Quick-Start for TCP and IP

draft-ietf-tsvwg-quickstart-02.txt

A. Jain, S. Floyd, M. Allman, and

P. Sarolahti

TSVWG, March 2006

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 sentence about multi-access links.
- Changed the "Reporting Approved Rate" option from a "Possible Use" in Appendix D to a required use in Section 3.1, to allow routers and receivers some protection against misbehaving senders.
- Clarifications and corrections in response to Bob Briscoe's feedback.
- Added rate encoding and one-way hash function from Bob Briscoe's document as alternate possibilities.
- Added Appendix E about Sections 1-3 of Bob Briscoe's document.

Report of Approved Rate:

- The 4-bit "Flags" field is now a Function field.
  - If the Function field is "1000", then the QS Option is a "Report of Approved Rate".
- The sender sends a "Report of Approved Rate" after receiving a Quick-Start Response. The Report might report an Approved Rate of zero.
- Routers may:
  - Ignore the Report of Approved Rate;
  - Use Report to check for misbehaving senders;
  - Use Report to keep track of committed Quick-Start bandwidth.

Clarifications and Corrections:

- Clarified that the approval of a Quick-Start request at a router does not affect the treatment of the subsequent arriving Quick-Start data packets.
- Clarified the phrase "incrementally deployable".
- Clarified semantics about additional rate.
- Made changes suggested in Section 5.1.3 of Bob's paper, including saying that the router should decrement the QS TTL by the same amount that it decrements the IP TTL (on the off chance that it decrements the IP TTL by more than one).
- Fixed nits.

Alternate Rate-encoding Function, Appendix A.2:

"Section 4.4 of [B05] suggests a mantissa and exponent representation for the Quick-Start encoding function. With e and f as the binary numbers in the exponent and mantissa fields, and with $0 <= f < 1$, this would represent the rate $(1+f)*2^e$. [B05] suggests a mantissa field for f of 8, 16, or 24 bits, with an exponent field for e of 8 bits. This representation would allow larger rate requests, with an encoding that is less coarse than the powers-of-two encoding used in this document."

One-way hash function as alternate QS Nonce, Appendix A.7.

- "An alternate proposal for the Quick-Start Nonce from [B05] would be for an n-bit field for the QS Nonce, with the sender generating a random nonce when it generates a Quick-Start Request. Each route that reduces the Rate Request by r would hash the QS nonce r times, using a one-way hash function such as MD5 [RFC1321] or the secure hash 1 [SHA1]. The receiver returns the QS nonce to the sender."

- "Because the sender knows the original value for the nonce, and the original rate request, the sender knows the total number of steps s that the rate has been reduced."

Appendix E on Sections 1-3 of Bob Briscoe's document.

- "[B05] proposes an alternate to Quick-Start where endpoints allocate rates to themselves. [B05] argues that adding rate allocation to interior routers is not the fundamentally correct direction."

Extra Viewgraphs on "Report of Approved Rate"

-

Routers using the Report of Approved Rate:

- If Report of Approved Rate reports a higher rate than router recently approved:
  - Router could deny future requests from this sender.
- If router sees Report of Approved Rate, and didn't see an earlier Quick-Start Request:
  - Either path changed, or sender is cheating.
  - In either case, router could deny future requests from this sender.

Routers using the
Report of Approved Rate, continued:

- If router sees a Quick-Start request, but doesn't see a Report of Approved Rate:
  - The QS Request was denied and dropped downstream; OR
  - The sender didn't send a Report of Approved Rate; OR
  - The Report was dropped; OR
  - The Report took a different path in the network.
- In any of these cases, the router could deny future QS Requests from this sender.

Old Viewgraphs:

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.

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?