Impact of High TCP's Initial Window in Congested Links

Balekaki Gerald Nathan[1] and Idris A. Rai[2]

[1] Makerere University
P.O.Box 7062
Kampala, Uganda
balekaki.gerald@gmail.com
[2] rai@cit.mak.ac.ug

Abstract. Recent studies proposed that the permitted TCP initial window be increased from between 2 and 4 segments to 10 segments estimated to about 15KB of data. The increase has been mainly motivated by accelerated Internet growth coupled with high speeds and penetration levels in the world today. Over 95% of Internet traffic are short flows and majority of the traffic is TCP. Previous work have already studied the performance of TCP with IW10, but are not conclusive on the benefits of IW10 for short flows under highly congested links. In this paper we considered congested links to be always fully utilized, which is majorly the case for many links in developing countries. We investigate the impact of IW3 and IW10 on short flows of 9,21 against 10,22 packets on a highly congested link with the aid of experimentation and emulation methods. Obviously our results show better performance in 9,21 packets than 10,22 packets despite very high congestion conditions. Our results also show the TCP with IW3 favours shorter flows of 9,10 packets in terms of shorter flow completion time and higher throughput, under a highly congested link as opposed to TCP with IW10 by previous work. While larger flows of 21,22 packets are slighly favoured by TCP with IW10 because of their bursty property which in turn overwhelm TCP with IW3.

Key words: initial window, highly-congested links, congestion patterns

1 Introduction

The congestion window limits the amount of unacknowledged data the TCP sender injects into a network to prevent overwhelming of the network with larger bursts of data traffic. The initial value of congestion window which is also known as the *initial congestion window*, is mainly used by the TCP during congestion control, both at the beginning of a new connection or after a timeout. The TCP's initial window value has been incremented overtime a decade age, from one segment to roughly four segments [1], since then the Internet has continued to experience accelerated growth of penetration to over 487 million unique IP addresses, higher broadband adoption levels [2], and huge deployment of heavy bandwidth applications such as *Google Earth, Youtube*, Web browsers like *Internet Explorer (IE8), Mozilla Firefox, Google Chrome* and many more browsers that had devised means of opening up multiple connections in order to increase speed of Web downloads. In order to cope up with the increasing demands of Internet speeds, there has been a proposal to increase the TCP's initial window from between 2 and 4 segments to 10 segments estimated at 15KB of data [3] such that much more data can be transmitted at connection startup and eventually will increase exponentially as required by standard slow start algorithm (SS) [4]. There is also an argument in [5] supporting the use of a high TCP initial window size of 10 segments. Despite all the benefits described in [5] IW10 is associated with, most of them have been observed in high-speed links but not congested links.

Recent measurement studies presented in [5], [6] have shown high initial window improves the response time of short flows by up to 10% on average thus improving the overall performance of many web services without risking congestion collapse. Additionally, for links such as high-speed links, high initial window has an additional benefit of improving link utilization. While this is intuitive for non-congested links, similar benefits can not be obviously acclaimed for highly congested links. For instance, the study in [5] also shows that the average response time may worsen if the high initial congestion window is used for browsers with multiple concurrent connections connected to low-speed links. Most of these studies support the use of high initial window but do not clearly show its impact on short flows in congested networks, which is the case in developing countries especially those largely depicted in Africa and South America are usually congested with a high number of users accessing a limited access link at the same time.

In this paper, we compared the performance in terms of completion time or latency, throughput for TCP with IW3 and IW10 for highly congeted links. We observe that short flows have shorter completion time which greatly result into throughput under a TCP with IW3 as opposed to previous work where by short flows are favoured by TCP with IW10. We again oppose previous work on the short flow sizes of 10 and 22 packets which may be biased against IW3. For instance, TCP with IW3 sending 10,22 packets, transmits an extra window of one additional window of one packet unlike for 9,21 packets. We therefore compared the performance of 10, 22 packets with the performance of 9, 21 packets. Obviously our results show better performance in 9,21 packets than 10,22 packets despite high congestion on the link. Shorter flows of 9,10 packets favors IW3 as opposed to IW10 stated earlier by previous work. While larger flows of 21,22 packets slighly favor IW10 because of their bursty property which in turn overwhelm IW3.

The remainder of the paper is as follows. In the next section we present related work. In Section 3 we present the experimental setup and define parameters. In Section 4, we discus the results, and finally conclude in Section 5., we conclude the paper.

