Design of Novel Optical Router Controller and Arbiter Capable of Asynchronous, Variable length Packet Switching

Haijun Yang, Venkatesh Akella, Chen-Nee Chuah, and S. J. Ben Yoo

University of California at Davis, USA

Abstract: This paper presents a novel optical router controller for asynchronous, variable-length packet switching with fast, fair, and wavelength-aware arbitration. The efficient hardware implementations are evaluated in terms of the area and delay for various switch sizes with the Xilinx XCV1000E field programmable gate array (FPGA) chip.

Keywords: optical router controller, arbiter, switch fabric

1. Introduction

Switching asynchronous, variable-length packets in the optical router data plane recently received strong interest [1-3], because it greatly simplifies the optical router system by eliminating the needs for packet synchronization and optical-electrical-optical (O/E/O) conversion. In addition, it reduces processing overhead and bandwidth loss due to the segmentation and assembly of variable-length Internet packets at the optical router node. At the same time, it also poses new challenges for designing the optical router controller to support asynchronous, variable-length packet switching. This paper proposes a new optical router controller incorporating fast, fair, and wavelength-aware arbiters to faciliate the switching of asynchronous, variable-length optical packets.

2. Optical Router Controller and Arbiter

2.1 Optical Router Data Plane and Control Plane

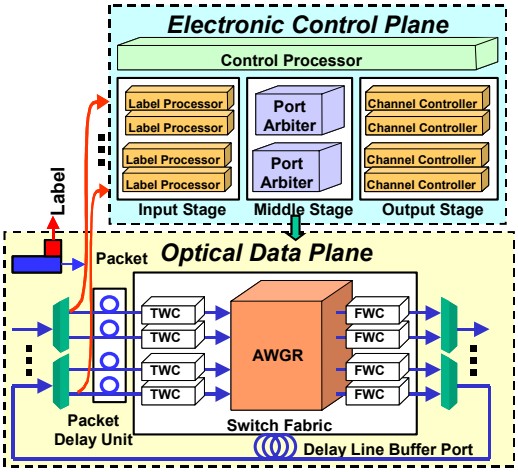

Figure 1 : Optical router data plane and control plane

Figure 1 shows the optical router with the data plane switch architecture [1-3] and the control plane optical router controller. The data plane is in all-optical domain, avoiding the overhead of O/E/O conversion for arrival packets, while the control plane is in the electrical domain to take use of mature electrical technologies for diverse control functions. When an optical packet is coming, its label is extracted and sent to the electronic control plane for processing, while its packet payload travels through the data plane packet delay unit to compensate the label processing time. The optical router controller makes a switching decision according to the label content, and instructs the data plane optical switch fabric to forward the packet to the desired output port.

2.2 Novel Optical Router Controller Architecture

Figure 1 illustrates the novel optical router controller structure, consisting of the label processor array, port arbiter array, channel controller array, and the control processor. Each input wavelength channel associates with one label processor for processing labels, and each output wavelength channel associates with one channel controller for sending out switch fabric reconfiguration signals and new labels. In addition, each output fiber port corresponds to one port arbiter for making switch decisions. The control processor computes the routing path and updates the forwarding tables. It also communicates with the network management plane. The proposed optical router controller consists of the input stage, middle stage, and output stage.

2.2.1 Input Stage: Asynchronous Label Processing

The input stage contains the label processor array. Figure 2 shows the label processor structure, consisting of the serial to parallel bit converter, preamble detector, label field analyzer, forwarding table search engine, channel manger, and request generator, etc. The label processor maintains two separate forwarding tables, storing the first preferred output path and the second preferred output path information for incoming packets, respectively.

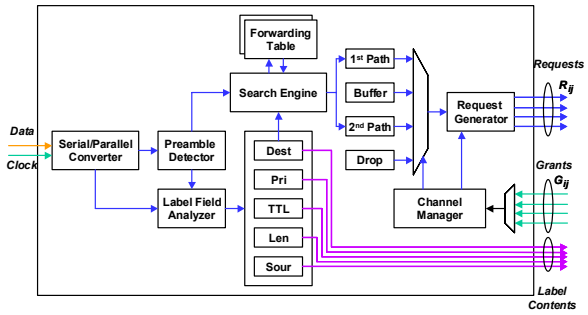

Figure 2 : Label processor structure

Processing Asynchronous Packets

To accommodate the asynchronous traffic in the data plane, the preamble detector of the label processor enables immediate recognition and processing of incoming

packet labels. Once the preamble detector detects a valid label, the label processor will start to analyze the label content, determine the incoming packet's first and second preferred output ports, and sends forwarding request to corresponding output port arbiters in the middle stage for switching decision making.

