An Enhanced Buffered Switching Node for a Data Vortex Interconnection Network

Assaf Shacham and Keren Bergman
Department of Electrical Engineering, Columbia University, New York, NY 10027; assaf@ee.columbia.edu

Introduction

The data vortex is a topology for optical packet switched interconnection networks designed to provide ultra-high bandwidth at very low routing latencies, as well as scalability to a large number of ports [1]-[2]. Data vortex based interconnection networks are therefore especially suitable for high performance computing applications [1] as well as for the implementation of storage and local area networks.

Typically based on semiconductor optical amplifier (SOA) gates, the original data vortex architecture utilizes bufferless photonic switching nodes, and thus uses deflection routing as the sole mean of contention resolution in the network. Adding buffering capacity to the switching nodes has been proposed in the past by the authors and a switching node based on a fiber delay line (FDL) has been demonstrated [3]. This design, however, led only to a minor improvement in the network performance. In this paper we revisit the problem and suggest an improved design for a data-vortex-compatible buffered switching node. Computer simulations show improved performance in networks comprised of the new switching node compared to the one reported in [3] and to the original data vortex [4]. A switching node is assembled using commercially available components and the feasibility of implementation is experimentally demonstrated.

Architecture Overview

A data vortex network is a multistage interconnection network, comprised of 2×2 bufferless photonic switching nodes. Optical packets progress from stage to stage according to their optically encoded header. When a contention prevents a packet from progressing to the next stage in its path, the packet cannot be buffered in the switching node and is transmitted to another switching node within the same stage, according to a practice known as *deflection routing* [5]. Additional switching nodes are placed at every stage to ensure the existence of deflection paths. A set of electronic cables connect the nodes, facilitating the transmission of signals encoding node availability information. The additional nodes make each stage essentially a circle and the network, therefore, has a cylindrical structure and is identified by two parameters: H (height), and A (angles). For more information on the data vortex topology the reader is referred to [1] and [2].

According the deflection routing scheme, packets continue to traverse each stage until they can progress to the next one, while preventing other packets from progressing and creating backpressure in the network. This undesired feedback effect may increase the congestion of the network, and is a well-known shortcoming of deflection routing networks [5].

In order to mitigate this undesired effect, we have proposed in [3] a new design for the switching node which included the addition of a recirculating FDL buffer to each switching node. Whereas that simpler design, dealing with a single packet per timeslot, had a lower hardware complexity, it failed to provide the desired performance boost, as was concluded in [3].

In the following section we describe a different approach to the buffered switching node: the node still contains an recirculating FDL buffer, but in the current design a larger photonic switching element is used which can deal with two packets at the same time.

Switching Node Functional Description

As defined in the data vortex architecture, the switching node has two input ports (*North* and *West*) and two output ports (*South* and *East*). While the *East* port is always available for packets, the *South* port may be blocked at some timeslots, as indicated by an electronic control signal, C_{IN} (1=blocked, 0=available). The node also generates a control signal, C_{OUT}, switching to logical 1 when a packet is routed *East*, to be used in the same manner by another node. At any timeslot the node may receive a single packet from one of its input ports. The optical header of the packet is extracted using optical couplers, WDM filters, and photodetectors and the packet is then forwarded to one of the output ports or, in case the requested output port is blocked, stored on the FDL buffer, to be used in the next timeslot. At the subsequent timeslot, for example, a new packet may be received and the node deals with the new packet and the stored packet at the same time. A schematic diagram of the node is given in Fig. 1.

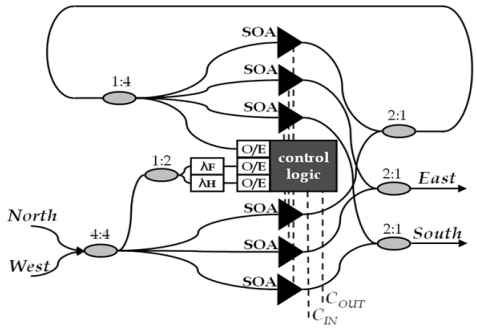

Fig. 1: Schematic diagram of the switching node

The decision rule which controls the optical switch is computed as follows: from the relevant bit in the packet's optical header, the requested output port (ROP) is derived. The ROP of each of the packets (ROP_{IN} for

the incoming packet and ROP_{FDL} for the previously buffered packet), and C_{IN} are used to compute a selected output port (SOP) for each packet (SOP_{IN} and SOP_{FDL}). The SOPs are used to control six SOA gates to forward each packet to its destination. The decision rule is described in table 1, where illegal and redundant logical states have been removed and combined, respectively.

Table 1: Switching node decision rule. *E*-East, *S*-South, ϕ-don't care condition.

