RFC1323bis – TCP Extensions for High Performance

Richard Scheffenegger (Editor)
David Borman
Bob Braden
Van Jacobson

RFC1323bis – since IETF85

- Feedback from three reviewers

- No change in technical content
  - last change in -02

- Editorial Changes (reviewer feedback)
  - shortened introduction
  - “changes” section is last appendix
  - Introduction updated to align with later sections
  - Word smiting
RFC1323bis – delta to RFC1323 (content)

- Window Scale option:
  - Clear guidance to implementers for corner cases (Window Reduction – Section 3.4)

- Timestamp option:
  - Updated wording to allow “late” Timestamp enabling (non SYN)
  - Clear guidance which TS values can update RTT
  - Edge cases in receiver TS processing
  - Explicitly address <ACK> with SACK option
  - Recommend TS in <RST>, but exclude from PAWS test
RFC1323bis – delta to RFC1323 (editorial)

- Formatting updated (using xml2rfc instead of noff)
  - Errata of 1323 addressed
  - Indentation of RFC1323 fixed
- Use of RFC2119 wording in normative sections
- New pseudo code appendix D
- Addressing lots of Nits mentioned over the years
  - oversights, word smiting
  - Proper variable names in appendices
  - Fixed internal references and added/updated external references
- TS Offset explicitly mentioned
  - To address potential information leak
Introduction (upcoming)

- Original conception of the text concerned WS, TS and SACK simultaneously. Updated to properly apply to the two options remaining in this document.

Old:

... Therefore, for each of the extensions defined below, TCP options will be sent on non-SYN segments only when an exchange of options on the SYN segments has indicated that both sides understand the extension.

New:

The window scale option negotiates fundamental parameters of the TCP session. Therefore, it is only sent during the initial handshake. Furthermore, the window scale option will be sent in a <SYN,ACK> segment only if the corresponding option was received in the initial <SYN> segment.

The timestamps option may appear in any data or ACK segment, adding 12 bytes to the 20-byte TCP header. It is recommended that this TCP option will be sent on non-SYN segments only after an exchange of options on the SYN segments has indicated that both sides understand the extension. ...
Next steps

- Technical changes are uncontested since -02
- Editorial changes only
  - important to clarify to implementers
- Current open editorial point
  - Change “Changes” section from chronological to partitioned technical/editorial, sorted by section

- Ready for WLC
  - Final discussions on specific wording during WLC
Thank you
Backup
Window Scale Retraction

- Expanded text to dedicated section 3.4
- Explicitly quoted section 4.2.2.16 of RFC1122 to describe the expected behavior.
Timestamp negotiation

- Allow late negotiation:

Old:

A TCP may send the Timestamps option (TSopt) in an initial <SYN> segment (i.e., a segment containing a SYN bit and no ACK bit), and may send a TSopt in other segments only if it received a TSopt in the initial <SYN> segment for the connection.

New:

A TCP MAY send the Timestamps option (TSopt) in an initial <SYN> segment (i.e., a segment containing a SYN bit and no ACK bit). Once a TSopt has been sent or received in a non <SYN> segment, it MUST be sent in all segments. Once a TSopt has been received in a non <SYN> segment, then any successive segment that is received without the RST bit and without a TSopt MAY be dropped without further processing, and an ACK of the current SND.UNA generated.
Timestamp RTTM processing

- Only reflect timestamp from last in-sequence data packet.
- Only process timestamp when new data is acknowledged.
- However, ACK loss may lead to increased RTT (first ACK in a series of duplicates lost)
- Presence of SACK option indicates that reordering/loss was present at the receiver, sender SHOULD ignore that RTT update.
Technical Changes between -01 and -02

- **Window Scale (WS):**
  - sec 2.4 window retraction – M. Mathis

- **Timestamp (TS):**
  - sec 3.2 removed text to allow potential in-session negotiation of TS – M. Mathis
  - sec. 3.3 explicitly excluding ACKs with selective acknowledgements (SACK) for round trip-time measurement (RTTM) processing – R. Scheffnenegger