

Sharing Small Optical Buffers Between Real-Time and TCP Traffic

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—Buffers in emerging optical packet routers are expensive resources, and it is expected that they would be able to store at most a few tens of KiloBytes of data in the optical domain. When TCP and real-time (UDP) traffic multiplex at an optical router with such small buffers (less than 50 KiloBytes), we recently showed that UDP packet loss can increase with increasing buffer size. This anomalous loss behaviour can negatively impact the investment made in larger buffers and degrade quality of service. In this paper we explore if this anomalous behaviour can be alleviated by dedicating (i.e., pre-allocating) buffers to UDP traffic. Our contributions within this context are two fold. First, we show using extensive simulations that there would seem to be a critical buffer size above which UDP benefits with dedicated buffers that protect it from the aggressive nature of TCP. However, for smaller buffers that are below this critical value, UDP can benefit by time-sharing the buffers with TCP. Second, we develop a simple linear model that quantitatively captures the combined utility of TCP and real-time traffic for shared and dedicated buffers, and propose a method to optimise the buffer split ratio with the objective of maximising the overall network utility. Our study equips designers of optical packet switched networks with quantitative tools to tune their buffer allocation strategies subject to various system parameters such as the ratio of traffic mix and relative weights associated with TCP and UDP 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. Due to the inherent complexity associated with storing and processing of optical signals, emerging optical routers are expected to have very limited buffer capacity. 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 tens of KiloBytes?

Today, core routers in the Internet follow the rule-of-thumb [1] guideline to size buffers, which states that the amount of buffering needed at an output interface is equal to its delay-bandwidth product. In other words, if the output capacity is C , and the average round-trip time of a TCP connection flowing through the router is RTT , then the buffer size is $B = RTT \times C$. This would ensure that the output link is

utilised 100% of the time. For a typical RTT of 250 ms and capacity C of 40 Gbps, the rule-of-thumb mandates a buffer size of 1.25 GigaBytes, which poses a considerable challenge to router design.

The use of such large buffers (implemented using a combination of SRAM and DRAM chips) complicates router design, increases power consumption, and makes them very expensive. Thus, researchers from Stanford University recently challenged [2] the rule-of-thumb, and further argued (using analysis, simulation and experiments) in [3], [4] that the buffer size can be reduced to as small as 20-50 KiloBytes in core routers and the Internet would continue to work just fine, albeit, with a slight reduction of 20-30% in link utilisation. For a comprehensive survey on the topic of router buffer sizing, we refer the interested reader to our recent survey paper [5].

This important result is indeed a step towards the use of optical routers in the Internet backbone. It is believed that the cost, scaling and power requirements for next generation routers can be successfully addressed by building and deploying optical routers in the Internet core. Although some researchers (e.g., University of Melbourne) are skeptical about optical switching replacing electronic switching [6], [7], recent advances such as [8] have demonstrated the capability of all-optical switching, and solid-state integrated opto-electronic storage devices are emerging [9] that can realise optical FIFO packet buffers by storing few tens of KiloBytes of data on-chip. Other researchers [10] discuss various recent developments in the area of photonic switching within the European e-Photon research groups. In particular, in Section 2 of their paper, they outline various technological advances that may well pave the way for OPS deployment in the future. Bell-Labs currently has a working prototype of an optical packet router called IRIS [11]. Further, [12] points to a web-seminar that was held in Dec. 2008 that highlighted a potential revival of the optical switching industry in 2009. These efforts suggest that optical packet switching may be a reality in the future.

Given the orders of magnitude difference in buffer size between electronic routers and optical routers, Internet service providers (ISPs) need to be convinced that a move to an optical core would not severely degrade the performance of their network. The recent studies on routers with very small buffers, henceforth referred to as optical routers, have focused

A short summarized version of this paper was presented at the Second Symposium on Advanced Networks and Telecom Systems (ANTS) 2008 in Mumbai, India, in December 2008.

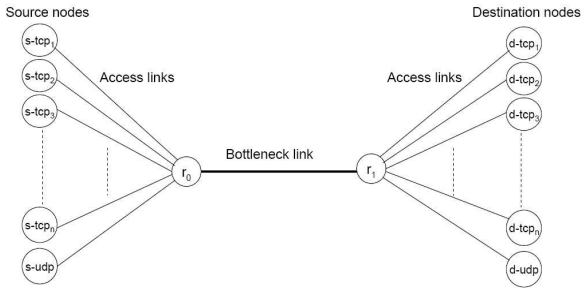


Fig. 1. *ns-2* simulation topology

largely on TCP performance, since TCP is the predominant transport protocol today. While the results are encouraging with TCP being able to realise acceptable 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, we believe that not only is it important to study TCP performance in small-buffered OPS networks, but UDP's performance also necessitates a deeper understanding [13], [14].