2 Related Work

The earliest proposal on increasing TCP's initial congestion window dates back more than a decade ago, [6] mainly to speed up the Internet. TCP initial window has remained unchanged since 2002 to only 4KB yet the use of a high initial window of 10 segments will benefit huge bandwidth applications. Recent work on larger initial window [3], [5] mainly focuses on increasing the TCP's initial window from four(4) segments to at least ten (10) segments. It is observed that high initial window is good for high-speed connections in reducing latency by over 10% on the network [5], [6]. Simulation studies on high TCP initial window have also shown an improvement in the response time of short flows which have led to the adoption of initial window size of 4KB that is less than 3 segments. However, simulations studies did not consider the realistic distribution of Internet traffic [7], and only considered the highest window size of 4KB.

The web has become much more popular now, leading to significant changes to Internet traffic distributions. The Internet traffic measurement studies have shown that Internet flow size distributions are highly skewed, this means it constitutes of many short flows (about 99% of the flows) that contribute to less than 50% of the bytes, and a tiny fraction of the largest flows (i.e less than 1 %) contributes to nearly half of the bytes [8], [9]. The observation of many short flows in the Internet is likely to persist due to the popularity of web and a plethora of existing and emerging web based applications or services. For instance, search service presents results using short flows, more than 90% of which are less than (10KB), and about 90% of HTTP objects from top 100 to 500 web sites are less than 16KB [5].

Web users naturally require low response time due to interactive nature of web browsing. However, the existing Internet's congestion control protocols do not consider the required urgency of short flows transfers, neither do they take into account the heavy-tailed flow size distribution of Internet traffic. Short flows are transferred during the slow start phase of the TCP protocol which is known to be dominated by connection's round-trip time. The newly established connection has no idea of congestion state on the path and the slow start phase is equivalent to probing the unknown path.

3 Experimental Setup

In this Section, we define our testbed on the simple dumbbell topology shown in Figure 1. We also recognize that this topology is a limited one, but the behavior of standard TCP on this topology is well studied and so it provides a natural starting point. Our setup consists of five (5) Linux boxes and two (2) Ethernet switches. The one (1) server in the middle uses Netem to emulate bandwidth and delay on the link. The other 4 servers are grouped into two (2) server/client pairs. Each pair may generate either a unidirectional long-lived flow using *Iperf* or a bidirectional short lived flow using *Netperf* software tools. Long-lived flows are used to emulate bulk data transfer like FTP while short-lived flows emulate

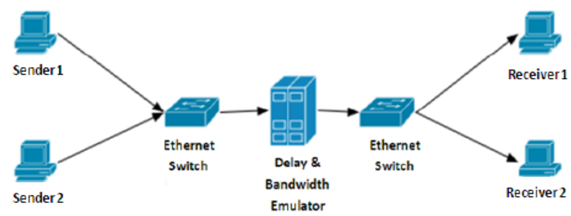


Fig. 1. Experimental Topology

short request/response traffic like Web. In our research, we use Network Emulation with NetEm [10] as one of the most efficient tool to emulate delay and bandwidth over a single link. Previous studies such as [10],[11] highly recommend the use of emulation in networking as a most effective method of evaluating real TCP implementations over very slow networks. In our experiments, we involve a number of tools used in evaluating TCP performance. However, we choose Wireshark measurement tool to capture data in traces as shown in Table 1.

