Demonstration of a Complete 12-Port Terabit Capacity Optical Packet Switching Fabric

Benjamin A. Small, Odile Liboiron-Ladouceur, Assaf Shacham, John P. Mack, and Keren Bergman
Department of Electrical Engineering, Columbia University, New York, New York 10027
bergman@ee.columbia.edu

Abstract: We report on the implementation of a complete 12-port Data Vortex optical packet switching fabric containing 36 fully-interconnected nodes. Correct routing behavior is verified for 14-channel WDM packets, and latencies below 60 ns are achieved.

OCIS codes: (060.2310) Fiber optics; (060.4250) Networks

1. Introduction

One of the most critical challenges for next-generation high-performance computing systems is the realization of an interconnection network capable of providing ultra-high bandwidth communications between thousands of computing and memory elements with minimal access latencies [1,2]. Recent work suggests that optical interconnection architectures may offer viable solutions for these systems, particularly when the full bandwidth available via dense WDM is utilized to achieve multi-terabit capacities per port, and latencies are minimized to the optical time-of-flight [3,4]. However, a key obstacle to implementing viable large-scale optical interconnection networks has been the lack of adequate random-access optical buffering, particularly for hybrid data structures such as DWDM. Furthermore, to maintain ultra-low access latencies in the optical domain, contention resolution and packet routing must occur with minimal processing. A recently proposed highly scalable optical interconnection network based upon the Data Vortex topology utilizes self-routing of individual packets and alleviates the need for central scheduling and processing [5,6]. Deflection routing eliminates the need for packet buffering and minimizes packet traffic congestion. The Data Vortex architecture's unique absence of internal optical buffering elements enables the routing of DWDM packet payloads while maintaining flexibility of extending the packet size by simply adding (or removing) wavelength channels.

2. Data Vortex architecture

The Data Vortex optical packet switching network architecture was designed specifically for realization in the optical domain, taking into consideration the difficulty of implementing optical buffering and complex optical logic [5,6]. The architecture is a fully implemented directed deflection routing topology composed of simple 2×2 switching elements, or nodes. The nodes are arranged in hierarchal levels or cylinders, each affixing an additional bit in the packets' destination address (Fig. 1), in a manner similar to Banyan network addressing [2].

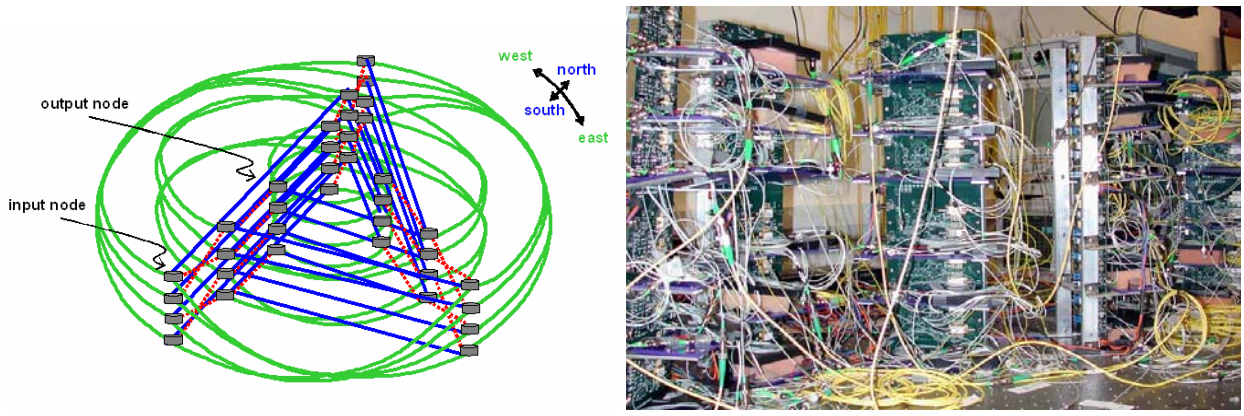

Fig. 1. Graphical depiction of the implemented 12×12 Data Vortex topology which contains 3 cylinders, each with 4 nodes in the height dimension and 3 nodes in the angular dimension. Photograph of implemented system (right pane).

The packets utilize DWDM, with particular wavelength channels designated for the header and routing information, and other wavelength channels for the packet payloads (Fig. 2). In fact, each routing bit is given its own

separate wavelength and extends over the packet duration, so that simple optical filtering can be used to demultiplex and decode the address information, minimizing the node processing latency.