Our recent work in [15], [16] shows via analysis and simulation that when TCP and UDP traffic multiplex at a bottleneck link optical router, UDP packet loss does not fall monotonically with increasing buffer size, as one would expect. Instead, in a certain region of buffer size (typically between 10-30 KB) UDP loss tends to increase with larger buffers. 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 to UDP traffic rather than sharing with TCP. This paper explores the pros and cons of such a buffering system.

Our main contributions in this paper are as follows. First, we show using *ns-2* simulations that when the total buffer size is greater than a critical value, UDP benefits with dedicated buffers that protect it from the aggressive nature of TCP. However, for buffer size that are below this critical value, UDP can benefit by time-sharing the buffers with TCP. Second, we develop a simple linear model that quantitatively captures the combined utility of TCP and real-time traffic for shared and dedicated buffers. We also propose a method to optimise what fraction of the total buffers should be dedicated to UDP and TCP with the objective of maximising the overall network utility. Our study equips designers of optical packet switched networks with quantitative tools to tune their buffer allocation strategies subject to various system parameters such as the ratio of the traffic mix and the relative weights associated with TCP and UDP traffic.

The rest of this paper is organised as follows. In Section II, we present simulation results and illustrate the anomalous loss behaviour when both UDP and TCP traffic share the total buffers. In Section III, we present simulation results for the dedicated buffers case and analyse what effect the buffer split ratio has on UDP and TCP traffic performance. We present a simple utility maximisation model in Section IV that determines the optimal buffer split ratio incorporating

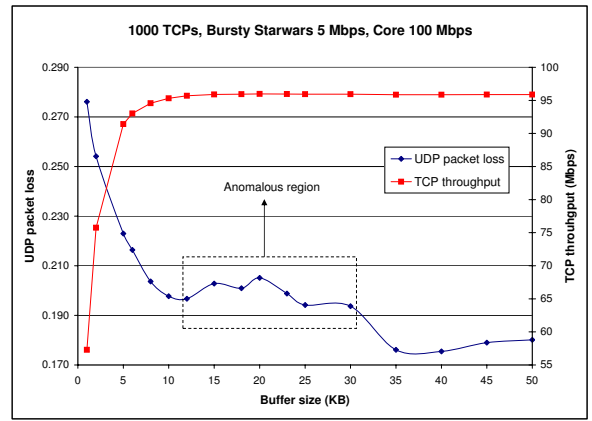


Fig. 2. Shared buffers case: UDP loss with the anomalous region

various system parameters. We conclude the paper in Section V and point to directions for future work.

II. SHARED BUFFERS

We use *ns-2* [17] 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 (variable bit-rate traffic) 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. Measurement based studies show that TCP is the predominant transport protocol, constituting in excess of 90% of the Internet traffic, with UDP accounting for about 5-10% [18], [19]. 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 bottleneck router r_0 is varied in terms of KiloBytes. There is only a single FIFO queue at r_0 and drop-tail active queue management is employed.

Fig. 2 shows the UDP loss and TCP throughput curves as a function of bottleneck buffer size. 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 loss drops again. However, in the range 12-30 KB, UDP loss does not decrease with buffer size. Clearly, this anomalous region should be 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 the bottleneck router 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 [15], [16]. Each KiloByte of extra buffering can add significantly to the cost of the optical switch [11], thus 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 appropriately.

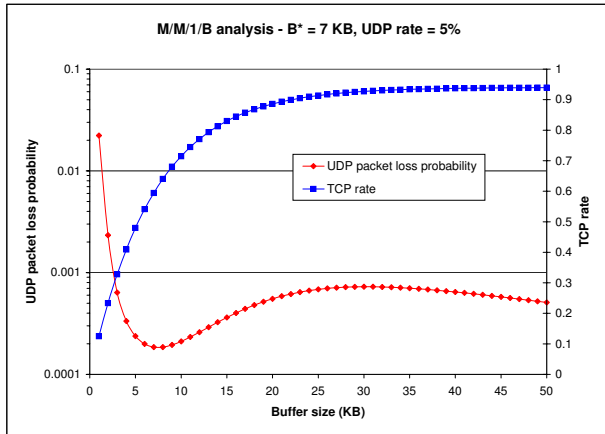


Fig. 3. Anomaly from the M/M/1/B Markov chain analysis

In [16], we developed a M/M/1/B Markov model for loss when both TCP and UDP traffic share the bottleneck buffers. For brevity, we do not reproduce our Markov chain but plot in Fig. 3 the UDP packet loss and TCP throughput as a function of bottleneck router buffer size. It can be observed from the figure that as the buffer size increases till about 8 KB, UDP loss falls monotonically. Further increase in buffer size increases UDP packet loss, showing that the model is able to predict the anomaly observed in simulations.