Table 1. Traces used in our analysis

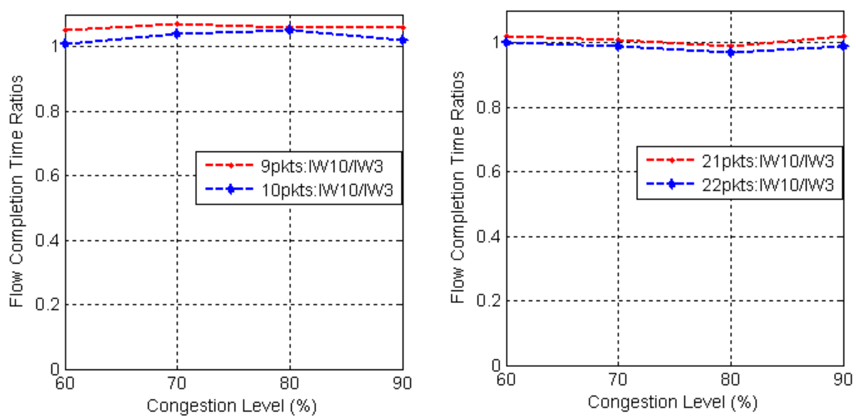
No.	Time	Source	Destination	Protocol	Info
1	0.084201	192.168.2.2	192.168.1.1	TCP	[ACK]Seq=1 Ack=25 Win=5824 Len=0
2	0.084493	192.168.1.1	192.168.2.2	TCP	[ACK]Seq=25 Ack=1 Win=5888 Len=1448
3	0.084552	192.168.2.2	192.168.1.1	TCP	[ACK]Seq=1 Ack=1473 Win=8704 Len=0
4	0.084628	192.168.1.1	192.168.2.2	TCP	[ACK]Seq=1473 Ack=1 Win=5888 Len=1448
5	0.084645	192.168.2.2	192.168.1.1	TCP	[ACK]Seq=1 Ack=2921 Win=11584 Len=0
6	0.184839	192.168.1.1	192.168.2.2	TCP	[ACK]Seq=2921 Ack=1 Win=5888 Len=1448
7	0.184932	192.168.1.1	192.168.2.2	TCP	[ACK]Seq=4369 Ack=1 Win=5888 Len=1448
8	0.18534	192.168.2.2	192.168.1.1	TCP	[ACK]Seq=1 Ack=4369 Win=14528 Len=0
9	0.185458	192.168.2.2	192.168.1.1	TCP	[ACK]Seq=1 Ack=5817 Win=17408 Len=0

3.1 Congestion Patterns

In this section, we define three congestion patterns i.e. 1) *Constant congestion*; where congestion rate is flat on the link and only induced by UDP background traffic. We also clearly state in our study that constant congestion is used to illustrate a constantly congested link. 2) *Varying congestion* is induced by TCP traffic only. 3) *Mixed congestion* is induced by mixed traffic such as TCP and UDP traffic. We also consider congestion level as a percentage of congestion on a link induced by each of the congestion patterns.

4 Experimental Results

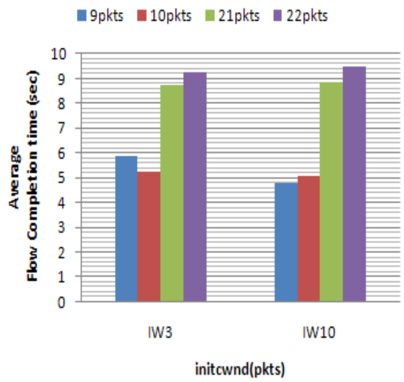
In this section, we looked at the performance comparison between IW3 and IW10 in terms of average flow response time(completion time), average throughput,and some impact on packet losses and timeout. We also compare the performance of IW10 with flows of 9, 21 packets to 10, 22 packets.



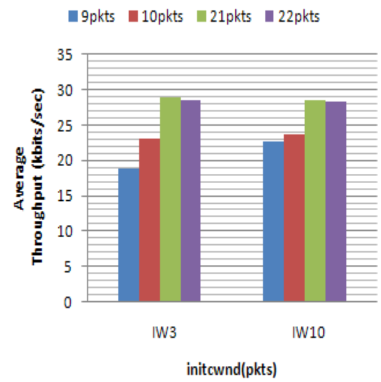
(a) Short flows 9,10 packets (b) Short flows of 21,22 packets

Fig. 2. Average flow completion time ratios under constant congestion

Figurel 2(a) showing ratios of completion time of flows with 9,10 packets. We observe that both curves are way beyond ratio of 1. This simply implies that IW10 completion times are much longer than IW3 in both flows of 9 and 10 packets. Therefore IW3 favours short flows of 9,10 packets in very high and constant congestion. We also obviosly note that short flows of 10 packets are more penalized than 9 packets, because that extra packets induces additional congestion which further leads to deterioration in performance. In figure 2(a), we notice a similar pattern. However, the ratios are just below or along 1. This signifies a very minimal performance difference between flows 21 and 22 packets. Ratios for 22 packets are actaully less than 1, which implies the completion times by IW10 are lower than IW3. Therefore larger flows of 21,22 packets transmits packets faster with IW10 as opposed to IW3 under congested link.



