Proposal for Updates to Guidance on Packet Reordering

draft-white-intarea-reordering

(GitHub link)

INTAREA – IETF 126

July 22, 2026

Greg White, Ingemar Johansson, Dibakar Das, Chris Box

Why is this important?

- Some widely used L2 links implement frame resequencing to enforce in-order delivery
- This is done without any awareness of L3/L4 context
- The result is the introduction of head-of-line blocking latency that potentially affects all traffic on the link
  - Regardless of whether the reordering at L2 resulted in reordering of any particular L3/L4 flow
  - Regardless of whether the L3/L4 flow(s) are sensitive to out-order arrivals (most aren't)

Example Case: Wi-Fi HoL Blocking

- Wi-Fi links introduce reordering due to link-layer retransmissions
- Transmission opportunities often carry an “aggregate” of L2 frames. Individual frames within the aggregate can be lost in transmission.
  - Frame loss percentage depends on RF conditions and MCS selection algorithm (in stable conditions typically between 1-10%)
- L2 retransmits the missing frames, but the retransmission involves another MAC contention attempt (can be 10s of ms).
  - e.g. with 7 other stations actively contending, the retransmission delay can be more than 75ms at P90, and over 400 ms at P99.
- All higher sequenced frames are held waiting for the retransmission in order to assure in-order delivery.

Draft status

- Draft-03 updates
  - Added detail on IPsec based on discussion on ipsec mailing list
- Draft-04 updates
  - Added discussion of IP translators (CG-NAT) and IP Fragmentation
  - Added statistics on this issue for Wi-Fi links
  - Significant re-work of proposed guidance text (next slide)

New text providing guidance:

Subnetwork implementers SHOULD avoid introducing unnecessary packet reordering. However, where packet reordering is an unavoidable consequence of mechanisms that improve overall performance or reliability (for example, packet striping across multiple links or link-layer retransmissions), the resulting packet reordering SHOULD generally be exposed to the receiving endpoint rather than hidden by subnetwork resequencing.

Internet endpoints are ultimately responsible for interpreting packet ordering according to the semantics of the transport protocol or application. A subnetwork generally lacks sufficient knowledge to determine whether a particular degree of packet reordering would have adversely affected the receiving transport protocol or application, but it does know that resequencing necessarily delays packets that could otherwise have been forwarded immediately.

Modern transport protocols and many applications already incorporate mechanisms to tolerate bounded levels of packet reordering. Performing resequencing within the subnetwork duplicates functionality that already exists at the endpoint, while introducing head-of-line blocking that increases latency and delay variation for packets that were not themselves reordered. Subnetwork resequencing SHOULD therefore only be employed where there is clear evidence that its benefits outweigh the additional latency and jitter that it introduces.

Next Steps

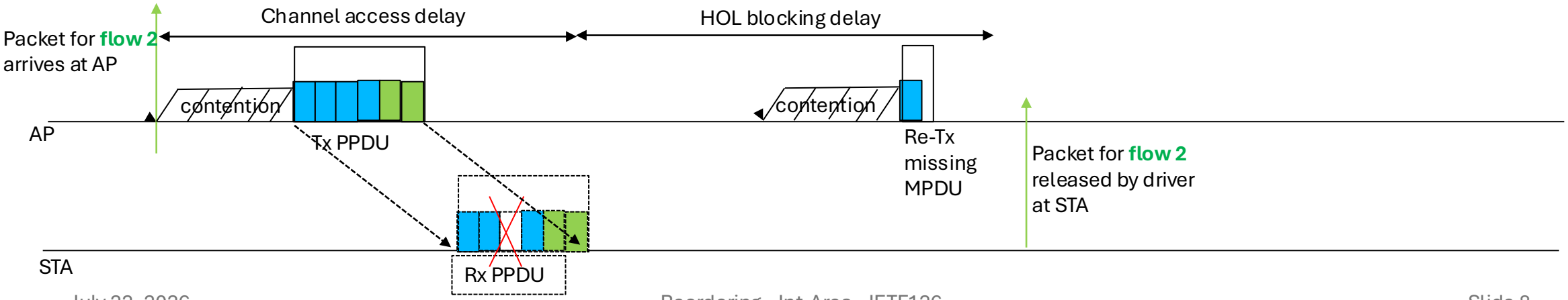
- Adoption Call?