2.2.2 Middle Stage: Fast, Fair, and Wavelength-Aware Port Arbiter Supporting Variable-length Packets

The middle stage consists of the port arbiter array and port manager array. Each output port is associated with one port arbiter, one port manager, and one output port packet length database. Figure 3 shows the port arbiter contains a mixed tree arbiter (MTA) and a shuffling-network based first-fit scheduler (FFS). The port arbiter receives forwarding requests from input label processors and determines which requests to grant based on the channel availability status provided by the port manager. The granted labels will be sent to the output stage for further processing.

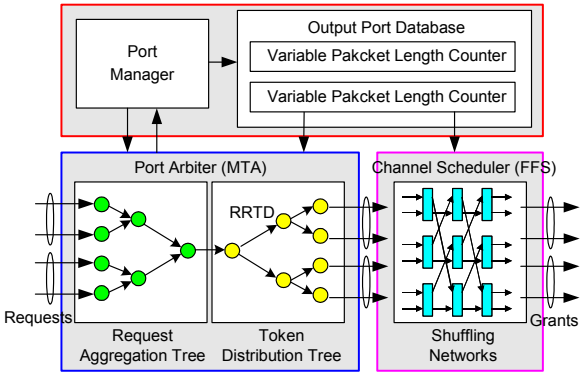

Figure 3 : Port arbiter and channel scheduler structure

Arbitration for Asynchronous Packet Traffic

For asynchronously arriving packets, ideally there is no need for arbiters, since they can be served in a first-come-first-serve way. However, in reality, it is challenging for the router controller to recognize the packet arrival order if packets arrive within the same clock cycle. As a result, arbiters are necessary to schedule optical packets, which arrive within the same arbitration period, in a batch to facilitate efficient hardware implementation.

Fast Arbitration by Dual Binary Tree Structure

The proposed arbiter structure MTA consists the request aggregation phase and the token distribution phase. Each phase is implemented with a binary tree data structure for fast arbitration. The request aggregation binary tree first aggregates forwarding requests from label processors, and then the reversed token distribution binary tree distributes the available output wavelength channels, *i.e.*, the tokens, among input forwarding requests.

Fair Arbitration by Round Robin Token Schedulers

Figure 3 shows the token distributuion binary tree consists of multiple distributed round robin token distributors (RRTD). The proposed arbiter achieves fair arbitration by fairly distributing available output wavelength channels to input forwarding requests. Each RRTD node in the token distribution tree distributes tokens to its child nodes in a round robin way to ensure the fairness.

Wavelength-Aware Arbitration by Shuffling Network Schedulers

FFS is a shuffling-network based output channel scheduler, which always schedules the packet to the first available output channel of its desired output port. The FFS simultaneously distributes multiple available output channels on the same output port to multiple incoming packets within one arbitration period. By this way, the arbitraion may fully take advantage of WDM channel knowledge in helping reduce packet contention cases.

Scheduling Variable-length Packets

Each port manager maintains a specific packet length database, storing the packet length information currently being transmitted on its associated output port to facilitate variable-length packet switching. The port manager tracks the availability of each output channel based on the database contents. The port arbiter uses such information to determine how many forwarding requests to grant in one arbitration cycle.

2.2.3 Output Stage: Asynchronous Switch Control and New Label Output

The output stage includes the output channel controller array. Figure 4 shows the output channel controller contains a wavelength lookup table to store the reconfiguration information for the data plane switch fabric, a programmable counter to control the time instant to send out control signals, and a new label generator to output new labels for label swapping and rewriting.

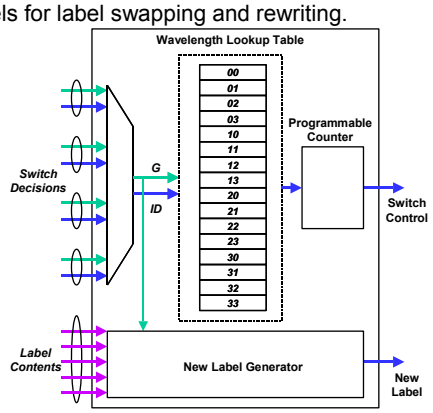

Figure 4 : Channel controller structure

Reconfiguration of Data Plane Switch Fabric to Adapt Asynchronous Traffic

To adapt the asynchronous packet traffic in the data plane, the output channel controller utilizes the programmable counter to reconfigure the data plane optical switch fabric at the desired time instant when the payload packets enter the data plane switch fabric. When the counter expires, the channel controller controls to reconfigure the data plane optical switch fabric, where a new connection is established between the switch fabric's input port and the desired output port immediately without bothering other existing connections.

3. Hardware-Efficient Implementation and Prototyping

3.1 Pipelined Implementations of the Port Arbiter

The port arbiter, the core module in the proposed optical router controller, is implemented on a Xilinx 1000E FPGA chip [4]. To increase the arbitration speed, the request

aggregation phase can be pipelined with the token distribution phase, leading to a 2-stage pipelined architecture. The token distribution phase can be further logically divided into two sub-phases: the inter-port distribution sub-phase and the intra-port distribution sub-phase, resulting in a more advanced 3-stage pipelined implementation architecture. We assume 8 input/output fiber ports. Figure 5 shows the un-pipelined (MTA), 2-stage pipelined (MTA-2S), and the 3-stage pipelined (MTA-3S) implementations. For each design, the fiber port carries 2, 4, 8, 16, and 32 wavelength channels, respectively. The clock period denotes the arbitration period, rate at which new input forwarding requests can be processed. The 3-stage pipelined design efficiently decreases the clock cycle and increases the system frequency.

