2013 IEEE 11th Malaysia International Conference on Communications
26th - 28th November 2013, Kuala Lumpur, Malaysia

Impact of Access Bandwidth on Packet Loss: A Flow-level Analysis

Muhammad Amir Mehmood[1], Nadi Sarrar[2], Steve Uhlig[3] and Anja Feldmann[2]
[1]Al-Khawarizmi Institute of Computer Science, UET, Lahore, Pakistan
[2]Technische Universität Berlin, T-Labs, Berlin, Germany
[3]Queen Marry University of London, London, UK
amir.mehmood@kics.edu.pk, nadi@net.t-labs.tu-berlin.de
steve@eecs.qmul.ac.uk, anja@net.t-labs.tu-berlin.de

***Abstract*—High-speed Internet connectivity is becoming more ubiquitous due to advances in access technologies. On the one hand, the increased access speed allows flows to consume the larger available bandwidth. On the other hand, large aggressive flows lead to more bursty traffic. In this paper, we study the impact of higher broadband access speeds on the resulting packet losses, and dissect those losses into their constituent flows. Based on tightly controlled experiments using open source tools, we show that as ISPs are increasing the access bandwidth of their residential customers, the effects on flow-level properties are non-trivial. Surprisingly, the average number of flows that participate in a packet loss burst remains small, as compared to the burst size. As expected, higher access bandwidth translates into larger packet loss bursts. Furthermore, large flows dominate packet losses, and even more so as access bandwidth increases. In contrast, irrespective of the access bandwidth, a few small flows systematically participate in loss episodes, and suffer significantly due to their limited size.**

***Index Terms*—Access networks, Flows, Burst losses.**

I. INTRODUCTION

The last two decades have seen a major evolution in the Internet access from 56Kbps modems to a wide variety of high-speed access technologies, e.g., cable, VDSL, and FTTH. The fierce competition among network operators to capture the largest possible fraction of a limited customer base leads to constantly upgraded Internet access products [1]. Irrespective of the available access bandwidth, a handful of applications seem to always manage to fully utilize it: from bulk transfer over FTP in the early 90's to HTTP dominance in the early commercial Internet, followed by the rise of P2P ten years ago and the recent comeback of HTTP and new bandwidth-hungry applications, e.g., IPTV and video streaming [4].

Although the traffic properties of the Internet have been widely studied in the past, predicting the impact of customer bandwidth upgrades on the quality of experience (QoE) of different applications is a surprisingly difficult task. The rapid increase in the bandwidth available at the access is fundamentally changing the Internet, where the bottleneck used to be at the edge. Nowadays, capacity in the core is not growing as fast as at the edge. To compensate for this, content providers (CDN) are moving content closer to the edge [4].

One aspect of the Internet that did not change much since its early days is that the majority of the traffic is transported over TCP. However, when TCP segments from various different applications are sent across the Internet, the different conditions they encounter impact packet loss and consequently flow performance. Among those conditions, the buffer size of IP routers plays a major role [2].

Due to the reactive nature of TCP congestion control, the way in which a set of flows participate in losses inside a router buffer is not well understood. The limited buffers inside routers that multiple flows have to share make losses inevitable. The bandwidth probing mechanism of TCP leads to packet bursts creating *loss episodes*. A loss episode is defined as a set of consecutive packets that are lost inside the buffer. Note that in this paper we focus on TCP traffic, given that it represents most of the traffic in the Internet.

Given the stringent requirements in terms of losses of some applications and the uncertainty in the accuracy of loss estimation techniques, we believe that it is important to investigate the loss process inside an aggregation router as experienced by individual flows when different access rates limit the bandwidth of individual flows. In this paper, we address experimentally general questions about the nature of losses inside an aggregation router buffer at the flow-level, including questions like "How does the structure of loss episodes change under different network conditions?" or "Which flows are responsible for which loss episodes?".

