Shared versus Dedicated Buffers for Real-Time Traffic in Optical Routers with Very Small Buffers

Arun Vishwanath and Vijay Sivaraman

School of Electrical Engineering and Telecommunications

University of New South Wales

Sydney, NSW 2052, Australia

Emails: {arunv@ee.unsw.edu.au, vijay@unsw.edu.au}

Abstract— **When TCP and real-time (UDP) traffic multiplex at an optical packet router with very small buffers (less than** 50 **KiloBytes), we recently showed that UDP packet losses can increase with buffer size. In this paper we explore if this anomalous loss behaviour can be alleviated by dedicating (pre-allocating) buffers to UDP traffic. We show that while dedicating buffers to UDP traffic protects it from TCP's greedy behaviour, it can also reduce the overall efficiency of buffer usage. Our simulations using trace data indicate that buffer sharing yields better performance when the total buffer capacity is below a critical value of around** 20 **KiloBytes, and dedicated buffers are desirable otherwise. Our study informs designers and operators of optical packet switched networks with very small buffering capacity of the relative merits of shared and dedicated buffers for real-time traffic.**

I. Introduction

One of the fundamental concerns in wide-scale deployment of optical packet switched (OPS) networks is the issue of buffering, which is required at a switch when two or more packets have to be transmitted on an output link, on the same wavelength, at the same time. Contending packets can be stored by circulating them in fibre delay lines (FDLs); however, the high speed of light warrants large fibre spools for even minimal buffering, thereby making optical switches very expensive. Solid-state optical storage devices are emerging [1], which can buffer up to a few dozen KiloBytes of data. With a view towards moving to an OPS Internet core, a worthwhile question to be asked is will the existing Internet continue to work if buffers at core routers are reduced to only a few dozen KiloBytes?

This question has generated a lot of debate [2], [3], with the majority of work focusing on sizing router buffers largely from a TCP point of view, since TCP is the predominant transport protocol today. While the results are encouraging with TCP being able to realise nearly 80% link utilisation with only about 10-20 KB of buffering, there are many real-time applications such as online gaming, IPTV, and VoIP that are gaining prominence in the Internet. Hence, not only is it important to study the performance of small-buffered OPS networks on UDP traffic [4], but it is equally important to study the joint performance of TCP and UDP as well.

Our recent study in [5], [6] shows via analysis and simulation that when TCP and UDP traffic multiplex at a bottleneck link buffer in an optical packet router, UDP packet loss does not fall monotonically with increasing buffer size. Instead,
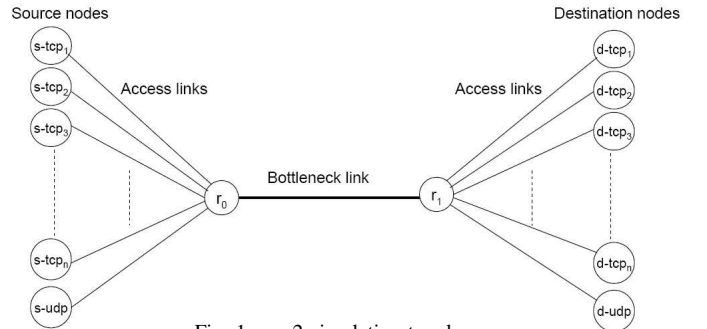

Fig. 1.   ns2 simulation topology

there is a certain contiguous range of buffer sizes where losses for UDP traffic increases with increasing buffer size. This anomalous loss behaviour for UDP traffic seems to occur due to its interaction with TCP's congestion control dynamics. It therefore seems reasonable to speculate that the anomaly can be mitigated by dedicating buffers for UDP traffic rather than sharing with TCP. This paper shows that dedicated buffers yield better performance only if the total buffer capacity is above a critical value, otherwise sharing is better for both UDP and TCP traffic.

II. Shared Buffers

We use $ns2$ [7] to simulate joint TCP-Reno and UDP performance on a dumb-bell topology shown in Fig. 1. We multiplex 1000 TCP sources along with a modified UDP trace from the movie Star Wars at router r_0. The UDP trace has an average rate of about 5 Mbps, constituting roughly 5% of the bottleneck link rate, which is operated at 100 Mbps. The propagation delay on the access links is uniformly distributed between $[1, 25]$ ms, while the bottleneck link (r_0, r_1) has a propagation delay of 50 ms. TCP and UDP packet sizes are 1000 Bytes and 200 Bytes respectively. Buffer size at the core router r_0 is varied in terms of KiloBytes. There is only a single FIFO queue at r_0 and drop-tail queue management is employed.

