Fine-grained P4 Measurement Toolkit for Buffer Sizing in Carrier Grade Networks

Swapna Buccapatnam
AT&T Labs, Middletown, NJ
sb646f@att.com

Yaron Koral
AT&T Labs, Middletown, NJ
yk216h@att.com

Simon T Tse
AT&T Labs, Middletown, NJ
st2196@att.com

Xiaoqi Chen
Princeton University
Princeton, NJ
xiaoqic@cs.princeton.edu

Steven A Monetti
AT&T Labs, Middletown, NJ
sm1818@att.com

John Tulko
AT&T Labs, Middletown, NJ
jt3114@att.com

Ken Duell
AT&T Labs, Middletown, NJ
kduell@att.com

Aswatnarayan Raghuram
AT&T Labs, Middletown, NJ
ar2521@att.com

Tzuu-Yi Wang
AT&T Labs, Middletown, NJ
tw503J@att.com

Shir Landau Feibish
Princeton University
Princeton, NJ
sfeibish@cs.princeton.edu

Jennifer Rexford
Princeton University
Princeton, NJ
jrex@cs.princeton.edu

Kathleen Meier-Hellstern
AT&T Labs, Middletown, NJ
km9378@att.com

Joe Stango
AT&T Labs, Middletown, NJ
js3861 @att.com

ABSTRACT

In the talk, we present analysis of measurements taken from the AT&T network, using the P4 Toolkit. We characterize the underlying traffic and give implications for carrier network engineering. Specifically, we show evidence and persistence of microbursts in utilization. We also show that these microbursts together with the buffer size determine the packet-drop rates experienced by a carrier grade network.

KEYWORDS

P4 Measurement Toolkit, Buffer Sizing, Microbursts, Inter-packet arrival times

1 INTRODUCTION

Network router buffers must be managed to avoid network congestion due to router queue buildup. For carrier grade networks, this task is challenging since there is little control over the carried traffic. While a carrier normally has enough capacity to handle all committed customer traffic, the carrier must also prepare for cases where many customers simultaneously burst traffic, creating high bandwidth utilization and buffer usage. This case differs from a dedicated data center or private network where all endpoints are controlled by the service provider and managed to avoid traffic surges (for example, Google's B4 network [1]).

Since controlling the traffic is not an option, network planning, along with network buffer management, is based on network traffic models that predict expected volumes for different traffic scenarios. One extreme example is the microburst, which is a short surge of traffic that peaks well above the outgoing link capacity. For microbursts, the network operator would like to know information about the actual connections that might have caused buffer buildup. Based on this information the operator can decide whether to dedicate larger buffers, throttle the microbursts, apply traffic shaping, or even contact the source that is generating the burst. Unfortunately, most existing network measurements cannot be provided at a timescale more accurate than once every few minutes. Hence, events such as microbursts which take place in a shorter timescale are invisible to the network operator.

The limitations described above raise the need for a different measurement paradigm that can push some of the analytics to the data plane and perform measurements in a much smaller timescale. Recent technological advancements introduced programmable switches that answer both requirements. In this talk, we propose a new measurement toolkit, based on P4 programmable switches [2], that is suitable for a carrier grade, network measurement requirement.

The architecture utilizes fine-grained measurement techniques, capable of performing efficient and flexible network measurements in a non-intrusive way. The toolkit supports modular detection capabilities as well as reporting logic that passes data to an external collector for further analysis (see Figure 1).

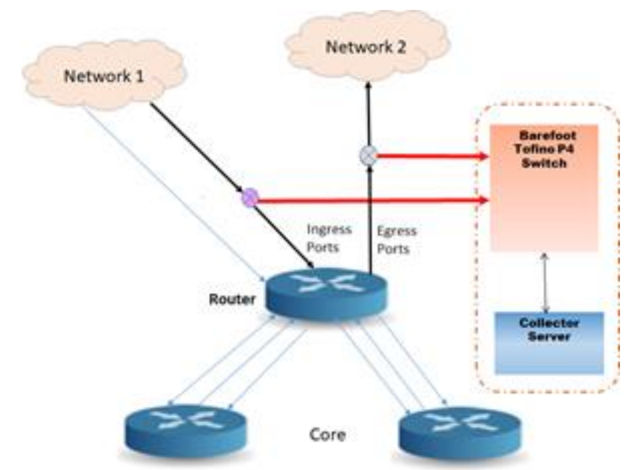

Figure 1: P4 Measurement Toolkit experiment setup

The toolkit attaches to the network using a set of optical tap devices as a bump-in-the-wire, which enables the creation of a duplicate packet stream where the original packet stream is uninterrupted, and the duplicate stream is delivered to the external measurement toolkit. To measure a single router, we use tap devices on incoming and outgoing links. All the tap outputs are directed as incoming links for our external measurement toolkit.