Based on experimentation using a high-resolution monitoring setup in a controlled lab environment, we study losses inside an aggregation router buffer and their distribution across flows. Our findings are the following:

- **Size of loss episodes:** Under limited access bandwidth, most loss episodes are made of a small number of packets. Under unlimited access bandwidth on the other hand, most loss episodes are made of a large number of packets.
- **Number of flows participating in loss episodes:** The number of flows participating in a loss episode does not grow as fast as the size (in packets) of the loss episode, irrespective of the access bandwidth. Most loss episodes involve only a limited number of flows.
- **Type of flows participating in loss episodes:** Most loss episodes are composed of packets generated by a wide variety of flow sizes, even though, on average they are

larger than global [1] flow size distribution. The higher the access bandwidth, the more large flows dominate large loss episodes.

- **Impact on small flows:** Flows of small size systematically participate in loss episodes and therefore suffer significantly in view of their flow size.

The rest of the paper is organized as follows. Section II discusses the related work. Section III describes our experimental methodology. Section IV presents our results. We discuss the impact of our results in Section V. Section VI summarizes the paper.

II. Related work

The current lack of understanding of the loss process inside a router buffer is startling given its importance in the performance of different applications, e.g., voice and video traffic [5]. The limited related work available about the loss process concentrates on its estimation, rather than on understanding how individual flows participate in losses.

In the past many researchers have tried to estimate the loss process and their implications on TCP. For example, Sommers et al. [13] measured and estimated packet loss characteristics in the presence of different traffic scenarios. They developed a tool, *BADABING*, which reports loss measurements with a greater accuracy than its predecessors [15]. In another work, Sommers et al. [14] analyzed delay and loss measurements from a SLA compliance perspective. Their main contribution is to find out congestion episode duration and their frequency of occurrence.

Papagiannaki et al. [9] performed a measurement study inside the SPRINT network by using passive monitoring to infer the occurrences of micro-congestion episodes at small time scales. They proposed a performance metric to track the time scales over which these micro-congestion episodes occur.

All of the above mentioned studies focused on loss episode durations, their frequency of occurrence and other aspects like correlations over time [15], [8], [7]. However, they did not study how different flows contribute to loss episodes nor how different flows participate in loss episodes across different network conditions, including different access bandwidth.

III. Methodology

One of the challenging tasks for understanding Internet traffic behavior as it occurs inside a router buffer is to build a configurable and flexible testbed that allows tightly controlled experiments. In this section, we present the components of our experimental testbed and details about how our traffic is generated and monitored.

A. Realistic Traffic Generation:

To reproduce conditions similar to what traffic experiences in the Internet, we need a realistic traffic generator. The most suitable tool available in the literature is Harpoon [12]. Harpoon is able to reproduce flow-level behavior consistent with Internet traffic characteristics. The two main parameters used for customizing Harpoon are distributions for flow sizes as well as flow inter-arrival times. Most flows in the Internet rely on closed-loop feedback [10]. Consistently, we use TCP flows for most of the traffic. Harpoon is configured for flow sizes according to a Pareto distribution. A Pareto distribution ensures a finite mean while also making sure that the generated traffic exhibits a realistic variability. We played with various values of the parameters of the Pareto distribution, and report results for $\alpha = 1.2$ and a mean of $\mu = 110KB$ that are close to what has been observed in Internet traffic traces[2]. For flow inter-arrival times, i.e., the times users wait between different requests, we choose an exponential distribution with a mean of $\mu = 1$. Other parameter values provide similar results.

B. Topology

The network topology we use is a classical dumbbell one as shown on Figure 1. This topology is simple while reproducing the flow multiplexing that takes place within aggregation routers in ISP networks. All network interfaces are 1 Gigabit Ethernet cards. The configurable network bottleneck is located between the NetFPGA router and the Dummynet delay emulator. Harpoon clients send requests to Harpoon servers. Using Dummynet [11], we add a delay of $150ms$ to every ACK packet sent by Harpoon clients to Harpoon servers. This delay enables us to emulate round-trip-times which are typical for WAN environments [6].

C. Monitoring

We opt for the NetFPGA platform with a DropTail queue as a router. The NetFPGA allows us to limit the buffer size arbitrarily and to gather highly accurate buffer statistics. We use the NetFPGA buffer monitoring facility to obtain very fine-grained packet loss time series, which we also use to validate our flow-level statistics. We vary the size of our bottleneck buffer inside the NetFPGA from as small as recommended by Appenzeller et al. [2], up to an order of magnitude more. The trends in our results do not differ across buffer sizes. In the remainder of the paper, we report results for a buffer size of 512 packets, which is much larger than recommended by Appenzeller et al.