Fig. 2 shows the UDP loss and TCP throughput curves as a function of bottleneck buffer size r_0. We can see that up to about 12 KB of buffering, UDP loss falls monotonically with increasing buffer size, while TCP throughput rises to saturation. Beyond 30 KB of buffering, UDP losses drop again. However, in the range 12-30 KB, UDP losses increase with buffer size. Clearly, this anomalous region should be
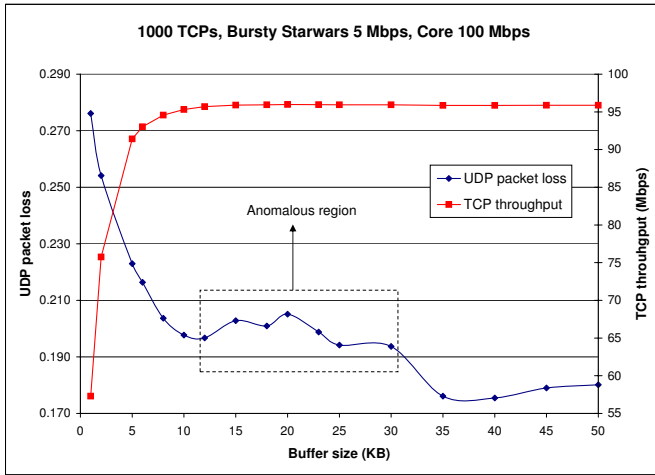
Fig. 2. UDP loss with the anomalous region

Fig. 3. UDP loss with varying % of dedicated buffers

Fig. 4. TCP throughput with varying % of dedicated buffers

avoided from UDP's point of view, since its performance in this region is no better than its performance with 12 KB of buffering.

A qualitative explanation of the anomaly is as follows. When the buffers at router r_0 are extremely small (say 1-10 KB), the congestion window of all the TCP flows remain extremely small as well. It is not allowed to grow beyond a few KiloBytes. As a result, the buffers are minimally used by TCP, and UDP has access to the entire buffer space for the most part. Put another way, UDP is able to "time-share" the buffers with TCP in this region. However, as buffers get bigger (say between 10-30 KB), a larger fraction of TCP flows are active since there is now room to increase their window sizes, causing UDP to "fair-share" the buffers with TCP, leading to lower effective buffers for UDP to use.

This transition between the time-share and fair-share components can lead to anomalous losses for UDP, as detailed in [5]. Since each KiloByte of extra buffering can add significantly to the cost of the optical switch, router designers and network providers should be cautious of the negative returns (in terms of UDP loss) on these extra buffers, and should size router buffers accordingly.

Since the anomaly arises from the interaction of TCP and UDP, it would seem that separating them out by pre-allocating or dedicating buffers would eliminate the anomaly and be beneficial in protecting UDP and reducing its losses.

III. DEDICATED BUFFERS

Having shown the anomalous behaviour that exists when TCP and UDP share buffers, we now study if dedicating buffers mitigates the anomaly. In the absence of knowing exactly how large commercial optical buffers will be, but knowing that they are constrained to be within a few dozen KiloBytes, we undertake two sets of simulations to study how dedicated buffering will affect their performance.

