An FDL-Based Photonic Switching Node for a Data Vortex Optical Packet Switched Interconnection Network

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

Abstract *Adding buffering capacity to the switching nodes in the data vortex topology is considered. Experimental feasibility is demonstrated on a prototype SOA-based switching node and the performance is studied using computer simulations.*

Introduction

The data vortex has been suggested as a topology for optical packet switched interconnection networks [1],[2]. Data vortex networks are promising to provide ultra-high bandwidth through wavelength parallelism, at very low routing latencies, and scalability to a large number of ports, making them especially suitable for high performance computing applications [1].

The data vortex architecture is a multistage interconnection network (MIN) comprised of bufferless photonic 2×2 switching nodes, where contentions are resolved using deflection routing. In this paper we introduce a modified switching node design which includes an optical buffer based on a fibre delay line (FDL). The new design tries to resolve contentions using the FDL instead of deflecting the packets to the undesired output ports.

Architecture Description

A data vortex network is an MIN, comprised of 2×2 bufferless switching-nodes. Packets are routed 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, deflection routing is employed and the packet is transmitted to another switching node within the same stage. To assure the existence of a deflection path additional switching nodes, organized in circles, are added at every stage and a set of control cables are connected carrying signals encoding node availability. A data vortex switch, therefore, has a cylindrical structure and is identified by two parameters: *H* – height, and *A* – the number of nodes along its circumference [1],[2].

According the deflection routing scheme, packets continue to traverse each stage until they can progress to the next one, thus creating backpressure in the network. This undesired feedback effect, may lead to over-congestion of the network, and is a well-known shortcoming of deflection routing networks [3].

In order to mitigate the unfavourable effects of deflection routing and to reduce the back-pressure in the system, modifications are introduced to the switching node design facilitating limited buffering of optical packets in the switching nodes. Structural adaptations of the topology are also required and include the introduction of a set of intra-stage control signals, added to the already existing inter-stage control signals [1]. The new intra-stage cables connect nodes that are connected through deflection fibres [1] carrying electronic signals in a direction opposite to the flow of data on the deflection fibres. Since a node may process no more than a single packet in each time slot, the signals traversing the intra-stage cables indicate the availability of the node for deflected packets. This is different from the original data vortex topology where every deflected packet was assured to have a node available for it.

Switching Node Design

The switching node (see Fig. 1) has two input ports (*North*, from the previous stage, and *West*, within the same stage) and two output ports (*South*, to the next stage and *East*, within the same stage), and is capable of processing a single packet in each timeslot, similarly to the switching node described in [4]. It receives two electronic control signals, driven by other switching nodes, indicating (logically high) when the output ports are not available to receive packets (*Ci0* for *South* and *Ci1* for *East*). The *Ci0* signal is the already existing inter-stage control signal and *Ci1* is the new intra-stage signal. The node also emits two output control signals (*Co0* and *Co1*) to send information about its availability to other nodes.

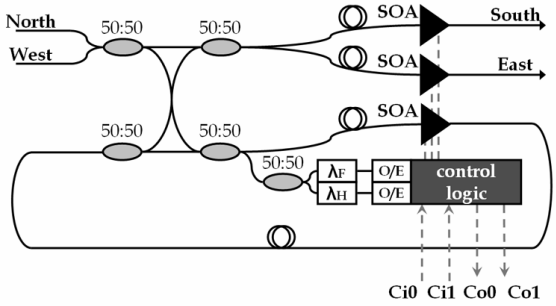

Fig. 1 - switching node block diagram

When a packet is received from either *North* or *West* port, it is directed to a set of passive optical couplers splitting its power to three semiconductor optical amplifier (SOA) gates and two wavelength filters followed by *p-i-n* photodetectors. The filters are used to extract two wavelengths, each carrying a single bit of information: *Frame*, denoting the packet presence, and *Header*, its address. *Frame* and *Header* are then detected and forwarded to electronic control logic. According to this information, along with *Ci0* and *Ci1*, the control logic decides whether the

