TCP ACK Rate Request (TARR) option

draft-gomez-tcpm-ack-rate-request-01

Carles Gomez
Universitat Politècnica de Catalunya

Jon Crowcroft
University of Cambridge

IETF 109 Virtual, TCPM WG, November 2020
Motivation

• Delayed ACKs
  • Intended to reduce protocol overhead
  • But may also contribute to suboptimal performance

• “Large” cwnd scenarios (i.e. cwnd >> MSS):
  – Saving up to 1 of every 2 ACKs may be insufficient
    • Performance limitations due to asymmetric path capacity
    • Computational cost and network load

• “Small” cwnd scenarios (i.e. cwnd up to ~1 MSS):
  – Data centers: BDP up to ~1 MSS
    • Delayed ACKs will incur a delay much greater than the RTT
  – Transactional data exchanges, or when cwnd decreases
    • Immediate ACKs may avoid idle times, allow faster cwnd growth
Status

• Related prior discussion
  • Sender control of TCP ACKs
  • Converged to defining a new TCP option serving two purposes:
    – Requesting a given ACK rate
    – Requesting an immediate ACK

• draft-gomez-tcpm-ack-rate-request-00
  • Presented at IETF 108

• Version -01
  • Aims to address the comments received during IETF 108
  • TARR option support announcement
Updates in -01 (I/III)

• Section 1
  • Expanded motivation for the need to use immediate ACKs in some IoT environments
    – Some devices with RAM only for a send buffer of 1 MSS
    – Due to Delayed ACKs, sender may need to wait for Delayed ACK timeout (and process the ACK) to transmit a new data segment

• Section 3
  • A TCP sender MAY indicate that it has a reordering tolerance of R packets by setting the Ignore Order field
    – In that case, a TCP receiver MUST continue to send one ACK every R data segments, even when reordering occurs
  • A TCP sender can request an immediate ACK for a data segment, and for the subsequent N data segments
Updates in -01 (II/III)

• Section 4. Option format:

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>01234567</td>
<td>89012345</td>
<td>67890123</td>
<td>45678901</td>
</tr>
</tbody>
</table>

  +------------------------------------------+-------------------+-------------------+-------------------+
  | Kind | Length | ExID |               |                 |
  +------------------------------------------+-------------------+-------------------+-------------------+
  | R    | IgnOrd | N    |               |                 |

  New

• Kind: experimental codepoints 253 or 254 (TBD)
• Length: 7 bytes
• Experiment Identifier (ExID): 0x00AC (to be requested)
• R: binary encoding of the R value
• Ignore Order: True/False
• N: subsequent data segments with immediate ACKs

**Question 1**: what about 7 bits for R, and 1 bit for Ignore Order?
Updates in -01 (III/III)

• Section 3. Announcing the TARR option
  • A TCP endpoint announces that it supports the TARR option by including the TARR option format in packets with the SYN bit set
    – R, Ignore Order, and N fields ignored

• Question 2: perhaps consider using a second codepoint, dedicated to just announce support of the option by means of a shorter 4-byte format (without R, Ignore Order, N)?
Thanks!

Questions? Comments?

Carles Gomez
Universitat Politècnica de Catalunya

Jon Crowcroft
University of Cambridge

IETF 109 Virtual, TCPM WG, November 2020