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, as studied next.

III. DEDICATED BUFFERS

In the absence of knowing exactly how large commercial optical buffers will be, but knowing that they are constrained to be within a few tens of KiloBytes, we undertake two sets of simulations to study how dedicated buffering will affect the performance of UDP and TCP.

In the first set, we fix the total buffers at the bottleneck router, 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 4, 5 respectively. For comparison, the plot also shows the performance of UDP and TCP when the total buffers are shared by the two.

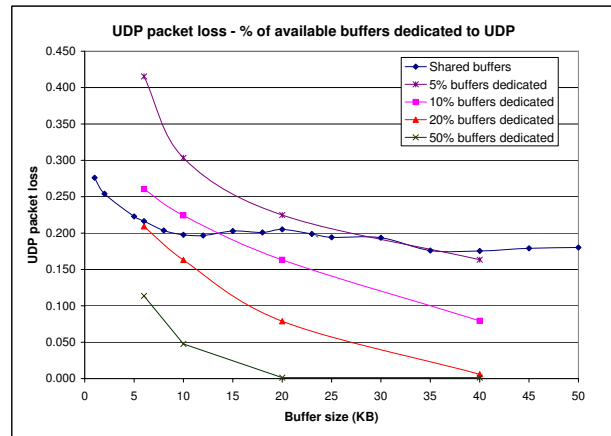


Fig. 4. UDP loss versus buffer size for different fractions of dedicated buffers

The top curve in Fig. 4 depicts UDP loss when 5% (proportional to the UDP rate which is 5%) of the total buffers is dedicated to it. It shows that up to a critical buffer size of around 35 KB, UDP performs worse with dedicated buffers when compared to shared buffers. As the fraction of allocation to UDP increases, e.g., 10%, 20%, this critical value of buffer size above which dedicating buffers to UDP is better reduces (to 15 KB and 5 KB respectively). First, this shows that the determination to share or dedicate buffers is non-trivial and depends both on the total buffer capacity and the split ratio. Second, allocation of buffers to UDP that is proportional to its traffic rate is insufficient for obtaining acceptable performance. While dedicating larger than fair-share of the total buffers to UDP will help reduce loss, it will impact upon TCP performance as discussed next.

Fig. 5 shows that when buffers are large (beyond 25 KB), TCP achieves near saturation throughput irrespective of whether the buffers are shared or dedicated. For smaller buffer sizes, TCP throughput with dedicated buffering is lower than with shared buffering by as much as 10%. Clearly, there is a tradeoff between UDP loss and TCP throughput; one comes at the cost of the other. These results suggest that

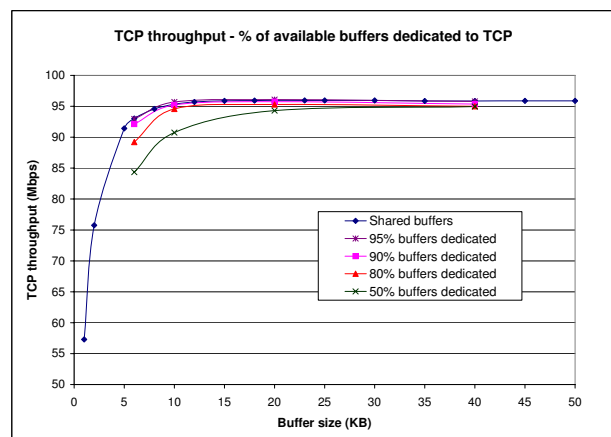


Fig. 5. TCP throughput versus buffer size for different fractions of dedicated buffers

if there is sufficient buffering available (greater than 20 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 6 and 7, 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 allocation varying from 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.

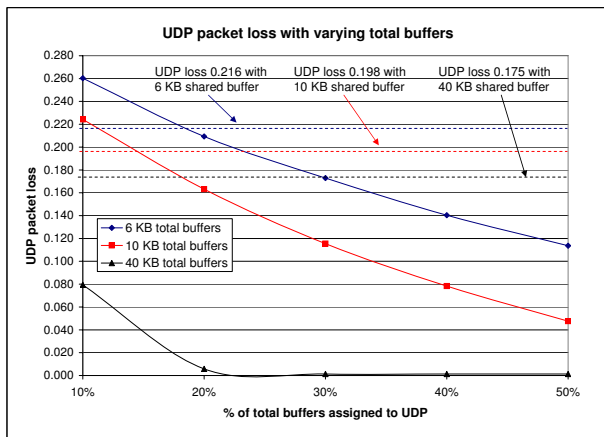


Fig. 6. UDP loss versus fraction of buffers assigned to UDP for different buffer sizes