INPUTS			OUTPUTS		
ROP_{IN}	ROP_{FDL}	C_{IN}	SOP_{IN}	SOP_{FDL}	C_{OUT}
–	–	ϕ	–	–	0
–	*S*	0	–	*S*	0
–	*S*	1	–	*B*	0
E	–	ϕ	*E*	–	1
E	*S*	0	*E*	*S*	1
E	*S*	1	*E*	*B*	1
S	–	0	*S*	–	0
S	–	1	*B*	–	0
S	*S*	0	*B*	*S*	0
S	*S*	1	*B*	*E*	1

An interesting observation can be made from table 1: Since a packet with ROP=*E* is never forwarded to the FDL, then ROP_{FDL}=*S* for any packet on the FDL, and there is no need to recover the packet header, just its existence on the FDL. This observation reduces the hardware complexity, as some of the detection hardware can be removed.

Performance analysis

One of the key metrics to evaluate the performance of a data vortex interconnection network is the acceptance rate: the fraction of successfully injected packets to the total number of injection attempts. The acceptance rate of the data vortex in three configurations is measured, using computer simulations: (1) original bufferless switching nodes [4]; (2) low-complexity FDL-based switching nodes [3]; and (3) the buffered switching node presented in this paper. All simulations are run for an 80-port data vortex (H=16, A=5).

In the simulations, another question is investigated: whether it is beneficial to limit the number of FDL hops in each node and use deflection routing to route packets exceeding this limit. The simulation results are plotted in Fig. 2.

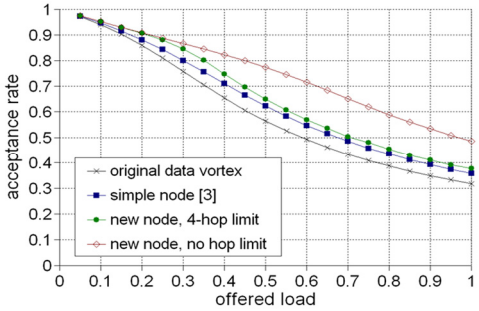

Fig. 2: Performance simulation results.

Two immediate conclusions can be extracted from Fig. 2: (1) a significant (nearly 50%) improvement in the acceptance rate, compared to the original data vortex, can be attained be using the new switching node; and (2) the highest acceptance rate is reached when no limit is set on the number of FDL hops in each node.

Experimental Switching Node

An experimental switching node comprised of a SOAs, photodetectors, and passive optical elements is assembled, according to the diagram in Fig. 1, to demonstrate the feasibility of implementation. The decision rule is implemented using a Xilinx complex programmable logic device (CPLD) controlling the SOAs.

Optical packets, 45 ns long, with 10 Gb/s payload and an optically encoded header, are injected into the switching node through one of the input ports (since a only a single packet may be received from both input ports in each timeslot, there is no loss of generality), and C_{IN} is toggled to emulate alternating availability of the *South* port. Both output ports of the switching nodes are directed to a communications signal analyzer (CSA) and the waveforms (packets, headers, and electronic control signals for inputs and outputs) are shown in Fig. 3. The node is configured so that when the header bit is logically high, the ROP is *South*, and *East* otherwise.

The arrows on Fig. 3 connect the same packets between the input and output ports. In this experiment, two packets are buffered when the *South* port is blocked.

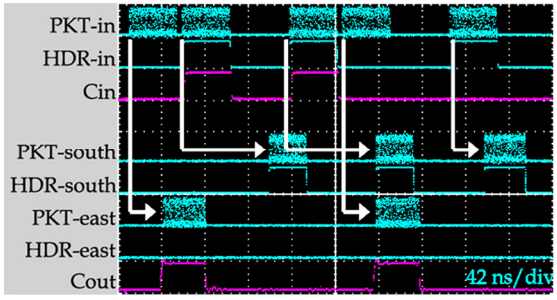

Fig. 3: CSA waveforms demonstrating correct switching node functionality and packet buffering.

Conclusions

The design of a buffered photonic switching node for a data vortex network is presented. Superior network performance is shown through simulations, and implementation is experimentally demonstrated.

This work was supported in part by the National Science Foundation under grant ECS-0322813 and by the U.S. Department of Defense under subcontract B-12-664.

References

[1] Shacham et al., *JLT*, 23 (2005), pp. 3066-3075.
[2] Yang et al., *JLT*, 19 (2001), pp. 1420-1426.
[3] Shacham et al. *ECOC* 2006, paper We3.P.138.
[4] Small et al., *PTL*, 17 (2005), pp. 1564-1566.
[5] W. J. Dally, *Principles and Practices of Interconnection Networks*, Morgan Kaufmann, 2004.