(a) Average flow completion time



(b) Average throughput

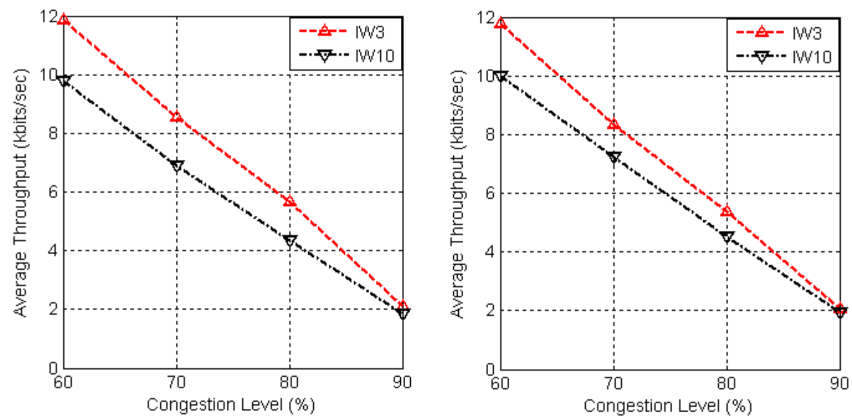
Fig. 3. Average flow completion time and throughput under varying congestion

Figure 3 evaluates performance of TCP with IW10 and IW3 under varying congestion on the link. This congestion pattern takes up approximately 60% on average, which makes the link less congested as compared to the constant or mixed congestion pattern. Short flows with 9 packets are much more favored by IW10 with a shorter completion time than flows of 10 packets. This effect greatly impacts on throughput whereby IW10 throughput of 9 packets is much higher than flows with 9 packets under IW3. For larger flows with 21 and 22 packets, we notice the performance is equally similar. However, IW3 flow completion values are lower for both 21 and 22 packets which results into a higher amount of throughput compared to IW10.

4.1 Throughput outperformance

We show the effect of using TCP initial windows of IW3 or IW10 subjected to higher levels of congestion of about 60% to 90% while observing the performance of short flows of sizes 9, 10, 21 and 22 packets.

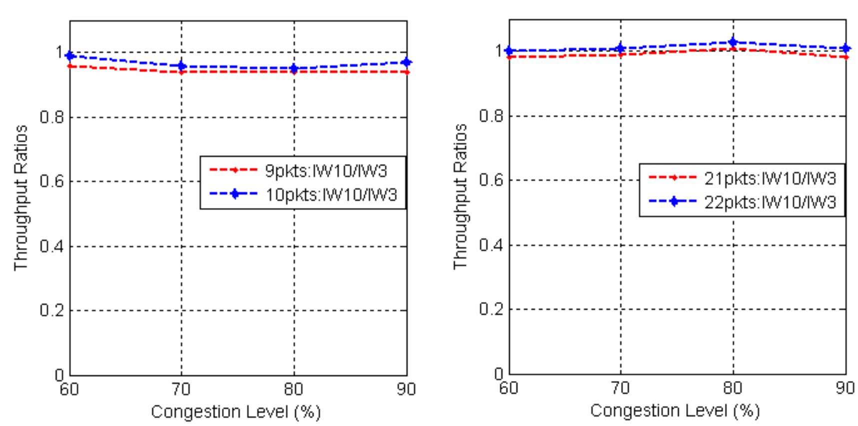
In Figure 4, we illustrate average throughput against congestion levels. As the congestion increases towards to 100% on the link, average throughput deteriorates to zero. This effect is clearly observed in Figure 4(a) and Figure 4(b) respectively. We further note that IW10 throughput of short flows of 9,10 packets reduces much faster than when using IW3. In very high congestion e.g at 90%, performance difference in terms of throughput while using IW10 or IW3 is equally very negligible. We finally conclude that average throughput of flows of 10 packets is much more affected than flows of 9 packets irrespective of the



(a) Short flows of 9 packets (b) Short flows of 10 packets

Fig. 4. Average throughput for shorter flows under mixed congestion

window size, simply because of the additional packets being transmitted. This finding renders previous work which only considered flows of 10 packets inclusive of the benefits of using high initial windows.