Fig. 6 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. 7 that when the total buffers is 6 KB, TCP throughput falls from 92 Mbps to about 84 Mbps when its share of buffers reduces 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. A SIMPLE UTILITY MAXIMISATION PERSPECTIVE

We now attempt to develop a simple utility maximisation based model that gives insights as to whether the total buffer capacity should be shared by both TCP and UDP traffic, or be split such that a fraction of the buffers is pre-allocated

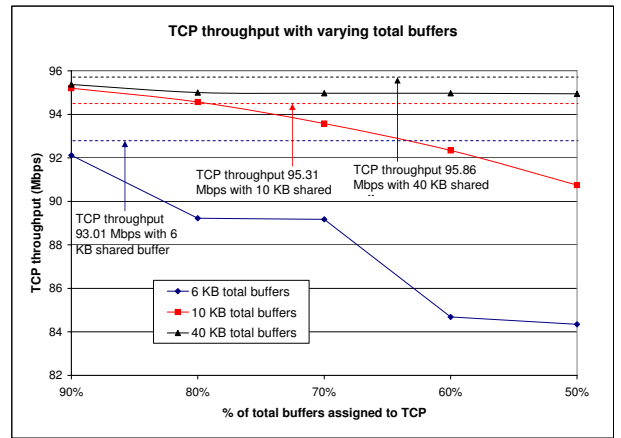


Fig. 7. TCP throughput versus fraction of buffers assigned to TCP for different buffer sizes

to UDP and the rest to TCP. From an ISP point of view, the decision would depend on the scenario that maximises the utility of the network. Intuitively, as more buffers are pre-allocated to UDP, which reduces the amount of buffers available to TCP, UDP loss would reduce, while at the same time TCP throughput would also reduce. This in turn has an implication on the network utility. We would like to capture this choice mathematically by a simple model. To this end, we first introduce the following notations.

- 1) Total buffer size is B KiloBytes.
- 2) λ_{tcp} is the TCP throughput. It is known [20, Sec. III], [3, Fig. 1] that TCP throughput increases fairly exponentially with buffer size. In other words, the probability that TCP leaves the buffer empty falls exponentially as buffer size increases. For example, Fig. 8 plots on log-scale the TCP empty buffer probability curve as a function of total buffer size B with 1000 TCP flows, and for two different sets of round-trip times. These parameters reflect the simulation settings used in Sections II and III. The fairly linear behaviour in the range 5 to 50 KB allows us to approximate the probability of an empty buffer as e^{-B/B^*} . Thus, TCP throughput is given by

$$\lambda_{tcp} \approx (1 - e^{-B/B^*}) \times \lambda_{tcp}^{sat} \quad (1)$$

where λ_{tcp}^{sat} is the TCP saturation throughput for very large buffer size.

- 3) B^* is a constant that depends on various parameters such as the number of TCP flows, round-trip times etc. We can infer B^* from the slope of the TCP empty buffer probability curve. If its slope is m , then $B^* = -1/m$. Fig. 8 allows us to estimate B^* as being 7 KB for the chosen set of simulation parameters.
- 4) x denotes the fraction of total buffers B dedicated to UDP traffic with $0 < x < 1$.
- 5) λ_{udp} is the fraction (rate) of UDP traffic. We set λ_{udp} to 0.05, i.e., 5% UDP, which is consistent with the simulation results shown previously, and the volume of UDP traffic in the Internet.
- 6) ρ is the utilisation of a M/M/1/K queueing system when

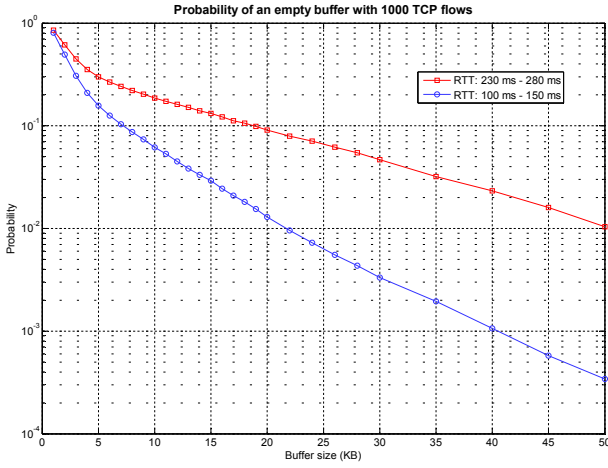


Fig. 8. TCP probability of empty buffer versus bottleneck router buffer size

buffers of size xB are pre-allocated to UDP traffic. The server capacity μ is normalised to unity. Thus, $\rho = \frac{\lambda_{udp}}{\mu} = \lambda_{udp}$. When traffic from a large number of sources multiplex at a bottleneck link, stochastic studies such as [21] have shown that the resulting aggregate traffic converges to a Poisson process. Hence, we employ this result and model UDP packet arrivals as a Poisson process with rate λ_{udp} . Also, we set $\lambda_{tcp}^{sat} = 0.94$ to ensure the stability of the system such that $\lambda_{tcp} + \lambda_{udp} < \mu$.

