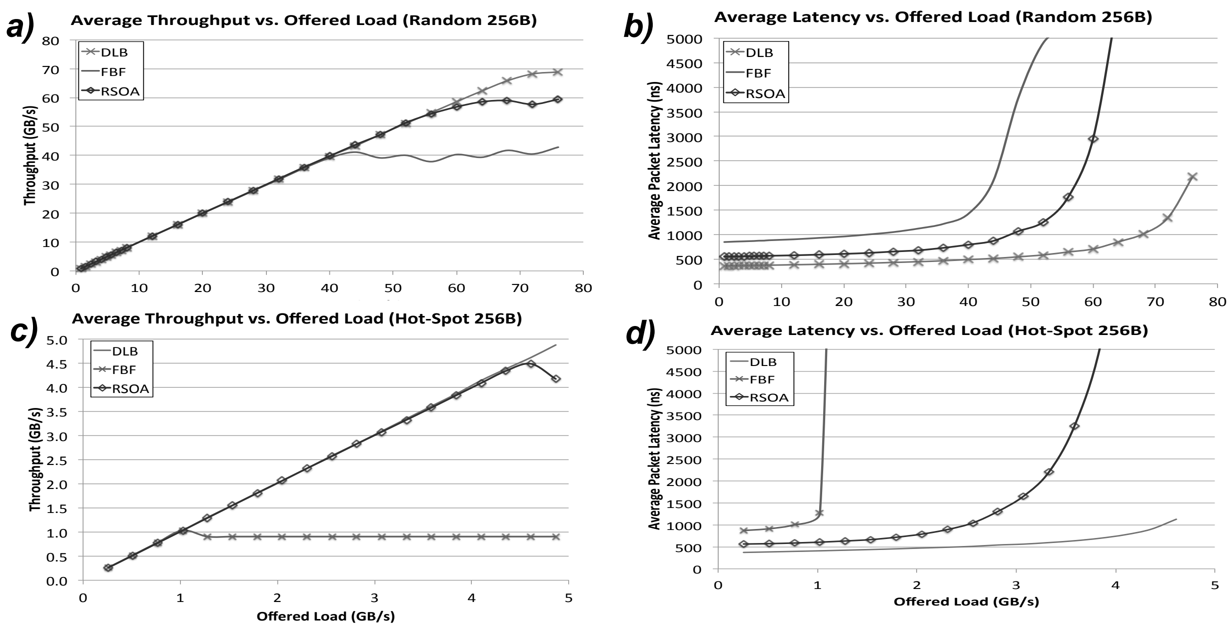

Figure 2. DLB, FBF, and RSOA (AO-TOKEN). a) Throughput vs. Offered Load for 256B Packets on Uniform Random Traffic. b) Average Packet Latency vs. Offered Load for 256B Packets on Uniform Random Traffic. c) Throughput vs. Offered Load for 256B Packets on Hot-Spot Traffic. d) Average Packet Latency vs. Offered Load for 256B Packets on Hot-Spot Traffic.

4. Conclusions

We studied the performance an AWGR-based optical interconnects with distributed all-optical contention resolution. Due to the limited space, additional details about the retransmission policy, performance dependence from the packet size, scalability in terms of port count and more details about DLB and FBF architectures used for comparison will be discussed at the conference.

5. References

[1] R. Hemenway, "Optical-packet-switched interconnect for supercomputer applications," *Journal of Optical Networks,* (2004).
[2] O. Liboiron-Ladouceur, "The Data Vortex Optical Packet Switched Interconnection Network," *Journal of Lightwave Technology*, July 2008.
[3] X. Ye, "DOS - A scalable Optical Switch for Datacenters," ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS), 2010.
[4] R.Proietti, "All-Optical Token Technique for Contention Resolution in AWGR-based Optical Interconnects," in *Conference on Lasers and Electro-Optics (CLEO)*, Paper CM2A.7, May, 2012.
[5] X. Ye*, et al.*, "Buffering and Flow Control in Optical Switches for High Performance Computing," *Optical Communications and Networking, IEEE/OSA Journal of,* vol. 3, pp. A59-A72, 2011.
[6] J.E. Simsarian, "Fast switching characteristics of a widely tunable laser transmitter," *IEEE Photonics Technology Letters,* August 2003

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