Before the
Federal Communications Commission
Washington, D.C. 20554

In the Matter of	)	
	)	
Inquiry Concerning Deployment of Advanced	)	GN Docket No. 22-270
Telecommunications Capability to All Americans	)	
in a Reasonable and Timely Fashion[1]	)	

Dave Taht
CEO, Bufferbloat.net

Members and supporters of
the Bufferbloat.net community

December 1, 2023

This letter is in response to a multiplicity of statements within this NOI, summarized by Chairwoman Rosenworcel as:

"... kicking off this inquiry to update our national broadband standard to better align it with the standards in pandemic-era legislation of 100 Megabits per second down and 20 Megabits per second up and also set a long-term goal for gigabit speeds."[2]

We argue that to best serve the people of the United States, the Commission should balance its near-term efforts on achieving internet resilience and **minimizing latency**, instead of only increasing "speed" or "bandwidth".

Calls for further bandwidth increases are analogous to calling for cars to have top speeds of 100, 500, or 1000 miles per hour. Without calling also for better airbags, bumpers, brakes, or steering wheels, (or roads designed to minimize travel delay), these initiatives will fail[3] (and are failing) to meet the needs of present and future users of the internet.

---

[1] https://docs.fcc.gov/public/attachments/FCC-23-89A1.pdf
[2] https://docs.fcc.gov/public/attachments/FCC-23-89A1.pdf Page 30
[3] https://blog.cloudflare.com/making-home-internet-faster/

Latency matters because it:

Impacts Real-time Applications
Applications like video conferencing, online gaming, and VoIP are sensitive to latency. High latency or variable latency (jitter) in these applications can cause delays, echoes, distortions, or freezes in video calls, gaming lag, and audio conversation interruptions, leading to a frustrating user experience.

Reduces Throughput and Efficiency
High latency can also reduce overall network throughput. It can take a very long time for a new flow to ramp up to parity with other flows, no matter how short, as a minimum for the TCP protocol is essentially 3 round trips to get any data at all. At large delays or jitter, a single packet loss on an established flow can trigger a burst of packet retransmits, rather than just replacing the lost packet. This results in slower page loads, possible changes in resolution (or interruption) of streaming video, a “Buffering” message, and many other noticeable effects. In extreme cases, “Congestion Collapse” can occur, where the network becomes so busy retransmitting lost data that it cannot transfer much new data, and retransmitting old data leads to more lost data, a death spiral that cannot end until users stop hitting the reload button and give up. Many networks experiencing persistent induced delays of over 250ms are essentially in this state.

Gives Variable User Experience
High latencies can cause highly variable and unpredictable network performance. Users may experience sporadic slowdowns and speed-ups, making the network feel unreliable and leading to a poor user experience.

1. We’re Looking in the Wrong Places

Regrettably, the NOI, with the exceptions of sections 26-29, essentially reads like an argument for ever more bandwidth - which will have very limited (or even no) additional value to end users. Bandwidth used to be a primary concern, but that is no longer the case. In addition, the OpenVault Report, cited in the NOI, tracks the habits of “Extreme Power Users”, and not the general population.

For example, a 10gbit link with 50ms latency is readily *outperformed* by a 10mbit link with 1ms latency for most interactive traffic. Increasing the “speed limit” of the link without actually making the road navigable at the higher speed is a waste of effort.

The changes in the use of the internet during the COVID pandemic made this more obvious. In addition to web browsing, we now make far greater use of VoIP, video-conferencing, and using employers’ online programs remotely. All these suffer when they have to deal with latency.

Increases in capacity over the past few years improved the responsiveness of most traffic, up to around 20 Mbps. From that point on the only way to improve responsiveness is to robustly and reliably reduce the latency, and especially the “latency under load” for the end-user. Given these facts, it seems advisable to not focus mainly on increasing the minimal access rates further, but on making sure internet access links offer robust and reliable low-latency service under all conditions.

Appendix A of our response expands on this, and describes a number of tests that ordinary users can run to see how well their link performs at various underlying latencies.

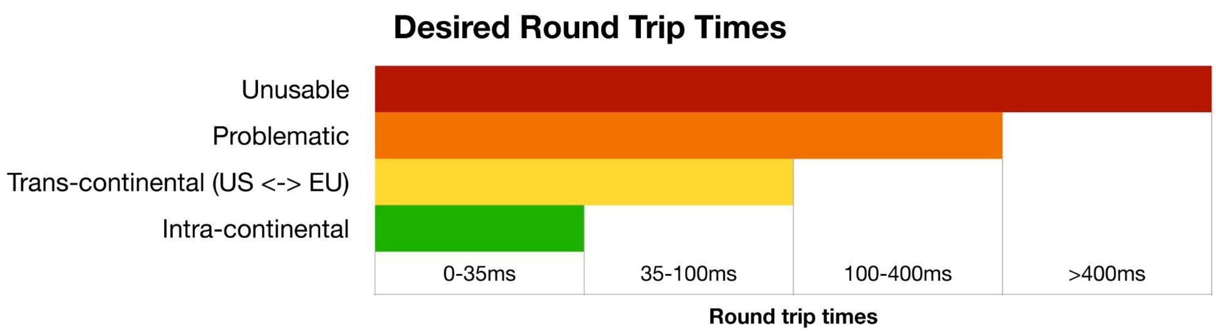
Appendix B describes various, more detailed, test results that demonstrate the bufferbloat problem, and the solutions for it.

Appendix C discusses how Network Neutrality and the bufferbloat problem intersect.

2. The Right Place starts with Reducing Round Trip Latency

The Broadband Technical Advisory Group (BITAG) produced a detailed report on the characteristics of all forms of internet traffic in January, 2022, with an even more detailed focus on the various network latency issues we are describing in this filing. We urge more readership of that report, in particular![4]

Adding bandwidth alone is inadequate to meet the recent demands on the internet for working from home, for better videoconferencing, gaming, and other interactive experiences.
Actual usage patterns today, minute to minute, are better measured in steady bytes per millisecond, not in megabits/second. Clearing a credit card, sending an instant message, and posting an email are trivial uses of bandwidth. Each and every keystroke typed into the Google search bar, or Google docs, generates a single packet (around 100 bytes) where latency and jitter are most apparent. There is more than just transmission delay and bufferbloat to cope with here, as those keystrokes engage much processing, searches, and database lookups that also have to complete consistently and rapidly for a good user experience.

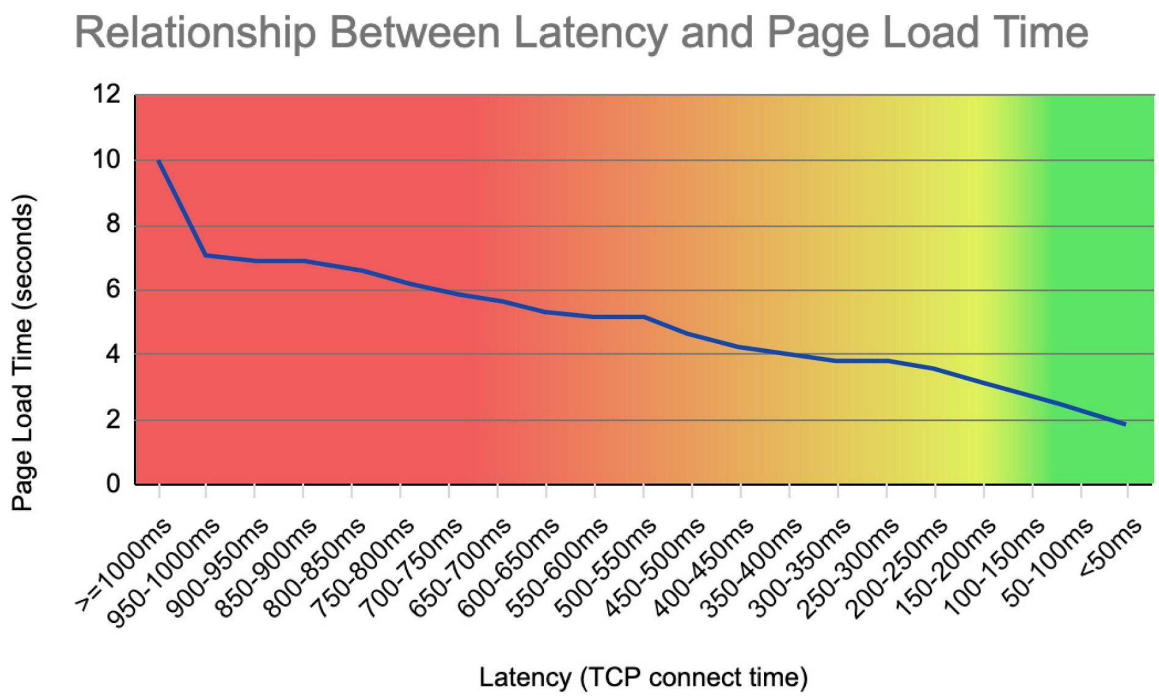


Reducing the network's round-trip latency, whether it be to a server in a data center, or between your phone service and your home service, is always worthwhile[5]. As for services that use data, reduced round trip time (RTT) always speeds up the initiation of service and controls like fast forward or reverse for streaming media.

