Demonstration of an FPGA Controller for Guaranteed-Rate Optical Packet Switching

Maryam Rezaee
Dept. ECE, McMaster University
Hamilton, Ontario, Canada

Ted H. Szymanski
Dept. ECE, McMaster University
Hamilton, Ontario, Canada

***Abstract*—A switch controller which achieves low-latency deterministic *'Guaranteed-Rate'* (GR) connections in packet switched networks is demonstrated on a 'Field Programmable Gate Array' (FPGA) device. A Software Defined Networking (SDN) control plane can configure the FPGA controllers to establish deterministic GR connections in a forwarding plane of IP routers or layer-2 packet switches. The use of deterministic GR connections can reduce queueing delays to negligible values, so that the end-to-end delays are reduced to the fiber latency. The switch controller can operate routers, switches and links at 100% loads, while simultaneously guaranteeing very low end-to-end queueing delays. A testbed consisting of 8 switches in a linear array is synthesized on an Altera Cyclone IV FPGA. An SDN control plane routes 128 traffic flows through the testbed to saturate the switches and links. Packets move through the forwarding plane at a clock rate of 52 MHz, transferring millions of packets per second, and statistics are recorded. The demonstration shows that all traffic flows in the testbed achieve deterministic GR service. The FPGA demonstration is accompanied by a video which visually illustrates packet forwarding in the network. The GR technology applies to IP routers and layer 2 packet switched networks, with thousands of nodes spanning arbitrary distances. The FPGA controller also applies to layer 2 MPLS (Multiprotocol Label Switching) and Ethernet switches, and to 'Optical Packet Switched' networks, to achieve deterministic GR services and to operate links at 100% loads.**

I. Introduction

Recently, a significant amount of research has been directed to reduce Internet latency [1]. Several international research projects such as the *BufferBloat* project (www.bufferbloat.net), the European Union *RITE project* ('Reducing Internet Transport Latency'), and the Swedish *READY* project ('Research Environment for Advancing Low Latency Internet') are attempting to reduce Internet latencies. A recent study by researchers at Akamai [2] argues for reducing Internet latency down to the fundamental limits of physics, i.e., the speed of light in fiber. According to [2], a fast Internet will offer a *truly transformative potential*, and enable new types of Internet services which were not possible previously.

In [3], [4], a new low-jitter GR technology suitable for establishing GR connections in Internet routers and layer 2 packet switches has been proposed. According to [3], [4], the use of GR connections will minimize packet buffering and queuing delays, and the end-to-end packet delays will be reduced to the fiber latency. The new GR technology can reduce the amount of buffering required in all-optical packet switches by several orders of magnitude, i.e., typically by a factor of 1000.

In this paper, a testbed for the switch controller and the GR technology is developed on an Altera DE2-115 board using an Altera Cyclone IV FPGA chip, as shown in Fig. 1. A linear array of 8 switch controllers and minimum-complexity switches/routers are implemented in the FPGA. An SDN control plane routes 128 traffic flows through the network, to effectively saturate every switch and link. A traffic rate matrix describes the GR traffic demands flowing through each switch. Each matrix is processed to determine several conflict-free deterministic transmission schedules for each switch. Each controller has several *Look Up Tables* (LUTs), which are used to store the precomputed schedules. Each controller uses these precomputed schedules to control the transmission of packets through one switch in real-time.

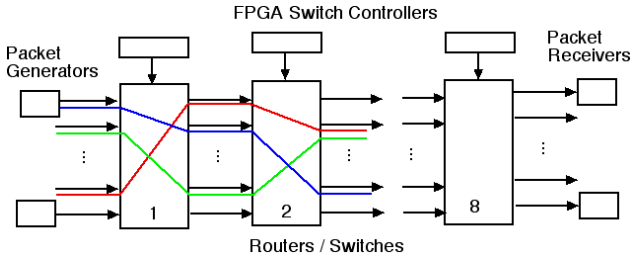

Fig. 1: The demonstration: A linear array of 8 routers/switches realized on an FPGA.

A. Guaranteed-Rate Services

A *Scheduling Frame* consists of 1024 time-slots. A fixed-sized packet can be transmitted in the network in one time-slot. An SDN control plane can establish GR end-to-end connections in the network, by reserving bandwidth for every connection, in every switch and every link along an end-to-end path. The bandwidth reservations are expressed in terms of time-slot reservations per scheduling frame. The scheduling algorithms in [4] are used to pre-compute deterministic GR transmission schedules, which are loaded into the *Lookup Tables* (LUTs) in the FPGA controllers. Each switch controller controls one switch for every time-slot in a scheduling frame, to provide every GR connection with its deterministic GR service.

Each router (or packet switch) consists of an *Input Queued* (IQ) switch design. Each controller contains several Lookup

Tables to store the precomputed deterministic schedules that control each switch. The Altera Cyclone FPGA has a limited amount of hardware resources, which limits our demonstration to 8 switches.

Each switch contains a *Routing-LUT*, which controls the state of the switch for each time-slot in a scheduling frame. In our demonstration, the Routing-LUT for each 4x4 switch will have $F = 1024$ rows, where each row contains a 12-bit vector containing four 3-bit control signals to control the multiplexers in one switch. (There are several other LUTs as well.)