Backup Slides

Wi-Fi HOL blocking study

Wi-Fi HOL Blocking: Background

- The wifi latency consists of channel access delay and the HOL blocking delay.
- We pursued this general demultiplexing problem in wifi-8; ended up only with minor enhancement to use additional TID to enable demultiplexing.
- At that time we presented sim results showcasing benefits.
- Goal of the study: use practical implementations to see whether implementations already fix this.

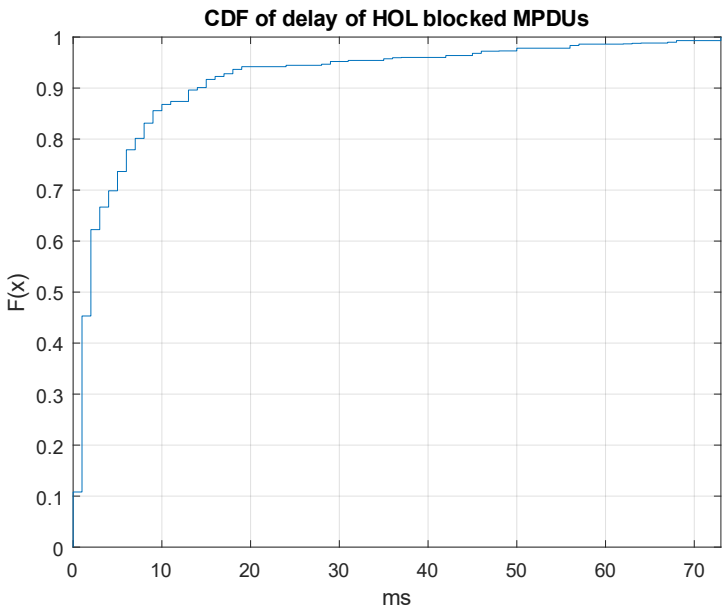


Goals

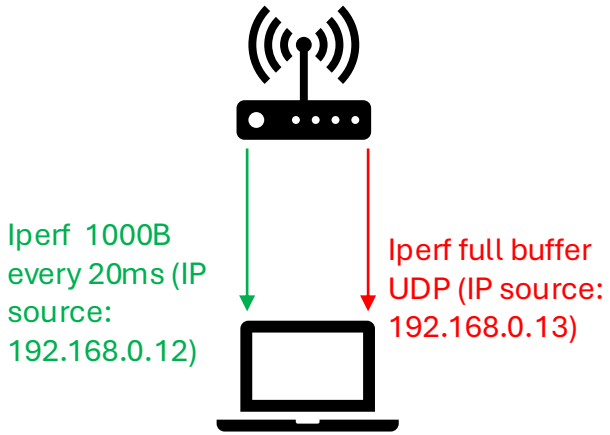
- Analyze how often the HOL blocking happens ? Has two parts:
  - How often an MPDU is lost (due to collision/channel error) in a PPDU ?
  - Whether the missing MPDUs belong to same flow/QUIC stream ?
    - Motivation: if there is dependency on the missing MPDU, then wont be useful to forward this MPDU up anyway.
- When it occurs, what is the delay ?

Scenario 1

- Motivation: heavy traffic mixed with small periodic traffic (e.g., audio/control) in presence of full buffer interference.
  - This could be unmapped QoS traffic or QoS traffic that got mapped to same TID.
- Observations: AP doesn't separate the traffic, significant blocking events, delay scale as function of network congestion..

