TCP over OBS - fixed-point load and loss

Craig Cameron, Hai Le Vu, JungYul Choi, Syed Bilgrami,
Moshe Zukerman and Minho Kang

C. Cameron, H. Vu, S. Bilgrami and M. Zukerman are with the ARC Special Research Centre
for Ultra-Broadband Information Networks (CUBIN), Department of Electrical and
Electronic Engineering, The University of Melbourne, VIC 3010, Australia. CUBIN is an
affiliated program of National ICT Australia (NICTA).
J. Choi and M. Kang are with the Optical Internet Research Center (OIRC), Information and
Communications University, 119, Munjiro, Yuseong-gu , Daejeon, South Korea.

c.cameron@ee.unimelb.edu.au, h.vu@ee.unimelb.edu.au, s.bilgrami@ee.unimelb.edu.au

m.zukerman@ee.unimelb.edu.au, passjay@icu.ac.kr, mhkang@icu.ac.kr

Abstract: The sending rate of commonly used TCP protocols is tightly coupled to packet loss within the network: a high rate of packet loss will cause a sender to slow down, thereby reducing the network load and decreasing subsequent packet loss rates. In this paper, we combine a widely verified source rate TCP model with an Optical Burst Switching (OBS) loss model, to find fixed-point input loads and loss rates for an OBS link carrying TCP traffic. In doing so, we show that if OBS networks are to be efficiently used to carry TCP traffic, many wavelengths with full wavelength conversion are required.

© 2005 Optical Society of America

OCIS codes: (060.4510) Optical communications.

References and links

1. A. M. Odlyzko, "Internet traffic growth: Sources and implications," in *Proceedings of SPIE Optical Transmission Systems and Equipment for WDM Networking II*, **5247**, 1–15 (2003).
2. K. Claffy, G. Miller, and K. Thompson, "The Nature of the Beast: Recent Traffic Measurements from an Internet Backbone," (1998), http://www.caida.org/outreach/papers/1998/Inet98/Inet98.html.
3. M. Fomenkov, K. Keys, D. Moore, and K. Claffy, "Longitudinal study of Internet traffic in 1998-2003," (2003), http://www.caida.org/outreach/papers/2003/nlanr/nlanr_overview.pdf.
4. M. Zukerman, E. Wong, Z. Rosberg, G. M. Lee, and H. L. Vu, "On Teletraffic Applications to OBS," IEEE Commun. Lett. **8**, 116–118 (2004).
5. J. Padhye, V. Firoiu, D. F. Towsley, and J. F. Kurose, "Modeling TCP Reno Performance: A Simple Model and Its Empirical Validation," IEEE/ACM Transactions on Networking **8**, 133–145 (2000).
6. S. Floyd and V. Paxson, "Difficulties in simulating the Internet," IEEE/ACM Transactions on Networking **9**, 392–403 (2001).
7. S. Gowda, R. K. Shenai, K. M. Sivalingam, and H. C. Cankaya, "Performance Evaluation of TCP over Optical Burst-Switched (OBS) WDM Networks," in *Proceedings of IEEE International Conference on Communications (ICC), Anchorage, USA* (Institute of Electrical and Electronics Engineers, New York, 2003), pp. 1433–1437.
8. X. Yu, C. Qiao, and Y. Liu, "TCP Implementations and False Time Out Detection in OBS Networks," in *Proceedings of IEEE INFOCOM, Hong Kong, China* (Institute of Electrical and Electronics Engineers, New York, 2004), pp. 774–784.
9. A. Detti and M. Listanti, "Impact of Segments Aggregation on TCP Reno Flows in Optical Burst Switching Networks," in *Proceedings of IEEE INFOCOM, New York City, USA* (Institute of Electrical and Electronics Engineers, New York, 2002), pp. 1803–1812.
10. Z. Rosberg, H. L. Vu, M. Zukerman and J. White, "Performance Analyses of Optical Burst Switching Networks," IEEE J. Sel. Areas Commun. **21**, 1187–1197 (2003).
11. H. L. Vu, A. Zalesky, E. W. M. Wong, Z. Rosberg, M. S. Bilgrami, M. Zukerman, and R. S. Tucker, "Scalable Performance Evaluation of a Hybrid Optical Switch," J. Lightwave Technol. **23**, 2961–2973 (2005).

