TCP ACK Rate Request (TARR) option

draft-gomez-tcpm-ack-rate-request-05

Carles Gomez
Universitat Politècnica de Catalunya

Jon Crowcroft
University of Cambridge

IETF 114 Philadelphia, TCPM WG, July 2022
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 immediate ACKs

• Since IETF 113: versions -04 and -05
  • Aim to address comments from (many thanks!):
    – The last IETF
    – The mailing list
Updates (I/III)

• Main format
  • OLD (-03):
    ```
    | Kind | Length | ExID |
    +-------+--------+-------+
    | R     | V|I    |
    +-------+--------+-------+
    Ignore Order: removed
    ```
  • NEW (-05):
    ```
    | Kind | Length | ExID |
    +-------+--------+-------+
    | R     | Reserved|
    +-------+--------+-------+
    Even length
    ```
    254
Updates (II/III)

• Main format
  • OLD (-03):
    - R carries binary encoding of ACK rate
    - R = 0: request of an immediate ACK
      - while keeping the ACK rate
    - Maximum value of R: 2047
  
  • NEW (-05):
    - Option 1: binary encoding, max R = 63
    - Option 2: mantissa, expon., max R = 1024
Updates (III/III)

• **Maximum value of R**
  - In -05, it is 2047
  - Questions on why $R > 63$ needed
  - Answers, with reasons given:
    - R values in the range of $\sim 100$ and even $\sim 1000$ justifiable in current scenarios
      - Assuming 1 Gbps, RTT up to $\sim 100$ ms
      - Rule of thumb of at least 4 ACKs per RTT
    - Making TARR useful for future scenarios (greater link rates)

• **Section 3:**
  - A TCP endpoint **SHOULD** announce TARR support in packets with SYN bit set
    - In some cases (e.g. SYN cookies used [RFC 4987]) TARR MAY be announced in packets subsequent to the SYN packet
    - Note: announcing TARR option support in the ACK (3WHs) not reliable
Running code

• Prototype implementation
  • Michael Tuexen
  • FreeBSD
Thanks!
Questions? Comments?

Carles Gomez
Universitat Politècnica de Catalunya

Jon Crowcroft
University of Cambridge

IETF 114 Philadelphia, TCPM WG, July 2022
Annex. In -03

• Two possible encodings for the R field:

  • OPTION 1:
    – Binary encoding of the requested ACK rate
    – The maximum value of R is 63

  • OPTION 2:
    – 4 leftmost bits represent a mantissa (m)
    – 2 rightmost bits represent an exponent (e)
    – The requested ACK rate is $R = (m+1) \times 2^{(2 \times e)}$
    – The maximum value of R is 1024