Web page load time is almost entirely bound by the round-trip time and latency, not by bandwidth[6], and it is rare that a typical web page will use up more than 20Mbits at any instant in time[7].

---

[4] https://www.bitag.org/latency-explained.php
[5] https://www.cs.cmu.edu/~./xia/resources/Documents/Singla_hotnets14.pdf Internet At the Speed of Light
[6] https://blog.cloudflare.com/making-home-internet-faster
[7] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4178804 Understanding the Metrics of Internet Broadband Access: How Much Is Enough?



Even web page load time is a poor metric[8], as most web pages frequently used are cached locally, and navigation within a website is often also measured only with a few hundred kilobytes of fresh data per click, the loading of which is dominated by (wait for it), the network round trip time and the time it takes to assemble the data from the servers. A typical web page download may require hundreds of round trips, so milliseconds can and do turn into whole seconds in the end.

3. Other Flaws in data

The speedtest regimes used by Measuring Broadband America to assess network quality have many flaws remaining. Perhaps the largest is that no official test measures the side effects of doing an upload *and* download performed at the same time, and very few[9] user-accessible tests do either. The frequently occurring situation where one user is doing a videoconference or game, while another does a big upload or large web page download is not covered by this dataset. In recent years many speed test tools have evolved to cover what is variously called "responsiveness", "working latency", and "ping under load" - exactly because this measure has such a large effect on how users perceive the quality of interactive use cases, but these only test up + interactive or down + interactive traffic. Further improvements in measurement methods will help prevent optimizing the internet for speed tests and other artificial metrics.

[8] LCP (Largest Contentful Paint) is preferred to page load time, as a metric. It engages the user by specifically measuring the rendering speed of the largest and most important content element on a web page. This metric provides a more user-centric perspective on performance, focusing on what matters most to visitors, which can lead to a better overall user experience.
[9] https://github.com/Zoxc/crusader/ tests latency under load for both up and down.

Other issues with the test regime include:

- Passive testing metrics for long-term connectivity, latency, and outages
- Latency only probed without load
- Lack of spatial diversity (measurement nodes ideally are located at either many major exchanges, and/or major ASs) that is user selectable or randomized
- Lack of automated bi-directional traceroute between client and server, or client and client
- Homogeneity of run time. Most tests are not run long enough to reliably cause the same effects on responsiveness as e.g. a multi GB download would do.
- Lack of testing of bursty/DASH(netflix)-like traffic
- Few Interconnectivity (Cybergeography) tests between ISPs

A “speedtest” typically tests 10 or more full-rate flows starting at the same time, usually completing in under 20 seconds, a situation that almost never occurs with real traffic. Projecting future bandwidth needs based on speed tests is merely optimizing the network for speedtests. Would you optimize a car to accelerate only in a straight line for 20 seconds?[10]

Our industry and regulatory and test regimes were caught unprepared by working from home and videoconferencing needs, and still have not caught up yet. Many current WFH applications actually need less bandwidth, not more, but have far, far more stringent demands for consistently low latency. Gaming requires consistent packet delivery within 20 ms intervals, videoconferencing 16 ms (at 60 FPS), and truly good VR and AR experiences 4-8ms.

[10] https://blog.cerowrt.org/post/speedtests/

4. Detailed Recommendations

To address some of these issues, we have a number of recommendations, starting with strongly suggesting better tests be used and developed for measuring the characteristics of real traffic.

4.1 Reducing Latency Under Load

We start with latency, as we described above. The endless drive for more and more bandwidth has obscured the key factor in actual experienced “speed”, in having right-sized buffers or smart algorithms for the offered load, which allows for all traffic types to co-exist.

The past few years have seen an explosion of new tests such as Waveform’s Bufferbloat test[11] Ookla’s “Loaded Latency”[12], SamKnows[13], CloudFlare’s[14], and Apple’s “NetworkQuality”[15]. Most of these sites also give recommendations as to how end-users could achieve low latency under load[16]. More people understanding what the new numbers being reported mean and what to do about them would be very good moving forward[17].

4.2 Adding Quality of Experience Improvements Universally

A core problem with aiming for bandwidth alone is in not also aiming for good multiplexing of different traffic types. The Internet’s edge devices today are almost universally over-buffered to a level far beyond what any theorist would recommend[18], making common interactive traffic far slower than necessary.

Ironically, adding larger memory buffers to home routers results in the user having to wait for messages with spinning wheels that say “buffering”, instead of making the routers faster. This particular blunder gave us the name “bufferbloat” for the problem, and spurred significant work in managing the underlying data transiting the device. Advancements in Fair Queuing (FQ) and Active Queue Management (AQM)[19] techniques have become widely available since 2012, available in newer edge equipment (or retrofitted to them with newer software), and are simple

---

[11] https://www.waveform.com/tools/bufferbloat The waveform bufferbloat test is very popular

[12] https://www.ookla.com/articles/introducing-loaded-latency OOKLA Introducing loaded latency

[13] For SamKnows: Methods for measuring jitter and latency are described in the last section of https://samknows.com/tests/latency-loss-jitter - Their previous methodology which was used by the FCC for MBA is mentioned on page 30 https://data.fcc.gov/download/measuring-broadband-america/2022/Technical-Appendix-fixed-2022.pdf

[14] https://speed.cloudflare.com/ Cloudflare Speedtest

[15] https://datatracker.ietf.org/doc/draft-ietf-ippm-responsiveness/ IETF draft on responsiveness

[16] https://www.ookla.com/articles/maximized-speed-non-gigabit-internet-connection Ookla

[17] https://evenroute.com/test-evaluation Understanding bufferbloat results

[18] https://gettys.wordpress.com/2018/02/11/the-blind-men-and-the-elephant/

[19] https://queue.acm.org/detail.cfm?id=2209336 The Codel AQM

to configure by the end-user or ISP. Although per device testing is needed - to vastly improve latency under load, nowadays, all that is needed for all DOCSIS 3.1 modems is the addition of a single line ("AQM 1") to the configuration file to enable the RFC8034[20] "PIE" AQM. Comcast deployed this a few years ago[21]. Cox has also deployed this with others in the works.

There are still many older DOCSIS 3.0 or even DOCSIS 2.0 modems extant which can be either retired or upgraded.

There is an enormously long "tail" of older deployed devices across all our edge technologies (LTE/5G/DSLI/fiber/FWA) that can either be upgraded or replaced to more modern software, which also generally brings benefits like IPv6 support and improved security.

Many advanced or "gaming" home/small business routers now include technology combining FQ with AQM, derived from RFC8290 (fq_codel or CAKE). The benefits to the gaming community of reducing lag in this way are widely known, and yet they apply to any household with more than one member attempting to use the internet at the same time[22], enabling simultaneous usage of all internet applications by multiplexing packets better and holding overall latencies low. RFC8290[23] has had many trade names - "Streamboost", "Anti-Bufferbloat", "Adaptive QoS", and "Smart Queues", to name a few. A common name used in the OpenWrt derived distributions is "Smart Queue Management" (SQM), or CAKE, which combines it with a shaper and classifier[24]. Configuring this is mildly more complicated:

tc qdisc add dev eth0 root cake bandwidth 20Mbit nat ack-filter [docsis|ethernet|dsl][25]

But relative to the other millions of lines of code in routers today, seems straightforward, although enabling automated provisioning of the correct values remains difficult across the thousands of brands of routers across the internet, and the many that are not upgraded, or upgradable.