- 7) L_{udp} is the UDP packet loss probability for the M/M/1/K system. This is well known from standard queueing theory results.
- 8) k_1 and k_2 are positive constants (i.e., relative weights) attributed to TCP throughput and UDP packet loss respectively, which will be used as coefficients in the utility maximisation function.

We can model our utility maximisation objective as a function that is a linear combination of utilities of TCP throughput and UDP packet loss. An ISP will in general be interested in maintaining high TCP throughput and low UDP packet loss, and hence it is reasonable to combine these two quantities to maximise the utility of the network. Although this model is highly simplified, it serves for our illustration. More complicated cost functions that take into account non-linear combination of TCP throughput and UDP packet loss can be employed in our model. However, this is beyond the scope of the paper. Continuing, if we denote the network utility as $U(x)$, then

$$U(x) = k_1 \cdot \left(\frac{\lambda_{tcp}}{\lambda_{tcp}^{sat}} \right) + k_2 \cdot \log \left(\frac{1}{L_{udp}} \right) \quad (2)$$

The first term of the expression denotes the utility due to TCP throughput normalised to the saturation throughput. In other words, the utility of the network increases as the TCP throughput increases. The second term denotes the utility due to UDP packet loss. ISPs typically operate their network by ensuring that loss rates remain in order of 10^{-3} . Hence, the

choice of using $\log(1/L_{udp})$ is motivated by the fact that a reduction in UDP loss from say 10^{-3} to 10^{-4} is more likely to increase the revenue not by a factor of 10, but by a linear constant, and hence proportional to the logarithm of packet loss. Thus, the utility of the network due to UDP traffic increases as UDP packet loss decreases.

Further, the exact expression for L_{udp} from an M/M/1/K queueing system is given by $\frac{\rho^{xB}(1-\rho)}{1-(\rho^{xB+1})}$. However, since ρ in our case is very small (0.05), for all practical purposes, we can ignore the denominator and $L_{udp} \approx \rho^{xB}(1-\rho)$.

Thus,

$$U(x) = k_1 \cdot \left[1 - e^{-(1-x)B/B^*} \right] + k_2 \cdot \log \left(\frac{1}{\rho^{xB}(1-\rho)} \right) \quad (3)$$

For given buffer size B , constants k_1 and k_2 , and utilisation ρ , we are interested in finding x that maximises the overall network utility. This can be obtained in closed-form by differentiating the above expression and equating it to 0. Therefore,

$$\begin{aligned} \frac{d}{dx} U(x) &= \frac{d}{dx} \left[k_1 - k_1 e^{-(1-x)B/B^*} \right] + \\ &\quad \frac{d}{dx} \left[-k_2 \log(\rho^{xB}(1-\rho)) \right] \\ &= -k_1 e^{-(1-x)B/B^*} \frac{B}{B^*} - k_2 B \log(\rho) \end{aligned} \quad (4)$$

Equating (4) to 0, and after some algebraic manipulation yields,

$$x = 1 - \frac{B^*}{B} \ln \left(\frac{k_1/k_2}{B^* \log(1/\rho)} \right) \quad (5)$$

Intuitively, (5) confirms that for a given B^* , ρ and total buffer capacity B , the fraction of buffers dedicated to UDP monotonically decreases with k_1/k_2 , the relative importance of TCP to UDP performance.

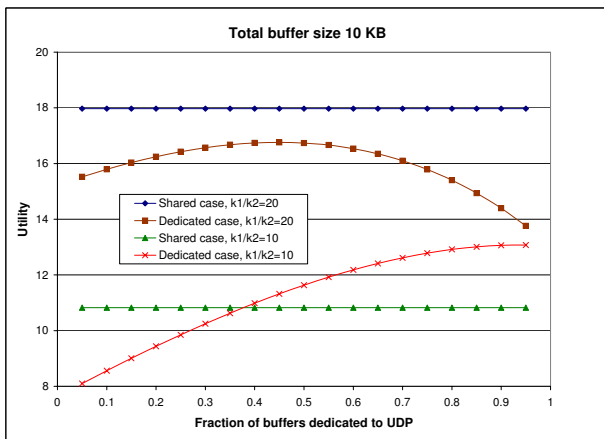
To see if $U(x)$ has a local maximum at x obtained from (5), we differentiate (4) w.r.t x to get the second derivative of $U(x)$. Thus,

$$\begin{aligned} \frac{d^2}{dx^2} U(x) &= \frac{d}{dx} \left[-k_1 e^{-(1-x)B/B^*} \frac{B}{B^*} - k_2 B \log(\rho) \right] \\ &= -k_1 e^{-(1-x)B/B^*} \left(\frac{B}{B^*} \right)^2 \end{aligned} \quad (6)$$

