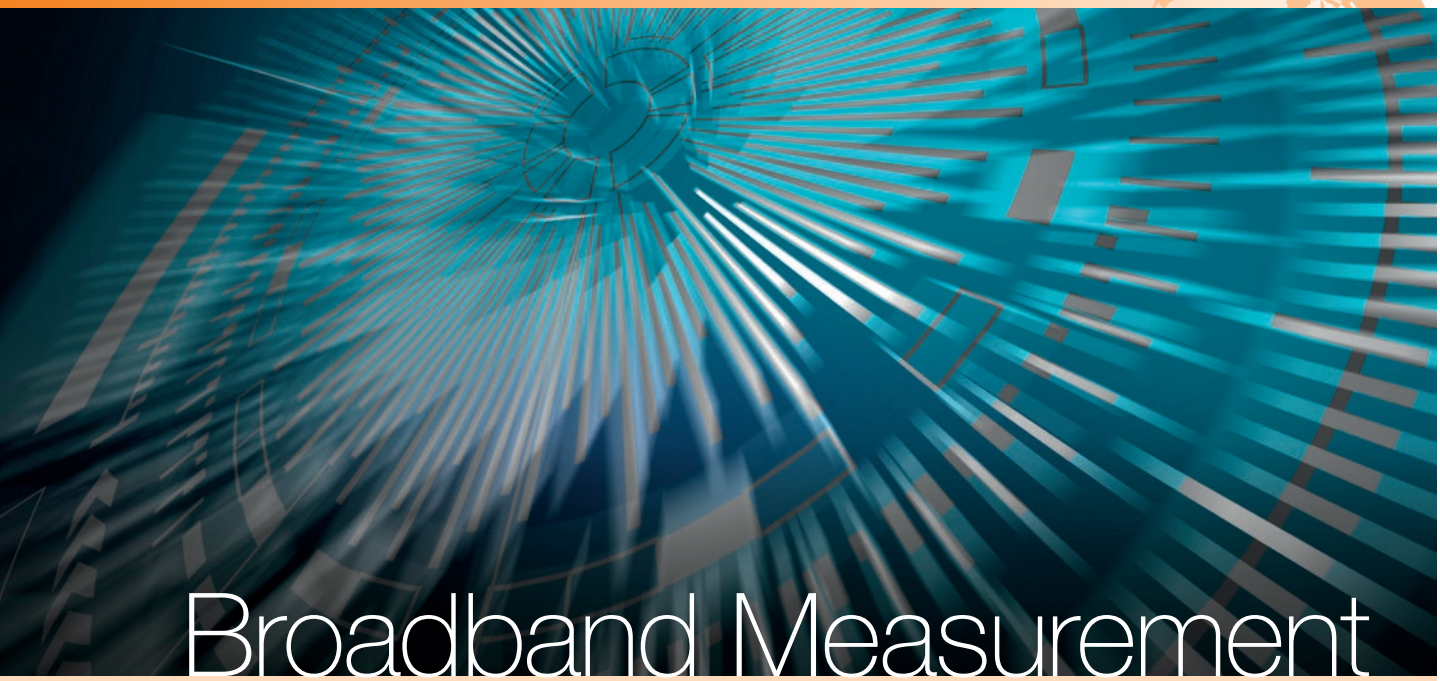

Broadband Measurement

needs to Shift from Speed to Experience

by Canopus Networks

Introduction

Speed is etched into our psyche – consumers like to buy broadband services from the "fastest" provider; there is a whole industry of speed testing services; governments invest millions of dollars in national speed monitoring programmes; and Internet Service Providers (ISPs) use speed test rankings in their advertising. We contend that with broadband speeds now approaching 100Mbps and beyond, speed is becoming an increasingly poor proxy for user experience.

We therefore need to evolve our thinking to put new regimes in place that measure user experience directly, rather than using speed as a proxy. Without this, investments will be misplaced, infrastructure will be tuned sub-optimally, pricing will not reflect value, and users will be frustrated. Luckily, new platforms are emerging in the market that use AI technology to measure user experience on streaming, gaming, conferencing and other applications accurately, cost-effectively and at scale.

Our misplaced obsession with speed

Access network speeds have grown by several orders of magnitude over the past three decades, from kilobits-per-second to gigabits-per-second. This has undoubtedly enabled users to enjoy streaming, gaming, conferencing and other online activities at far superior levels of experience than ever before. However, beyond a certain limit, user perception of speed starts saturating, as shown in Fig.1 below taken from