The architecture of the toolkit consists of three modules: (1) Anomaly Detection Element; (2) Reporting Logic Element; and (3) State Snapshot for History Capture Element. The first module monitors the packet streams entering the toolkit at line rate and computes parametric values used to detect anomalous events that are of interest to the network operator. The second module contains the decision logic for determining what and when stored information in the measurement toolkit is distributed to a collector and analysis engine. The third module maintains ongoing snapshots of the system's state [3], e.g. a running history of metadata of the various packet streams delivered to the measurement toolkit. Whenever a report is required, the current state is "dumped" and sent to an external collector for further analysis.

We employed a Barefoot Tofino P4 programmable switch in our implementation of the toolkit, enabling fine-grain measurements of a sub-millisecond timescale, which suffices to capture a microburst.

2 CARRIER NETWORK MICROBURST MEASUREMENT & BUFFER SIZING TEST

Before our P4 Measurement Toolkit was available, we had conducted an experiment to characterize the arrivals to various buffers in AT&T networks. The measurements were performed at various Core, Aggregation, Cluster and Edge routers with the network. The specific measurements involved measuring and resetting high watermark buffer counters every 5 seconds. We were able to observe jumps in buffer byte usage indicative of microbursts. The following figure shows queue utilization over a 24-hour period in a carrier network switch highlighting transient congestion events with more than double the average queue length.

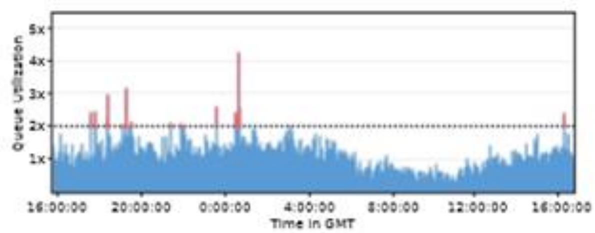

Figure 2: Queue utilization in a carrier switch

We analyzed the time-series of these measurements to compute the Hurst exponent [4, 5], which characterizes the burstiness of traffic. We observed that even arrivals to high utilization, core-to-core links exhibited high burstiness as characterized by Hurst exponents significantly larger than 0.5. This is counterintuitive to earlier work that argues that such burstiness is broken down or aggregated to smoother traffic under multiplexed conditions.

This study motivated the more fine-grained study using our P4 Measurement Toolkit to understand how this burstiness in traffic impacts the router buffers. To this end, we designed an experiment to obtain small timescale measurements of packet delay and inter-packet arrival times at one of our routers. The goal of the experiment is to capture detailed traffic history during a microburst in order to profile the traffic to aid in buffer sizing modeling.

3 RESULTS

In the talk, we present measurements taken by the P4 Measurement Toolkit, to identify the existence of microbursts, characterize the underlying traffic, and give implications for carrier network engineering. Analytical outputs include probability distributions of packet time delay, ingress port bytes, and inter-packet arrival times. We use stochastic modeling of delay and arrival processes to establish lower bounds on the buffer size needed to achieve low losses. In addition, we establish self-similarity and burstiness of carrier grade traffic using finer timescale data.

REFERENCES

[1] Jain, Sushant, et al. "B4: Experience with a globally-deployed software defined WAN." ACM SIGCOMM Computer Communication Review. Vol. 43. No. 4. ACM, 2013.
[2] Bosshart, Pat, et al. "P4: Programming protocol-independent packet processors." ACM SIGCOMM Computer Communication Review 44.3 (2014)`.
[3] Chen, Xiaoqi, et al., "Fine-Grained Queue Measurement in the Data Plane." CoNEXT 2019
[4] H. E. Hurst, "Methods of Using Long-Term Storage in Reservoirs", Proc. of the Institution of Civil Engineers, Part I, 519-577, 1955.
[5] Leland WE, Taqqu MS, Willinger W, Wilson DV. On the self-similar nature of Ethernet traffic (Extended Version). IEEE/ACM Trans. Netw., vol. 2, no. 1, Feb. 1994, pp. 1-15.