D. Data capture

The advantage of our analysis stems from the integrated monitoring at various layers at a fine time resolution. We capture packet traces at both the ingress and egress ports of the NetFPGA router. By comparing both traces, we are able to pinpoint missing packets along with transport layer information, e.g., TCP sequence numbers, as well as timing information about when the drop occurred. We are therefore able to reconstruct the whole dynamics of any given TCP flow. We run each experiment for 15 minutes and cut 2.5 minutes from both beginning and end to ensure that a stable behavior of the flows is considered for the results.

[1] We use term global to represent overall statistics of the data

[2] α is a shape parameter/tail index and μ is the scale parameter of the Pareto distribution

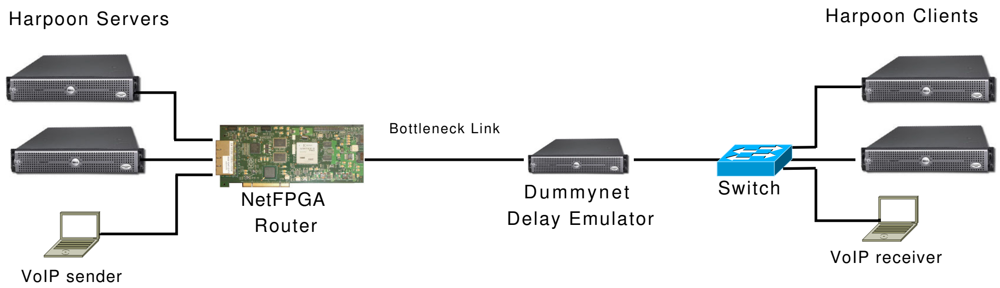


Fig. 1. Experimental setup.

E. Network Bottleneck

To ensure that the only bottleneck in our setup is the router buffer of the NetFPGA card, we set the maximum TCP receive window to 20MB. Note that such high value of TCP recieve window is used to ensure that the throughput of TCP flows is not limited by the advertised receive window [3]. Furthermore, this allows us to focus on network bottlenecks as experienced by the end users in the Internet. All experiments use TCP New Reno to control the size of the TCP congestion window with SACK enabled.

F. DSL bandwidth

To emulate limitations at the customer access in the download direction, as happens in practice, we limit the bandwidth of individual TCP flows to currently popular and future high-speed DSL rates, e.g., 8 and 16 Mbps (ADSL) or 32 Mbps (Cable and VDSL). We also run experiments with unlimited bandwidth, which corresponds to FTTH/very high speed Internet access. We use Dummynet pipes to restrict the bandwidth of individual TCP flows.

G. Load

Because load in the Internet may vary from one link to another, we create scenarios with different network conditions. We rely on two load levels, *low* and *high*, realized by changing the number of parallel Harpoon sessions [3] on our clients. To determine the necessary number of Harpoon sessions, we run the experiments without link capacity limitations. The low load corresponds to a mean link utilization of around 50%. Once the load exceeds a link utilization of 50%, one can expect degradations in the quality of service, e.g., increased delay and packet loss. Therefore, we choose the high load scenario in such a way that the resulting utilization will be close to the link capacity, similar to a congested link.

IV. Characterization of Loss Episodes

In this section, we characterize loss episodes by studying their sizes and their composition in terms of constituent flows, across different access bandwidth limitations (8, 16, 32 Mbps, and unlimited). We start by explaining how to identify the boundaries of a loss episode. We then provide some distributional statistics about the observed sizes of loss

[3]A Harpoon session is equivalent to flows generated by an Internet user.

