Analysis of the phenomenon of several slow start phases in TCP

Chadi Barakat and Eitan Altman
INRIA, 2004 route des Lucioles, 06902 Sophia Antipolis, France
Email : {Chadi.Barakat, Eitan.Altman}@sophia.inria.fr

1. INTRODUCTION

TCP is known to send bursts of packets during its slow start phase due to the fast window increase and the ACK-clock based transmission [1, 2, 4]. If buffers in network routers are not well dimensioned to absorb these bursts, they will overflow prematurely before reaching the capacity of the network or even before reaching a slow start threshold set to less than this capacity. By capacity of the network we mean the maximum number of packets that can be fit between the source and the destination. The result is an underestimation of the network capacity and a deterioration in performance. The problem has been studied in [1, 4] in case of the Tahoe version where it has been shown that, for a buffer size smaller than one third the bandwidth-delay product, TCP requires two consecutive slow start phases to reach the congestion avoidance phase. Recall that with Tahoe, the window is set to one packet and a slow start phase is called upon every loss detection. According to the model introduced in [1, 4], only two consecutive slow start phases could occur. This seems logical since the first slow start probes the network for the correct slow start threshold and the second slow start uses this new threshold to get in congestion avoidance without losses. However, simulations of TCP Tahoe in presence of small buffers [1] have shown the possibility of three consecutive slow start phases before congestion avoidance which results in a further deterioration in performance. In our work (see [2] for details), we analyze this later problem. A refinement of the model used in [1, 4] is required. Our main contribution is an understanding of how three consecutive slow start phases can occur. A necessary condition to see such phenomenon is a buffer size less than one seventh the bandwidth-delay product. As a function of the buffer size, the phenomenon appears and disappears in a periodic manner. An appearance of the phenomenon results in up to 10% performance deterioration.

2. OVERVIEW OF THE ORIGINAL MODEL

Consider a simple model for the network where only the bottleneck bandwidth μ and the available buffer in the bottleneck router are considered. Let B be the size of this buffer in terms of packets and let T be the two-way propagation delay. A TCP Tahoe connection with an infinite volume of data to send is running across this network. The key point in the analysis of the slow start phase of TCP is the calculation of the window at which packets are dropped during slow start assuming that the slow start threshold is set to an infinite value. This is called the overflow window and it is denoted by W_B. See [3] for an exact calaculation of W_B. A comparison between the slow start threshold and W_B permits to know whether losses appear or not during a slow start phase. For a slow start threshold W_{th}, losses don't occur if W_{th} is set less than W_B. In case of $W_B < W_{th}$, the buffer B overflows during slow start at a window W_B and this overflow is detected later at a window $W_D = min(W_{th}, 2W_B)$. The window is set to one packet and a new slow start phase is called with a threshold $W_D/2$. In principle, $W_D/2$ is always less than W_B and a buffer overflow during the second slow start could not occur.

3. REFINEMENT OF THE MODEL

The inability of the previous model to explain losses during the second slow start phase comes from the fact that it makes the assumption that the window increase during this phase is similar to that during the first phase. But, this is not the case due to the packets correctly received and queued in the receiver buffer when the second slow start is called. Indeed, a buffer overflow during slow start causes many losses from the current window which takes TCP many round trip times to recover from them. When the buffer overflows at the end of the first slow start, the receiver buffer contains many packets correctly received separated by gaps which correspond to losses and which need to be filled in order for the receiver buffer content to be acknowledged. During the time TCP is trying to fill these gaps, it transmits some packets which have been already received. This results in less ACKs sent to the source, thus in a slower window increase. It may sometimes result in a quick sliding of the window and the generation of long bursts of packets. The window increase becomes similar to the one of the first slow start phase once all the gaps in the receiver buffer are filled.

Call a slow start phase where there is no gaps in the receiver buffer a *normal* slow start phase. During the second slow start, a return to the normal slow start mode happens if W_{th} is less than twice W_B and it is easy to show that in this case losses cannot occur during the second slow start. The problem exists when W_{th} is higher than twice W_B. We follow the window evolution during the second slow start phase and

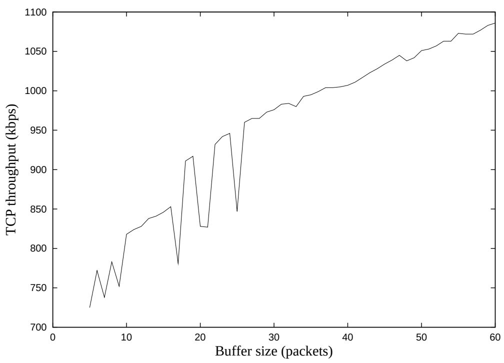

Figure 1: The cyclic appearance of triple slow start

we see the possibility of a buffer overflow in the round trip time where the last gap in the receiver buffer is filled. But, this overflow doesn't always occur. It requires that some condition be satisfied. We state the required conditions,

