TCP Behavior in Sub-Packet Regimes

Jay Chen
New York University
jchen@cs.nyu.edu

Janardhan Iyengar
Franklin and Marshall College
jiyengar@fandm.edu

Lakshminarayanan Subramanian
New York University
lakshmi@cs.nyu.edu

Bryan Ford
Yale University
bryan.ford@yale.edu

ABSTRACT

Many network links in developing regions operate in the *sub-packet regime*, an environment where the typical per-flow throughput is less than 1 packet per round-trip time. TCP and other common congestion control protocols break down in the sub-packet regime, resulting in severe unfairness, high packet loss rates, and flow silences due to repetitive timeouts. To understand TCP's behavior in this regime, we propose a model particularly tailored to high packet loss-rates and relatively small congestion window sizes. We validate the model under a variety of network conditions.

Categories and Subject Descriptors: C.2.2 [Computer-Communication Networks]: Network Protocols.

General Terms: Experimentation, Measurement, Performance.

Keywords: TCP, congestion control, low bandwidth networks.

1. INTRODUCTION

Existing congestion control schemes such as TCP-NewReno, TFRC and many others assume the fair-share bandwidth of a flow is at least 1 packet per round-trip time (RTT). Surprisingly, there exists a large number of low-bandwidth network environments in the developing world with high levels of network sharing where this assumption does not hold [6, 3].; we define such an environment as the *sub-packet regime*. While heavy sharing among users leads to the sub-packet regime, web browsers can exacerbate the problem by spawning several TCP connections per web request. Over the past decade, the average size of web pages and the number of objects per page has grown at a faster rate than the growth in connectivity.

The sub-packet regime has not been a traditionally important region of operation for network flows, and as a result this space has remained relatively unexplored. The concept of a sub-packet regime arises in prior work in the context of understanding the behavior of TCP in the face of many competing flows [5, 8, 4].

This paper proposes an analytical model to characterize the equilibrium behavior of TCP in the sub-packet regime. Our model is a simpler variant of a full Markov model for TCP operating in traditional regimes [2], but gives more careful attention to modeling repetitive timeouts, an extremely common state experienced by TCP flows in sub-packet regimes. Since Markov models are inherently not suited to keep memory in the state transitions, modeling repetitive timeouts is not straightforward (since one needs memory of the previous timeout value). We address this problem by determining aggregate transition states which both capture the memory effect while significantly reducing the number of states. Using extensive analysis, we show that our model accurately predicts the stationary distribution of a TCP flow across different states using few aggregate states. Our model can be used by network middle-boxes in practice to enhance TCP performance and fairness in sub-packet regimes in a non-intrusive manner.

2. OUR MODEL

We build a simple model particularly tailored for analyzing the behavior of TCP-NewReno in sub-packet regimes with high-packet loss-rates and with relatively small average congestion windows. The main purpose of this model is to analyze the *stationary distribution* of a set of TCP flows, which provides a detailed characterization of the state of a TCP connection.

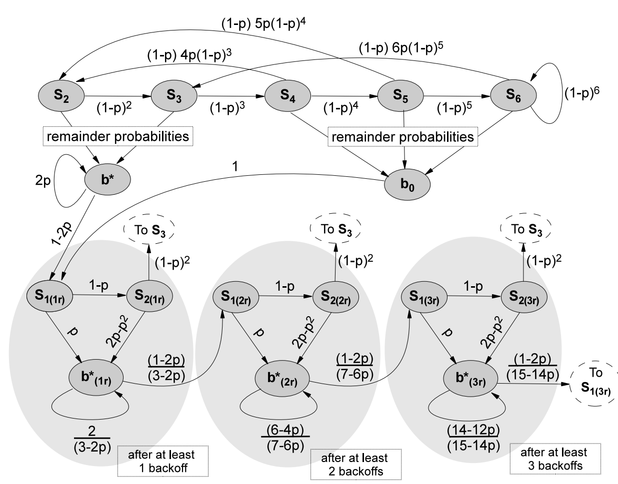

Figure 1: The Full Model for a max congestion window, $W_{max} = 6$.

SIGMETRICS'11, June 7–11, 2011, San Jose, California, USA.
ACM 978-1-4503-0262-3/11/06.

Our overall model is described in Figure 1 for a maximum congestion window, $W_{max} = 6$, and can be easily extended to larger W_{max}.. Our model is built around three assumptions. First, we assume that most TCP flows operate in small window sizes (less than W_{max}). Second, we assume that all TCP flows experience medium to high loss rates in sub-packet regimes. Third, given small congestion windows where TCP packets are more spaced-out, we can model packet losses using a single loss parameter, p.

There are several key insights we draw from this model. First, all the transition probabilities in this model are modeled using a single parameter p, the packet-loss probability at the bottleneck link. Second, under conditions where p is roughly a constant, the stationary distribution probability is only dependent on p and is independent of RTT. Third, we observe a shift in the stationary distribution beyond $p = 0.1$ where the probability of repetitive timeouts significantly increases thereby lowering the effective throughput of a TCP flow.