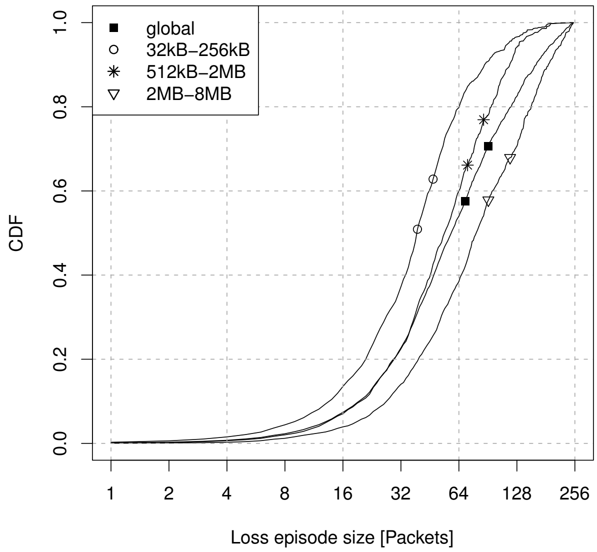


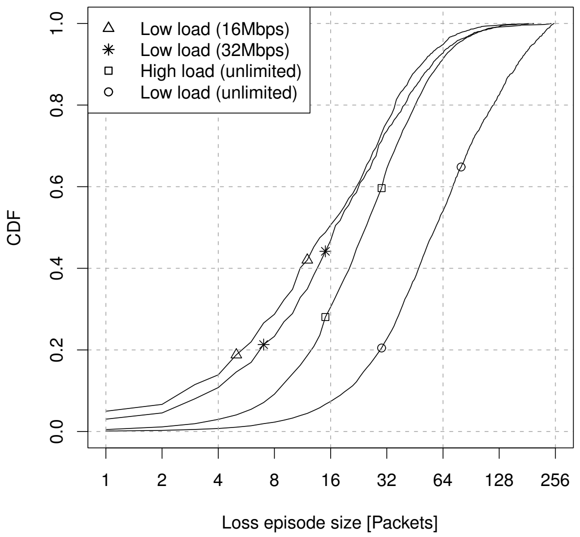
Fig. 3. CDF of loss episode sizes for different flow sizes (low load, x-axis = log scale)

episodes. In the second part of this section, we dissect loss episodes into their constituent flows and study the participation of different flows within the loss episodes.

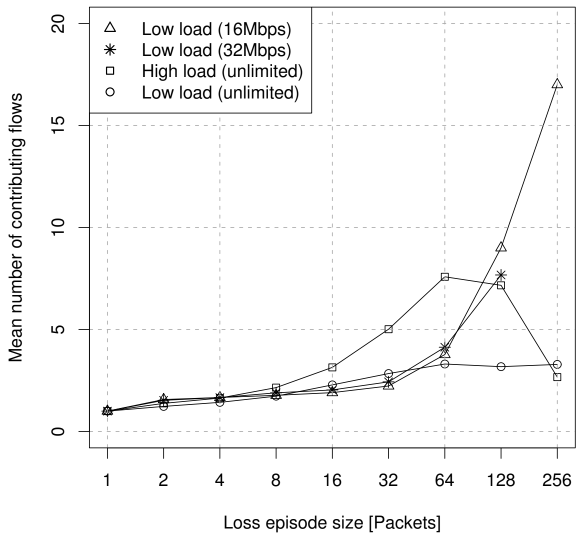
A. Size of loss episodes

When examining the packet loss time series of the NetFPGA buffer, we notice that large bursts of lost packets are often separated by a single packet that is successfully forwarded by the NetFPGA card. This phenomenon leads to an underestimation of the size of loss episodes. Therefore, we stitch together two loss episodes that are separated by a single successfully delivered packet. Using more than one packet during the stitching does not significantly change the loss episode size distribution.

The resulting distributions of loss episode sizes is shown in Figure 2(a) for low and high load scenarios with an unlimited access rate, and for low load with 16 and 32 Mbps access rates. The x-axis on Figure 2 uses a log scale to ease the visual comparison. The log scale gives more visual weight to the small loss episode sizes. The y-axis of Figure 2(a) gives the



(a) CDF of loss episode sizes weighted by number of contributed packets



(b) Number of flows contributing to loss episodes

Fig. 2. Characteristics of loss episodes (x-axis = log scale).

