PkbTCPLimitModels

TCP Limitation Models

Mathis Equation

The Mathis equation states that, for a small loss rate (less than 1%), the maximum achievable throughput of a TCP connection is limited by:

Rate <= (MSS/RTT)*(1 / sqrt{p})

where MSS is Maximum Segement Size

RTT is Round Trip Time as measured by TCP

p is the probability of packet loss

Padhye et al Equation

(Extract from Les Cottrell's 'Throughput versus loss')

An improved form of the above formula that takes into account the TCP initial retransmit timer and the Maximum TCP window size, and is generally more accurate for larger (> 2%) packet losses, can be found in: . The formula is given below (derived from eqn 31 of Padhye et. al.):.

if w(p) < wmax

Rate = MSS * [((1-p)/p) + w(p) + Q{p,w{p}}/(1-p)] /
(RTT * [(w{p}1)](Q{p,w{p}}*G{p}*T0)/(1-p))

otherwise:

Rate = MSS * [((1-p)/p)+ wmax+Q{p,wmax}/(1-p)] /
(RTT * [0.25*wmax+((1-p)/(p*wmax)+2)] + (Q{p,wmax}*G{p}*T0)/(1-p)])

Where:

We have assumed the number of packets acknowledged by a received ACK is 2 (this is b in the Padhye et. al. formula 31)
wmax is the maximum congestion window size
w{p} = (2/3)(1 + sqrt{3*((1-p)/p) + 1} from eqn. 13 of Padhye et. al. substituting b=2

Q{p,w} = min{1,[(1-(1-p)3)**(1+(1-p)3)**(1-(1-p)(w-3))] /
[(1-(1-p)w)]}

G{p} = 1+p+2*p2+4*p3+8*p4+16*p5+32*p6 from eqn 28 of Padhye et. al.
T0 = Initial retransmit timeout (typically this is suggested by RFCs 793 and 1123 to be 3 seconds).
Wmax = Maximum TCP window size (typical default for Solaris 2.6 is 8192 bytes)

If you are tuning your hosts for best performance then also read Enabling High Performance Data Transfers on Hosts and TCP Tuning Guide for Distributed Application on Wide Area Networks. Also The TCP-Friendly Website summarizes some recent work on congestion control for non-TCP based applications in particular for congestion control schemes that maintain the arrival rate to at most some constant over the square root of the packet loss rate.

Proposed Deprecation

Recently, some of the authors of the initial TCP model have proposed abandoning these closed-form "macroscopic" models. Reasons include that some of the original assumptions, such as sufficient buffer space in routers, are no longer tenable, and that the models don't fit for promising new congestion control approaches such as BBR.

References

The macroscopic behavior of the TCP congestion avoidance algorithm, Mathis, Semke, Mahdavi & Ott in Computer Communication Review, 27(3), July 1997

Modelling TCP throughput: A simple model and its empirical validation, J. Padhye, V. Firoiu, D. Townsley and J. Kurose, in Proc. SIGCOMM Symp. Communications Architectures and Protocols Aug. 1998, pp. 304-314

Deprecating the TCP Macroscopic Model, M. Mathis, J. Mahdavi, in Computer Communications Review, 49(5), October 2019

– TobyRodwell - 12 Aug 2006
– Simon Leinen—14 Oct 2019