What's a reasonable buffer size for TCP?

Niels Möller

August 12, 2002

Introduction

We consider the TCP-flow over a bottle-neck router. What buffer size should that router have? If it has a too small buffer, it can't flatten out the sawtooth-like TCP-stream, and the bottleneck link will be under utilized. If it has a too big buffer, it will unnecessarily increase the end-to-end roundtrip time making the connection less responsive to the user. Furthermore, large roundtrip times, like all delays, can be expected to degrade performance of the control algorithms used in TCP.

The goal is to derive approximativly optimal values for the buffer size, or rules-of-thumb, expressed in parameters such as link speeds and ebd-to-end roundtrip time.

Notation

Notation in this paper: $r(t)$ is the rate of the input to the router. $q(t)$ is the size of the queue, measured in utits of time. In otherwords, $q(t)$ is the time a packet will spend inside the router, if it arrives at time t. s is the constant sending rate of the router, which is also the end-to-end bandwidth. The maximum transfer unit (MTU) is denoted by M.

We look at the following case, which is intended to model the worst case: At time t, data is received at a rate just above s, and the buffer is full. Then, the incoming rate is lahved to $s/2$ and increased again slowly according to the TCP rules, until it reaches s again. In order to keep the output link saturated, we want the queue to stay non-empty all the time, but we want it to be almost empty at the time the input rate reaches $r(t) = s$.

The "slow increase" of TCP is moddled as follows:

$$\begin{cases} r(0) = s/2 \\ r'(t) = \dfrac{M}{R(t)} t \end{cases} \tag{1}$$

where R(t) is the current end-to-end roundtrip time. We neglect the details of the TCP estimation of RTT, and assume that it knowns the true values.

As long as the queue is non-empty, the queue length satisfies

$$\begin{cases} q(0) = q_{\max} \\ q'(t) = \dfrac{r(t)}{s} - 1 \end{cases} \tag{2}$$

Constant RTT model

For a start, let's consider the constant RTT case, i.e. $R(t) = R$, for some constant R. This makes sense if the delay in the router is smaller than in the rest of the network. We get

$$r(t) = \frac{s}{2} + \frac{M}{R}t$$

We get $r(t)$ as a straight line with slope M/R, and $sq(t)$ equals the triangular area above that line and and below s. We get

$$q_{\max} = \frac{Rs}{8M}$$

So we need a queue of one eighth of the bandwidth-delay product divided by the maximum transfer unit.

Unfortunately, this formula is not very useful. For example, consider a connection over ethernet, width 10Mbit/s bandwidth, a roundtrip time of 20ms, and an MTU of 1500 bytes. The bandwidth-delay product is about 25Kbyte, and 8MTU is about 12 Kbytes, so the conclusion is that we should buffer about 2 seconds of data. And of cource, this invalidates the assumption of a constant RTT of 20ms.

So we really need to take the effect of the queue length on the RTT into account.

Improved model

So let's instead assume that the RTT of the rest of the network is constant (which is reasonable, as we shouldn't have any significant queues except just before the bottleneck link), and that the total RTT is given by

$$R(t) = R_0 + q(t) \tag{3}$$

where R_0 is constant. Remember that q is measured in units of time. We now get a system of differential equations,

$$\begin{cases} r(0) = \dfrac{s}{2} \\ q(0) = q_{\max} \\ r'(t) = \dfrac{M}{R_0 + q(t)}t \\ q'(t) = \dfrac{1}{s}r(t) - 1 \end{cases} \tag{4}$$

From the equation for q' we get $r = sq' + s$, and differentiating yields $r' = sq''$. We can substitute that into the equation for r' and get

$$q'' = \frac{M/s}{R_0 + q(t)}t \tag{5}$$

If we substitute $x(t) = t/(R + q(t))$ and $A = M/s$, we get the second order equation

$$tx'' + 2x' - Ax = 0$$