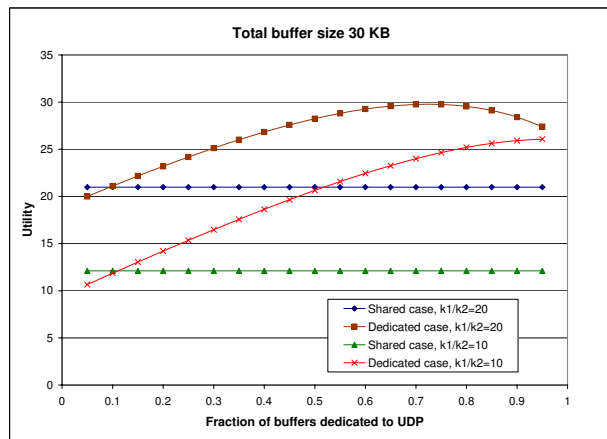
Substituting (5) in (6) and simplifying results in

$$\frac{d^2}{dx^2} U(x) = k_2 \frac{B^2}{B^*} \log(\rho) \quad (7)$$

Since RHS of (7) < 0 (because $\log(\rho) < 0$) it is clear that the function $U(x)$ has a local maximum at x obtained from (5).



(a) Utility for 10 KB total buffer capacity



(b) Utility for 30 KB total buffer capacity

Fig. 9. Network utility as a function of fraction of buffers dedicated to UDP for different buffer sizes and k_1/k_2 ratios

A. Impact of fraction x of buffers dedicated to UDP

Having obtained an optimal split ratio x , we are now ready to investigate the impact of various parameters on the utility maximisation objective. First, we are interested in analysing how the network utility changes as a function of the fraction of total buffers dedicated to UDP (i.e., x). We plot this in Fig. 9 for two sets of total buffer size (10 and 30 KB) and two choices of the ratio k_1/k_2 (10 and 20). The horizontal lines in the figures denote the network utility for the shared buffers case, which is obtained by substituting in (2) λ_{tcp} and L_{udp} derived from the Markov chain analysis shown in Fig. 3. The bottom two curves in Fig. 9(a) denote the network utility for the shared and dedicated cases respectively, when the total buffer capacity is 10 KB and $k_1/k_2 = 10$. For this choice of relative weights between TCP and UDP, the dedicated case results in higher network utility than the shared case when the fraction of buffers dedicated to UDP exceeds 40%. However, for the same relative weights but for a larger buffer size of 30 KB, the bottom two curves in Fig. 9(b) indicate that the dedicated case yields higher utility when at least 12% of the total buffers are pre-allocated to UDP.

Looking next at what happens when the weight of TCP is doubled, i.e., for $k_1/k_2 = 20$, the top two curves in Fig. 9(a) indicate that when the total buffer capacity is severely constrained, it is better to operate in the shared case, but as the total buffer capacity increases (from 10 to 30 KB), operating in the dedicated case results in higher utility, as shown by the top two curves in Fig. 9(b).

B. Impact of relative importance k_1/k_2 of TCP to UDP traffic

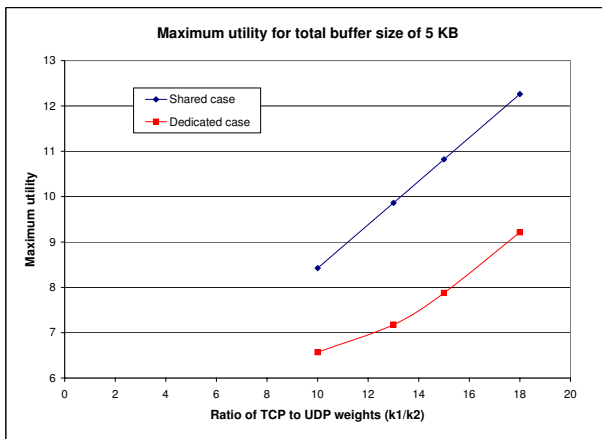
In this section, we are interested in analysing how the maximum network utility changes as the ratio k_1/k_2 changes. Since TCP is the predominant traffic in the Internet and the primary source of revenue for an ISP, we assign higher weights to TCP relative to UDP. We plot in Figures 10(a)-10(d) the maximum network utility (for both shared and dedicated buffers case) as a function of the ratio k_1/k_2 for four different choices of total buffer size (5, 10, 15 and 20

KB). We are only interested in those ratios of k_1/k_2 that result in $0 < x < 1$, which is an optimal choice of x such that xB buffers are assigned to UDP traffic and the rest to TCP traffic. For this reason, we will ensure that k_1/k_2 stays within the region $B^* \log(1/\rho) \leq k_1/k_2 \leq B^* \log(1/\rho)e^{B/B^*}$.

