

Updating TCP to support Rate-Limited Traffic draft-ietf-tcpm-newcwv-00

TCPM WG

IETF-86

G. Fairhurst A. Sathaseelan & R. Secch

Background: New-CVV key concepts

- Definition of pipeack
 - Pipeack is the TCP sustained rate during data-limited periods
 - Pipeack is calculated from ACKs that acknowledge new data
 - not *FlightSize*
- While **$cwnd/2 < pipeack < cwnd$** , the cwnd is “validated”
 - Cwnd is increased using normal TCP rules
- While **$pipeack \leq cwnd/2$** , the cwnd is non-validated
 - Cwnd is frozen
 - More cautious cwnd reduction upon loss in NVP
 - Cwnd halved after 5min of low path utilization

Patch for Linux kernel 2.6

- Linux loadable kernel module (LKM)
 - Sender-only modification *congestion_ops* hooks in TCP processing flow
- Based on Reno *congestion_ops*
 - After idle event (CA_EVENT_TX_START)
 - Timeout (CA_EVENT_FRTO)
 - In sequence ACK (CA_EVENT_FAST_ACK)
- New variables for congestion control
 - pipeACK (uint32_t)
 - Timer timestamps (uint32_t): 5 min timer, pipeack sampling
 - May not require actual timers (just the timestamp)

Currently open problems...

To address with tests/experiments

- Pipeack estimation
 - Initialization of pipeack large in latest draft (00)
 - How to reset **pipeack** when resuming after a packet loss?
 - Do we need smoothing filters?
 - Can pipeack sampling be relaxed (e.g. every few RTTs)?
 - How to address corner cases: Small RTTs (<1ms), large RTTs (>10s), TX interrupted (from cwnd validated to idle)
- Is there negative impact of NewCWV bursts?
 - Is packet pacing even feasible? (pass)
- How to reset **cwnd** after loss during NVP?
 - Current reset based on “Flightsize”
 - Not addressing end-of-burst loss (cwnd after tail loss)

Related documents/issues

- Interoperation with TFO over Ethernet ?
- Pipeack vs IW initialisation?
 - Pipeack initialisation/resetting shall consider IW
 - Impact of Linux *rwind* autotuning mechanism
- Estimation of pipeack during FR
 - Can reuse “pipe” as in RFC3517?

Next steps

- Patch to Linux
- Experience with use
 - anyone interested in experimenting?
- End-of-burst loss effects on cwnd.