There are also now multiple middlebox makers making it possible for ISPs to inexpensively manage their bufferbloat and latency under load via these algorithms without having to upgrade their entire physical plant to run this code directly. [26]

After a quality queue management system is installed at the device and at the ISP, misplaced user demand for more bandwidth essentially vanishes, support calls about bad VOIP,

---

[20] https://datatracker.ietf.org/doc/rfc8034/ DOCSIS-PIE AQM. See also RFC8033

[21] https://www.reddit.com/r/HomeNetworking/comments/r7cfvg/comcast_delivers_90_working_latency_reductions_by/

[22] https://circleid.com/posts/20130418_bufferbloat_demo_see_how_much_faster_internet_access_can_be/

[23] https://datatracker.ietf.org/doc/rfc8290 FQ_Codel. See also RFC8289 (Codel)

[24] https://www.bufferbloat.net/projects/cerowrt/wiki/Smart_Queue_Management/

[25] https://man7.org/linux/man-pages/man8/tc-cake.8.html CAKE manual page (dsl more complicated)

[26] LibreQoS.io, Preseem, Bequant, and Paraqum all leverage RFC8290-derived technologies.

videoconferencing, or gaming stop, and the internet gets better in subtle ways difficult to measure with a short test. Users report improvements in smoothness and are less frustrated with bursts of lag. Better testing for useful features (such as per device fairness) needs to be tested in a lab, at device certification time, rather than in the field.

Originally, there was much fear, uncertainty, and doubt in the ISP industry[27] that misguided lawyers might regard these improved intelligent packet drop and scheduling algorithms as a violation of network neutrality. In fact packet drop, is a necessary part of the underlying structure of the internet [28], and "fair queuing" is a common multiplexing-improvement technique dating back to the 1990s[29].

Since PIE is readily available for DOCSIS and fq_codel the default[30] in all modern Apple and Linux gear, and used by so many routers already, it is hoped by the authors of this document that the FCC might aid in accelerating deployment of these technologies across the edges of the internet, possibly by setting minimum standards of required responsiveness under load/working conditions.

4.3 End-to-End Congestion Control Improvements

There have been vast improvements in end-to-end congestion controls also, over the past decade, notably packet pacing, and new technologies called BBR [31][32] and LEDBAT [33] that are increasingly deployed. One subtle interaction that few perhaps have noticed, is that nowadays system updates for Windows, Apple, and Linux devices all take place transparently in the background without affecting foreground activities in the network all that much. Only the gaming industry still seems to be using congestion controls that negatively affect the user experience nowadays while (admittedly) downloading hundreds of gigabytes of data.

Microsoft Windows also has quietly upgraded its client TCP stacks to RFC9406 to lighten the impact on slow upload bandwidths. Netflix has improved its packet pacing implementations to make room for other applications.

In a more rapidly developing world, these improvements to congestion controls would also make home users doing uploads more comfortable with their internet performance. Certainly upload

[27] https://blog.cerowrt.org/post/net_neutrality_customers/ Fear of faster queueing.
[28] https://ee.lbl.gov/papers/congavoid.pdf Congestion Avoidance and control, Jacobson & Karels. 1988
[29] http://www2.rdrop.com/users/paulmck/scalability/paper/sfq.2002.06.04.pdf SFQ, 1990
[30] https://blog.cerowrt.org/post/state_of_fq_codel/ State of fq_codel, June 2022
[31] https://queue.acm.org/detail.cfm?id=3022184 Congestion based Congestion Control
[32] https://www.lk.cs.ucla.edu/data/files/Kleinrock/Internet%20congestion%20control%20using%20the%20power%20metric%20LK%20Mod%20aug%202%202018.pdf - Keep the pipe full, but no fuller, Kleinrock, L., 2018
[33] https://www.sciencedirect.com/science/article/pii/S1389128622001888 LEDBAT++

speeds need to be improved to 20Mbit and beyond, but fixing the bufferbloat, reducing the impact of uploads vs other more interactive traffic, should happen also.

4.4 Improving Wi-Fi

The FCC, despite many public investments into 5G, has had no (few?) programs to promote deeper understanding or research of Wi-Fi, which is the last few dozen feet of the internet connection that 97% of Americans get their broadband from. However, Wi-Fi Issues of interference and range, however, occur often and are often misunderstood by the public to be the ISP's fault. It is in general better, in Wi-Fi, to use less wide channels that do not interfere, and improve bandwidth at range. A speedtest of a Wi-Fi connection a few feet from the access point is often wildly different than one from the next room over. At higher bandwidths from the ISP, the bufferbloat shifts to the Wi-Fi, and with more devices online usability of Wi-Fi becomes an increasing factor.[34] In mandating ever higher bandwidths and not range, Congress seemingly forgot about the inverse square law, and too few understand Wi-Fi's multi-device scaling limitations[35].

4.6 What Forms of Fiber?

One question that is never asked, is what form of fiber do you intend to deploy?

"Active Fiber Ethernet" scheduling delay is measured in nanoseconds, and the hardware available to anyone at Best Buy and commonly used within and to connect data centers. Bandwidths of 100-400 Gbit are readily available for it today. GPON is presently limited to multiplexing 25 Gbit, across customers, and has an intrinsic 250 microsecond (µs) scheduling delay. For cable, scheduling delays are in the 2-6ms (DOCSIS 4.0-LL gets closer to 1 millisecond (ms)). Many forms of wireless also get down below 250µs. DSL is typically 20ms (interleaving), Starlink is presently 40ms of intrinsic delay, and aiming for 20ms. These inherent baseline latencies dominate most day to day, minute to minute, uses of the internet, not the bandwidth. The much lower latency of both forms of fiber is responsible for much of the perception that "fiber is faster". That said, it is the total path latency that matters. Starlink may have that intrinsic 40ms delay but strives to make the next hop to multiple CDNs short. A fiber network with a 40ms path to the data will be outperformed by a cable network with a 10ms path.

Active fiber is far better for "transit" than GPON, as it has vastly more capacity, also. GPON has advantages over active fiber in many other scenarios in terms of simplicity and power.

[34] https://apenwarr.ca/diary/wifi-data-apenwarr-201602.pdf Google Fiber's 2016 research
[35] https://www.youtube.com/watch?v=Rb-UnHDw02o&t=1544s An explanation of WiFi Aggregation

4.7 We Lack Resiliency

During an internet outage productivity plummets. On-line commerce ceases. A view towards improving mean time to repair (MTTR) and mean time between failures (MTBF), and degrading gracefully during disasters and other interruptions of service, is needed. Improvements in topology, such as meshy-ness, or ring-like designs, are better than carrying all packets back to Washington, or relying on metrics such as pure "passings". That kind of reliability was one of the original aims of the Internet, NSFnet, and ARPAnet.

Diverse, duplicate, and redundant paths are needed throughout any network to simplify MTTR. A break in a cable or wireless connection with a backup path can be repaired at leisure, one with a single point of failure must be repaired immediately.

4.8 Replacing Aging, Insecure Firmware

The NOI has an additional broad scope of "Whether advanced telecommunications capability is being deployed to all Americans in a reasonable and timely fashion."

Multiple other nations have established programs to roll out the IPv6 protocol suite across their populations[36]. Ours has not. How can the Digital Divide be crossed without sufficient IP addresses? With native IPv6, VOIP and video-conferencing, transit latency can go from 300ms to a few ms due to native P2P, without reliance on an IPv4 TURN server located many miles away from the endpoints.

Router software often lags 7-12 years behind current developments. Laptop computers and cellphones have evolved to a steady set of constant security patches and upgrades delivered over the air (OTA), most edge, router, and IoT devices have not. Perhaps the FCC device certification processes can be improved.

4.9 Capacity Ratios Are Important

The focus is typically on download capacity (speed). However, the ratio of download to upload is also essential. A minimum is necessary to allow the download to achieve its full capacity, which is typically around 20% of download rates.

The growth in cloud-based storage systems (OneDrive, Google Drive, DropBox, etc.) and services that synch content across devices (iCloud Photos, Google Photos, etc.) along with security cameras with online storage and control, the upload path has much greater demands on it than in the past, and the trend is growing. Also, emerging tech such as VR and telehealth demand more uplink, specifically very low-latency uplinks, to work as expected.

[36] https://isif.asia/infrastructure-ipv6/ Is one example of an IPv6 upgrade grant process.

Therefore, we recommend a baseline ratio as close to 5:1 for download vs. upload capacities as the technology allows and a requirement that the upload has an AQM enabled to ensure low latency. In the long term, a goal of a 2:1 ratio would be good.

We recognize that some technologies are limited in what ratios they can achieve. Still, the goal should be to maintain sufficient upload to allow downloads to reach their max and have additional capacity and quality for the use cases listed.

5. Conclusion

Network bandwidth (aka speed) is merely one dimension of end-user quality of experience on the Internet. While it's an important element of that experience, the FCC should begin to focus on working latency as a critical aspect of end-user quality that is even more important than bandwidth. Today's applications are not typically bandwidth-limited - and are often 'adaptive' to varying bandwidth - but are limited significantly by working latency.