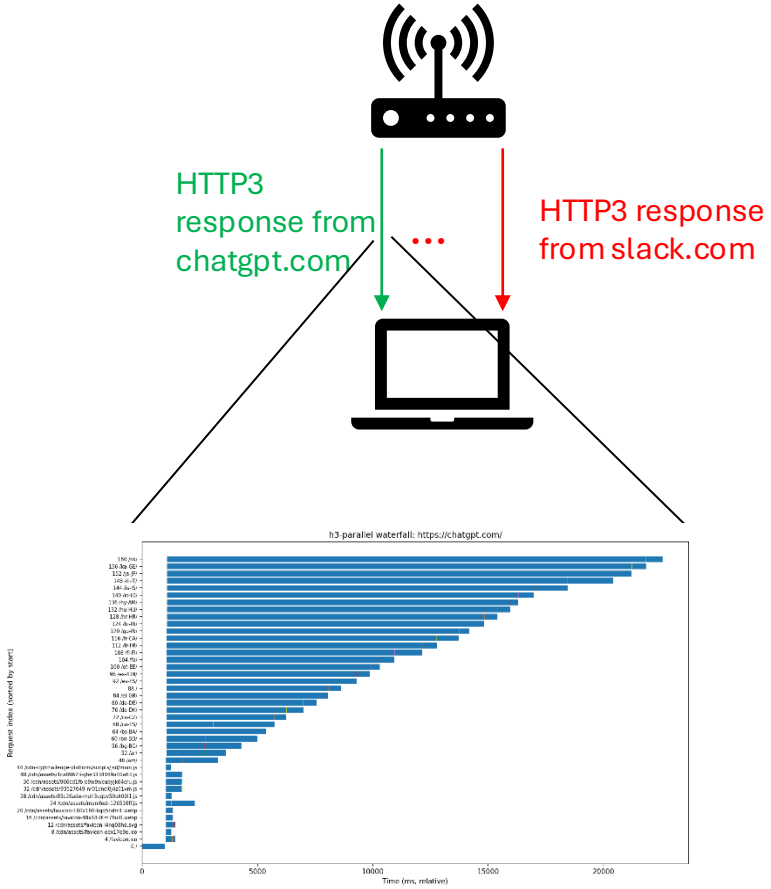


src_ip	Total blocked packets	Blocked by diff stream	Blocked pct
192.168.0.12	1063	1038	97.648
192.168.0.13	178826	3402	1.902

