The quiet plan to make the internet feel faster

A few months ago, I downgraded my internet, going from a 900Mbps plan to a 200Mbps one. Now, I find that websites can sometimes take a painfully long time to load, that HD YouTube videos have to stop and buffer when I jump around in them, and that video calls can be annoyingly choppy.

In other words, pretty much nothing has changed. I had those exact same problems even when I had near-gigabit download service, and I'm probably not alone. I'm sure many of you have also had the experience of cursing a slow-loading website and growing even more confused when a "speed test" says that your internet should be able to play dozens of 4K Netflix streams at once. So what gives?

Like any issue, there are many factors at play. But a major one is latency, or the amount of time it takes for your device to send data to a server and get data back — it doesn't matter how much bandwidth you have if your packets (the little bundles of data that travel over the network) are getting stuck somewhere. But while people have some idea about how latency works thanks to popular speed tests, including a "ping" metric, common methods of measuring it haven't always provided a complete picture.

The good news is that there's a plan to almost eliminate latency, and big companies like Apple, Google, Comcast, Charter, Nvidia, Valve, Nokia, Ericsson, T-Mobile parent company Deutsche Telekom, and more have shown an interest. It's a new internet standard called L4S that was finalized and published in January, and it could put a serious dent in the amount of time we spend waiting around for webpages or streams

to load and cut down on glitches in video calls. It could also help change the way we think about internet speed and help developers create applications that just aren't possible with the current realities of the internet.

Before we talk about L4S, though, we should lay some groundwork.

a:hover]:shadow-highlight-franklin dark:[&>a:hover]:shadow-highlight-franklin [&>a]:shadow-underline-black dark:[&>a]:shadow-underline-white">That sounds like a zombie enemy from *The Last Of Us*

Right? But to understand what buffer bloat really is, we first have to understand what buffers are. As we've touched on already, networking is a bit of a dance; each part of the network (such as switches, routers, modems, etc.) has its own limit on how much data it can handle. But because the devices that are on the network and how much traffic they have to deal with is constantly changing, none of our phones or computers really know how much data to send at a time.

To figure that out, they'll generally start sending data at one rate. If everything goes well, they'll increase it again and again until something goes wrong. Traditionally, that thing going wrong is packets being dropped; a router somewhere receives data faster than it can send it out and says, "Oh no, I can't handle this right now," and just gets rid of it. Very relatable.

While packets being dropped doesn't generally result in data loss — we've made sure computers are smart enough to just send those packets again, if necessary — it's still definitely not ideal. So the sender gets the message that packets have been

dropped and temporarily scales back how its data rates before immediately ramping up again just in case things have changed within the past few milliseconds.

That’s because sometimes the data overload that causes packets to drop is just temporary; maybe someone on your network is trying to send a picture on Discord, and if your router could just hold on until that goes through, you could continue your video call with no issues. That’s also one of the reasons why lots of networking equipment has buffers built in. If a device gets too many packets at once, it can temporarily store them, putting them in a queue to get sent out. This lets systems handle massive amounts of data and smooths out bursts of traffic that could have otherwise caused problems.

a:hover]:shadow-highlight-franklin dark:[&>a:hover]:shadow-highlight-franklin [&>a]:shadow-underline-black dark:[&>a]:shadow-underline-white”>I’m paying for gigabit internet — how could I still have latency issues?

This is the trick of internet service provider, or ISP, marketing. When users say they want “faster” internet, what they mean is that they want there to be less time from when they ask for something to when they get it. However, internet providers sell connections by capacity: how much data can you suck back at once?

a:hover]:shadow-highlight-franklin dark:[&>a:hover]:shadow-highlight-franklin [&>a]:shadow-underline-black dark:[&>a]:shadow-underline-white”>Bit versus byte
Talking about the amount of time it takes to download files brings up another problem with how internet services are marketed. Usually, we think of file sizes in terms of bytes — a song is 10 megabytes, and a movie is 25 giga*bytes.* But ISPs

rate connections in *bits*.

If you miss the distinction, you'd be forgiven for thinking that a service plan that gives you a gigabit per second would let you download a movie in 25 seconds. However, bits are eight times smaller than bytes – one giga*bit* (Gb) is equivalent to 125 mega*bytes* (MB), or 0.125 gigabytes (GB). So that movie is going to take over three minutes to download instead, assuming perfect conditions.

(By the way, give yourself a prize if you realized that a lowercase b versus an uppercase one is how you distinguish between the two units in their abbreviated forms.)

There was a time when adding capacity really did reduce the amount of time you spent waiting around. If you're downloading a nine-megabyte MP3 file from a *totally legal* website, it's going to take a long time on 56 kilobit per second dial-up – around 21 and a half minutes. Upgrade to a blazing-fast 10Mbps connection, and you should have the song in less than 10 seconds.

But the time it takes to transfer data gets less and less noticeable as the throughput goes up; you wouldn't notice the difference between a song download that takes 0.72 seconds on 100Mbps and one that takes 0.288 seconds on 250Mbps, even though it's technically less than half the time. (Also, in reality, it takes longer than that because the process of downloading a song doesn't *just* involve transferring the data). The numbers matter a bit more when you're downloading larger files, but you still hit diminishing returns at some point; the difference between streaming a 4K movie 30 times faster than you can watch it versus five times faster than you can watch it isn't particularly important.

