Updates on Windows TCP

Praveen Balasubramanian, Matt Olson, Yi Huang

Microsoft

Quick recap

- Recent TCP advancements on nearly 800 million+ devices running Windows 10
  - IW10
  - RACK + TLP
  - CUBIC
  - TFO (Limited websites, Aggressive fallback)
  - LEDBAT++ (Limited scenarios)
  - Reduced Delayed ACK timeout default
  - Higher ABC limit
- Server 2016 and 2019 deployment for online services at the edge
  - CUBIC
  - TFO

Improved Slow Start

- HyStart
  - Collect 8 RTT samples per round (previously 1 sample when timestamps off)
  - Delay Increase algorithm only
    - "Inter-packet arrival" algorithm has issues due to ACK compression

    Eta = min(8ms, max(2ms, LastRTT / 16))

- Limited Slow Start after HyStart exit
  - HyStart causes premature exit when RTT fluctuates (e.g. Wifi or congestion)
  - LSS based on RFC 3742
  - For each arriving ACK in slow-start

    K = int(cwnd /(0.25 * ssthresh)
    cwnd += int(bytes_acked / K)

- HyStart and LSS currently limited to just the initial slow start

RACK updates

- Compliant with draft-ietf-tcpm-rack-04
- In conjunction with traditional Dup ACK based loss detection

- Following optional portions not implemented:
  - DSACK based dynamic reorder window
  - "Reordering setting" timer
  - Optimization to sort sent packet list in time order
  - Ability to retransmit lost retransmits

- Request for moving forward with WGLC

Lower InitialRTO

- Windows default for InitialRTO was 3 seconds
- Total SYN retransmissions were capped to 2 resulting in 21 seconds total timeout
- Lowered InitialRTO to 1 second by default
- Total timeout kept at 21 seconds for app compat reasons
- The only TCP option removed for SYN retransmissions is TFO if it was attempted

Q&A