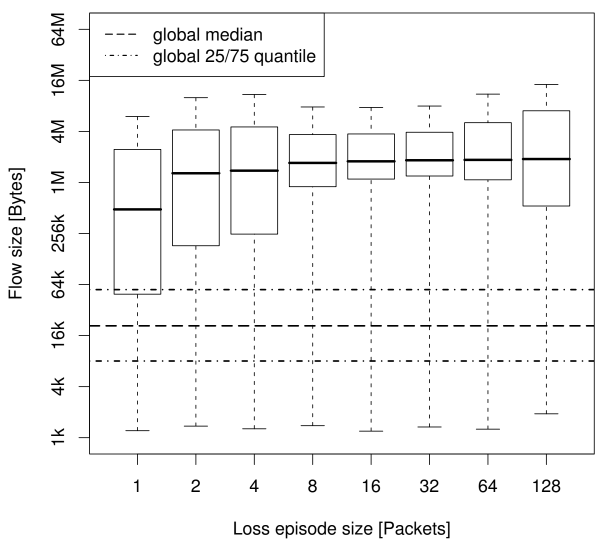
cumulative fraction of lost packets that belong to loss episodes of a given size, so that the figure shows what fraction of the total packet loss takes place within loss episodes of a given size.

With access bandwidth limitations, loss episode sizes are smaller than without bandwidth limitations. More than 40% of packet losses occur within loss episodes smaller than 16 packets when the access bandwidth is limited, against less than 30% under high load and less than 10% under low load when access bandwidth is not limited. More than 90% of the packet losses occur within loss episodes smaller than 64 packets under access bandwidth limitation or high load. Large loss episodes are thus rare under high load or under limited access bandwidth. When the load is low and flows are not limited by the access bandwidth on the other hand, more than 40% of the losses take place within loss episodes larger than 64 packets. The large loss episodes illustrate how bursty TCP can become when no bandwidth limitation shapes individual flows.

To summarize this section, we make our first key observation related to the size of loss episodes: *Under limited access bandwidth, most loss episodes are made of a small number of packets. Under unlimited access bandwidth on the other hand, most loss episodes are made of a large number of packets.*

B. Structure of loss episodes

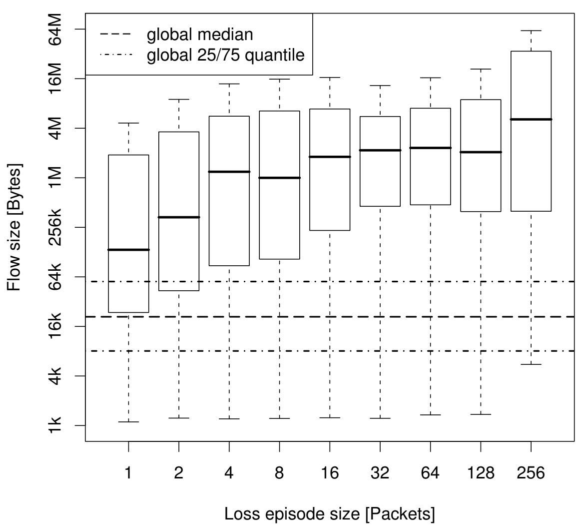
After looking at the distribution of loss episodes sizes, we take all individual packets of a loss episode and map them to their respective TCP flows. This loss episode decomposition enables us to study (i) how many different flows are contributing to a particular loss episode, and (ii) which flows are dominant in loss episodes of a given size.



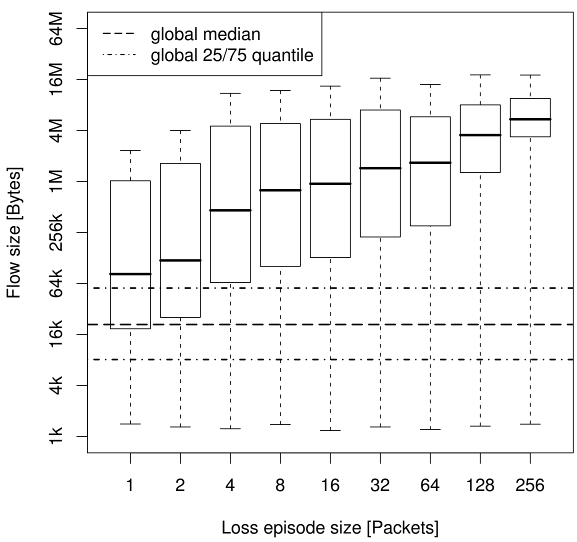
(a) 8 Mbps, low load

Fig. 4. Participation of flows in loss episodes, normalized by the flow contribution in number of packets (log-log scale).

