Reconsidering draft-dukkipati-tcpm-tcp-loss-probe-01

N. Dukkipati, N. Cardwell, Y. Cheng, M. Mathis

tcpm@IETF 87, July 2013.

Re-thinking loss recovery

3 DUPACKs => fast recovery

Otherwise after RTO => cwnd = LW then slow start

Many performance tweaks but it does not work (well) due to many reasons
1. Tail loss is common
2. False timeout badly hurts latency
3. Shortening RTO fuels false timeouts
4. Bufferbloat and mobile make RTT estimation challenging
5. Upping LW makes network less stable
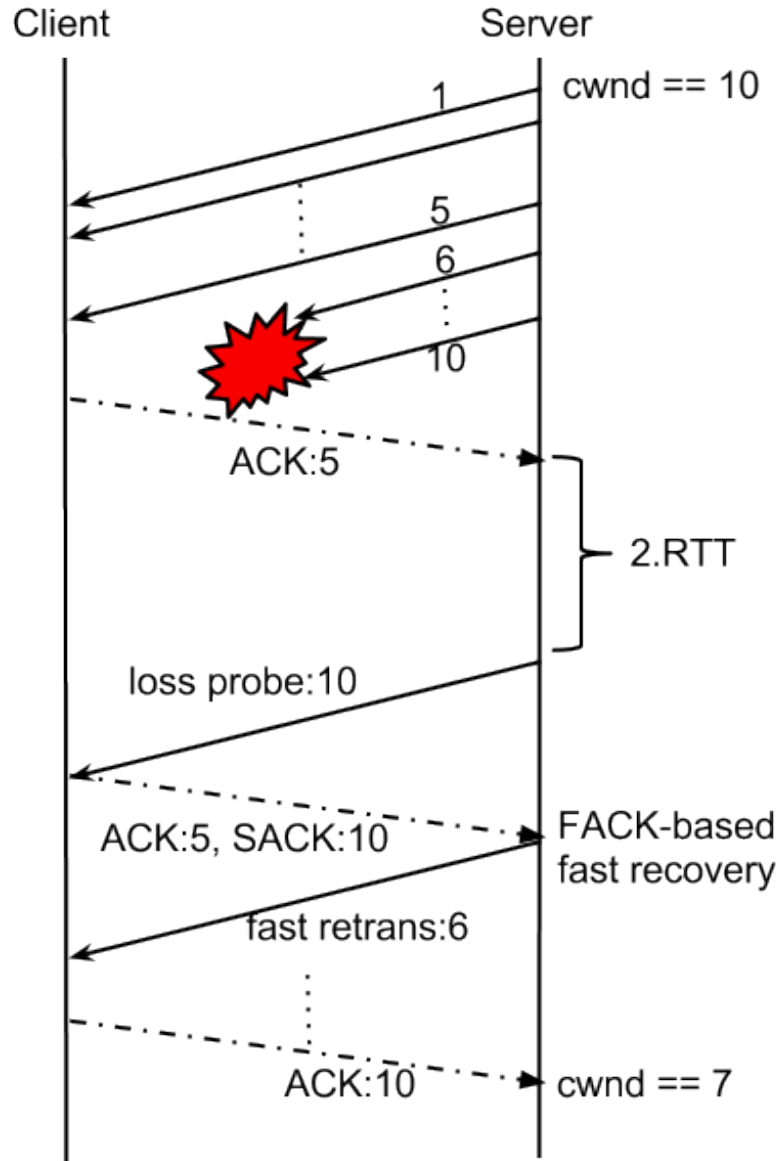
TCP loss probe: is network really busted?

On timeout send a new packet (or the last unacked one), instead of first unacked one

If timeout is short enough (1.5 RTT) and the packet is acked, no need to reduce cwnd to 1

On second timeout, behave like current old timeout

It reduces latency and is stable against heavy congestion

TLP example

New TCP loss recovery

1. 3 DUPACKs for fast recovery
2. After 1.5 RTT send a probe
3. After X RTT (>= 1sec) declare network is really congested and slow start again

That's it

Details are in
`draft-dukkipati-tcpm-tcp-loss-probe-01`

or maybe we should merge RFC 6298, 3517, 5827, TLP?