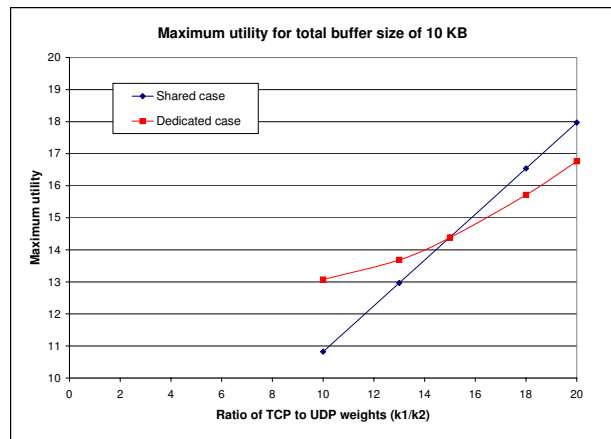
The following can be observed from the figures. First, (Fig. 10(a)) shows that when the total buffer capacity is extremely limited at 5 KB, it is beneficial for both UDP and TCP to share the buffers in order to maximise the network utility. For all admissible values of k_1/k_2 , the optimum utility for shared buffers is significantly higher than for dedicated buffers. Second, Fig. 10(b) shows that for an intermediate buffer size (10 KB) the ratio k_1/k_2 largely governs the maximum utility. For small k_1/k_2 , namely when TCP has relatively lower importance than UDP, utility is maximised by dedicating buffers to UDP to protect it against TCP traffic. However, as k_1/k_2 increases, i.e., TCP gains in importance, overall utility benefits from allowing TCP access to the entire buffers by sharing with UDP. As buffer size increases, such as to 15 KB in Fig. 10(c), the difference between the shared and dedicated case diminishes, particularly when the utility is dominated by TCP. When buffer size is 20 KB and larger, Fig. 10(d), dedicating buffers to both classes yields higher utility than shared buffers, corroborating our earlier results shown in Fig. 9(b).

V. CONCLUSIONS AND FUTURE WORK

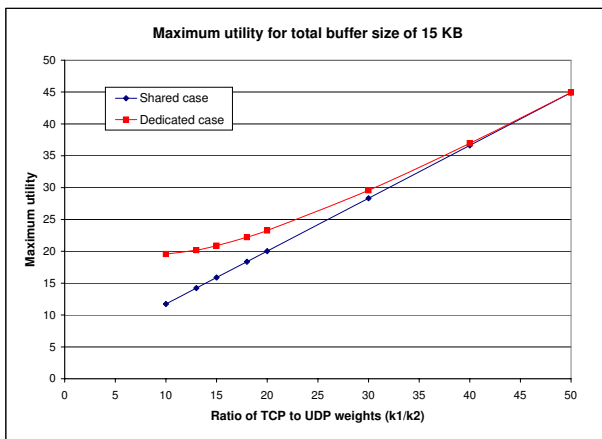
In this paper, we asked if TCP and UDP traffic should share bottleneck link buffers at OPS nodes with very limited buffering - sharing can result in more efficient buffer usage, but can lead to anomalous loss performance for UDP due to the greedy nature of TCP. The answer we found is not a definitive yes or no. It largely depends on how much buffering is available as well as the relative importance of TCP and UDP traffic. Using extensive simulations and a simple network utility maximisation objective we showed that there are three regimes: if the total buffer capacity is small (5 KB or less), it is in general better for TCP and UDP to share the buffers in order to maximise the utility.



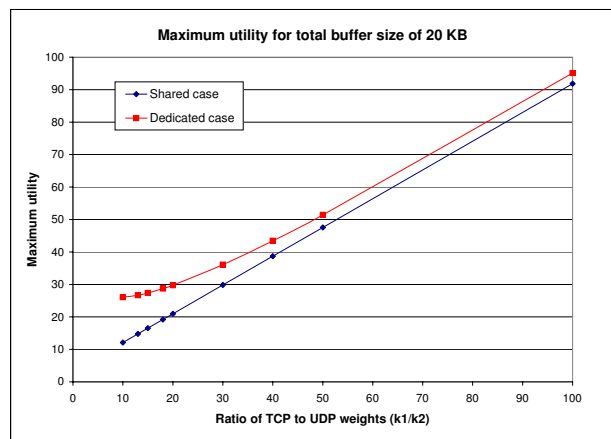
(a) Maximum utility for 5 KB total buffer size



