Characterizing the Impact of Active Queue Management on Speedtest Measurements

Anonymous

No Institute Given

Abstract. Present day speed tests measure peak throughput, but often fail to capture the user-perceived responsiveness of a network connection under load. Recently, platforms such as Ookla's Speedtest.net and Cloudflare have introduced metrics such as "latency under load" or "working latency" to fill this gap. Yet, the sensitivity of these metrics to basic network configurations such as Active Queue Management (AQM) remains poorly understood. In this work, we conduct an empirical study of the impact of AQM on speed test measurements in a laboratory setting. Using controlled experiments, we compare the distribution of throughput, latency, and latency under load measurements across different AQM schemes, including CoDel, FQ-CoDel and Stochastic Fair Queuing (SFQ). On comparing the results with a baseline of no AQM, we find that [TS: add the main punchline here.] These results highlight the critical role of AQM in shaping how emerging latency metrics should be interpreted, and underscore the need for careful calibration of speed test platforms before their results are used to guide policy or regulatory outcomes.

Keywords: Speed Tests · Active Queue Management · Responsiveness · Bufferbloat

1 Introduction

Internet performance has historically been summarized using a single number: "speed" [1, 5, 3]. Despite the widespread utility, the user perceived Quality of Experience (QoE) for many applications (e.g., video-conferencing, gaming, cloud collaboration) is governed less by peak bandwidth and more by latency under load. To address this, measurement providers have recently begun introducing "latency under load" (LUL) or "responsiveness" metrics, which attempt to capture how queuing delays increase during simultaneous upload and download activity.

However, the interpretation and use of these metrics has not been standardized. For instance, Ookla defines "working latency" as the increase in round trip time (RTT) under load compared to the unloaded RTT, measured during a speed test [2]. Apple uses a different metric, called round-trips per minute (RPM) under load, which counts the number of round trips completed during a fixed time interval while the connection is saturated [6]. Further, these tests have been known to discard outliers that often correspond to glitches that users typically notice during real-time applications such as video conferencing and streaming [2]. As a result, users and regulators are left with incomplete pictures of what causes an Internet connection to be unresponsive, and how it can be mitigated. A central, unanswered question is how old metrics such as throughput and latency, and new metrics such as LUL and RPM, behave in the presence or absence of active queue management (AQM) algorithms such as FQ-CoDel, which were explicitly designed to maintain low latency under load [4].

In this paper, we investigate how the distribution of modern speed test measurements shifts when an AQM is deployed. Rather than reporting only mean throughput, the 90^{th} percentile or median latency, we analyze full distributions: the tails, the spikes, and metrics similar to "glitches per minute" [2] that are most relevant to real-time applications. Our goal is to empirically characterize the difference between unmanaged queues and AQM-enabled network paths, and to highlight how this difference is (or is not) reflected in widely deployed measurement platforms. By doing so, we aim to inform both test designers and network operators of the gaps between the status quo of Internet measurement and the actual experience of end-users.

[TS: Methods summary and contributions go here.]

References

1. Bauer, S., Clark, D.D., Lehr, W.: Understanding broadband speed measurements. TPRC (2010)
2. Dave Täht: The new features and flaws of speedtest.net, ookla and cloudflare. https://blog.cerowrt.org/post/speedtests/, accessed: 2025-10-01
3. Feamster, N., Livingood, J.: Internet speed measurement: Current challenges and future recommendations (2019), https://arxiv.org/abs/1905.02334
4. Høiland-Jørgensen, T.: Bufferbloat and beyond
5. Midoglu, C., Wimmer, L., Lutu, A., Alay, O., Griwodz, C.: Monroe-nettest: A configurable tool for dissecting speed measurements in mobile broadband networks (2018), https://arxiv.org/abs/1710.07805
6. Paasch, C., Meyer, R., Cheshire, S., Hawkins, W.: Responsiveness under Working Conditions. Internet-Draft draft-ietf-ippm-responsiveness-07, Internet Engineering Task Force (Jul 2025), https://datatracker.ietf.org/doc/draft-ietf-ippm-responsiveness/07/, work in Progress