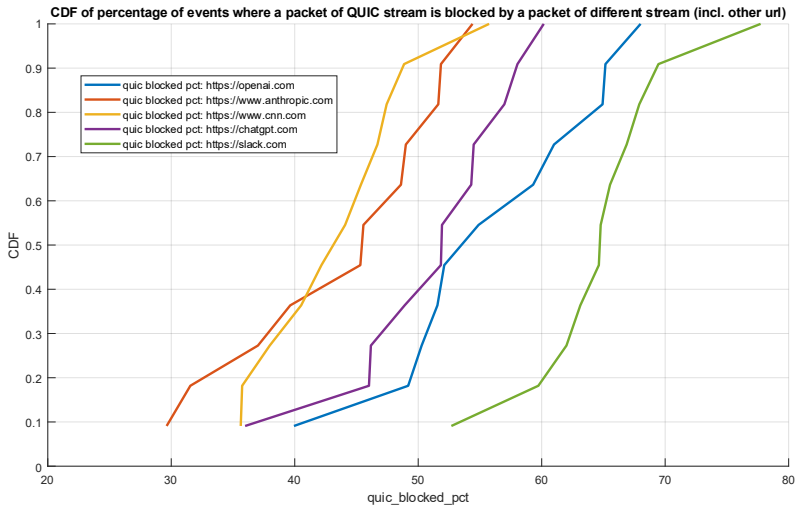


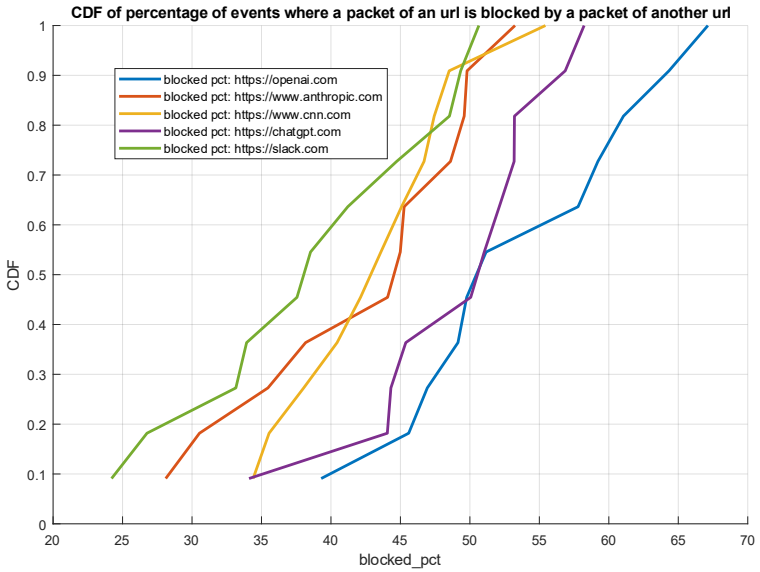
Scenario 2

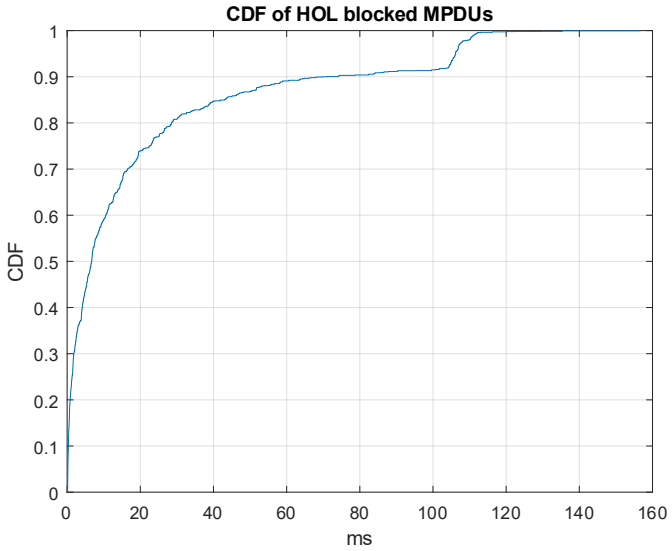
- Motivation: QUIC streams in presence of interference as well as other flows.
- Setup: Use Aioquic to initiate multiple parallel QUIC connections to several sites that support HTTP3.
  - For each website try to fetch resources in upto 30 parallel QUIC streams.
  - Sites used (all together): cnn.com, anthropic.com, chatgpt.com, openai.com, slack.com, google.com
  - Besides common webbrowsing, this could also model multi-agent orchestration workflows that requires HTTP .
  - 10 iterations



Scenario 2 results







- **When MPDUs suffer from HOL blocking,**
  - significant instances when blocking happens due to a different flow.
  - The impact of the delay depends on the network congestion.
- **Note: user experience for HTTP session deteriorates after ~400ms**

Backup Slides

Recap of reordering draft status

Quick Recap

- IETF122 (March 2025): slides
  - Background
  - Existing IETF guidance
  - Goals
  - Potential benefits of guaranteeing in-order-delivery
  - Potential downsides
  - Current state of L3+ protocols
  - Current state of L2 technologies

Background

- L2 technologies can have link characteristics that result in packet *reordering*, e.g.:
  - Link Aggregation
    - When aggregated links have different latency or capacity
  - L2 retransmissions
    - Common in wireless links
- Several L2 standards introduce requirements for *resequencing* to restore the frame sequence by delaying early-arriving frames
  - Is this what IP protocols want?

Existing IETF guidance

- RFC3819 §15
  - The Internet architecture does not guarantee that packets will arrive in the same order in which they were originally transmitted; transport protocols like TCP must take this into account.
  - [L2 links] should try to avoid packet reordering whenever possible, but not if doing so compromises efficiency, impairs reliability, or increases average packet delay.

  - But…
    - … reordering does come at a cost with TCP as it is currently defined … the TCP sender needlessly reduces the congestion window and performance suffers.
    - Note that every header compression scheme currently standardized for the Internet requires in-order packet delivery on the link between compressor and decompressor.

Goals

- Provide up-to-date information for L2 technology SDOs on the impact that packet reordering & resequencing have on IP protocols.
- Enable SDOs to make informed decisions for future standards.

Quick Recap

- IETF122 (March 2025): slides
  - Background
  - Existing IETF guidance
  - Goals
  - Potential benefits of guaranteeing in-order-delivery
  - Potential downsides
  - Current state of L3+ protocols
  - Current state of L2 technologies

The gist

- Latency is increasingly important
- Guaranteeing in-order-delivery:
  - increases latency
  - adds cost and complexity to L2
  - provides no value in many cases
- RFC3819 sounds dire

Let’s fix this!

- At least update/remove the text about TCP and RoHC…

- Can we come up with something better?
  - Better guidance for L4 protocol developers on how much reordering they might need to tolerate?
    - e.g. what if existing L2 networks stopped resequencing packets?
  - Better guidance for L2 protocol developers on how much reordering they can create?
    - e.g. what remaining sensitivities are there in existing protocols/implementations?

- How do you quantify reordering anyway?
  - RFC4737 provides several different metrics