Stateless Reverse Traceroute

Rolf Winter

### The problem

<table>
<thead>
<tr>
<th></th>
<th>Router</th>
<th>IP Address</th>
<th>Time 1 (ms)</th>
<th>Time 2 (ms)</th>
<th>Time 3 (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>routerA.aug.net-a.com</td>
<td>(10.10.0.1)</td>
<td>1ms</td>
<td>2ms</td>
<td>1ms</td>
</tr>
<tr>
<td>2</td>
<td>routerB.muc.net-a.com</td>
<td>(10.10.0.2)</td>
<td>5ms</td>
<td>6ms</td>
<td>12ms</td>
</tr>
<tr>
<td>3</td>
<td>routerC.fra.net-a.com</td>
<td>(10.10.0.3)</td>
<td>11ms</td>
<td>21ms</td>
<td>14ms</td>
</tr>
<tr>
<td>4</td>
<td>routerD.fra.b-net.com</td>
<td>(20.20.0.1)</td>
<td>340ms</td>
<td>320ms</td>
<td>350ms</td>
</tr>
<tr>
<td>5</td>
<td><a href="http://www.example.com">www.example.com</a></td>
<td>(20.20.0.2)</td>
<td>345ms</td>
<td>310ms</td>
<td>360ms</td>
</tr>
</tbody>
</table>
The problem
Meet reverse traceroute

- Uses a new ICMP request to trigger a reverse traceroute
- One request per TR packet
Meet reverse traceroute

- A regular TR packet is sent (UDP, ICMP or TCP)
- Fields for load-balancing can be controlled

![Diagram showing reverse traceroute and routers](image)
Meet reverse traceroute

- For that single probe, an ICMP response is sent back
Stateless - how?

- Stateless: all the information needed has to be put into the probe packets
  - For IPv4, the ICMP Time Exceeded packet only guarantees to contain the original IP header plus the 64 bits following it
- ICMP probes (as an example, applies similarly to UDP and TCP*):
  
  +---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  | Type=8 | Code=0 | Checksum (Flow) |
  +---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  | Identifier (Query) | Sequence (Probe) |
  +---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  | *Timestamp |
  +---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  | Payload |
  +---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

- RTT estimations are not guaranteed to work

  *well, TCP needs 28 bits - instead of 16
And on today’s internet?

- RIPE Atlas, publishes archives of their measurement data, or rather metadata about the actual measurements
  - We filtered that data to only contain IPv4 ICMP Time Exceeded packets without extensions to reliably infer the true return size
  - Removed duplicates with the same (payload size, IP address) tuple
  - Data isn’t perfect: e.g. sometimes it seems that answers are bigger than the actual probe → filter data that seems inconsistent away, too

- In sum, we looked at 200 GB (compressed) archives

<table>
<thead>
<tr>
<th>Response size [bytes]</th>
<th>8</th>
<th>12</th>
<th>20 - 28</th>
<th>32 - 39</th>
<th>44</th>
<th>&gt;=48</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of responses</td>
<td>48.6%</td>
<td>0.1%</td>
<td>0.4%</td>
<td>0.5%</td>
<td>0.4%</td>
<td>50%</td>
</tr>
<tr>
<td>Stateless</td>
<td>Stateful</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identifies all routers (that send ICMP time exceeded packets today)</td>
<td>Identifies all routers (that send ICMP time exceeded packets today)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can estimate the RTT in about 50% of all cases</td>
<td>Can always estimate the RTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is restricted to trace towards the requestor</td>
<td>Can potentially trace to some other host on the internet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Extensions

- There is some amplification: a client request triggers a single traceroute probe being sent AND also a reverse traceroute response
- Solution: make the client send more bytes using a padding extension
Next steps

- Ask for WG adoption
  - We did the measurements
  - We did the implementation
  - The document is in good shape (for an individual document)

- If the document gets adopted
  - Fold the stateful document back into this one and just name it “Reverse Traceroute”
  - The party offering the reverse traceroute function should be able to pick stateful over stateless (this is where the state is kept)