Quick-Start for TCP and IP

draft-ietf-tsvwg-quickstart-01.txt
A. Jain, S. Floyd, M. Allman, and P. Sarolahti
TSVWG, November 2005

This and earlier presentations::
www.icir.org/floyd/talks
QuickStart with TCP, for setting the initial window:

• In an IP option in the SYN packet, the sender's desired sending rate:
  – Routers on the path decrement a TTL counter,
  – and decrease the allowed sending rate, if necessary.

• The receiver sends feedback to the sender in the SYN/ACK packet:
  – The sender knows if all routers on the path participated.
  – The sender has an RTT measurement.
  – The sender can set the initial congestion window.
  – The TCP sender continues using normal congestion control.

• From an initial proposal by Amit Jain
Changes since last IETF:

- Added a 30-bit QS Nonce (feedback from Guohan Lu and Gorry Fairhurst).
- Significantly revised the section on IP tunnels and on IPsec AH (feedback from David Black and Joe Touch).
- Added a section about "Possible Uses for the Reserved Fields".
- General editing (feedback from Gorry Fairhurst and Martin Duke).
To do:

- **Delete the sentence in Section 4.6.2 about a retransmitted SYN packet using a different Initial Sequence Number.**
- **Respond to feedback from Bob Briscoe.**
The 30-bit QS Nonce:

- Initialized by sender to a random value.
- If router reduces Rate Request from K to K-1, router resets related bits in QS Nonce to a new random value.
- Receiver reports QS Nonce back to sender.
- If Rate Request was not reduced in the network below K, then the lower 2K bits should have their original random value.
The 30-bit QS Nonce:

<table>
<thead>
<tr>
<th>Bits</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bits 0-1</td>
<td>Rate 15 -&gt; Rate 14</td>
</tr>
<tr>
<td>Bits 2-3</td>
<td>Rate 14 -&gt; Rate 13</td>
</tr>
<tr>
<td>Bits 4-5</td>
<td>Rate 13 -&gt; Rate 12</td>
</tr>
<tr>
<td>Bits 6-7</td>
<td>Rate 12 -&gt; Rate 11</td>
</tr>
<tr>
<td>Bits 8-9</td>
<td>Rate 11 -&gt; Rate 10</td>
</tr>
<tr>
<td>Bits 10-11</td>
<td>Rate 10 -&gt; Rate 9</td>
</tr>
<tr>
<td>Bits 12-13</td>
<td>Rate 9  -&gt; Rate 8</td>
</tr>
<tr>
<td>Bits 14-15</td>
<td>Rate 8  -&gt; Rate 7</td>
</tr>
<tr>
<td>Bits 16-17</td>
<td>Rate 7  -&gt; Rate 6</td>
</tr>
<tr>
<td>Bits 18-19</td>
<td>Rate 6  -&gt; Rate 5</td>
</tr>
<tr>
<td>Bits 20-21</td>
<td>Rate 5  -&gt; Rate 4</td>
</tr>
<tr>
<td>Bits 22-23</td>
<td>Rate 4  -&gt; Rate 3</td>
</tr>
<tr>
<td>Bits 24-25</td>
<td>Rate 3  -&gt; Rate 2</td>
</tr>
<tr>
<td>Bits 26-27</td>
<td>Rate 2  -&gt; Rate 1</td>
</tr>
<tr>
<td>Bits 28-29</td>
<td>Rate 1  -&gt; Rate 0</td>
</tr>
</tbody>
</table>
IP Tunnels and Ipsec AH:

• **Quick-Start is compatible with IPsec AH.** (The Integrity Check Value covers the right things.)
• **There are some tunnels that are not compatible with Quick-Start** (Section 6.2):
  – This refers to tunnels where the IP TTL is not decremented before encapsulation;
  – Therefore, the TTL Diff is not changed;
  – The sender can falsely believe that the routers in the tunnel approved the Quick-Start request.
  – This will limit the possible deployment scenarios for Quick-Start.
Possible Uses for the Reserved Fields:

- Reporting Approved Rate.
- Report of Current Sending Rate.
- Request to Increase Sending Rate.
- RTT Estimate.
- Informational Request.
- Use Format X for the Rate Request Field.
- Do Not Decrement.
From Feedback from Bob Briscoe:

• Clarify Experimental status.
• Clarify router requirements for judging a link to have been underutilized.
• Add description of possible alternatives:
  – for QS nonce;
  – for an expanded range for the rate request;
  – for an alternate encoding for the rate request;
But don’t change the current proposal.
From Feedback from Bob Briscoe:

- Problems with untrusted senders:
  - Add “Reporting Approved Rate”?
  - The Quick-Start Option in QS data packets would report the approved rate request, along with the QS Nonce returned with that rate request.

- Add a standardized timeout for rate requests?
  - Rate requests are only valid at the sender if the response is received within N seconds?

- Add error codes from routers to end nodes?
  - Using one of the reserved bits, and the Rate Request or QS Nonce field?
Slides from last time:
Section 3.6: A Quick-Start Nonce?

• There are four unused bits in the IP option -
  – Use them for a Quick-Start Nonce?
• Some times the receiver knows the original rate request R.
• Goal of QS Nonce: discourage receivers from lying about the value of the received rate request.
• Mechanics:
  – Sender sets QS Nonce to a random value.
  – When a router reduces the approved rate request, it sets the QS Nonce to a new random value.
  – Receiver reports back value to sender.
  – If no routers reduced the rate request, then the QS Nonce should have its original value.
• Should we add this to the spec?
Feedback from Joe Touch about IP tunnels:

• **Tunnels that aren’t part of the forwarding path** don’t decrement the inner header’s IP TTL.
• The decrement is supposed to occur **before encapsulation**.
• **IPsec tunnels** need to be addressed; they typically drop IP options.
• Some tunnels **decrement the IP TTL by more than once**, to emulate the hopcount of the underlying path.