Our model is different from and extends previously proposed models of Padhye *et al.* [7], Fortin-Parisi *et al.* [2]. The fundamental problem with using a generic Markov model for capturing TCP behavior is the state space explosion. One of the advantages in the sub-packet regime, however, is that the state space is constrained and may be accurately captured using appropriate state transitions. Our model focuses on high loss rates and captures exponentially increasing silence periods due to repetitive timeouts, the dynamics of which are not captured in detail in prior work.

3. VALIDATION AND APPLICATIONS

We validate the model using ns2 simulations of TCP flows operating in sub-packet regimes. TCP-SACK is used at the endpoints and we run the simulations for varying levels of contention on a bottleneck link resulting in varying loss scenarios. For each simulation, we measure the packet loss rate p and also determine the distribution of the individual flows across different congestion window states. Figure 2 shows the model's predicted probability distribution for varying loss rate p, overlaid with results from simulations where we measure and plot probability against observed loss rate. For this simulation set, the flows all have a propagation RTT of 200ms, the bottleneck capacity is 1 Mbps, and the bottleneck link is equipped with an RTT's worth of buffer (50 packets, at 500 bytes per packet).

Note that "0 sent" is the sum of probabilities for all the b* states in the model, and similarly "1 sent" and "2 sent" represent the sums of the S_1 states and S_2 states, respectively. Simulation results agree well with our model at loss rates greater than or equal to $p = 0.1$. We note that the simulations slightly differ from our model for $p < 0.1$ for the following reason: under lower loss rates flows grow to window-sizes larger than 6; hence we need to compute the stationary distributions for larger values of W_{max} to get a more accurate distribution match. We also ran simulations under a variety of link bandwidths, variable propagation RTTs, and under RED and SFQ AQM schemes, and obtained similar agreement with the model. A more detailed analysis of the model, validation of the model and how it can be applied in practice can be found in our technical report [1].

Applications: Our model can be applied in a variety of ways at a middle-box to both predict the status and behavior of a flow as well as to potentially design non-intrusive

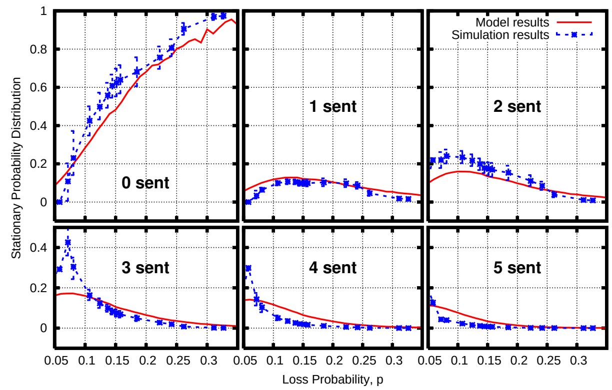

Figure 2: Stationary probabilities from the model and from simulations for 6 of the sending states that a TCP connection can be in. Error bars show 10th and 90th percentile flow values.

middle-box solutions to enhance performance in sub-packet regimes. Given the aggregate loss rate at the bottleneck, the model currently gives us the probability of finding a flow in one of several states. Similarly, the distribution can also be used to estimate the fraction of flows that are currently in timeout states on a pathologically-shared link. For a single flow, the model can be used as a mechanism for estimating the probability of hitting a timeout state by simply estimating the RTT of the flow and observing the number of packets within each epoch. Using this information, one can design middle-box queue management routines to reduce the possibility of flow timeouts. More details are outlined in our technical report [1].

4. REFERENCES

[1] Chen, J., Iyengar, J., Subramanian, L., and Ford, B. TCP Behavior in Sub-packet Regimes. NYU Technical Report, Nov 2010. `http://www.cs.nyu.edu/~lakshmi/subpacket.pdf`.

[2] Fortin-Parisi, S., and Sericola, B. A Markov model of TCP throughput, goodput and slow start. *Performance Evaluation 58*, 2-3 (2004), 89 – 108. Distributed Systems Performance.

[3] Huerta, E., and Sandoval-Almazán, R. Digital literacy: Problems faced by telecenter users in Mexico. *Information Technology for Development 13*, 3 (2007), 217–232.

[4] Juan, C. P., Agrelo, J. C. S., and Gerla, M. Using back-pressure to improve tcp performance with many flows. In *IEEE INFOCOM* (1999).

[5] Morris, R. Tcp behavior with many flows. In *ICNP '97: Proceedings of the 1997 International Conference on Network Protocols (ICNP '97)* (Washington, DC, USA, 1997), IEEE Computer Society, p. 205.

[6] Oyelaran-Oyeyinka, B., and Adeya, C. N. Internet access in africa: empirical evidence from kenya and nigeria. *Telemat. Inf. 21*, 1 (2004), 67–81.

[7] Padhye, J., Firoiu, V., Towsley, D., and Kurose, J. Modeling tcp throughput: a simple model and its empirical validation. In *SIGCOMM '98: Proceedings of the ACM SIGCOMM '98 conference on Applications, technologies, architectures, and protocols for computer communication* (1998).

[8] Qiu, L., Zhang, Y., and Keshav, S. Understanding the performance of many TCP flows* 1. *Computer Networks 37*, 3-4 (2001), 277–306.