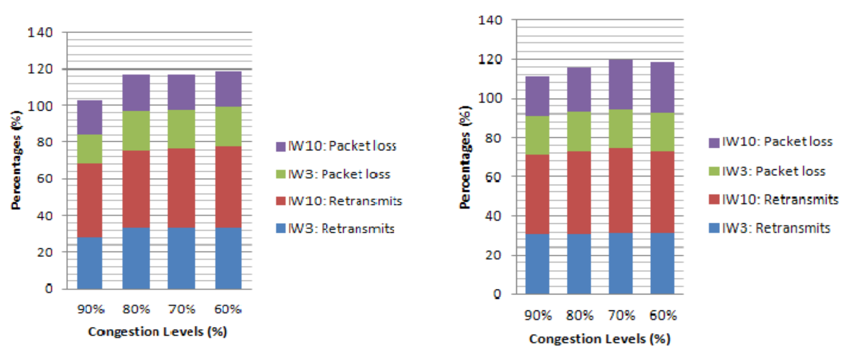
(a) Short flows of 9,10 packets (b) Short flows of 21,22 packets

Fig. 5. Throughput ratios for flows under constant congestion

Figure 5(a), illustrates the througput ratios against congestion levels for shorter flows of less than 10 packets. We observe that ratios are less that 1 irrespective of the flows size. This implies that shorter flows using IW3 transmits much greater throughput as opposed to IW10, under very high and constant congestion. However, we also note that throughput of flows with 9 packets is perferably higher than that in 10 packets. This is because IW10 is too aggressive and drops more packets as compared to IW3. In Figure 5(b), we considered larger flow sizes of 21,22 packets. We also observe a similar pattern but IW3 and IW10 throughput performance pattern is very minimal, thats why the ratios are close to one. Finally we note 21 packets are more favored while using IW3 than 22 packets, this is simply because IW3 will send an extra window while transmitting 22 packets unlike 21 packets. Where as IW10 is peferrable more suitable transimitting bursty flows carrying many packets of about 21 and 22 packets

4.2 Packet losses, retransmissions and timeouts

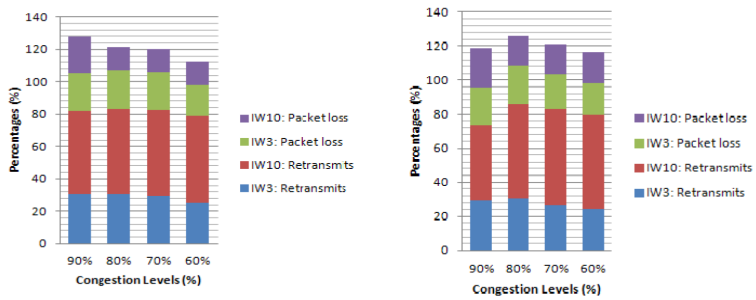
In this Section, we also determined the impact of high initial windows on flows of various sizes under heavy congestion.



(a) Short flows of 9 packets (b) Short flows of 10 packets

Fig. 6. Packet loss and retransmits for shorter flows under mixed congestion

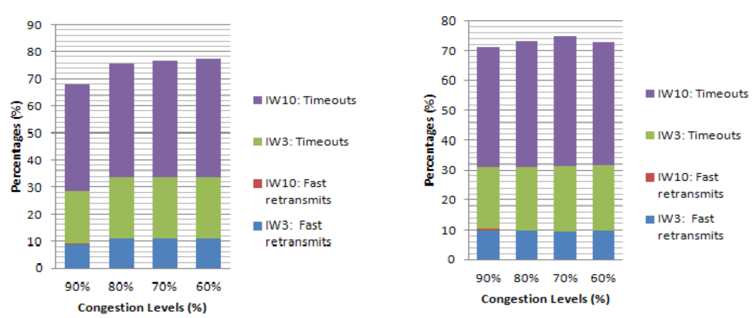
In Figure 6, we observe packet losses are much more in short flows carrying 10 packets than 9 packets. This is due to extra congestion induced by that extra packet sent. We also note IW10 drops more packets than IW3 under short flows of 10 packets as opposed to TCP with IW10 which drops less packets than TCP with IW3 under short flows of 9 packets as congestion increases. This is because flows with 9 packets transmitted under IW3 suffer much more resistance as compared to IW10 which is less resistant since most of the packets are carried through the pipe within one window. We also observe that IW10 retransmissions are more than IW3 retransmits for flows less than 10 packets.