12. H. L. Vu and S. Hanly, "A Study of TCP Performance and Buffer Occupancy Over a Fading Wireless Link," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM), Texas, USA* (Institute of Electrical and Electronics Engineers, New York, 2001), pp. 3478–3482.
13. J. Mahdavi and S. Floyd, "TCP-Friendly Unicast Rate-Based Flow Control," Technical note sent to the end2end-interest mailing list, Jan. 8, 1997.

1. Introduction

In the last twenty years, the Internet has changed from a university-based research network to a ubiquitous communication medium that enables a diverse range of useful applications, including email and the World Wide Web. Within the USA, the amount of Internet data traffic surpassed that of voice traffic several years ago and continues to grow rapidly, approximately doubling every year since 1997 [1]. Assuming this growth continues, it is expected that current electronic network architectures and protocols will soon be unable to carry such large amounts of traffic.

This expectation has fuelled interest in all-optical technologies that do not require optical signals to be converted to electronic form at each hop within the network. Given that advances in Wavelength Division Multiplexing (WDM) have enabled hundreds of wavelengths to be sent down a single fiber, optical-electronic-optical (OEO) conversion of every wavelength on every fiber may become infeasible due to cost, power and space constraints. It is therefore expected that all-optical networks will replace current networks at some point in the future.

OBS has been proposed as a future high-speed switching technology for all-optical networks that may be able to efficiently utilize extremely high capacity links without the need for data buffering or optical-electronic conversions at intermediate nodes. Packets arriving at an OBS ingress node that are destined for the same egress OBS node and belong to the same Quality of Service (QoS) class are aggregated and sent in bursts. At intermediate nodes, the data within the optical signal is transparently switched to the next node according to forwarding information contained within a control packet preceding the burst. At the egress node, the burst is subsequently de-aggregated and forwarded electronically. Unlike classical circuit switching, contention between bursts may cause loss within the network.

A trace taken by the Cooperative Association for Internet Data Analysis in 1998 showed that 95% of the bytes, 90% of the packets, and 80% of the flows on the examined link used variants of the Transmission Control Protocol (TCP) [2]. While the uptake of new applications using User Datagram Protocol (UDP), including streaming media and grid computing, seems to have increased over the last few years, measurements from 2003 show that TCP applications, including peer-to-peer file-sharing and the world wide web, continue to dominate [3]. Assuming that this dominance will continue, at least in the near future, it is crucial to be able to analyze the performance of TCP over new optical network architectures and protocols.

The sending rate of a TCP source is tightly coupled to packet loss within the network: a high rate of packet loss will cause a sender to slow down, thereby reducing the network load and decreasing subsequent packet loss rates. If all endpoints are transferring large files using TCP, the range of feasible input loads collapses to a single fixed operating point. Therefore, a useful result in this case is not traditional load vs. loss but expected TCP input rate vs. network parameters, such as round-trip delay and wavelengths/link.

In this paper, we link theoretical models that estimate loss in OBS networks as a function of input load [4] and TCP sending rate (input load) as a function of loss [5]. From these two models, we find the fixed point - the steady state input rates of TCP sources over an OBS link.

2. OBS and TCP Models

It is important to note that the main contribution of this paper is not performance analysis of a particular OBS or TCP variant but the linkage between TCP and OBS. We choose simple models to clearly demonstrate this link.

2.1. OBS Network Topology

Burst loss probability in OBS networks is tightly coupled to transient traffic matrices and the network topology. Choosing a single definitive matrix and topology is an intractable problem, as explained in [6]. In this paper, we try to limit the dimensionality of the parameter space, collapsing the network to a single link. While doing so may limit the direct applicability of our results, we are able to clearly isolate the theoretical process, which is the main contribution of our paper, from distracting OBS and TCP minutiae. There have been many simulations of TCP over OBS, including [7, 8, 9], but to the best of the authors' knowledge, none find the fixed point input TCP load or loss. In addition, the fixed point approach can be applicable to situations where the simple single OBS model is replaced by other models. For example it can be replaced by the network model of [10], by the single node optical hybrid switch model [11] and even by a wireless system model [12].