Theorem : Let n_B be the integer number satisfying,

$$2^{(n_B-2)} < B \leq 2^{(n_B-1)}.$$

Let n be the smallest integer number satisfying,

$$\left(2^{(n_B-1)} - B + 2\right)\left[\left(\frac{3}{2}\right)^{n+1} - 1\right] > B - 1.$$

We get three consecutive slow start phases if and only if

$$W_{th} > 2W_B,$$

and

$$\left(2^{(n_B-1)} - B + 2\right)\left[2\left(\frac{3}{2}\right)^{n} - 1\right] > 2B - 1.$$

◇

Thus, in contrast to the double slow start phenomenon seen in the case of $W_B < W_{th}$, a necessary condition for the triple slow start phenomenon is that W_B is less than $W_{th}/2$. Using the values for W_B and W_{th} found in [2, 4], this corresponds to a buffer less than one seventh the bandwdith-delay product. The condition for the double slow start phenomenon corresponds to a buffer less than one third the bandwidth-delay product.

Let us study the feasibility of the second condition. We divide the set of B satisfying $W_{th} > 2W_B$ into intervals in a way that all the B in the same interval give the same integers n_B and n. We vary then B inside an interval and we study the previous condition on the boundaries. We show [2] that the condition is always satisfied on the left boundary whereas it cannot be satisfied on the right boundary. This means that while increasing B inside an interval, the triple slow start phenomenon appears at the beginning and disappears at the end. Given that the intervals are contiguous, we conclude that the phenomenon appears and disappears in a periodic manner while changing B.

4. SIMULATION

We simulate with `ns` a scenario where $\mu = 1.5$ Mbps, $T = 560$ ms (e.g. a transfer through a satellite link) and packets

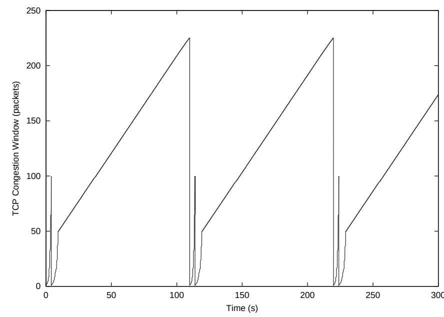

Figure 2: The case of B=19 Packets

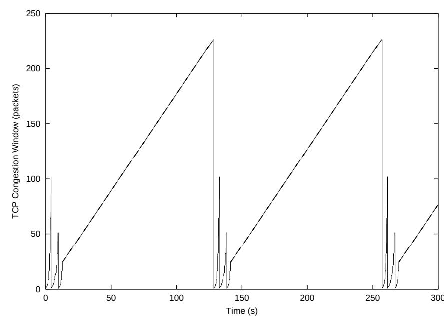

Figure 3: The case of B=20 Packets

of 512 bytes. According to these parameters, a buffer size smaller than 26 packets is required to see the triple slow start problem and a buffer size less than 64 packets is required to see the double slow start one. We run the simulation for a long time and we plot the average throughput of the connection as a function of the buffer size. Figure 1 shows well the periodic deterioration in performance at small buffers due to the appearance and disappearance of losses during the second slow start phase. An amazing thing is that, in some cases, an increase in the buffer size deteriorates the performance instead of improving it. We see this in Figures 2,3 where we show the window as a function of time. An increase in B by one packet makes the problem of three consecutive slow start phases appear and reduces then the performance.

5. REFERENCES

[1] E. Altman, et al., "Performance Modeling of TCP/IP in a Wide-Area Network", *IEEE Conference on Decision and Control*, Dec 1995.

[2] C. Barakat and E. Altman, "Analysis of TCP in Networks with Small Buffering Capacity and Large Bandwidth–Delay Product", *INRIA Research Report N=3574*, Dec 1998.

[3] C. Barakat and E. Altman, "Performance of Short TCP Transfers", *Networking 2000 (Performance of Communication Networks) conference*, May 2000.

[4] T.V. Lakshman and U. Madhow, "The performance of TCP/IP for networks with high bandwidth-delay products and random loss", *IEEE/ACM Transactions on Networking*, Jun 1997.