In the first set, we fix the total buffers at r_0, and for each value of the total buffer size, we in turn dedicate 5%, 10%, 20% and 50% of it to UDP. In each case, TCP gets the remaining fraction of the total buffers. The resulting UDP loss and TCP throughput curves are shown in Figures 3, 4.
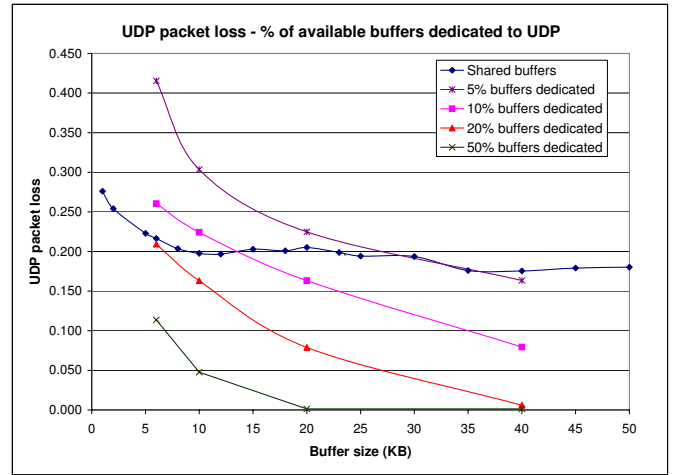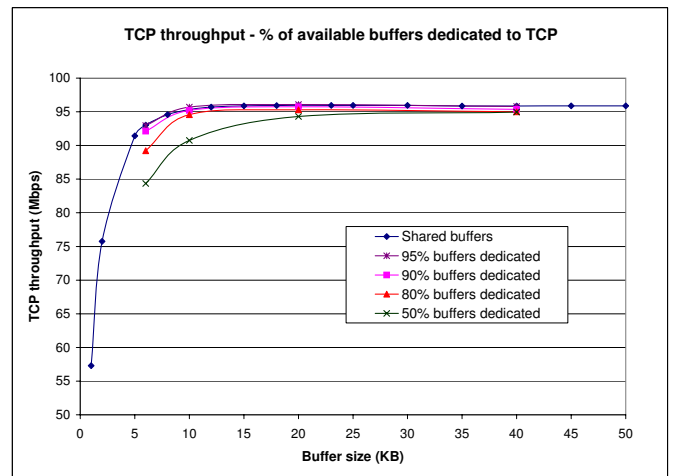
For comparison, the plot also shows the performance of UDP and TCP when the total buffers are shared by the two.

The top curve in Fig. 3 depicts UDP loss when 5% of the total buffers is dedicated to it. It illustrates the significance of deciding how much buffers UDP should get. The figure shows that up to about 35 KB of total buffers, UDP suffers from higher losses, indicating that it is actually better off sharing the buffers with TCP rather than having dedicated buffers.

The figure also shows that unless the total buffer size exceeds about 15 KB, it is better for UDP to share the buffers with TCP even with 10% of the total buffers being dedicated to it. Of course, dedicating larger percentage of the total buffers to UDP will reduce its losses significantly, this can be clearly seen from the figure.

While it is interesting to see UDP's performance, it is also important to analyse how TCP throughput is affected. Fig. 4 shows that the TCP throughput is fairly identical between the shared and dedicated buffering schemes beyond 25 KB of total buffer size. Up to that point however, TCP throughput with dedicated buffering is lower than with shared buffering by as much as 10%. Clearly, there is a tradeoff between
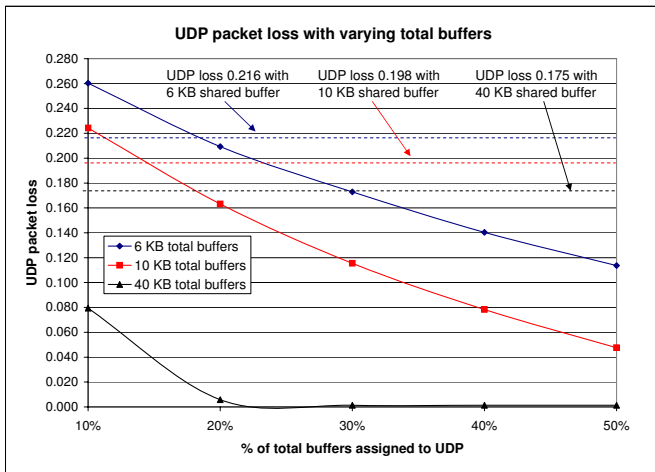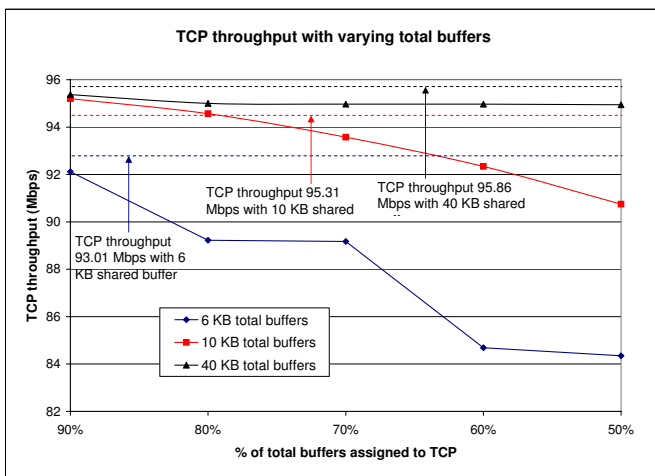
Fig. 5. UDP loss with varying buffer sizes

Fig. 6. TCP throughput with varying buffer sizes

UDP loss and TCP throughput; one comes at the cost of the other. These results suggest that if there is sufficient buffering available (greater than 25 KB or so), then both TCP and UDP benefit by having dedicated buffering. If however the amount of buffering is low, then it seems better for both TCP and UDP to share the buffers.