For the remainder of this paper, we analyze a simplified single-link topology with F physical optical fibers with W wavelengths per fiber, giving a total of K output wavelengths. We assume full wavelength conversion for values of $W > 1$: any input wavelength can be switched to any output wavelength. In the case of no wavelength conversion, K is equal to F. We consider a finite number of buffers, M, each with a finite number of TCP sources, N, such that the total number of TCP sources equals $N * M$. The number of buffers, M, is chosen such that the bottleneck and subsequent loss will be in the OBS network, not at the buffers. This implies a many-to-one mapping between buffers and egress nodes: all packets queued in a common buffer are destined for a common egress node, however there may be multiple buffers that map to a single egress node to meet QoS requirements for certain traffic streams. A graphical representation of this single-link topology is shown in Fig. 1.
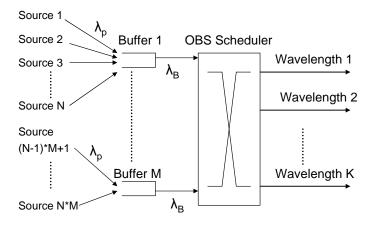

Fig. 1. Simplified single-link topology

2.2. OBS On-Off Source Traffic Model

On first glance, it is tempting to use TCP traces from real networks in order to obtain a realistic traffic characterization. However, OBS networks have not yet been deployed and even if they were, the resulting traces would be *shaped* [6]. A trace reflects the conditions in the network at the time the connection was measured. Due to the feedback incorporated in the rate control mechanisms of TCP, traces cannot be reused in another context as the sources would have behaved differently. Instead, models that characterize source behaviour must be employed either to generate the traffic for a simulation or to produce theoretical results. Simulations can be extremely difficult, especially with something as complex as a large, high-speed optical communication network [6]. This paper presents theoretical results that serve as a base for further complex simulations.

We consider a finite number of TCP sources, $N * M$, and a finite number of TCP sinks, M. Each TCP source generates an independent data stream destined for a unique TCP sink. All traffic destined for a common TCP sink is buffered together, such that there are M independent buffers, each buffering N streams. Packets from each stream arrive in accordance with a Poisson process with an arrival rate of λ_p and a fixed service rate of μ_p.

Each T seconds, the corresponding buffered packets are aggregated into a burst and passed on to the scheduler for possible transmission. Bursts depart in accordance with an On-Off process, with the burst departure rate, λ_B and burst service rate, μ_B defined by the following equations:

$$\lambda_B = 1/T. \tag{1}$$

$$\mu_B = \frac{\mu_p}{NT\lambda_p}. \tag{2}$$

The burst input load is then

$$\rho_B = \frac{\lambda_B}{\mu_B} = \frac{N\lambda_p}{\mu_p} = N\rho_p. \tag{3}$$

The average *on* and *off* periods are shown in Fig. 2. Note that the inverse of the average *on* period, $\hat{\mu}_B$, is equal to μ_B, while the inverse of the average *off* period, $\hat{\lambda}_B$, is equal to $1/(T-1/\mu_B)$. We assume that the probability of no packets arriving within the interval T to be extremely low and we therefore ignore this case. In Section 4, we calculate an example and verify this assumption.
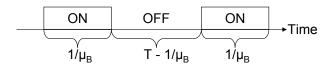

Fig. 2. On-off model mean values.

2.3. OBS Loss Model

We use the framework from [4] to directly calculate the burst loss probability for a single link OBS network from the *on/off* parameters, $\hat{\mu}_B$ and $\hat{\lambda}_B$. This framework does not model the effect of Fiber Delay Lines (FDLs), however a more complex framework linking loss and load can be substituted if needed.

