Congestion Control

Goal: Provide an end-to-end feedback mechanism which causes senders to adapt to a fair share of the bottleneck link.

What is the bottleneck link?

What is the maximum bandwidth achieved by a TCP connection?

RF2581 - M. Allman et. al. "Congestion Avoidance and Control" - Jacobson, Kars and Sigcom 1988

References:

References:

\[
\text{window} = \text{MIN}(\text{window, advertised})
\]

 congestion control is the problem of adapting the window window constrains the sending rate window constrains the sending rate

\[
\frac{RTT}{\text{window}} = BW
\]

...
Congestion Control - some basic definitions

cwnd

the effective congestion window

ssthresh

the current minimum bound on a reasonable congestion window

congestion event

a loss which occurs within the timeframe of the current

congestion window

convergence of packets don't inject new data until we are fairly certain

that a packet has left

acknowledged

packets in flight the number of packets that have been sent but not yet

sent through the current window

congestion threshold the current minimum bound on a reasonable congestion window
"Slow start" is actually exponential:

- Increase by 1 for every ACK
- Set cwnd to 1, and ssthresh to $\infty$
- Initial window size may be greater than number of available buffers

How should we start sending?
Congestion Avoidance

- On a timeout, set $cwnd = 1$.
- Linear: For each ACK increase $cwnd$ by $\frac{2}{\sqrt{I}}$
- Set $ssthresh = \frac{cmd}{I}$ on a timeout.
- Once $cwnd = ssthresh$ probe the network more slowly.
- Linear, increase multiplicative increase.

Congestion Avoidance
Ack Pacing

- limiting the window isn’t enough to stop bursts from occurring
- each (non-duplicate) ack advances the window by one segment
- this naturally smooths out transmissions to the bottleneck capacity
- slow start is necessary to start pacing
- so idle connections restart in slow start
fast retransmission

set cwnd equal to ssthresh + 3

set ssthresh to \( \frac{c}{2} \)

we can retransmit that segment immediately after 3 duplicate acks indicate that a segment was missing

single losses are catastrophic for performance
RTT Estimation

Determines how responsive the moving average is, is called the gain and where $\alpha$ is the gain and

\[
\text{RTT}_{\text{new}} = \alpha \cdot \text{Measurement} + (1 - \alpha) \cdot \text{RTT}_{\text{old}}
\]

is called the gain and determines how responsive the moving average is. If larger than the real RTT, we have excessive retransmission, if smaller than the real RTT, performance suffers.

What timeout interval should we use?