Thus, the FCC should prioritize network latency. This is the most critical next step needed to ensure that end-users are well-positioned to succeed in the digital age. Without addressing latency, constraints will be placed on the development & adoption of augmented reality, virtual reality, cloud gaming, video conferencing, work at home, remote learning, and other interactive applications. Focusing solely on bandwidth is essentially focusing on a service attribute that is necessary but insufficient. The test tools and data sets of the early Internet no longer represent the needs or actual demands on the Internet by normal users.

Congress and the FCC should not be considering raising the broadband standard beyond 100/20 mbps at this time without consideration towards the inherent latency under load and RTT of such networks, and the technology (including software) that already exists to mitigate it. With those taken care of, we indeed would have a better internet for all.

Sincerely,

Dave Taht CEO, Bufferbloat.net USA dave.taht@gmail.com	Jonathan Foulkes ex-CEO Evenroute (IQrouter) USA nobloat@jonathanfoulkes.com
David Collier-Brown, Canada. davecb@spamcop.net	Karl Auerbach CTO - InterWorking Labs USA (IWL - https://iwl.com/) - karl@iwl.com
Richard E Brown Chief Doc Guy, Bufferbloat.net USA richb.hanover@gmail.com	Leonard Kleinrock Professor, Computer Science, UCLA Internet Pioneer lk@cs.ucla.edu

Larry Press, lpress@csudh.edu	Frantisek (Frank) Borsik Slovakia frantisek.borsik@gmail.com
Sebastian Moeller Germany sebastian.moeller0@gmail.com	W. Daniel Ezell Owner/Curriculum Architect Chronos Academy, Larkspur, CA USA Daniel@chronos.academy
Bill Woodcock Executive Director, Packet Clearing House (USA) woody@pch.net	Robert Chacón CEO JackRabbit Wireless USA robert.chacon@jackrabbitwireless.com
Rohan Muscat Head of IT and Engineering, Flavorite group Australia Rohan.muscat@flavorite.com.au	Elijah Zeida elijah@airbridgebroadband.com Network Engineer AirBridge Broadband USA
Sam Mulvey, Station Manager KTQA-LP, USA, sam@ktqa.org	Larry Gadallah DevOps Engineer Impinj Inc. lgadallah@gmail.com
Brian Clapper CEO, ardentex bmc@ardentex.com	Oleg Kutkov Embedded systems and electronics engineer. AKA "Dr. Starlink" Ukraine contact@olegkutkov.me
Gary E. Miller Developer Rellim gem@rellim.com	James R. Maynard III, Embedded Systems and Network Engineer, Fairmont, MN jaymaynard@gmail.com
Perry E. Metzger Engineer and Computer Scientist perry@piermont.com	Ted Turner Cloud Solutions Architect, NC ted@turner.org
Stefan Hellermann Network Engineer, Germany mail@stefan-hellermann.de	Lognaden Velvindron Cyberstorm.mu logan@cyberstorm.mu
Mark Atwood, Principal Engineer, Seattle WA, mark.atwood@gmail.com	

Randolph Resnick, founder, Visions Under Construction and Resmo Communications, rmr@resmo.net	James Blessing, Head of Systems Development, Freedom Fibre UK james.blessing@freedomfibre.com
Sascha Meinrath Director, X-Lab Palmer Chair in Telecommunications Penn State University	Eric S. Raymond, co-founder and President Emeritus of the Open Source Initiative
Toke Høiland-Jørgensen "Dr WiFi" toke@redhat.com	Brough Turner Founder, netBlazr Inc, USA broughturner@gmail.com
Pete Heist, Test Engineer, Some Congestion Experienced Project, pete@heistp.net	William Allen Simpson Internet over dialup, cell phones, and fiber; Internet security; founded an ISP. william.allen.simpson@gmail.com
Tanya Weiman Internet User USA tanya@panacea42.com	Geoff Goodfellow ex-SRI & RadioMail.net USA geoff.goodfellow@iconia.com
Christopher Hawker Network Engineer, Sydney Australia chris@thesysadmin.au	Steven Stolen stolens@acm.org
Mark Steckel President PhillyWisper.net USA	James Troutman, co-founder & Director, Northern New England Neutral Internet Exchange (NNENIX), USA, jlt@nnenix.net
William Strang, strangwp@gmail.com Darrell Budic OHG Networks, LLC USA Internet Services & Consulting budic@ohgnetworks.com	Jianping Pan University of Victoria, Canada pan@uvic.ca
Robert Oliver Dassault Systèmes robert@olivers.us	Eric Miller, Principal Network Architect GiGstreem eric.miller@gigstreem.com
Ant Pruitt bufferbloattofcc@antpruitt.com	David Woodhouse Principal Engineer, Cambridge. dwmw2@infradead.org

Catherine O. Raymond, Esq. USA Longtime Internet user cathyr19355@gmail.com	Zachary Biles Owner, Northern Wireless Minnesota, USA zach.biles@northern-wireless.com
Daniel Marks May Mobility, Inc Ann Arbor, Michigan daniel.marks@maymobility.com	Jinwei Zhao University of Victoria, Canada clarkzjw@uvic.ca
Scott Miller Marcom Solutions LLC Orcas, WA USA scottmi@gmail.com	Alistair Woodman NetDEF, Inc San Jose, CA USA awoodman@netdef.org
Dennis Hock Retired Telecom Network Engr Casco, MI USA hockde47263@gmail.com	Tracy Rosenberg Executive Director Media Alliance tracy@media-alliance.org
Sudheesh Singanamalla, Internet Security and Measurement Researcher, Ph.D. Candidate Paul G. Allen School of Computer Science and Engineering University of Washington, Seattle, WA, USA sudheesh@cs.washington.edu	Goran Mahovlić
Lachlan Chapman Owner/Founder Aonet Broadband NZ Lachlan@aonet.nz	Spencer Sevilla, spencer.builds.networks@gmail.com
Ike Nassi Adjunct Professor Computer Science and Engineering, UC Santa Cruz. Software Systems Architect, frmr SVP Apple, EVP SAP, + 4 startups, nassi@nassi.com	Thomas M. Croghan Network Engineer Podcast Host "The Brothers WISP" USA tcroghan@lostcreek.tech
Mike Hammett CTO DNA Communications USA mhammett@dna-communications.com	Matthias Strubel bisQuITbox GmbH, Idstein, Germany m.strubel@bisquitbox.de
Paul Holcomb Network Engineer, FOSS Developer / Trainer CPOINT US locations: (Chicago, O‘ahu) mail@noptys.com	Daniel Denson Montana Express Networks Montana, USA dan@densontech.com

Nitinder Mohan Senior Researcher,	Technical University of Munich, Germany mohan@in.tum.de www.nitindermohan.com

APPENDIX A - Some example experiments

Experiment 1 - Injecting Latency

If there is any one issue about actual internet performance that the authors would like more deeply understood, worldwide, it is how latency affects the internet regardless of the amount of bandwidth available. Remarkably there are few tools available to most that can readily show the impact of added latency (delay) to a given connection. In the Linux world, delay is readily added to a given connection via the “netem” tool. A simple test would include setting the rate below the provided rate by the ISP, (so as to take control of the queue) and add delay.

tc qdisc add dev your_device root netem rate 18mbit delay 100ms

Then try a variety of real traffic types, especially web, voip, videoconferencing, and gaming. Only the user experience for video streaming will be largely unaffected by a 100ms delay.
Try 1000ms!

Apple includes in their Developer tools, the “Network Link Conditioner”[37], which also allows you to set delay like this. Android has tools for this too, see “Throttly”, for one.

While perhaps not as eye opening as Richard Feyman’s demonstration of the real cause of the Challenger accident, being able to quickly and reliably demonstrate the damage of network delay to yourself… or to a person claiming that bandwidth is everything, is useful!

Experiment 2 - Bufferbloat and Latency under Load Tests

There are many bufferbloat tests available that end users can run. In fact, over the past few years, every “speedtest” site is now reporting on it in some way - in rough order of authors’ preference there are the waveform bufferbloat test, the Ookla speedtest.net application, cloudflare’s speed.cloudflare.com, the long version of fast.com.

The concluding figure reported by most of these is a median, average, 75th percentile or taken from the later part of the test, which obscures the true impact of the jitter induced by every transfer that lasts long enough. Observe the numbers jumping around during the test! Few of these tests run long enough to do rigorous testing of something else while they are running. Several allow you to set them to run for longer than 20 seconds, optionally.

