TCP ACK Rate Request (TARR) option

draft-gomez-tcpm-ack-rate-request-02

Carles Gomez
Universitat Politècnica de Catalunya

Jon Crowcroft
University of Cambridge

IETF 110 Virtual, TCPM WG, March 2021
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

• Versions -00 and -01
  • Presented in IETF 108 and IETF 109, respectively

• Version -02
  • Aims to address the comments received in IETF 109
  • Main option format size reduced by 1 byte
  • Content added in the Security Considerations section
Updates in -02 (I/III)

- Two formats (same Kind value, e.g. 253)
  - To announce support of the option
Updates in -02 (II/III)

- Reduced the (main) format size by 1 byte:
  - Old format (7 bytes):

```
  0  1  2  3
0 1234567 89012345 67890123 45678901
```

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

```
  0  1  2  3  4  5  6  7  8  9  0  1  2  3  4  5  6  7  8  9  0  1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
```

```
+---------------------------+---------------------------+---------------------------+
| Kind  | Length | ExID     |                          |
+---------------------------+---------------------------+---------------------------+
| R    | I      | N        |                           |
+---------------------------+---------------------------+---------------------------+
```
Updates in -02 (III/III)

• Section 6. Security considerations
  
  • An attacker impersonating a legitimate sender may communicate a bad R value to a receiver
    – A too high R value in a “small” cwnd scenario
    – A too low R value in a “large” cwnd scenario
  
  • TLS does not protect TCP headers...
  • One approach could be using IPSec
IANA considerations

- Post-02, IANA has allocated ExID value 0x00AC for the TARR option

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x00AC</td>
<td>TCP ACK Rate Request</td>
<td>[draft-gomez-tcpm-ack-rate-request-02]</td>
</tr>
<tr>
<td>0x0348</td>
<td>HOST_ID</td>
<td>[RFC7974]</td>
</tr>
<tr>
<td>0x0A0D</td>
<td>Autonomous System Compensation</td>
<td>[draft-donnerhacke-linktax]</td>
</tr>
<tr>
<td>0x0CA0</td>
<td>TCP Capability Option</td>
<td>[draft-boucadair-tcpm-capability-option]</td>
</tr>
<tr>
<td>0x0ED0</td>
<td>Extended Data Offset</td>
<td>[draft-ietf-tcpm-tcp-edo]</td>
</tr>
<tr>
<td>0x454E</td>
<td>TCP-ENO [2]</td>
<td>[RFC8547]</td>
</tr>
<tr>
<td>0x5323</td>
<td>Service Number</td>
<td>[draft-touch-tcpm-sno]</td>
</tr>
<tr>
<td>0x75EFFFFEE</td>
<td>Timestamp Interval</td>
<td>[draft-trammell-tcpm-timestamp-interval]</td>
</tr>
<tr>
<td>0xACCE</td>
<td>AccECN Experimental Option</td>
<td>[draft-kuehlwein-tcpm-accurate-ecn]</td>
</tr>
<tr>
<td>0xE2D4C3D9</td>
<td>Shared Memory communications over RMDA protocol</td>
<td>[RFC7609]</td>
</tr>
<tr>
<td>0xF989</td>
<td>Fast Open (current and new implementations SHOULD use option 34)</td>
<td>[RFC7413]</td>
</tr>
<tr>
<td>0xF990</td>
<td>Low Latency</td>
<td>[draft-wang-tcpm-low-latency-opt]</td>
</tr>
</tbody>
</table>
Yoshi’s review of -02

• Provide guidance on when/how to use the TARR features
  • Setting the ACK frequency (via R): once in a connection lifetime?
  • If reordering is not supported, but Ignore Order (I) set to True, it will take time to detect losses
  • How to set N
  • Ability to set R = 0 for the entire connection?
    – N value to represent “INFINITY”?

• Editorial clarifications
WG adoption ?
Thanks!

Questions? Comments?

Carles Gomez
Universitat Politècnica de Catalunya

Jon Crowcroft
University of Cambridge

IETF 110 Virtual, TCPM WG, March 2021