The disconnect between our internet "speed" (usually what people are referring to is throughput – the question is less about how fast the delivery truck is going and more about how

much it can carry on the trip) and how we experience those high-bandwidth connections becomes apparent when simple webpages are slow to load; in theory, we should be able to load text, images, and javascript at lightning speeds. However, loading a webpage means several rounds of back-and-forth communication between our devices and servers, so latency issues get multiplied. Packets getting stuck for 25 milliseconds can really add up when they have to make the journey 10 or 20 times. The amount of data we can move at one time through our internet connection isn't the bottleneck — it's the time our packets spend shuffling between devices. So, adding more capacity isn't going to help.

a:hover]:text-black [&>a:hover]:shadow-underline-black dark:[&>a:hover]:text-gray-e9 dark:[&>a:hover]:shadow-underline-gray-63 [&>a]:shadow-underline-gray-13 dark:[&>a]:shadow-underline-gray-63″>*Apple tested L4S on a typical network and saw a massive improvement to round-trip traffic times. More on that later.*

a:hover]:shadow-highlight-franklin dark:[&>a:hover]:shadow-highlight-franklin [&>a]:shadow-underline-black dark:[&>a]:shadow-underline-white”>Okay, but how does it do that? Is it magic?

No, it's not magic, though it's technically complex enough that I kind of wish it were, because then, I could just hand-wave it away. If you *really* want to get into it (and you know a lot about networking), you can read the specification paper

on the Internet Engineering Task Force’s website.

L4S lets the packets tell your device how well their journey went

For everyone else, I’ll try to boil it down as much as I can without glossing over too much. The L4S standard adds an indicator to packets, which says whether they experienced congestion on their journey from one device to another. If they sail right on through, there’s no problem, and nothing happens. But if they have to wait in a queue for more than a specified amount of time, they get marked as having experienced congestion. That way, the devices can start making adjustments immediately to keep the congestion from getting worse and to potentially eliminate it altogether. That keeps the data flowing as fast as it possibly can and gets rid of the disruptions and mitigations that can add latency with other systems.

a:hover]:shadow-highlight-franklin dark:[&>a:hover]:shadow-highlight-franklin [&>a]:shadow-underline-black dark:[&>a]:shadow-underline-white”>What can’t L4S do?

Well, it can’t bend the laws of physics. Data can only travel so fast, and sometimes it has to go a long way. As an example, if I were trying to do a video call with someone in Perth, Australia, there would be, at the *very* least, 51ms of latency each way – that’s how much time light takes to travel in a straight line from where I live to there, assuming it’s going through a vacuum. Realistically, it’ll take a bit longer. Light travels a bit slower through fiber optic cables, and the data would be taking a few extra hops along the path, as there isn’t actually a direct line from my house to Perth, as far as I’m aware.

This is why most services that aren’t dealing with real-time data will try to cache it closer to where you live. If you’re watching something popular on Netflix or YouTube, chances are you’re getting that data from a server relatively close to where you live, even if that’s not anywhere close to those companies’ main data centers.

There’s nothing L4S can do about that physical lag. However, it *could* keep much additional lag from being added on top of that.

a:hover]:text-black [&>a:hover]:shadow-underline-black dark:[&>a:hover]:text-gray-e9 dark:[&>a:hover]:shadow-underline-gray-63 [&>a]:shadow-underline-gray-13 dark:[&>a]:shadow-underline-gray-63″>*How to turn on L4S for testing on an iPhone.*

a:hover]:shadow-highlight-franklin dark:[&>a:hover]:shadow-highlight-franklin [&>a]:shadow-underline-black dark:[&>a]:shadow-underline-white”>How can I tell if L4S will make my internet better?

That’s a great question. The biggest indicator will be how much latency you’re already experiencing in everyday life. As I mentioned before, ping is sometimes used to measure latency, but just finding your average ping won’t necessarily tell you the whole story. What really matters is what your ping is when your network is taxed and what it spikes to.

Thankfully, some speed test apps are starting to show this

data. In May 2022, Ookla added a more realistic overview of latency to Speedtest, which is one of the most popular tools for seeing how fast your internet is. To see it, do a test, then tap “detailed result,” and look at the “responsiveness” section. When I did one, it told me my ping when pretty much nothing else was going on was 17, which seems pretty good. But during the download test, when I was actually using my connection, it spiked as high as 855 milliseconds — that’s almost an entire second, which would feel like an eternity if I were, say, waiting for a webpage to load, especially if it gets multiplied several times during the communication’s round trips.

(I invite anyone who’s used dial-up to tell me how soft I am and to reminisce about the days when every website took 10 seconds to load, uphill in the snow both ways.)

If you only ever do one thing on the internet at a time and use sites that barely anyone else uses, then maybe L4S won’t do much for you if and when it finally arrives. But that’s not a realistic scenario. If we can get the tech onto our increasingly busy home networks that we use to visit the same sites as everyone else, there’s a possibility it could be a quiet revolution in the user experience of the web. And once most people have it, people can start developing apps that couldn’t exist without ultra-low latency.

https://www.theverge.com/23655762/l4s-internet-apple-comcast-latency-speed-bandwidth