There are three types of customers: (1) busy (bursts that are being transmitted), (2) free (unused input link), and (3) blocked (bursts that are being dumped). Let $\pi_{i,j}$ be the steady state probability where $i(0 \le i \le K)$ is the number of busy customers and $j(0 \le j \le M - K)$ is the number of frozen customers (sources who transmit blocked bursts). To ensure that the steady state of the Markov chain exists, we restrict values of N, λ_p and μ_p such that $\rho_b < 1$.

We now have the following steady state equations:

$$[(M-i-j)\hat{\lambda}_B + (i+j)\hat{\mu}_B]\pi_{i,j} = (M-i+1-j)\hat{\lambda}_B\pi_{i-1,j}$$
$$+ (j+1)\hat{\mu}_B\pi_{i,j+1} + (i+1)\hat{\mu}_B\pi_{i+1,j}. \tag{4}$$

and

$$[(M-K-j)\hat{\lambda}_B + (K+j)\hat{\mu}_B]\pi_{K,j} = (M-K+1-j)\hat{\lambda}_B\pi_{K-1,j}$$
$$+ (j+1)\hat{\mu}_B\pi_{K,j+1} + (M-K+1-j)\hat{\lambda}_B\pi_{K,j-1}. \tag{5}$$

and a normalization equation,

$$\sum_{i=0}^{K}\sum_{j=0}^{M-K} \pi_{i,j} = 1. \tag{6}$$

The input load is given by

$$T_o = \sum_{i=0}^{K}\sum_{j=0}^{M-K}(M-i-j)\hat{\rho}_B\pi_{i,j}, \tag{7}$$

the carried load is given by

$$T_c = \sum_{i=0}^{K}\sum_{j=0}^{M-K} i\pi_{i,j}, \tag{8}$$

and the burst blocking, or equivalently loss, probability is obtained by

$$B = \frac{T_o - T_c}{T_o}. \tag{9}$$

By solving these steady state equations, we can find the input load, the carried load and the subsequent loss.

2.4. TCP Model

To calculate the TCP sending rate per source, S [pkt/s], for a particular loss rate, p, and end-to-end Round Trip Time, RTT, we use the simple model for a saturated source from [13]:

$$S(p) = \frac{\sqrt{1.5}}{RTT}\frac{1}{\sqrt{p}}. \tag{10}$$

We assume a fixed packet service rate of μ_p, and fix the TCP packet size at 1KB. Furthermore, loss events for parallel sources are assumed to be independent.

3.   Combining the TCP and OBS Models

We find a fixed point for the input load of TCP over OBS networks by successively applying the calculations from Section 2.3 and Equation 10 until sufficient convergence is reached. The relationship between the equations is outlined in Fig. 3 and consists of four main steps - setting the input rates per source (TCP), calculating the corresponding burst distributions (OBS), calculating the resulting loss (OBS), calculating the expected sending rate for this loss (TCP) and
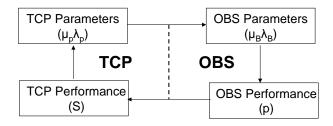
Fig. 3. Calculating fixed point of TCP input load and OBS loss

then using this sending rate as the new input load. We assume that the burst loss probability is equal to the packet loss probability.

A more intuitive way of representing the relationship between TCP and OBS is to overlay both loss and load graphs as in Fig. 4. The intersection point corresponds to the fixed point solution.
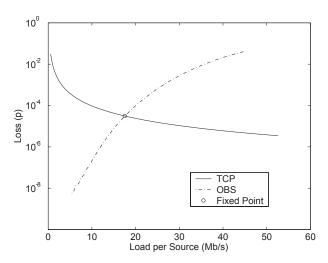

Fig. 4. Graphical method to find fixed point loss (M=16,K=10,RTT=0.1s).

4.   Numerical Results

We calculated fixed point loss and input loads over a range of numbers of buffers, M and output wavelengths, K, using the simplified topology of Fig. 1, Additional parameters were fixed: $N = 100$ sources per buffer, TCP round-trip time = 0.1s, output capacity per wavelength = 10Gb/s, TCP packet size = 1KB and timer value $T = 1$ms.

