Low Extra Delay
Background Transport

draft-ietf-ledbat-congestion-00.txt

Stanislav Shalunov 〈shalunov@bittorrent.com〉

IETF 76, Hiroshima
LEDBAT WG
November 13, 2009
WG draft out

- previously known as draft-shalunov-ledbat-congestion
- draft-ietf-ledbat-congestion
- addressed non-congestive constraints
Non-congestive constraint

• sender may be constrained by something other than congestion — maybe the app

• might not experience congestion for a long time

• congestion control as specified would keep climbing, reaching absurdly high values

• if non-congestive constraint is later removed, sender will burst out traffic
Non-congestive constrained (cont.)

• the burst will cause congestion
• the sender will back off to sanity
• the hiccup is unnecessary
• sender never should make assumptions far beyond probed rate
Flight size

- addresses the non-congestive constraint problem
- amount of data currently sent, but not acked
- section 4.15 + pseudocode changes
- keep the congestion window on a leash
  - $\text{cwnd} \leq \text{ALLOWED\_INCREASE} + \text{TETHER} \times \text{flight\_size()}$
Flight size (cont.)

- ALLOWED_INCREASE and TETHER are constant
- ALLOWED_INCREASE MUST be at least 1 packet
- ALLOWED_INCREASE SHOULD NOT be more than 3 packets
- TETHER MUST be greater than 1
- TETHER SHOULD be between 1.25 and 2
Flight size
implementation status

• Deployed and running in all BitTorrent products that use uTP
Known doc issues

• framing is out of scope
• late comer’s advantage
• fairness between LEDBAT flows
• choice of parameter values
• go over pseudocode and spec and check for loose ends — please help, everyone
Framing is out of scope

• Actual specification of any wire representation is out of scope
• Must have a timestamp and a way to carry delay back
• Might be used with a UDP-based framing, TCP (with important details), or any transport mechanism
Late comer’s (non)-advantage

• Model:
  • a connection builds up target worth of queue
  • second connection comes, and not seeing that, starves new one
  • Dips in queue enough to measure base delay
  • Even if they were not, second connection would see the base once the first backs off
Fairness between LEDBAT flows

- Random redistribution
- Add random noise to delay measurements
- Some connections will lose, some grab
- Constrained random walk, ratio converges
Choice of parameter values

- TARGET MUST be 25 milliseconds (relax?)
- GAIN MUST be 1 MSS/RTT (relax?)
- $2 \leq BASE\_HISTORY \leq 10$
- 1 packet $\leq$ NOISE\_FILTER $\leq$ cwnd/2 (SHOULD)
- 1 $\leq$ ALLOWED\_INCREASE $\leq$ 3 (packets)
- $1.25 \leq$ TETHER $\leq$ 2
Careful proofreading

• Implementors read most carefully
  • Step forward

• Everyone, please read, particularly section 5
• Nits to mailing list
QUESTIONS?