For the second set, in Figures 5, 6, we fix the total buffer size and plot the UDP loss and TCP throughput as a function of the percentage of buffers dedicated to UDP. Again, TCP is allocated the remaining fraction of the total buffers. The simulation is repeated for three sets of buffer sizes - 6, 10 and 40 KB, with UDP being allocated 10% to 50% of the total buffers in each case. For comparison, the figures also indicate via horizontal lines the corresponding UDP loss and TCP throughput when the buffers are shared by the two.

Fig. 5 shows that when the total buffer size is 10 KB or lower, UDP needs more than 15-20% of the total buffers dedicated to it in order to ensure that its losses are lower when compared with the corresponding scenarios in the shared buffer case. With larger total buffers however, only about 10% of buffers dedicated to UDP suffices to ensure that its losses are significantly lower. What is also important to note is how TCP throughput varies. We note from Fig. 6 that when

the total buffers is 6 KB, TCP throughput falls from 92 Mbps to about 84 Mbps when its share of buffers varies from 90% to 50%. This a decline of about 9% in throughput. It is also interesting to note that when the total buffers is 10 KB, it is better if TCP has 80-90% of the buffers dedicated to it as opposed to sharing the total buffers with UDP. This is because at such small buffer sizes, TCP's window is not allowed to grow beyond a few packets. This is further hindered by the presence of UDP, and it seems better for TCP to have lower dedicated buffers than having larger shared buffers. However, with much larger buffers, both TCP and UDP are benefited by dedicated buffering.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we asked if TCP and UDP traffic should share bottleneck link buffers at OPS nodes – sharing can result in more efficient buffer usage, but can lead to anomalous loss performance for UDP due to the greedy nature of TCP. The answer we found is not a definitive yes or no. It largely depends on how much buffering is available, and this is unclear at present since all-optical packet routers are not a commercial reality yet. Our results indicate that if the total buffers is about 20 KB or lower, it is better for TCP and UDP to share the buffers and face the anomaly. On the other hand, if bigger buffers are available, it is beneficial for both TCP and UDP to have dedicated buffering, thus alleviating the anomaly. These results inform network operators that it is very important to dimension their router buffer sizes carefully.

Several aspects of the problem will be explored as part of our future work. First, we can model the dedicated buffering system as an M/D/1/B queue and quantify analytically the losses for TCP and UDP traffic. These results can be compared with our results from the M/M/1/B model that we developed for the shared buffering scenario [6]. Second, it will be interesting to consider TCP flows that constitutes a mix of short-term (mice) and long-term (elephants) flows. Finally, we also plan to conduct extensive simulations with various other TCP versions and UDP traffic traces. An extended version of this paper will discuss these issues in greater detail.

REFERENCES

[1] H. Park, E. F. Burmeister, S. Bjorlin and J. E. Bowers, "40-Gb/s Optical Buffer Design and Simulations," *Proc. Numerical Simulation of Optoelectronic Devices (NUSOD)*, USA, Aug 2004.
[2] G. Appenzeller, I. Keslassy and N. McKeown, "Sizing Router Buffers," *Proc. ACM SIGCOMM*, USA, Sep 2004.
[3] M. Enachescu, Y. Ganjali, A. Goel, N. McKeown and T. Roughgarden, "Routers with Very Small Buffers," *Proc. IEEE INFOCOM*, Spain, Apr 2006.
[4] V. Sivaraman, H. ElGindy, D. Moreland and D. Ostry, "Packet Pacing in Short Buffer Optical Packet Switched Networks," *Proc. IEEE INFO-COM*, Spain, Apr 2006.
[5] A. Vishwanath and V. Sivaraman, "Routers With Very Small Buffers: Anomalous Loss Performance for Mixed Real-Time and TCP Traffic," *Proc. IEEE International Workshop on Quality of Service (IWQoS)*, Netherlands, June 2008.
[6] A. Vishwanath, V. Sivaraman and G. N. Rouskas, "Are Bigger Optical Buffers Necessarily Better?," *Proc. IEEE INFOCOM Student Workshop*, USA, Apr 2008.
[7] The Network Simulator - ns-2, http://www.isi.edu/nsnam/ns/