The model introduced in Section 2.2 assumes that the probability of no packets arriving in time T is very close to zero and therefore can be ignored. Packets arrive according to independent Poisson processes, therefore the probability of no packet arriving in time T from N independent streams, each with rate λ_p, is equal to $exp(-(NT\lambda_p))$. For an input load of 1.6Mbps ($\lambda_p = 200$ packets/s), the probability of no packets arriving in 1ms is very small; approximately 2.1×10^9.

Figure 5 shows the fixed point input TCP load per source versus the number of output wave-lengths, K, for a range of numbers of buffers, M. Across all values of M, the fixed point load increases exponentially as extra wavelengths become available, however, for large values of M, such as $M = 96$, the incremental benefit of adding wavelengths is much less than for smaller values of M. As discussed in Section 2.3, values of $M < K$ are outside the scope of the paper as M is chosen such any bottleneck occurs in the OBS link, not at the buffers.
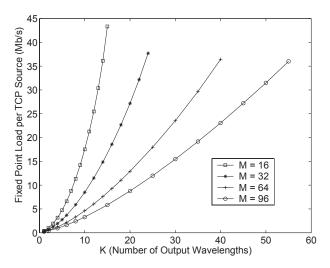

Fig. 5. Input load per source vs. number of output wavelengths.

Figure 6 shows the fixed point loss for the loads corresponding to the fixed point load in Fig. 5. As the number of buffers, M, increases, the number of output wavelengths, K, required to give a constant loss also increases but at a slower rate than the change in M. For example, to achieve a loss rate of 10^{-4}, only 8 wavelengths are needed for $M = 16$, but 22 wavelengths are needed if $M = 96$.
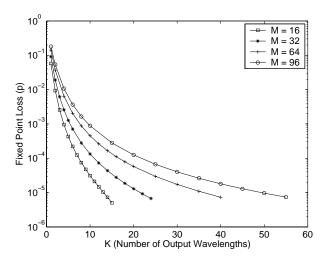

Fig. 6. Loss vs. number of output wavelengths.

Both results show that OBS scales well as the number of output wavelengths, K, increases.

Indeed, for values of K less than 5, each TCP source is restricted to rates of under 5Mbps with high loss of over 10^{-3}. As the number of output wavelengths increases beyond 20, all values of M give a reasonable loss of less than 10^{-4}, suggesting that OBS networks require large numbers of output wavelengths to support TCP traffic.
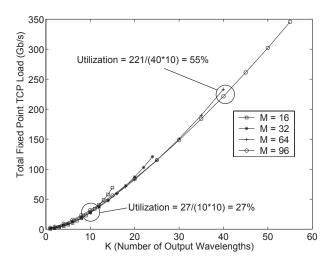

Fig. 7. Total input load vs. number of output wavelengths.

However, when summing over all input sources ($N * M$), the total input rate is roughly independent of the number of buffers M, as shown in Fig. 7. This result enables an OBS network to be dimensioned to achieve a specific total fixed-point input rate. Furthermore, for low values of loss, Fig. 7 also gives an estimate of utilization. For example, for $K = 40$ and $M = 96$, the total fixed point TCP input load is 221 Gbps or a utilization of $221/(40 * 10) \approx 55\%$. Similarly, for $K = 10$ and $M = 32$, the corresponding utilization is $27.2/(10 * 10) \approx 27\%$. Note that $K = FW$, where F is the number of physical fibers comprising the link and W is the number of wavelengths per link. In the case of full wavelength conversion, to achieve $K = 40$, the number of physical fibers, F, can be quite small, for example $F = 2$, however in the case of no wavelength conversion, the number of physical fibers, F, must equal the number of wavelengths, K. Therefore, this result shows that full wavelength conversion with many wavelengths per link results in much higher efficiency than no wavelength conversion, especially for large number of wavelengths per physical fiber, W.

5. Conclusion

We introduced a fixed-point method that incorporates TCP's feedback mechanism and can be used to compute the expected fixed-point steady-state TCP input rate over a single Optical Burst Switched optical link. As the number of wavelengths per link increased, the fixed point TCP rate increased, implying that large numbers of wavelengths are required for OBS networks to efficiently carry TCP traffic.

Acknowledgments