draft-khademi-alternativebackoff-ecn-03

N. Khademi, M. Welzl, G. Armitage, G. Fairhurst

TSVWG @ IETF 95
6 April 2016

Context

- The assumed sender behavior in RFC3168 is "...end-systems should react to congestion at most once per window of data (i.e., at most once per round-trip time),.."
  - Scalable TCP / DCTCP / L4S are a different context

Problematic text in RFC 3168

The indication of congestion should be treated just as a congestion loss in non-ECN-Capable TCP. That is, the TCP source halves the congestion window "cwnd" and reduces the slow start threshold "ssthresh".

Upon the receipt by an ECN-Capable transport of a single CE packet, the congestion control algorithms followed at the end-systems MUST be essentially the same as the congestion control response to a *single* dropped packet. For example, for ECN-Capable TCP the source TCP is required to halve its congestion window for any window of data containing either a packet drop or an ECN indication.

Motivation

- This RFC3168 rule means poor utilization when we use small queues (halving requires a BDP for 100% utilization)
- Receiving ECE tells the sender that the queue was probably small, especially using modern AQM algorithms; this rule prevents us from using this knowledge

➔ We think this rule is harmful.

ABE: draft history

- draft-khademi-alternativebackoff-ecn-03
  - Presented in TCPM @ IETF Prague and Yokohama, then discussed in ICCRG
  - Changes sender reaction to ECN
  - Proposes new RFC3168 language

- Next steps
  - We plan to split the draft into a PS update to RFC3168, and a short draft to recommend an appropriate congestion response

Question

- Adopt an upcoming a PS update to RFC3168 in TSVWG?

Draft new text

If the sender receives an ECN-Echo (ECE) ACK packet (that is, an ACK packet with the ECN-Echo flag set in the TCP header), then the sender knows that congestion was encountered in the network on the path from the sender to the receiver. This indication of congestion could be treated in the same way as a congestion loss, however reception of the ECN-Echo flag MUST produce a reduction in FlightSize of at least 15% (roughly the reduction achieved by multiplying FlightSize with 0.85). This reduction can be less than the reduction had the flow experienced loss. The reduction needs to be sufficient to allow flows sharing a bottleneck to increase their share of the capacity.

An ECN-capable network device cannot eliminate the possibility of loss, because a drop may occur due to a traffic burst exceeding the instantaneous available capacity of a network buffer or as a result of the AQM algorithm (overload protection mechanisms, etc [RFC7567]). Whatever the cause of loss, detection of a missing packet needs to trigger the standard loss-based congestion control response. This explicitly does not update this behaviour.