Towards a Highly Scalable Network Tester

Murali Ramanujam, Noa Zilberman

University of Cambridge

firstname.surname@cl.cam.ac.uk

ABSTRACT

High end networked-systems have quickly climbed from a throughput of gigabits/sec to terabits/sec, and are approaching petabits/sec. Alas, network testing equipment has not scaled: it remained either low throughput or extremely expensive. With suitable network testing equipment either not at scale or too expensive even for commercial vendors, systems may be released without proper testing and validation. We propose a methodology for large scale testing of networked systems, based on using a low-cost, open source network tester and a commodity switch. Our approach is scalable, open source, and accurate, and can be adapted to a variety of networking equipment, widely available to users.

1 INTRODUCTION

The demand for high end networked-systems has been steadily rising, but equipment to test such devices has remained high cost and out of reach of most academic environments. High end networked-systems today achieve over 10Tbps [2] at the Top of Rack (ToR) switch, and scale up to 922 Tbps [3]. This is in stark contrast to the equipment available to academic networking researchers, often using software-based traffic generators or are lucky to have access to a few 10Gbps ports.

The problem becomes even more profound when one considers the variety of network interconnects and port's form factors supported by high end platforms: ranging from 10GE SFP+, through 100GE using QSFP28 to emerging 400GE solutions. Researchers and developers rarely have access to the entire range of interconnects.

Network testing equipment has always been a trade off between cost and performance — specialised hardware based equipment offers the best throughput, number of ports and precise configuration capabilities. However, in addition to being prohibitively expensive this commercial testing equipment is also closed and proprietary.

It is possible to trade-off and achieve a comparatively cost-effective and efficient solution as evidenced by projects such as OSNT [1] and FLOWer [6]. In these situations, we often see somewhat specialised hardware, or even commodity server NICs, merged with open source software efforts — for example, MoonGen [5]. However, these projects have throughput limitations that do not match the benchmarking needs of next generation high throughput platforms and applications. Software based packet generators, providing more flexibility, rely on commodity hardware, making the speed, precision and number of ports quite restrictive.

Testing large scale networking equipment is challenging even using the most cutting-edge network-testing equipment. For example, three separate network-testing systems were required in order to test a single Arista 7500E, a switch with $288 \times 40GbE$ ports [7]. The question then becomes *how can we efficiently scale up network testing equipment to handle the ever-increasing capacities of networked-systems?*

In this work, we propose **Scalable Network Tester** (SNT), a network tester midway between low-cost, limited functionality network testers and very high-cost, full-functionality network testers. SNT achieves this trade-off by combining a commodity switch with an open source network tester to provide high throughput, capable of saturating 100+GbE links, while remaining affordable for academic research groups.

2 APPROACH

In SNT, traffic generation and monitoring is conducted by the low cost, open source OSNT [1] platform, while traffic management and amplification is done in a commodity switch [4].

2.1 OSNT

Open Source Network Tester (OSNT) [1] is a traffic generator and capture system available on the NetFPGA-SUME platform [8]. OSNT allows generation and capture of different packet sizes, and manipulation of the bitrate on four 10GE ports simultaneously. OSNT also provides methods for coordinating multiple generator/capture systems, and supports 6.4ns timestamp resolution. While OSNT provides quality flow differentiation methods, it is limited by its hardware platform to four 10GE ports, which are not sufficient to thoroughly test large scale devices.

2.2 Commodity Switch

We use the Delta AGC7648A [4] (Fig 1), an open compute project (OCP) ToR switch with deep buffers offering 100GE uplink ports. The total I/O bandwidth is 800G and the oversubscription ratio is up to 1.35:1. The switch supports Traffic Management and quality of service (QoS) mechanisms, including classification, marking, queuing, policing, and scheduling. Hierarchical Quality of Service (QoS) allows for traffic shaping distribution based on bandwidth available/consumed

from other flows. Ingress/egress policing is made possible through standard 2-rate, 3-colour policing and marking.

2.3 The Design of SNT

SNT combines OSNT with the commodity switch, harnessing sophisticated Traffic Management and QoS functions from both the switch and OSNT. SNT can efficiently saturate high capacity links with scripted scheduling and rate-limiting capabilities. SNT allows full rate packet blasting combined with fine-grained traffic shaping, applicable to many testing scenarios. Feedback is provided from the DUT to the source, allowing reproducible latency measurements under loads of highly selective traffic flows of varying characteristics.