Each node contains two SOAs, only one of which is turned on at a time, effectively directing the packet out through either the east or south port; couplers are used to appropriately divide the optical packet (Fig. 2). The SOAs are controlled by a routing logic implemented with high-speed electronics. The header information required for the routing decision is optically filtered from the packet and directed to receivers, which then drive the logical circuitry. The circuitry can also process an incoming deflection signal and generate an outgoing deflection signal [7]. The electronic and optoelectronic components are integrated on a PCB, so the entire routing process at a node is completed in just 11.7 ns.

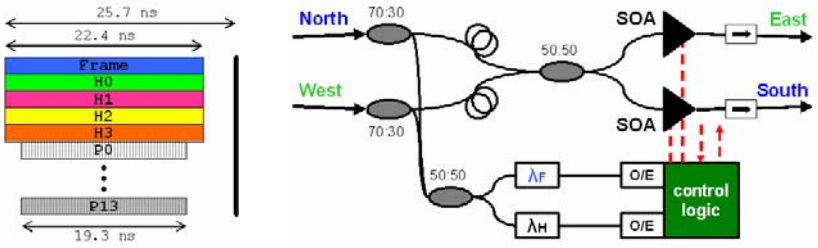

Fig. 2. Diagram of the DWDM header-parallel encoded packet format (left pane). Schematic of a Data Vortex switching node (right pane), containing fiberoptic couplers, filters (λ), photodetectors (O/E), and isolators (arrows).

3. Implemented switching fabric

The implemented switching fabric contains 36 individual 2×2 switching elements, 12 of which can be used as combined input terminal nodes, and 12 of which can be used as combined output terminal nodes. Each of the 12 output ports can be addressed with four header bits (only the words {0000,0100,1000,1100} are not used). The first cylinder of the implemented Data Vortex architecture utilizes the frame wavelength (C27, 1555.8 nm) and the *H0* header wavelength (C53, 1535.0 nm); the second cylinder again uses the frame with *H1* (C55, 1533.5 nm); and the final cylinder matches the last two header wavelengths *H2* (C33, 1550.9 nm) and *H3* (C58, 1531.12 nm). The payload channels span the entire ITU C-band, from 1531.9 (C56) nm to 1564.7 (C16) nm (Fig. 5).

The 14 payload channels (*P0* through *P13*) are modulated with a 2^9–1 PRBS at 10 Gbps and decorrelated with a length of optical fiber by approximately 450 ps/nm. The continuous data stream is then gated by an SOA to form packets which are injected into the switching fabric. The frame and header channels are modulated independently and are coupled with the gated payloads to form the complete header-parallel packets (Fig. 3). These packets are each 22.4 ns long with 19.3 ns (336 bytes) worth of payload information, and are separated from each other by a deadtime of 3.2 ns (Fig. 2). The total node latency is 15.8 ns. Furthermore, in order to maintain synchronous deflection routing within the whole Data Vortex system, the latencies of the paths between nodes must be set according to specific rules [8]. The fiber pigtails within the cylinders (east direction) are therefore 204 cm long (9.9 ns), and the pigtails between cylinders (south direction) are 74 cm (3.6 ns).

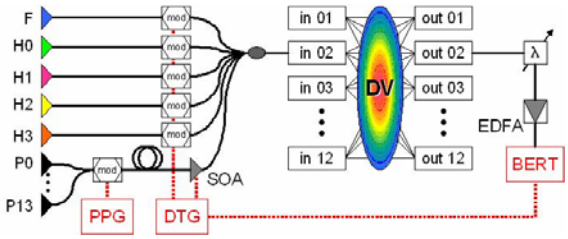

Fig. 3. Schematic of the experimental setup containing the implemented Data Vortex switching fabric and the necessary input signal modulation (PPG: 10-GHz pulse pattern generator; DTG: low-speed data timing generator).

Upon exiting the switching fabric, packets are directed to an optical filter and a pre-amplifying EDFA, then to an optical receiver module and a bit error rate (BER) tester. The BERT is gated to analyze only the packet duration, and it is synchronized with a packet traffic sequence which contains a known path and payload bit sequence (Fig. 3).

Our current setup allows for two input ports to be loaded simultaneously and independently. Packets are generated within fixed 25.7 ns time slots, which corresponds to the maximum time-of-flight latency between two adjacent nodes (within the same cylinder). Whereas the packet duration is fixed, the total packet size can be varied by the addition (or subtraction) of payload channels.