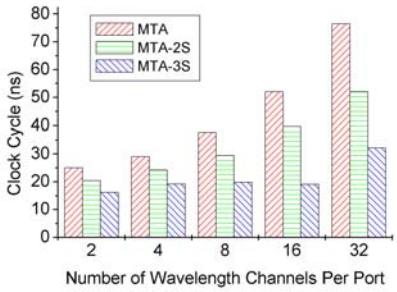

Figure 5 : Pipelined port arbiter implemenations

Table I shows the FPGA implementation results of the proposed port arbiter on the Xilinx XCV1000E FPGA chip. The area is shown in Xilinx Virtex slices and 4-input lookup tables (LUTs). Note that XCV1000E has 24576 LUTs [4], so we can easily fit the 8 arbiters required in our design on one XCV1000E chip.

TABLE I
FPGA AREA AND TIMING RESULTS

Design	Slices	LUTs	Clock Cycle (ns)
16×16	239	435	21.608
32×32	639	1148	23.905

3.2 FPGA Prototyping of Optical Router Controller

We further prototyped the complete optical router controller on the same FPGA chip with 4 input/output fiber ports, each carrying 4 wavelength channels. The controller consists of 16 label processors, 4 port arbiters, 4 port managers, 16 channel controllers, and the glue logic. Figure 6 shows the register-transfer-level (RTL) schematic of the optical router controller generated by Xilinx ISE tool.

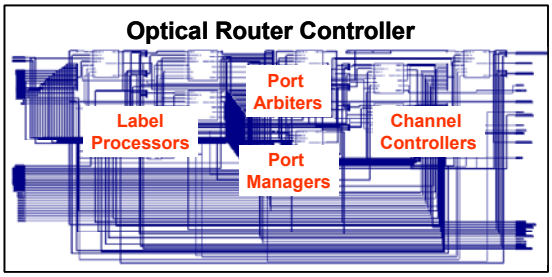

Figure 6 : Optical rotuer controller RTL schemeatic

Table II shows the FPGA implementation results of individual modules. The area is shown in Virtex slices and 4-input lookup tables (LUTs). The optical router controller uses 3886 slices or 6304 LUTs, about 25% of the total resource available on Xilinx XCV1000E.

TABLE II
FPGA AREA AND TIMING RESULTS

Design	Slices	LUTs	Clock Cycle (ns)
Label Processor	85	91	9.974
Port Arbiter	407	708	10.885
Port Manager	163	246	8.087
Channel Controllers	37	44	5.594

Figure 7 shows the contributions of major modules to the total area the optical router controller takes in terms of Xilinx Vertex slices, where the port arbiter array take the largest portion of 38% of the total controller area.

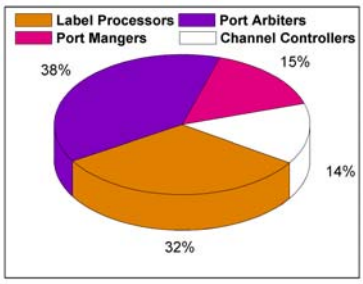

Figure 7 : Contribution to the total FPGA area of controller

4. Conclusion

This paper proposes a new optical router controller to support asynchronous, variable-length packet switching in the data plane without O/E/O conversion. The proposed controller and arbiters facilitates fast, fair, and wavelength-aware optical switch fabric arbitration. The hardware-efficient implementations of the port arbiter and the novel optical router controller are evaluated in terms of the area and the delay values for various switch sizes on a Xilinx XCV1000E FPGA chip.

5. Acknowledgment

This work was supported in part by NSF CISE 0335301 and 0435529.

6. References

[1] Z. Pan, *et al* "*Demonstration of variable-size packet contention resolution and packet forwarding in an optical-label switching router*," IEEE Photonics Technol. Lett., vol. 16, no. 7, July 2004.

[2] F. Xue, *et al*, "*Design and experimental demonstration of a variable-length optical packet routing system with unified contention resolution*," IEEE/OSA Journal of Lightwave Tehcnology, vol. 22, no.11, pp.2570-2581, Nov. 2004.

[3] Z. Pan, *et al* "*Experimental demonstration of variable-size packet contention resolution and switching in an optical-label switching router*," IEEE/OSA OFC 2004, paper FO4, Los Angeles, California, (2004).

[4] Xilinx, "*Virtex-E 1.8V FPGA complete data sheet*", 2006. http://direct.xilinx.com/bvdocs/publications/ds022.pdf