SNT will expose customisation controls to the user, including Traffic Management options, specifying the characteristics of traffic emerging from each SNT port. SNT assumes that the commodity switch is properly validated, and some of its properties can (and were) validated using OSNT. The switch can also be used as an interface converter, e.g., traffic returning from 100GE port on the DUT can connect to the switch, and sent back to OSNT using a 10GE port.

3 PROPOSED USE CASES

SNT can act as a **Packet blaster** to undertake large scale testing. By combining the rate control features of OSNT with the multicast capabilities of the switch, SNT can saturate 100+GbE links. SNT can also provide **Traffic management capabilities**, such as per-port and per-flow rate limiting, shaping and scheduling in order to test the behaviour of the device under test when exposed to selective traffic flows combined in varying fashion. Because of the feedback link from the device under test (DUT) to the traffic source (OSNT), SNT is capable of measuring the **latency** of the DUT incurred when handling multiple traffic flows of varying characteristics concurrently, as well as packet drops.

Fig 2 demonstrates an experiment using SNT: OSNT connects to the switch through $2 \times 10GE$ ports. Port 1 is sending traffic at a set rate, while port 2 sends traffic at full line rate. Traffic from port 2 is amplified within the switch, e.g., using multicast to all ports (in this case, $46 \times 10GE + 6 \times 100GE$), and sent to a DUT. Traffic from port 1 is sent to one output port (e.g., port 49, a 100GE port). Traffic manipulation is done on port 48: the output rate from the port is limited and shaped, and weighted scheduling is applied, e.g., 10% to the test flow (full 10GE), and 90% to the amplified traffic flow. With the DUT now being bombarded by (up to) 800GE traffic, one output from the DUT (with the monitored flow) is fed back to OSNT to test the latency and of this flow and check for packet drops.
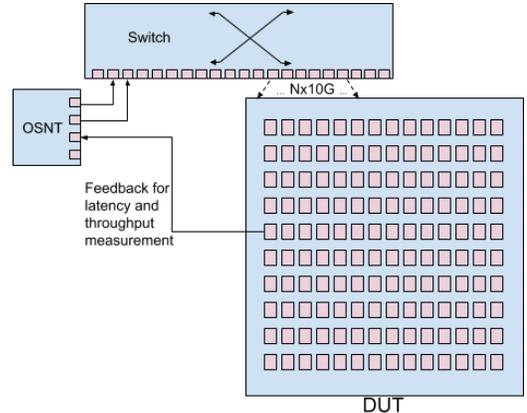

Fig. 1. SNT configured on the switch.

Fig. 2. SNT - Traffic amplification & flow monitoring

4 ACKNOWLEDGEMENT

REFERENCES

[1] G. Antichi et al. Osnt. *IEEE Network*, 28(5):6–12, Sep 2014.
[2] Broadcom. Tomahawk3, May 2018. https://www.broadcom.com/products/ethernet-connectivity/switching/strataxgs/bcm56980-series.
[3] Cisco. The aggregate switching capacity of the cisco crs multishelf system reaches upto 922 tbps, May 2018. https://www.cisco.com/c/en/us/products/routers/carrier-routing-system/models-comparison.html.
[4] Delta Networking, Agema. Agc7648a, May 2018. http://www.agemasystems.com/UserFiles/files/AGC7648A-rev2-datasheet.pdf.
[5] P. Emmerich et al. Moongen: A scriptable high-speed packet generator. In *Proceedings of the 2015 Internet Measurement Conference*, IMC '15, pages 275–287, New York, NY, USA, 2015. ACM.
[6] P. Emmerich et al. Flower - device benchmarking beyond 100 gbit/s. In *2016 IFIP Networking Conference (IFIP Networking) and Workshops*, pages 109–116, May 2016.
[7] I. Lippis Enterprises. *Arista 7500E Data Center Switch Test Report.* 2014.
[8] N. Zilberman, Y. Audzevich, G. A. Covington, and A. W. Moore. NetFPGA SUME: Toward 100 Gbps as research commodity. *IEEE MICRO*, 34(5):32–41, Sept. 2014.