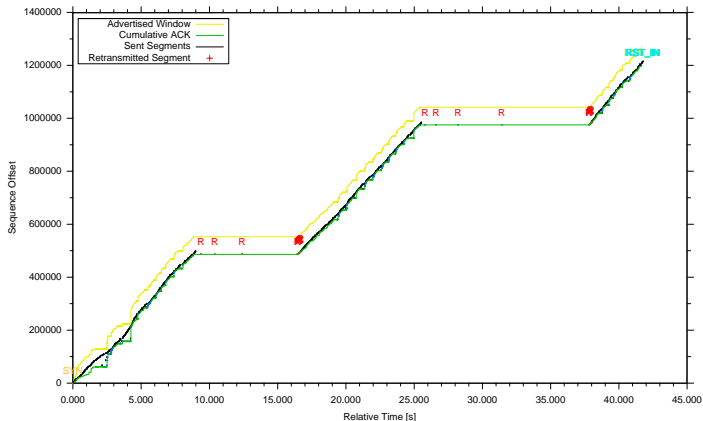# Make TCP more Robust to Long Connectivity Disruptions

Alexander Zimmermann and Arnd Hannemann

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

74th IETF Meeting - San Francisco, USA
March 20, 2009

The Problem of Long Connectivity Disruptions

TCP does not have any means to know when connectivity is re-established

Connectivity Disruption Indication

Queue of the router experiencing the link outage is

- ▶ *Deep enough*: cause variation in delay
  ⇒ Eifel [RFC3522], F-RTO [RFC4138]
- ▶ *Not deep enough*: drops the packets, discard according route
  ⇒ Notifies TCP source about the dropped packets via ICMP
  destination unreachable messages of code 0 (net unreachable)
  or code 1 (host unreachable) [RFC1812]

Conclusion

- ▶ ICMP unreachable messages of these codes are evidence that
  packets were dropped due to a link outage and not due to
  congestion

ICMP messages as Connectivity Disruptions Indication

Basic idea

▶ Revert unnecessary back-offs of RTO in case arriving ICMP messages at TCP source witness no congestion (at least on the very part of the path which was traveled by both, the TCP segment and the corresponding ICMP message)

Issues

▶ Do not ignore congestion indication from the network
▶ ICMP messages do not necessarily operate on the same timescale as the packets eliciting them [RFC1812]
▶ ICMP messages are subject to rate limiting [RFC1812]

Connectivity Disruption Reaction

- ▶ Upon expiration of RTO, set CWND to one segment, retransmit `SND.UNA` and back-off RTO [RFC2581, RFC2988]
  - ▶ Wait for ICMPs or ACKs to arrive. If nothing arrives (RTO expires again), repeat retransmission and back-off again
- ▶ If ACK arrives, proceed with usual recovery (Slow-Start)
  - ▶ Ignore ICMPs arriving during recovery, as reverting RTO back-off is of no use anymore
- ▶ If ICMP arrives, check sequence number
  - ▶ If not on retransmission, ignore it
  - ▶ Otherwise, for first retransmission, ignore it, else, halve current RTO, i.e., revert one back-off
    - ▶ If new RTO is already expired, send immediately,
    - ▶ Otherwise, wait according to new RTO

Evaluation
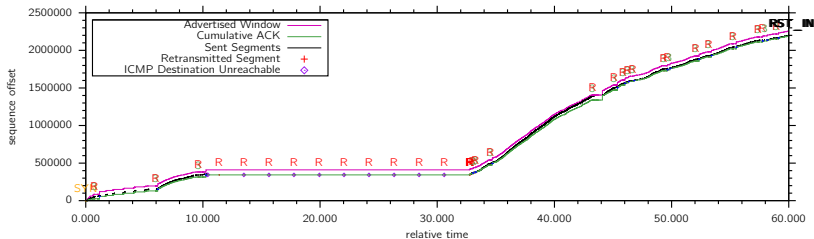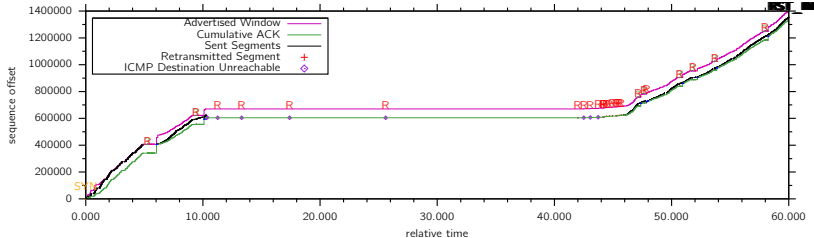
Discussion

Features

- ▶ Obey Congestion Control Principles
  - ▶ TCP cannot decide by receiving ICMP messages reporting on the original window if the whole window was dropped because of link outage or at least one segment due to congestion
  - ▶ After route repair, scheme falls back to usual back-off in case there is congestion along the path
  - ▶ Won't probe for route repair faster than slowest TCP can send
- ▶ Sender side only modification

Next steps

- ▶ Any interest from the WG?
- ▶ Use TCP Timestamps to dissolve the ambiguity problem