(b) Maximum utility for 10 KB total buffer size



(c) Maximum utility for 15 KB total buffer size



(d) Maximum utility for 20 KB total buffer size

Fig. 10. Maximum utility for shared and dedicated buffers case for different ratios of k_1/k_2 and total buffer size

When buffer size is large (20 KB or more), it is beneficial for TCP and UDP to each have dedicated buffering, thus preventing their interaction which led to the anomalous UDP loss. In the intermediate range of buffer size (around 10 KB), dedicating buffers is desirable if UDP's performance is relatively more important, while sharing buffers performs better if TCP's importance dominates. These results inform network operators that apportioning buffers to real-time and TCP traffic is therefore non-trivial, and needs to take into account the buffering technology as well as revenue model.

As part of our future work, we plan to investigate the impact of more realistic utility models as our maximisation objective. Also, rather than pre-allocating a fraction x to UDP and the rest to TCP, we could employ a threshold-based buffering system that allows a fraction to be dedicated to each traffic class with the rest being shared between them.

REFERENCES

- [1] C. Villamizar and C. Song, "High performance TCP in ANSNet," *ACM SIGCOMM Computer Communications Review*, vol. 24, no. 5, pp. 45-60, 1994.
- [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] N. Beheshti, Y. Ganjali, M. Ghobadi, N. McKeown, and G. Salmon, "Experimental study of router buffer sizing," *Proc. ACM/USENIX Internet Measurement Conference*, Greece, Oct. 2008.
- [5] A. Vishwanath, V. Sivaraman and M. Thottan, "Perspectives on router buffer sizing: Recent results and open problems," *ACM SIGCOMM Computer Communications Review Editorial Zone* vol. 39, no. 2, pp. 34-39, Apr 2009.
- [6] R. S. Tucker, R. Parthiban, J. Baliga, K. Hinton, R. W. A. Ayre and W. V. Sorin, "Evolution of WDM optical IP networks: A cost and energy perspective," *IEEE/OSA Journal of Lightwave Technology*, vol. 27, no. 3, pp. 243-252, Feb 2009.
- [7] R. S. Tucker, "Optical packet switching: A reality check," *Elsevier Optical Switching and Networking*, vol. 5, pp. 2-9, Mar 2008.
- [8] M. L. Masanovic, V. Lal, J. S. Barton, E. J. Skogen, J. A. Summers, L. Rau, L. A. Coldren, and D. J. Blumenthal, "Widely-tunable monolithically-integrated all-optical wavelength converters in InP," *IEEE/OSA Journal of Lightwave Technology*, vol. 23, no. 3, pp. 1350-1362, Mar 2005.
- [9] 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.
- [10] C. Raffaelli et al., "Photonics in switching: Architectures, systems and enabling technologies," *Computer Networks*, vol. 10, no. 52, pp. 1873-1890, Jul 2008.
- [11] P. Bernasconi et al., "Architecture of an integrated router interconnected spectrally (IRIS)," *Proc. IEEE High Performance Switching and Routing*, Poland, Jun 2006.

- [12] "The optical switching revival: 2009 forecast," http://www.lightreading.com/webinar.asp?doc_id=28851
- [13] V. Sivaraman, H. ElGindy, D. Moreland and D. Ostry, "Packet pacing in short buffer optical packet switched networks," *Proc. IEEE INFOCOM*, Spain, Apr 2006.
- [14] V. Sivaraman, H. ElGindy, D. Moreland and D. Ostry, "Packet pacing in small buffer optical packet switched networks," *IEEE/ACM Transactions on Networking*, Oct 2009 (To appear).
- [15] 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, Jun 2008.
- [16] A. Vishwanath, V. Sivaraman and G. N. Rouskas, "Considerations for sizing buffers in optical packet switched networks," *Proc. IEEE INFOCOM*, Brazil, Apr 2009.
- [17] The Network Simulator - ns-2, <http://www.isi.edu/nsnam/ns/>
- [18] E. Brosh, G. L.-Sharon and Y. Shavitt, "Spatial-temporal analysis of passive TCP measurements", *Proc. IEEE INFOCOM*, USA, Mar. 2005.
- [19] "Packet traces from WIDE Internet backbone," <http://tracer.csl.sony.co.jp/mawi/samplepoint-F/2009/>
- [20] L. Andrew et al., "Buffer sizing for nonhomogeneous TCP sources," *IEEE Communications Letters*, vol. 9, no. 6, pp. 567-569, Jun. 2005.
- [21] J. Cao and K. Ramanan, "A Poisson limit for buffer overflow probabilities," *Proc. IEEE INFOCOM*, New York, USA, Jun 2002.