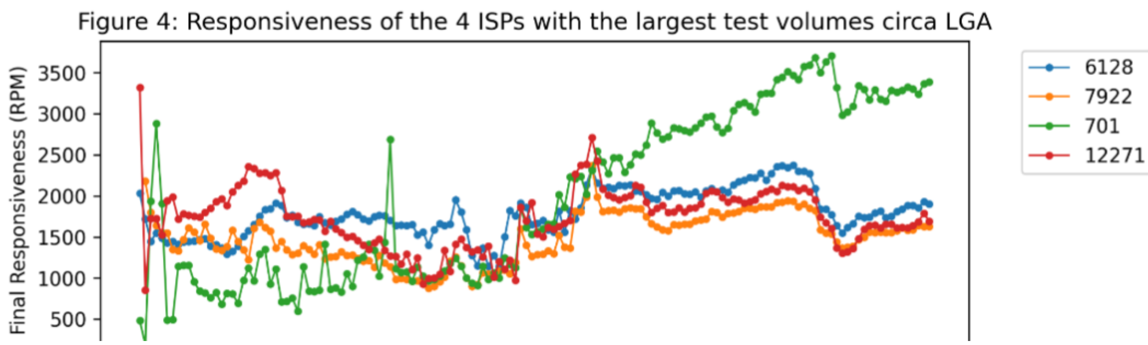
There are numerous widely available tools such as “iperf2” and “flent” that can generate more detailed tests and plots. The key thing to do during any tests of this sort, is put the load on the system, and then attempt to do any form of interactive traffic at the same time.

[37] https://tarikdahic.com/posts/using-network-link-conditioner-to-simulate-bad-network-conditions-on-ios-and-macos/

APPENDIX B BEFORE/AFTER COMPARISONS

1) DOCSIS-PIE Deployment

This independently produced chart from the m-labs dataset shows the improvement in responsiveness from Comcast's deployment of DOCSIS-PIE in the NYC area.[38]
The chart shows responsiveness (measured in round-trips-per-minute, higher is better) increasing from 2010 at the left edge of the chart to 2022 at the right. Comcast further published in-depth study in 2021[39] of a direct field comparison between a non-AQM'd modem and an AQM'd one, where the algorithm was not working correctly on one brand of device and they had a golden opportunity to do science.

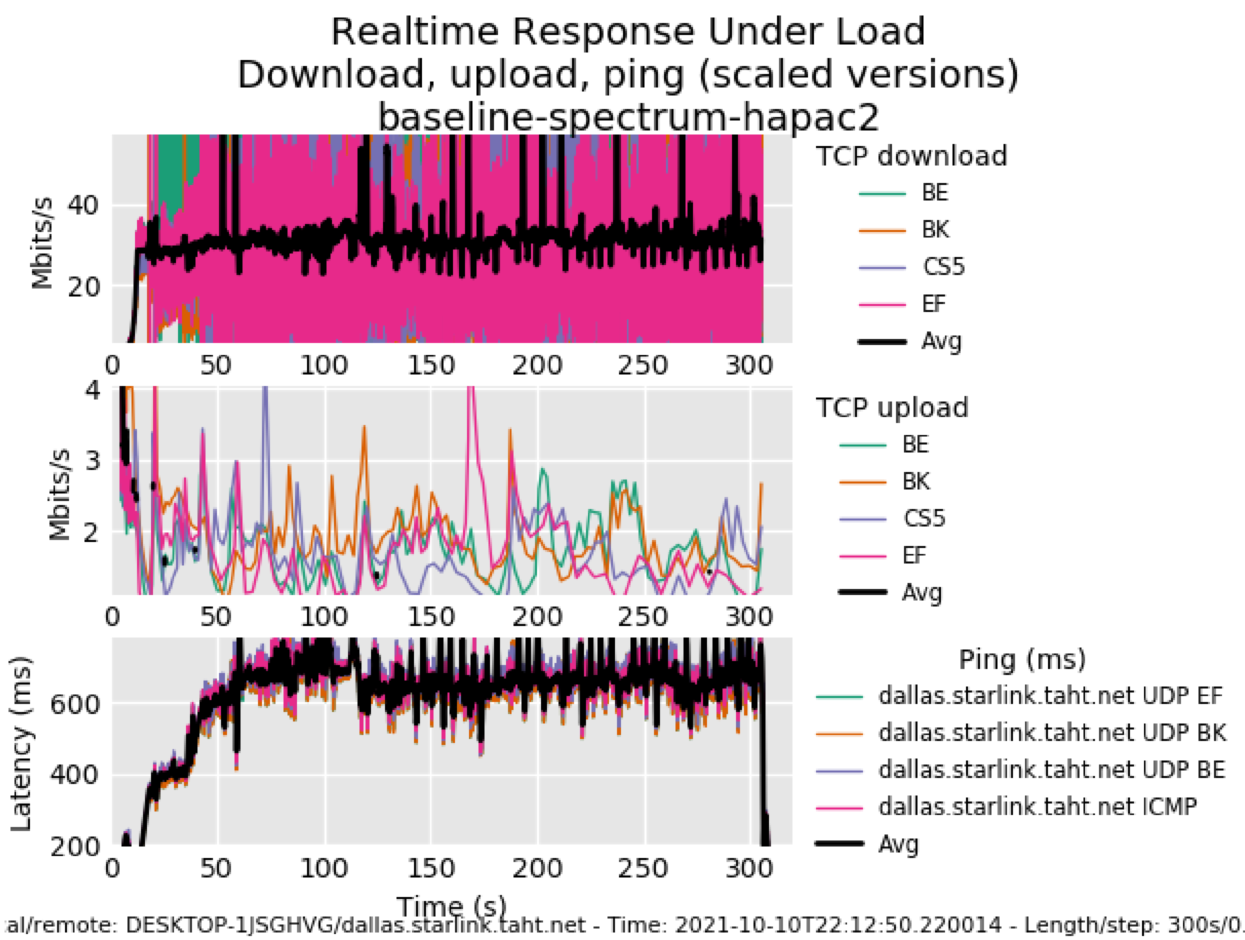


38 https://datatracker.ietf.org/meeting/interim-2021-mnqeuws-01/materials/slides-interim-2021-mnqeuws-01-sessa-matt-mathis-preliminary-longitudinal-study-of-internet-responsiveness-00.pdf IETF IAB workshop preliminary study

39 https://www.reddit.com/r/HomeNetworking/comments/r7cfvg/comcast_delivers_90_working_latency_reductions_by/ Public comments on the Comcast study of PIE vs non-PIE AQM

2) **A DOCSIS cable modem without bufferbloat mitigations (Spectrum Cable, 2022, 100/20mbps service)[40] using RRUL**

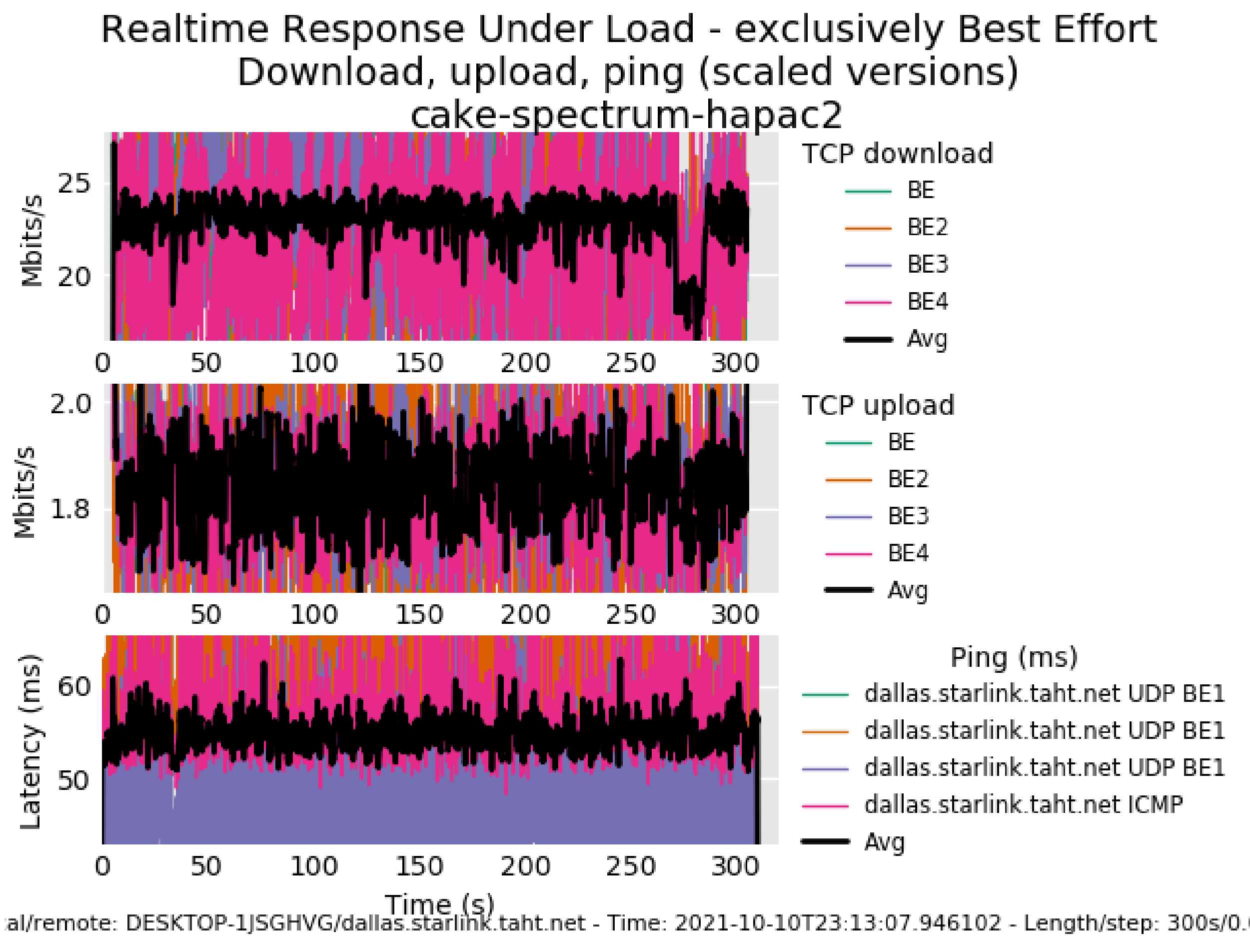
The RRUL (Real-time Response Under Load) Test is part of the Flent test suite. It provides a more stringent test than most current test regimes. RRUL tests a network's ability to manage upload and download simultaneously (60 seconds is the default) by creating multiple parallel download and upload connections (default, four) in each direction and measuring latency. Each test creates three charts: Download speed, Upload speed, and Latency. The black lines are the average for each.
Observations:
* The four download connections averaged about 25mbps each, using the full 100mbps
* The four upload connection speeds (shown by the multicolored plots) were wildly chaotic. They all fall below the expected 5mbps per connection.
* The latency climbed under this workload to over 700ms, rendering the network almost completely unusable for new web page downloads, voip, or videoconferencing, while the up + downloads are still running. Even a power user would be very inconvenienced by this behavior.