Figure 2(b) shows the mean number of affected flows as a function of the loss episode size in packets. We expected that the number of flows would increase linearly with the loss episode size. Surprisingly, the curves appear very flat. This means that a limited number of flows participate in loss episodes, whatever their size. Due to the shaping of the flows imposed by the 16 Mbps access bandwidth, a larger number of flows have to contribute to make larger loss episodes. This



(a) 32 Mbps, low load



(b) Unlimited access, low load

Fig. 5. Participation of flows in loss episodes, normalized by the flow contribution in number of packets (log-log scale).

larger number of flows is very small compared to the size of the loss episode though. Note that loss episodes larger than 64 packets are rare under an access bandwidth of 16 Mbps (see Figure 2(a)): they represent about 5% of the packet losses only.

This leads us to our second key observation related to the number of flows participating in a loss episode: *Most loss episodes are composed of packets from very few flows. The number of flows participating in a loss episode grows much more slowly than the size of the loss episode*.

The second observation does not account for interactions between flows. Given the difference in the sending rates of different flows [16], we would expect that large flows account for the majority of the losses and participate in large loss episodes, while the small flows sneak through router buffers without incurring much losses. To confirm this intuition, we plot on Figure 3 the CDF of the loss episode sizes for specific ranges of flow sizes and contrast it to the one of the total traffic, under low load and unlimited access bandwidth. Large flow sizes, e.g., between 2 and 8 MBytes, tend to participate in larger loss episodes compared to smaller flow sizes. However, small flow sizes unexpectedly participate in large loss episodes as well.

Figure 3 tells how different flow sizes sample loss episode sizes. However, it does not provide insight into how often small flows suffer from losses that are actually mostly caused by large flows. To quantify the relative contribution of flows of a given size to loss episodes, we plot on Figure 5 the distribution of flow sizes that participate in loss episodes of a given size, under low load and for the 3 different access bandwidth limitations (8 Mbps, 32 Mbps, and unlimited). Note that we present results for 8 Mbps instead of 16 Mbps in order to better show the trends with different access bandwidths. Both axes of Figure 5 are logarithmically binned: loss episode sizes on the x-axis and flow sizes on the y-axis. The boxplots of Figure 5 are normalized by the number of packets that flows of a given size represent in the loss episodes of a given size. Figure 5 therefore accounts for the relative importance of flows of a given size into loss episodes of a given size. Each plot of Figure 5 also provides the quantiles 25, 50, and 75 of the global flow size distribution, to help comparing the flow sizes that participate in loss episodes to the global flow size distribution.

Overall, we see that all boxplots lie well above the lines indicating the main quantiles of the global flow size distribution. Whatever the size of the loss episode, losses involve mostly large flows, much larger than most of the flows of the global flow size distribution. Smaller sized loss episodes tend to involve smaller flows compared to large sized loss episodes, whatever the access bandwidth.

The main difference between limited and unlimited access bandwidth lies in the range of flow sizes that participate in the loss episodes. Recall that under access bandwidth limitations, most loss episodes are smaller than 64 packets, so that a wide range of flow sizes from 64 Kbytes up to several Mbytes participate in loss episodes. Under unlimited access bandwidth, most loss episodes are larger than 64 packets. Packet losses then occur mostly within flows larger than 1 MBytes. When the access bandwidth is very low, e.g., 8 Mbps as on Figure 4(a), the boxplots tend to flatten out and the different loss episode sizes sample a more similar range of flow sizes.

This makes our third key observation related to the type of flows participating in loss episodes: *Most loss episodes are composed of packets generated by a wide variety of flow sizes, even though larger on average than those from the global flow size distribution.*

978-1-4799-1532-3/13/$31.00 ©2013 IEEE

Surprisingly, across all loads, we observe that loss episodes of all sizes contain packets from very small flows. These small flows suffer significantly from such losses given their very small size. This leads us to our fourth key observation related to the impact of losses on small flows: *Flows of small size systematically participate in loss episodes and therefore suffer significantly in view of their flow size.* Given that small flows are made of a limited number of packets, they are more susceptible to suffer from time-outs than large flows. The latter can benefit from fast retransmit.

V. Discussion

Higher access speeds for different heterogeneous broadband access technologies are enabling users to access diverse applications through different end devices. In this paper, we have presented a simulation study with different access speeds and discussed how such higher speeds can potentially affect Internet traffic loss patterns. We show that flow-level packet loss analysis is critical to understand end-to-end performance perceived by the end users. Application designers should therefore account for such potential high losses seen by different flows.