Each IQ consists of four *Virtual Output Queues* (VOQs). In our FPGA demonstration, one M9K *Embedded Memory Block* (EMB) is used to synthesize each IQ with 4 VOQs. (The VOQs are arranged as circular FIFO queues.) Our FPGA demonstrator is limited by the capacity of the Cyclone FPGA. A M9K memory block can only store 128 packets in total, so the capacity of each VOQ is limited to 32 packets.

In the testbed, the size of each packet is relatively small (13 bytes), so that more packets can be stored in the FPGA. Every packet has a *Flow-Label* in its packet header, i.e., 2 bytes which uniquely identify the flow. Each packet generator module in Fig. 1 contains a Flow-LUT, which stores the flow-labels of the new packets which are generated. In our demonstration, there are 128 unique flows and 8 bits are sufficient to identify the flow-label. In each packet, 4 bytes are used to store the birth time of the packet, so that packet delays can be measured.

II. Test Setup and Results

An SDN control plane routed 128 traffic flows through the linear array of switches in Fig. 1. Every IQ receives on average 32 traffic flows. The average link load is 99.2%, a very heavy loading where every switch and every link are essentially saturated. Counters are synthesized at the packet generators and the packet receivers in Fig. 1, to record statistics on the number of transmitted packets. The packet ages and the distribution of the packet ages, were also recorded in hardware.

The testbed was clocked at 52 MHz. Tests confirm that no packets are ever dropped, and that the IQ size is limited to $\leq$16 packets on average. The tests show that each flow f receives its GR of traffic in a very fair manner during each scheduling frame. At any time t where $1 \leq t \leq F$, each flow gets a pro-rated fraction of t/F of its guaranteed rate of traffic (i.e., $R(f)$ time-slot reservations per scheduling frame). Four pairs of LED displays are available on the DE2-115 FPGA board, to show these results visually. These LEDs are used to display the percentage of received traffic for four randomly-selected flows. The LEDs verify that every flow receives deterministic GR service, and that the service is very fair, as the counters increase in unison. To illustrate the movement of packets visually, a video which shows the switch states and the movement of packets through the network will be available.

The number of packets queued in every IQ was measured for every switch. The maximum IQ size is about 16 packets, and the maximum number of packets queued in any switch is about 64 packets (to support 128 traffic flows). Recall that all switches and all links are effectively 100% loaded. Even at such high loads, the demonstration indicates that every flow receives deterministic GR end-to-end service, with very low queueing delays, and with very low buffer sizes.

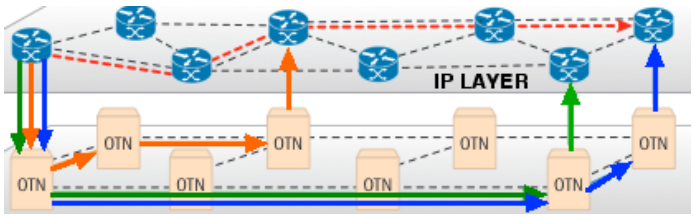

Fig. 2: A packet-switched Optical Transport Network, bypassing many layer 3 routers, and supporting IP over DWDM.

Figure 2 illustrates an *Optical Packet Switching (OPS)* underlay network supporting IP over *Dense Wavelength Division Multiplexing (DWDM)*. The optical packet switches are configured to provide deterministic GR transport connections (IP tunnels) between routers. The IP tunnels bypass layer 3 processing, thereby reducing latency and energy use. The end-to-end delays are effectively reduced to the speed of light in fiber.

A. Speed-up of FPGA versus Software

Running the FGPA testbed for 100 frames takes 25.6 msec on the Cyclone IV E FPGA with the frequency of 52 MHz. Simulating the same design for the same period of time in Modelsim-Altera takes 7980 sec (more than 2 hours). The FPGA demonstration achieves a speedup of about 311,718 times compared with software simulation.

III. Conclusions

The demonstration confirms that the FPGA controllers can provide deterministic GR service to 128 competing traffic flows, at very high link loads (99.2%), while reducing the end-to-end delays to the fiber latency. The low-cost FPGA controllers can handle IP routers, layer-2 electronic switches and optical packet switches, with hundreds of IO ports, with aggregate rates of 10s of Tbps. An SDN control plane can configure the controllers to achieve deterministic GR services in electrical and optical packet switched networks, while achieving 100% link loads and reducing end-to-end delays to the fiber latency.

References

[1] M. Ford, "Workshop Report: Reducing Internet Latency 2013", ACM SIGCOMM CCR, Vol. 44, No. 2, April 2014

[2] A. Singla, B. Chandrasekaran, P.B. Godfrey, B. Maggs, "The Internet at The Speed of Light", ACM Hotnets 2014, Oct. 2014, LA, USA

[3] T.H. Szymanski and M. Rezaee, "An FPGA Controller for Deterministic Guaranteed-Rate Optical Packet Switching", IFIP/EEE IM ManFi Workshop, Ottawa, May, 2015.

[4] T.H. Szymanski, "An Ultra Low Latency Guaranteed-Rate Internet for Cloud Services", IEEE Trans. on Networking, at IEEExplore, Oct. 2014

[5] Y.A. Vlasov, "Silicon-CMOS Integrated Nano-Photonics for Computer and Data-Communications Beyond 100G", IEEE Comm. Mag., Feb. 2012