[40] https://forum.mikrotik.com/viewtopic.php?t=179307 Configuring cake for mikrotik

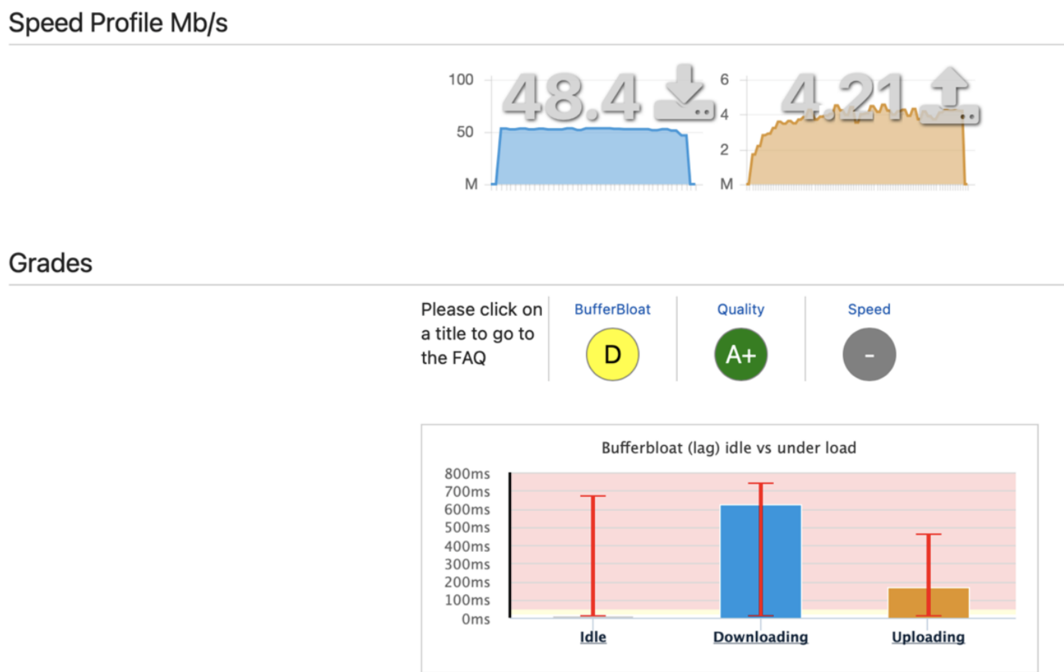
3) **The RRUL test was used with a third party device, Mikrotik with CAKE - derived from RFC8290.** The charts show:
* Average download is somewhat below 25mbps, its expected share of a 100 mbps link
* Average upload is about 2mbps
* Most importantly, the latency is one-tenth (60msec vs 700msec) of the previous test.
This transforms a 100/20 link from unusable to quite responsive.

Smart users with better gear from pfsense, Openwrt, Riverbed, Evenroute, Ubiquti, Merak, Mikrotik, and many more have been applying techniques like this for a decade. Although this is free (and simply better) software, techniques like this are only beginning to roll out in ISP supplied equipment. It would be interesting to learn how many power users have invested the time to implement these FQ and AQM technologies.

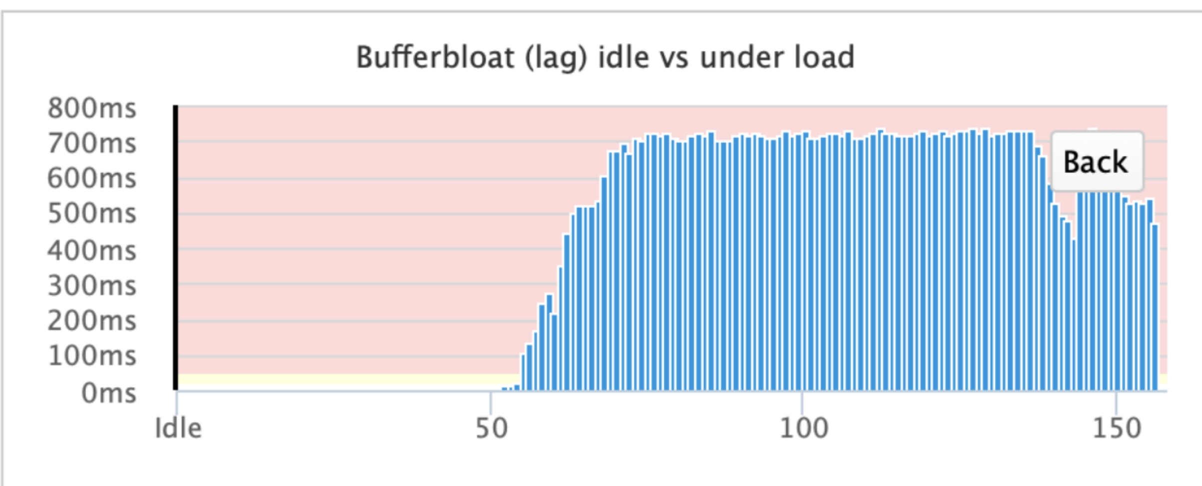


4) Testing a DOCSIS 3.0 modem with 50 / 5 Mbps service using dslreports.net

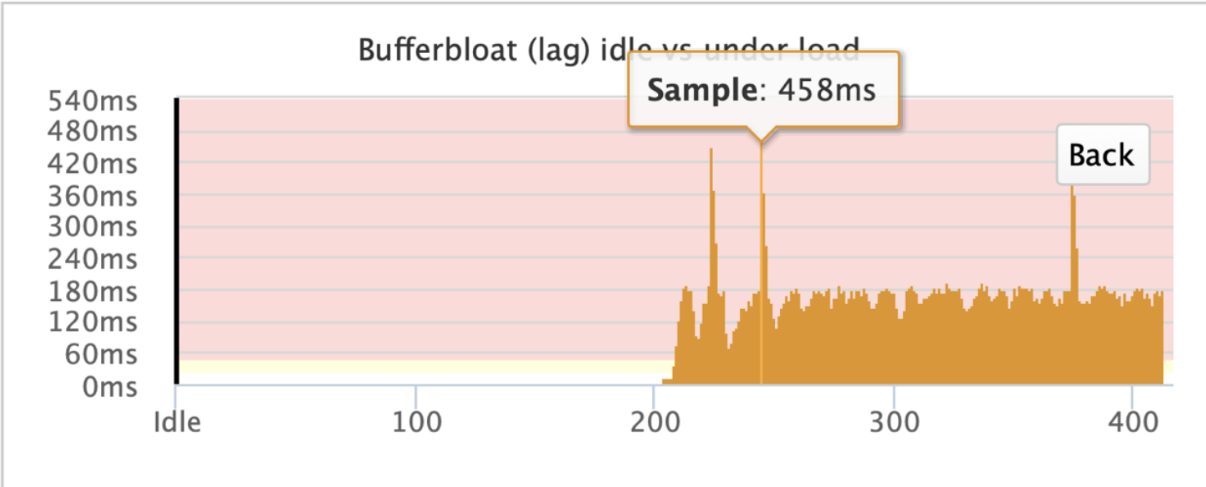
This exhibits serious bufferbloat directly from the modem, with an average download latency of >500ms, and very poor upload latency as well:



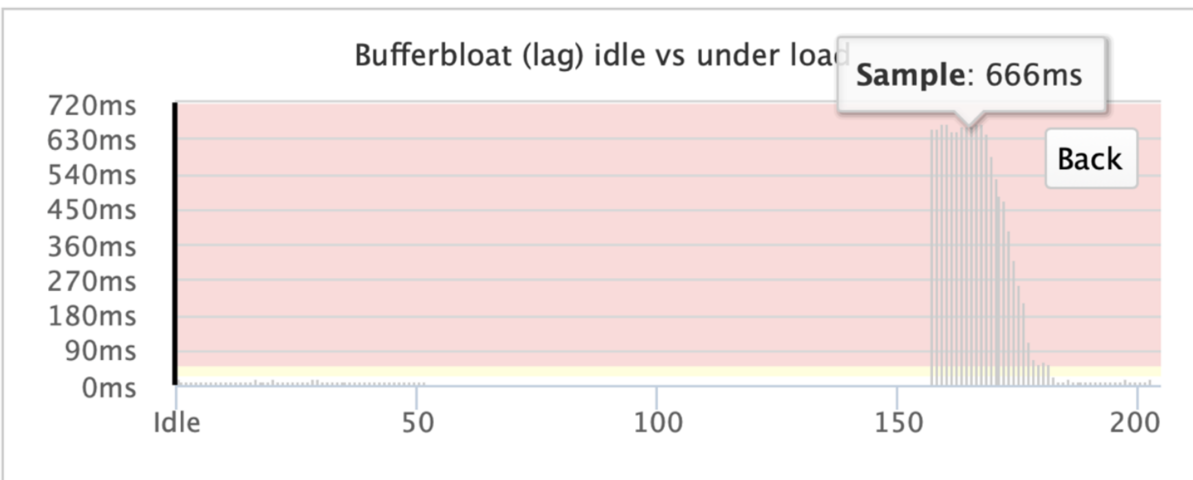
This detailed view of latencies building up during the download test phase illustrates the bufferbloat phenomenon well, as we see them quickly climb to over 700ms and stay there for most of the duration:



Upload latency detail is likewise bad, with spikes approaching half a second:

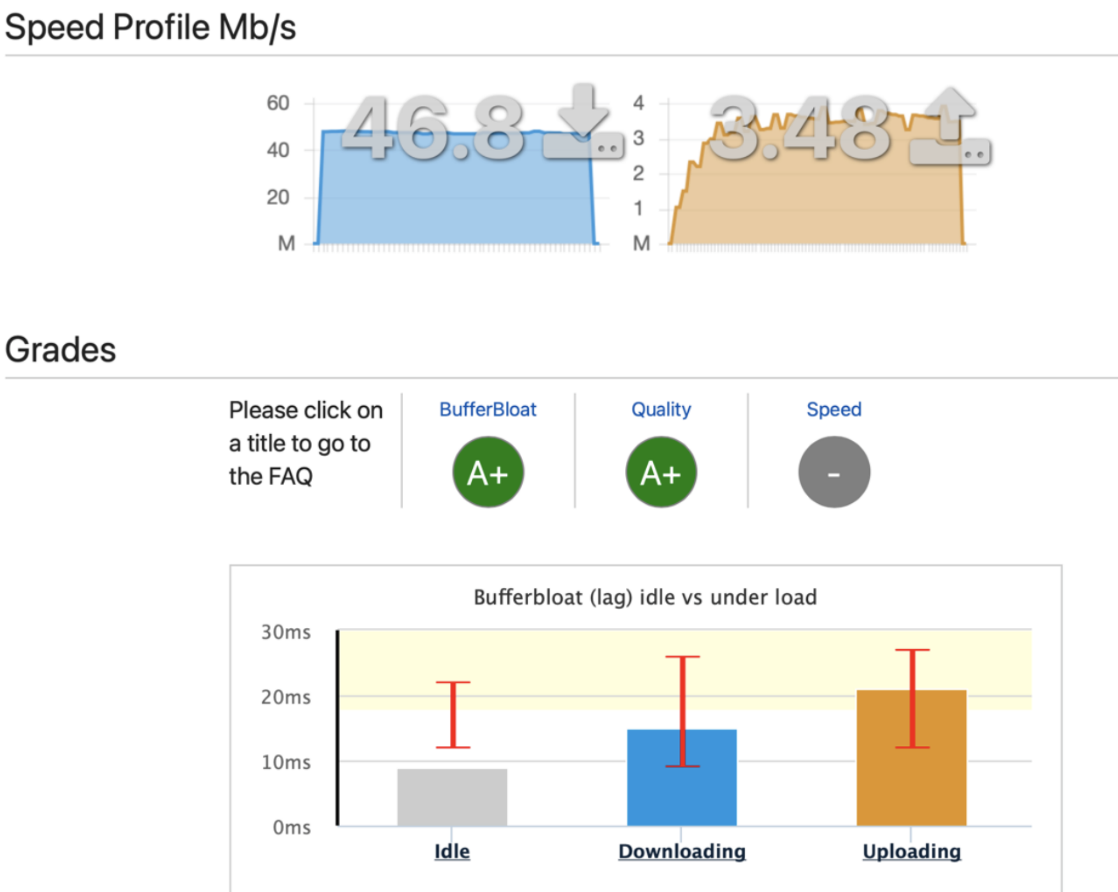


But to truly see how bad the upload lag really is, we look at the post-test idle time latencies, and we see that the buffers are so overrun, we have over a second of >600ms delays on upload. This destroys the usability of many interactive applications.



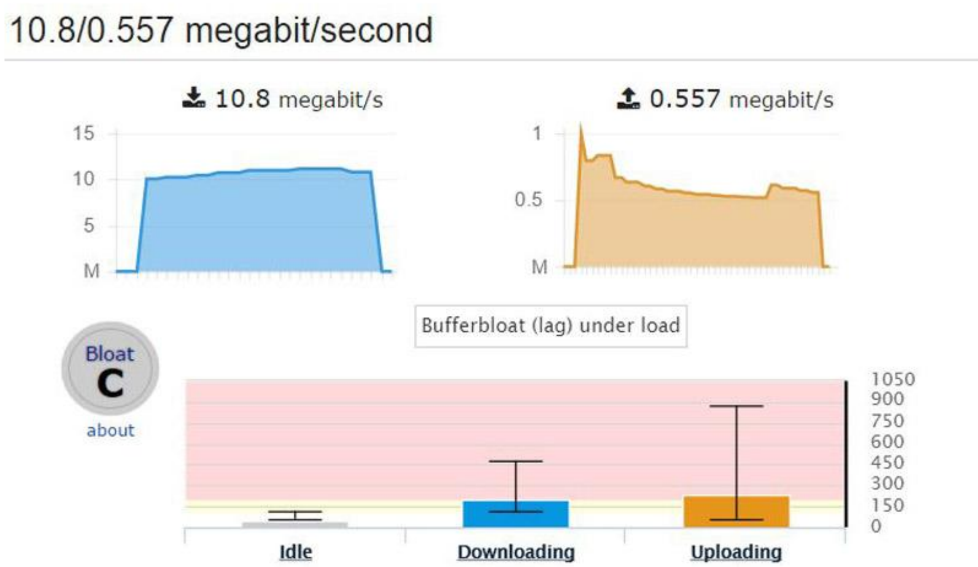
The same DOCSIS modem, supplemented with a router with SQM enabled.

The same dslreports.net test facility using a router with SQM enabled shows latency is well-controlled - well below 30ms in this case, with even with slightly lower transfer rates.