Our results can provide guidelines for operational improvements in the Internet engineering. We note that to derive our results we use different assumptions regarding traffic workloads commonly present in the Internet, however the reality may be even more complex than we could simulate in our testbed setup.

We believe that new applications with totally different traffic characteristics can result in varied loss patterns, therefore studies such as ours must be repeated to find out its potential impact on services. Moreover, due to space limitations, we have not considered different Quality of Service (QoS) mechanisms in this paper. We plan to incorporate different QoS mechanisms and their impact on the Internet loss patterns in our future work.

VI. Summary

In this paper, we have studied packet losses inside a router buffer on a per-flow basis, under different access capacities. Through extensive experimentation, we have studied how many flows contribute to loss episodes, their size, and which flows participate in losses.

Our results have led us to several new observations about the way TCP flows interact to create loss episodes: most loss episodes are made of packets generated by a limited number of flows. Loss episodes are not only contributed by large flows, but a variety of flow sizes. Finally, we observed that flows of small size systematically participate in loss episodes and therefore suffer significantly in view of their flow size. Across all results, we observed that increasing access capacity makes loss episodes larger and gives more importance to the large flows in the loss process.

Furthermore, our work calls for techniques that would mitigate the impact of losses on small flows, a few of which systematically suffer from very high losses that will hamper their QoE. We also plan to investigate the correlation between access bandwidth limitations and the loss rate.

References

[1] OECD. OECD Broadband Portal, 2013. http://www.oecd.org/sti/broadband/oecdbroadbandportal.htm.

[2] G. Appenzeller, I. Keslassy, and N. McKeown. Sizing router buffers. In *Proceedings of ACM SIGCOMM*, 2004.

[3] N. Beheshti, Y. Ganjali, M. Ghobadi, N. Mckeown, and G. Salmon. Experimental study of router buffer sizing. In *Proceedings of ACM IMC*, 2008.

[4] C. Labovitz, S. Lekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian. Internet Inter-Domain Traffic. In *Proceedings of ACM SIGCOMM*, 2010.

[5] Yi J. Liang, John G. Apostolopoulos, and Bernd Girod. Analysis of packet loss for compressed video: Does burst-length matter? In *Proceedings of IEEE ICASSP*, pages 684–687, 2003.

[6] G. Maier, A. Feldmann, V. Paxson, and M. Allman. On dominant characteristics of residential broadband Internet traffic. In *Proceedings of ACM IMC*, 2009.

[7] H. Nguyen and M. Roughan. On the correlation of internet packet losses. In *Proceedings of ATNAC*, 2008.

[8] H. Nguyen and P. Thiran. Network loss inference with second order statistics of end-to-end flows. In *Proceedings of ACM IMC*, 2007.

[9] K. Papagiannaki, R. Cruz, and C. Diot. Network performance monitoring at small time scales. In *Proceedings of ACM IMC*, pages 295–300, 2003.

[10] R. Prasad and C. Dovrolis. Measuring the congestion responsiveness of Internet traffic. In *Proceedings of PAM*, 2007.

[11] L. Rizzo. Dummynet: a simple approach to the evaluation of network protocols. *ACM CCR*, 1997.

[12] J. Sommers and P. Barford. Self-configuring network traffic generation. In *Prodeedings of ACM IMC*, 2004.

[13] J. Sommers, P. Barford, N. Duffield, and A. Ron. Improving accuracy in end-to-end packet loss measurement. In *Proceedings of ACM SIGCOMM*, 2005.

[14] J. Sommers, P. Barford, N. Duffield, and A. Ron. Accurate and efficient SLA compliance monitoring. In *Proceedings of ACM SIGCOMM*, 2007.

[15] M. Yajnik, S. Moon, J. Kurose, and D. Towsley. Measurement and modelling of the temporal dependence in packet loss. In *Proceedings of IEEE INFOCOM*, 1999.

[16] Y. Zhang, L. Breslau, V. Paxson, and S. Shenker. On the Characteristics and Origins of Internet Flow Rates. In *Proceedings of ACM SIGCOMM*, 2002.