packet is routed to the *South* output port, the *East* output port or is buffered on the FDL. *Co0* and *Co1* signals are also generated in this process.
The routing decision rule is as follows:

```
If (Frame==1) Begin
 if (Header=Href)&&(Ci0==0) then
        South_SOA=ON;
 elsif (Header≠Href)&&(Ci1==0)
        East_SOA=ON;
 elsif (Ci1==1)||(FDLcnt<CNT_MAX)
        FDL_SOA=ON;
 else
        East_SOA=ON;
End
Else All_SOAs=OFF
```

```
Co0 = East_SOA && Ci1;
Co1 = FDL SOA;
```

where *Href* is a local value configured in the node according to its location in the network [1] and *FDLcnt* is the number of FDL hops the current packet has taken in the node. This decision rule can be defined in shorthand as follows: *the packet is routed to the destination encoded in its Header bit unless that port is not available (according to the relevant Ci signal). In that case the packet is buffered on the FDL. If the FDL hop-count exceeds a limit, and the East port is available, the packet is routed East.*
The decision rule is implemented using a Xilinx complex programmable logic device (CPLD) which controls the SOA gates. The optical packets are delayed on a length of a fibre whose latency is matched to the electronic processing latency so that they reach the SOAs exactly when one of them is switched on to pass the packet. The SOA also amplifies the packet to compensate for the losses in the coupler network.
It is important to note that to accommodate the propagation time of the intra-stage control signals, and their processing in the nodes, the FDL latency must match two system slot times, so every packet forwarded to the FDL is delayed by two slots.

Experimental Demonstration
To demonstrate the correct functionality of the constructed switching node, an experimental setup is constructed. Packets comprised of *Frame* and *Header* wavelengths along with a high speed optical payload are generated for both input ports, and electronic signals drive the *Ci0* and *Ci1* ports. In this experiment *Href* is set to logical 1 and CNT_MAX=2. In the demonstrated pattern, the first packet is forwarded from *West* to *South*, and the second packet from *North* to *East*. The third packet (from *West*) has to be routed *South* but both outputs are blocked, as indicated by *Ci0* and *Ci1*, so the packet is buffered, delayed for two slots and is then routed south.

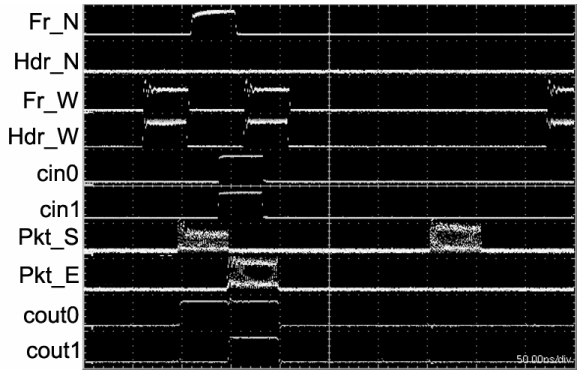

Fig. 2 - waveforms of input and output signals in functional demonstration experiment

Performance Study
The effect of adding 2-slot long FDLs to the data vortex nodes, and modifying the electronic control plane to accommodate the additional signal is studied using computer simulations. An 80-port data vortex switch (*A*=5, *H*=16) is simulated under uniform Bernoulli traffic. The results are plotted in fig. 3. A small (approximately 5%) increase in the acceptance rate is achieved consistently under several load scenarios.

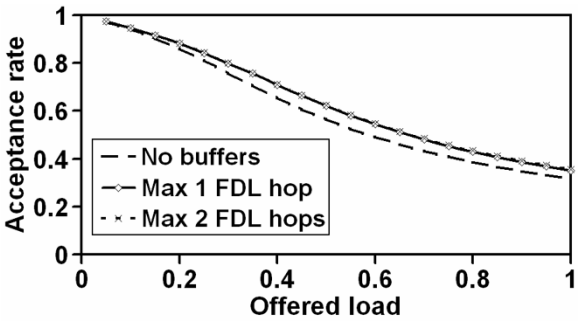

Fig. 3 - simulation results: acceptance rate vs. load under uniform Bernoulli traffic.

Conclusions
Modifying the data vortex switching node to include internal buffering capacity is studied and feasibility of implementation is shown through an experimental demonstration. A performance study shows a minor performance improvement under uniform traffic, suggesting that the virtual buffering mechanism of the original data vortex performs well in this scenario. Future studies of more adverse traffic flows such as bursty or bit-reversal traffic may yield a larger difference between the original and modified architectures.

References

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