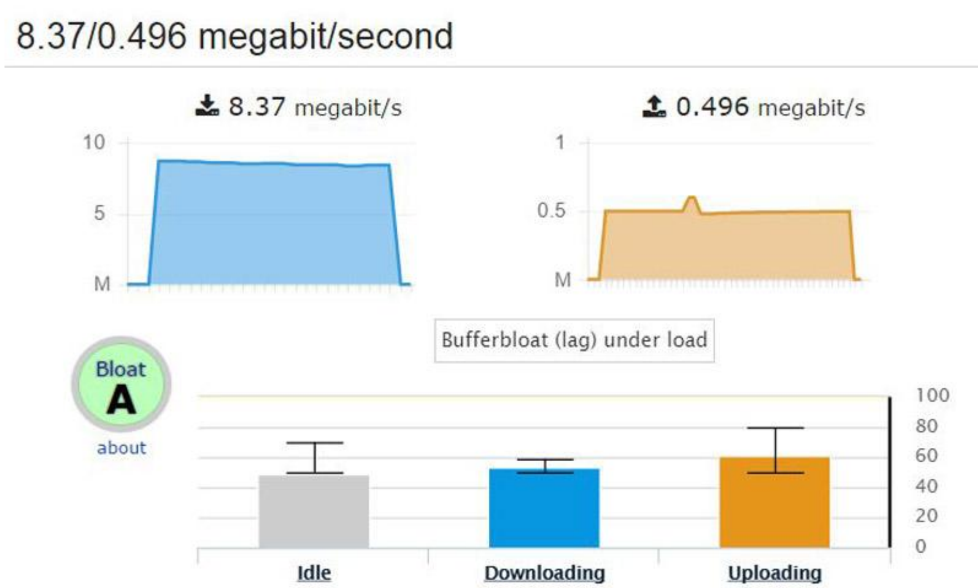


4) Testing a 10Mbps/768kbps DSL line with dslreports.net

This first test shows >800ms of lag on the upload, with an average close to 200ms, this line struggled with any form of interactive applications. The ISP-provided modem/router had much worse results, hitting over a second of delay on upload and scoring an F for Bufferbloat (not depicted).

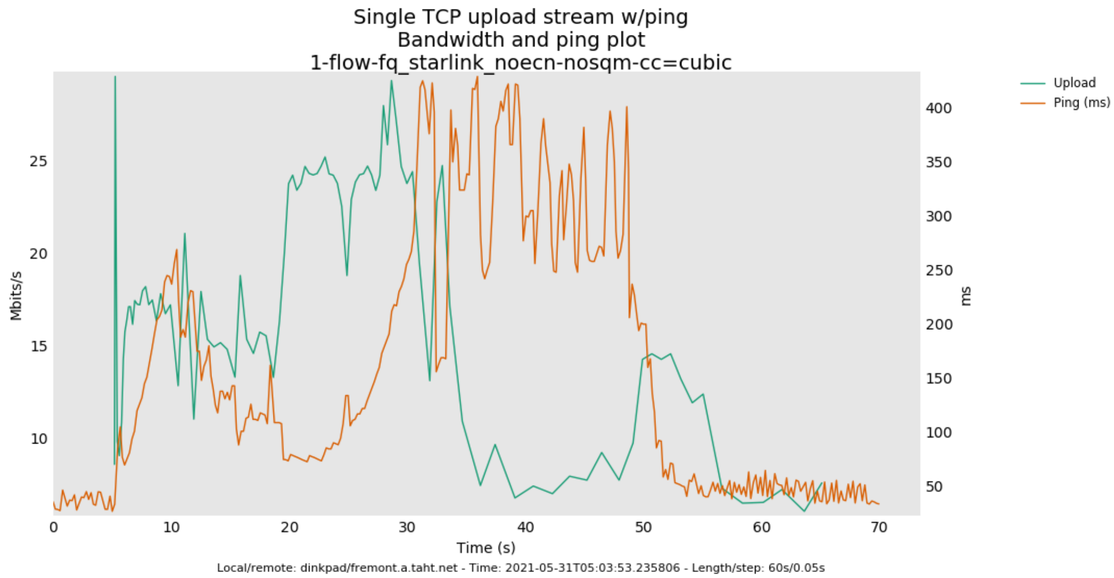


A second test, after enabling SQM on the router, now shows an A for bloat, as latencies under load are similar to the idle latency, and variation in latencies is minimal. Although the achieved speeds (download and upload) are slightly decreased, interactive applications are now very responsive, even though capacity is small.



5) Plot of Starlink's upload behaviors as of May 2021[41] (it has improved since).

Most speedtest facilities test for less than 20 seconds. Because Starlink updates its bandwidth every 15s…or any measurement interval here, what is the average bandwidth? What is the average latency?



Note that this is not Starlink's malfeasance but their best effort to compensate for orbital dynamics and offered load. Under most circumstances starlink seems to be doing a better job in compensating for variable bandwidth than 5G, and under most circumstances web page loading is fine. They can do better….

[41] https://docs.google.com/document/d/1puRjUVxJ6cCv-rgQ_zn-jWZU9ae0jZbFATLf4PQKblM/edit

APPENDIX C - Network Neutrality history

The concept of Network Neutrality dates back to Tim Bray's 2003 book defining the phrase. The internet at the time, however, was completely incapable of meeting a legal definition of common carriage for all traffic types, yet NN started becoming an issue with the rise of the Bittorrent protocols, which used up all bandwidth both up and down at the residence, and exposed problems with overbuffering, losing the room to also carry VOIP. To say this problem was ill-understood at the time, is an understatement. "Mistakes were made" in finding ways to cope with it, while outside that bubble, more mistakes were made in overbuffering all of the next generation of networking gear, often by multiple orders of magnitude. It was not until 2010[42] when Jim Gettys coined the word "Bufferbloat", that progress began to be made on solving the underlying problem that was at the root of much network misbehavior.

APPENDIX D - Networks are complicated

Networks are complicated. People who know how to design and diagnose complex ones are scarce. A problem at any point in a network—at the server, across the internet on the backhaul, at the customer shaper, at the CPE, or the WiFi, degrades the user experience, and without users with experience…

"Thank YOU guys. I literally spent multiple days walking my friend through installing OpenWrt on a gl-inet router, configuring several APs, connecting up powerline networking equipment and configuring cake in 2 or 3 spots on his network (both at the router towards the WAN and across the powerline segments to keep them from saturating). His wife was teaching anatomy classes, needed a camera on her face, a camera on her desk where she was showing images and drawing diagrams, and I think a camera on a microscope. All that just for her. They had something like a 600/60 DOCSIS connection, and in addition to all the interactive teaching she was doing, he was in video meetings all day and both their kids in online school, plus they were living with his parents during the stay-at-home period. so there were maybe 6 or so devices video streaming at any given moment throughout the day. The full download capacity of that internet connection was close to 10x the powerline speed... Expanding to 3 APs, ethernet over powerline, and cake to limit all the bottlenecks including over the powerline (which at best was getting something like 70Mbps reliably...) with diffserv4 and prioritization and network segmentation and all that stuff... they could seamlessly walk through the house roaming from AP to AP with everyone getting 0 latency issues after a week of configuration.... Before all that work no-one could understand any of her lectures at all and she was literally in danger of losing her job." [43]

---

[42] https://gettys.wordpress.com/2010/12/07/bufferbloat-and-network-neutrality-back-to-the-past/
[43] https://forum.openwrt.org/t/nftables-and-qos-in-2021/112013/530 OpenWrt Support forum

The sadness we feel for every teacher in every school that needed to summon technical resources like this to do their job during COVID, and couldn’t, is immense. Crossing the Digital Divide requires building bridges over multiple hurdles in education and expertise. Standards calling for recent software meeting modern requirements on home routers could greatly reduce this kind of hurdle.

GLOSSARY

BUFFERBLOAT - Bufferbloat is the undesirable latency that comes from a router or other network equipment buffering too much data. This buffering, meant to smooth out traffic, instead leads to increased delays if not managed carefully.

LATENCY - Latency is the delay (generally measured in milliseconds) between a user's action and the network's response. The terms “latency” or “Lag” is typically used as a proxy for round-trip time or ping times. Shorter is always better.

RESPONSIVENESS - A measure of network quality: a more responsive network will work and “feel” better. Responsiveness is generally measured as “round-trips per minute”. Responsiveness and latency are reciprocal measures: a higher number is better.

LATENCY UNDER LOAD - The latency of an un-loaded (“idle”) link is typically quite low - a small number of milliseconds. But if the buffering of any of the devices in the path between the endpoints is not well controlled, latency during upload or downloads can increase ten-fold or 100-fold. This is “latency under load”, generally caused by bufferbloat.

BANDWIDTH, SPEED, DATA RATE, CAPACITY – These are used interchangeably for the rate data is transferred in bits per second, (megabits/second, etc.)