TCP Extended Option Space in the Payload of a Supplementary Segment

draft-touch-tcpm-tcp-syn-ext-opt-00
Jul'14, IETF 90 - Toronto

Joe Touch, USC/ISI
Bob Briscoe, BT (presenter)
Ted Faber, USC/ISI

Overview

- Two separate problems
  - Extended data offset (EDO)
    - **proposed as PS**
    - Simple option to extend the data offset post-SYN
    - Easy to enable/disable system-wide or per-conn.
  - this doc **SYN extended option space (SYN-EOS)**
    - **proposed as Experimental**
    - More complex
    - Independent mechanism
    - Extends SYN using a supplemental segment
    - Successful SYN-EOS implies EDO

(Similar) Motivation

- SYN option space use increasing
  - More options
  - Larger options (TCP-AO, MPTCP, TFO)
  - Current desire to combine large options
- Current use
  - SYN - typical total 19B
    - SACK-ok (2), Timestamp (10), Window Scale (3), MSS (4)
    - **Combining new options as well won't fit 40B limit**
    - **TCP-AO (16), MPTCP (12), TFO (6-18)**
- Negotiation
  - Trivial after initial SYN
  - **Cannot extend SYN in a single, backward-compatible segment**

Key Components

- Two approaches in one doc
  - WG to eventually decide which goes forward (or both)
  - Both add a supplemental segment
  - Extra option space in payload of supplemental segment
  - Large amount of effective SYN option expansion
- Shared properties
  - Initial SYN includes SYN-EOS request
  - Upgraded servers delay SYN/ACK until both segments are received
  - One SYN/ACK with SYN-EOS option ACKs both segments
  - Backward compatible
  - Robust to random loss, duplication, reordering

Two solutions

Out-of-Band (OOB)

- Supplemental segment:
  - (!SYN && !ACK) flags
  - Same ISN, addrs, ports
- Features
  - Looks like out-of-band data
  - RFC793 requires it be silently dropped
  - Fate sharing with initial SYN for both port block and redirection
- Flaws
  - Leaks through SYN-blockers
  - Fails if first through NAT

Dual-SYN (DS)

- Supplemental segment:
  - Second initial SYN (SYN-C)
  - Same addrs & dest. port, but different source port
  - Additional Conn. ID (CID) to match to initial SYN-D
- Features
  - Client resets a 2nd SYN/ACK from legacy server
  - Traverses firewalls
  - Blocked by SYN-blockers
  - Any NAT traversal order
  - Could traverse split connections
- Flaws
  - Lack of host & path fate sharing for port block or redirection

Shared Issues

- Order of processing options
  - Some options MUST be in original space, processed before merging the segments (TCP-AO)
  - MAY be a need to replicate some options (see draft for risks)
  - Process merged segments and their option space in a specific order (see draft)
- Interaction with EDO
  - SYN-EOS negotiation implies EDO is available after initial SYN
  - If EDO fallback is desired when SYN-EOS fails, EDO request needs to occur in initial SYN option space
- Interaction with SYN Cookies
  - Feasible – see draft
- Possible caching as an optimization
  - OOB: MAY consider caching supplement received before initial SYN
  - DS: SHOULD cache second SYN state if received before initial SYN
  - Security issues with caching
- Meddleboxes ;-)
  - Some issues that affect all options:
    - NAT/NAPT (compatible)
    - DPI (false positive or negative) – see draft for proposed partial solution
    - Block or remove EOS (failsafe)

Earlier Alternatives (focus on SYN)

- LO/SLO (long options/SYN long options)
  - SLO extends SYN
    - Prolongs 3-way handshake (3WHS) for extra segments
    - SYN/ACK can't enter ESTABLISHED after 3WHS
      - because later segment options may be rejected
    - Either wait for 5WHS or add complex state management
- LOIC (long options by invalid checksum)
  - Dual SYNs, the second with invalid checksum
  - Second SYN won't traverse a NAT
    - checksum will fail or be revised as correct
- 4-way handshake (Borman 5/22 TCPM post)
  - First SYN asks to support long options
  - Server reply is a SYN using long options (reverses connection)
  - Adds 1 RTT to client-side data latency
  - Not compatible with directional options (TFO) or parameterized options
  - Requires server expose entire option capability list to clients

Current status

- Not posted for Toronto deadline
  - {to be submitted by end of July | just submitted }
- Request for feedback on draft-touch-tcpm-tcp-syn-ext-opt-00
  - Messy, but may be the only way out
  - OOB vs. DS vs. both?
  - Feedback on details and issues, please
- Open issues
  - DS
    - Distinguishing the two SYNs
    - Length of CID field
  - OOB
    - Verification of NAT/firewall traversal
- Individual draft, intended as TCPM Experimental
  - No known IPR
  - Too early to consider call for adoption ;-)
  - Adoption call on list soon?