Making TCP more Robust to Long Connectivity Disruptions (TCP-LCD)

Alexander Zimmermann and Arnd Hannemann

Department of Computer Science, Informatik 4
RWTH Aachen University, Ahornstr. 55, 52074 Aachen, Germany

77th IETF TCPM Meeting - Anaheim, CA
March 16, 2010
Motivation

Problem of Long Connectivity Disruptions

Observation

- Disruptions in e2e path connectivity which last longer than one RTO cause suboptimal TCP performance

Problem statement

- TCP interprets segment loss as a sign of congestion
  ⇒ Means to detect loss: DUPACKs and RTO
- RTO case: (repeated) backoff(s) of the retransmission timer
- Deferred detection of connection re-establishment since TCP has to wait until next RTO before retransmit again
Solution for Long Connectivity Disruptions: TCP-LCD

Disruption Indication

▶ Disambiguate true congestion loss from non-congestion loss caused by long connectivity disruptions
▶ Exploit standard ICMP destination unreachable messages during timeout-based loss recovery

Disruption Reaction

▶ Connectivity disruption loss: undoing one RTO backoff if an ICMP unreachable message reports on a lost retransmission ⇒ Enables prompt detection when connectivity is restored
▶ Congestion loss: Retaining std. timeout-based loss recovery

74th IETF meeting – San Francisco

- First presentation of the algorithm
- Exciting interest by the WG
- Valuable comments from Tim Shepard, and Joe Touch
- No negative comments
draft-zimmermann-tcp-lcd-01

Changes from previous draft version

- Updated algorithm’s motivation: Section 2
  - Congestion versus Non-Congestion Events/Loss
  - In-line with RFC 4653 (TCP-NCR)
- Added basic idea of the algorithm: Section 4.1
- Algorithm update: Section 4.2
  - Restructuring (suggestions Tim Shepard)
  - Removing of special case (first ICMP after RTO)
  - BACKOFF_CNT variable was introduced so it is no longer possible to perform more reverts than backoffs
- Expanded discussion: Section 4.3
  - Expanded discussion according to the algorithm changes
  - Try to clarify the “Wrapped sequence numbers” problem (comments Joe Touch)
draft-zimmermann-tcp-lcd-02

Changes from previous draft version

- Algorithm update: Section 4.2 (comments Ilpo Jarvinen)
  Based on observations made during the Linux implementation
  - Instead of reverting RTO by halving it when an ICMP arrives, we recalculate it with help of the Backoff\_cnt variable
  - Fix issue that occurred when the RTO backed off but is bounded by a maximum value

75\textsuperscript{th} IETF meeting – Stockholm

- Comments Joe Touch: How handle false positive/negatives?
- No negative comments
- Queued for mailing list discussion if pick upped as WG item
Changes from previous draft version

- Incorporated feedback/reviews submitted by Ilpo Jarvinen, Pasi Sarolahti, and Joe Touch
- Extended and reorganized discussion: Section 5
  - Heavily extended “Wrapped sequence numbers” discussion (based on Joe’s comments).
  - Extended “Retransmission Ambiguity” section
  - Influence of packet duplication (Ilpo’s comments)
- An interoperability issues section was added: Section 7
  - ICMPv6, IP Tunnels, ECN
  - ...
Status quo

- TCP-LCD is part of Linux kernel since 2.6.32
- All feedback has been positive
- We consider draft ready (modulo minor language updates)
- Oct'10: Submit document to the IESG for Experimental

Next Steps

- WGLC in the next few weeks?
- ...