Experimenting with Programmable Routers in Real Networks

Neda Beheshti*, Sara Bolouki*, Yashar Ganjali†, Brandon Heller*,
Nick McKeown* and David Underhill*

*Stanford University, Stanford, CA, USA
{nbehesht, sbolouki, brandonh, nickm, dgu}@stanford.edu

†University of Toronto, Toronto, ON, Canada
yganjali@cs.toronto.edu

Abstract

We have deployed a network of fully programmable routers in the Internet2 backbone. Each router, built on the NetFPGA platform [1], can process four ports of line-rate Gigabit Ethernet traffic. The routers are interconnected by a full mesh of dedicated links over the Internet2 backbone. Our goal is to demonstrate how such a network of routers can enable real-time experiments that are not possible with commercial routers. In particular, we will demonstrate the tradeoff between a router's buffer size and its forwarding performance, as seen by both the end-user and the network operator. Attendees will be able to interactively set the buffer size while monitoring the real-time occupancy of the buffers, drop rate, and throughput with a resolution of 8ns.

Categories & Subject Descriptors: C.2.1 [Computer Communication Networks]: Network Architecture and Design; C.2.6 [Computer Communication Networks]: Internetworking

General Terms: Design, Experimentation, Measurement

Demonstration

Figure 1 shows the topology of our experimental network of programmable routers in the Internet2 backbone. Each router consists of a rack-mounted Linux PC (Dell 2950 server) and a NetFPGA PCI card. The NetFPGA board contains an FPGA, four Gigabit Ethernet ports, SRAM and DRAM for packet buffers and lookup tables. It can be programmed to act as a router, Ethernet switch, network interface card, or virtually any other type of packet processor. The NetFPGA was originally developed for teaching networking hardware and the NetFPGA website [1] offers several standard reference designs, gateware, software and courseware.

This demonstration explores the consequences of reducing router's buffer size from $RTT \times C$ to $RTT \times C / \sqrt{N}$ to tiny buffers of about 20 packets. We set the output link rate of a NetFPGA router in the backbone network so as to make it the bottleneck for the flows.

The router reports the time series of its buffer occupancy, which shows the well-known TCP sawtooth.

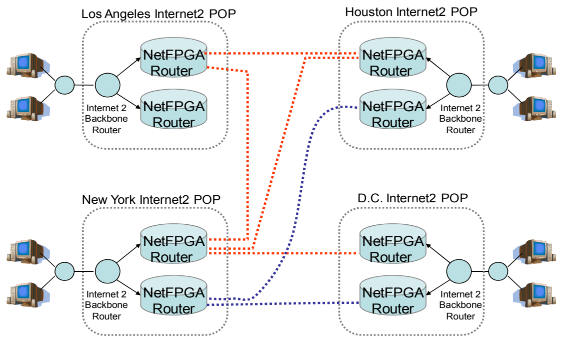

Figure 1: Experimental Network Topology

Attendees will be able to vary the size of the buffer from the bandwidth-delay product to zero, and see that the throughput degrades as the buffering is reduced. With an increasing number of TCP flows, a progressively smoother sawtooth will be observed. The attendees can then vary the size of the buffer to see that the throughput degrades only when the buffer is smaller than the bandwidth-delay product divided by the square-root of the number of flows [2].

Next, we will explore the consequences of reducing buffer size to just 10-20 packets. Theory and simulation [3] suggest that the throughput of the bottleneck link falls very slowly as we reduce the buffer size so that $B < RTT \times C / \sqrt{N}$. By the time the buffer size is just 20 packets, the throughput falls by only about 20%. When the buffer size falls below 20 packets, the throughput drops rapidly to zero [3]. These results hold only when the traffic is relatively smooth; e.g. when traffic enters the network on a relatively slow link before being routed to a faster backbone link. Bursts are broken when packets are spread out in the transition from a slow access link to a fast backbone link.

References

[1] The NetFPGA project. http://www.netfpga.org/.
[2] G. Appenzeller, I. Keslassy, and N. McKeown, Sizing router buffers. SIGCOMM 2004, pages 281–292, New York, NY, USA, 2004, ACM Press.
[3] M. Enachescu, Y. Ganjali, A. Goel, N. McKeown, and T. Roughgarden, Part III: Routers with very small buffers. ACM/SIGCOMM Computer Communication Review, 35(3):83-90, July 2005.