4. Results

In order to confirm that the entire switching fabric works correctly, we used packet traffic patterns which verify each node, routing pathway, and deflection condition. An example of such a pattern is shown below (Fig. 4). The signal waveforms verify that the four-bit address encoding causes the packet to be routed to the correct destination node while avoiding collisions with other packets within the network. The pattern consists of eight packets, three of which are destined to output ports #2 (0010), #4 (0101) and #9 (1011). The packets are shown emerging from their intended ports with latencies of 56.4 ns (three node hops), 107.8 ns (five hops) and 82.1 ns (four hops), respectively.

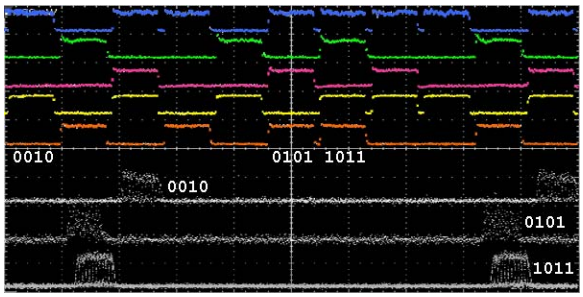

Fig. 4. The frame and header signals of the eight packets, followed by the emerging packets' payload waveforms.

Bit error rate measurements are performed on all 14 channels after a complete routing path of three node hops, confirming that the routing processes of the switching fabric do not affect the payload information. The 10-Gbps channels span almost the entire C-band (Fig. 5), confirming that the entire conventional DWDM spectrum could be used in this switching fabric. All channels were observed to have BERs of 10^{-9} or better for a 2^9–1 PRBS.

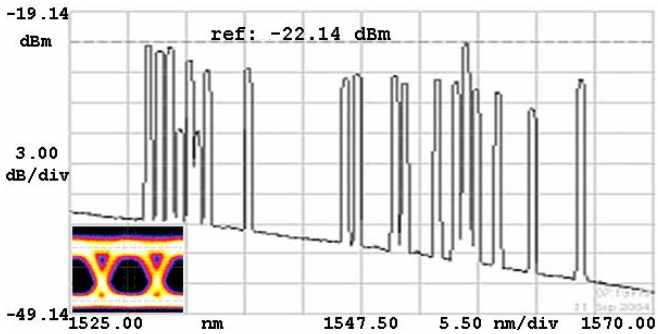

Fig. 5. Spectrum of the 14-payload channel 5-header optical packets (the frame and header wavelengths have higher peak powers). Eye diagram of the payload channel at 1564.7 nm confirms the measured BER of 10^{-9} at 10 Gbps (inset).

5. Conclusions

We have presented the first complete implementation of the Data Vortex optical packet switching architecture in a 12-port system containing 36 constructed switching nodes. The unique absence of internal buffers enables the utilization of practically the entire C-band capacity in each port since the DWDM packet payload is transparently routed through the network. Importantly, the demonstration also confirms that ultra-high bandwidth packets can be delivered to 12 independent output ports with minimal access latencies. The latency of the shortest input-to-output path through the interconnection network was measured to be less than 60 ns.

This work was supported by the Department of Defense (subcontract B-12-644) and in part by the National Science Foundation (ECS-0322813).

6. References

[1] *The Future of Supercomputing: An Interim Report*, NRC, National Academies Press, 2003.
[2] W.J. Dally and B. Towles, *Principles and Practices of Interconnection Networks*, San Francisco: Morgan Kaufmann, 2004.
[3] J. Gripp *et al.*, "Optical switch fabrics for terabit-class routers and packet switches," J. Opt. Netw. **2** 243-54 (July 2003).
[4] F. Masetti *et al.*, "High-speed high-capacity packet-switching fabric: a key system for required flexibility and capacity," J. Opt. Netw. **2** 255-65 (July 2003).
[5] C. Reed, "Multiple Level Minimum Logic Network," U.S. Patent 5 996 020 (30 Nov 1999).
[6] Q. Yang *et al.*, "WDM Packet Routing for High-Capacity Data Networks," J. Lightwave Technol. **19** 1420-26 (Oct 2001).
[7] A. Shacham *et al.*, "An Ultra-Low Latency Routing Node for Optical Packet Interconnection Networks," LEOS 2004 WM2 (Nov 2004).
[8] Q. Yang *et al.*, "Traffic Control and WDM Routing in the Data Vortex Packet Switch," IEEE Photonics Technol. Lett. **14** 236-38 (Feb 2002).