(a) Short flows of 21 packets (b) Short flows of 22 packets

Fig. 7. Packet loss and retransmits for larger flows under mixed congestion

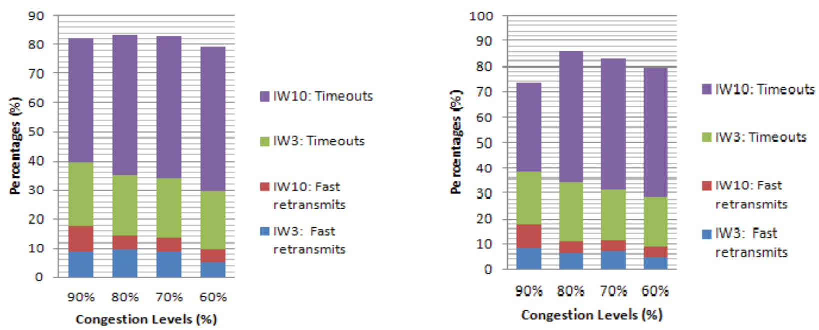
In Figure 7, short flows with 21 and 22 packets exhibit TCP with IW10 quite more packet losses than IW3 as congestion increases. Finally, we note that IW10 retransmissions are massively greater than IW3 retransmissions for short flows of 21 and 22 packets. This implies that many times, we experience packets being retransmitted more than twice especially when TCP detects severe packet losses due to congestion.



(a) Short flows of 9 packets (b) Short flows of 10 packets

Fig. 8. Average throughput for shorter flows under mixed congestion

In Figure 8, under short flows less than 10 packets, we observe that IW10 timeouts are twice as much as IW3 timeouts. This is mainly depicted by multiple retransmissions (retransmit a packets more than twice) from severe packet losses and delayed acknowledgments which is very common under very highly congested network.



(a) Short flows of 21 packets (b) Short flows of 22 packets

Fig. 9. Timeouts and fast retransmits under mixed congestion

Figure 9 showing the effect of IW3 and IW10 of short flows of 21 and 22 packets on on retransmissions and timeouts, indicate a similar pattern as in Figure 9 above. However, the difference comes in fast retransmissions where by flows with less than 10 packets record very minimal performance percentage compared to larger short flows of 21 and 22 packets which records IW10 with smaller percentage (about 5%) as opposed to being negligible as for 10 packets but IW3 records roughly about 10% under bursty flows.

5 Conclusion

We finally present the general conclusion of the research. We observed that IW10 is a poor performance initial window for flows less than 10 packets on a link that is experiencing congestion of about 80% and beyond. Short flows of 9,10 packets experience shorter completion time which greatly result into higher throughput under a TCP with IW3 which is contrary to previous work which recommends TCP with IW10. However, under a non-congested or slightly congested of about 60% congestion level and below, TCP with IW10 is very good for these flows as shown by previous work.

Secondly, we again deduce that results by short flow sizes of 10,22 packets are biased especially using TCP with IW3 rather that short flows with 9,21 packets. we have noted that results flows of 9,21 packets show better performance than 10,22 packets despite high congestion on the link. It is at this point that we recommend an automated or designated initial window value as network congestion changes.

References

1. M. Allman, S. Floyd, and C. Partridge, "Increasing TCP's Initial Window," sep 1998, iETF RFC 2414.
2. A. T. Inc, "State of the Internet," Cambridge, MA, First Quarter 2010, report.
3. H. K. J. Chu and D. Nandita, "Increasing TCP's Initial Window," mar 2010, iETF Internet Draft: draft-hkchu-tcpm-initcwnd-00.txt.
4. J. F. Kurose and K. Ross, *Computer Networking: A Top-Down Featuring the Internet*. Addison- Wesley, jul 2004.
5. N. Dukkipati, T. Refice, Y. Cheng, J. C. T. Herbert, A. Agarwal, A. Jain, and N. Sutin, "An Argument for Increasing TCPs Initial Congestion Window," in *Proc. ACM SIGCOMM Computer Communication Review*, vol. 40, no. 3, jul 2010.
6. M. Allman, C. Hayes, and S. Ostermann, "An Evaluation of TCP with Larger Initial Windows," in *Proc. ACM SIGCOMM Computer Communication Review*, vol. 40, no. 3, jul 1998.
7. K. Poduri and K. Nichols, "Simulation Studies of Increased Initial TCP Window Size," sep 1998, iETF RFC 2415.
8. V.Paxson and S.Floyd, "Wide-Area Traffic: The Failure of Poison Modeling," in *Proc. IEEE/ACM Transactions on Networking*, vol. 3, no. 3, jun 1995.
9. C. Williamson, "Internet Traffic Measurement," Nov 2001, university of Calgary.
10. S. Hemminger, "Network Emulation with NetEm," apr 2005, open Source Development Lab.
11. T. Henderson and R. Katz, "Transport Protocols for Internet-Compatible Satellite Networks," in *Proc. IEEE Journal on Selected Areas of Communications*, aug 1999.