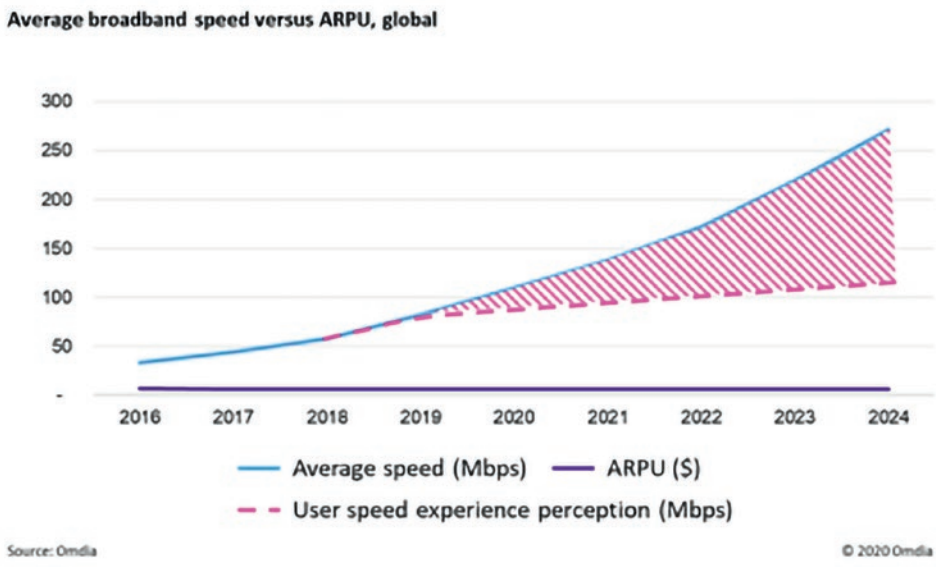

Fig. 1: The graph shows that for Mbps, increases above 100 have diminishing User Experience Perception returns. ARPU is not moving. How well spent is the money it costs to upgrade from 100Mbps to 300Mbps or 1000Mbps? (Source: Omdia)

[1], and higher speeds do not necessarily contribute to a better experience on most of the popular applications prevalent today. Nevertheless, the coupling between speed and experience is etched strongly into our psyche, and this blind faith leads consumers, governments and the industry at large to make decisions that are not always in the best interest.

To pick Australia as an example, speed is not lacking: around 73% of the National Broadband Network (NBN) services are at 50Mbps or higher today, which is sufficient to support 4K video, cloud gaming and virtual reality streams for most households. Interestingly, Australian consumers have collectively bought over 400Tbps on their speed plans, while actual peak traffic tops out at around 25Tbps, implying that only about a sixteenth of the speed paid for is used in aggregate. Consumption in terms of volume however has been steadily growing due to increasing time spent online and increasing network demands from applications — the average Australian household now consumes 450GB per month on their broadband plan. Our Internet usage now resembles a marathon runner gradually adding more and more miles to their training distances, rather than a sprinter reaching higher and higher top speeds.

Some of the inherent problems with using speed measurement as a proxy for user experience have been articulated in our research paper [2], and are summarised below:

- Speed tests are 10-second sprints which produce synthetic blasts of traffic, whereas most modern applications like streaming video, gaming, and conferencing regulate their behaviour. Measuring the former does not say much about the latter.
- Speed results are susceptible to test conditions such as server distance, test duration, TCP variant & threads, which can lead to biased results and wrong inferences.
- Speed tests are intrusive on broadband (especially wireless and high-speed) networks due to the spurts of high load they can impose on the contended infrastructure .
- Incentivising network operators to optimise their networks for speed test superiority (e.g. by tuning shaper buffers) can create perverse outcomes as it can degrade user experience for latency-sensitive applications such as gaming and conferencing.

Shifting focus to application experience

Broadband measurement needs to evolve to become meaningful, accurate and pervasive. Meaningful to the students attending online lessons over Zoom, Teams or other platforms; to the fans streaming live sport; to the gamers participating in eSport tournaments; and to those working from home; accurate so averaging does not smooth or wash out conference glitches, video spinning wheels and gaming lag spikes that may be last only a few seconds or less but can ruin the user experience; and pervasive so it can scale across millions of broadband services, rather than drawing inferences from very small sample sizes.

Measurement needs to be performed for popular applications in their natural operating environment instead of under synthetic conditions, and via passive observation making it risk-free to network operations and without imposing additional network load. Further, this has to be done with high accuracy, capturing every sub-second degradation in each stream, and at national scale in a very cost-effective manner. The measurement regime should be able to assess the experience for every subscriber. In what follows we illustrate how experience can be captured and depicted meaningfully for the major application categories including video streaming and gaming.

Fig 2: (a) Compares ISPs by percentage of time delivering the maximum video streaming experience; (b) Compares ISPs in terms of the average number of seconds of freeze experienced per hour of video viewing.

Video Streaming Experience

A user streaming video, such as, Netflix, YouTube, Twitch or Prime cares about two things: (a) Is my video playing at the best possible resolution for my device? E.g. SD, HD, HDX/4k; and (b) Will my viewing session be free of any 'video freezes' (spinning wheels)?

Fig. 2(a) shows ISPs being compared in their ability to serve video at the highest possible resolution during busy hour. ISP3 is the highest serving 96.2%, and ISP5 the lowest serving 80.1% of streaming hours in the highest resolution. Fig. 2(b) shows how ISPs can be compared in video freeze events during busy hours. Again ISP3 is the best averaging 36.7 seconds of freeze per hour of streaming, and ISP5 is the worst at 68.0 seconds. These meaningful comparisons help users select the ISP to best suit their streaming needs.

Gaming Experience

Most online games are latency sensitive - speed is rarely an issue. A latency variation (also known as "lag spike" or "jitter") of a few tens of milliseconds in CS:GO, a shooting game, can lead to gunshots not taking effect, or "teleportation" in the game. Fig 3(a) shows how ISPs can be compared on 95 percentile jitter values for the CS:GO games played over their network. Fig. 3(b) shows how gaming experience across popular shooter games (CS:GO, Fortnite, CoD, Apex Legends) can be categorised as low, medium, or high and compared across ISPs. Similar comparisons can be done for other genres such as sports, role-play and strategy.

Measuring Experience in a Scalable, Accurate and Cost-Effective Manner

A new AI-based approach is emerging that is agnostic to encryption and can be implemented at terabit scale, leveraging the fact that applications have unique behaviour patterns on the network in the form of a "heartbeat" or "pulse".

As an illustration, consider the typical network behaviour of a Netflix video-on-demand stream in terms of bit-rate every 100ms (Fig. 4a): once the client playback buffer is filled, the stream periodically fetches a "chunk" of video, manifesting in the form of a network activity spike every 4-8 seconds, and remains

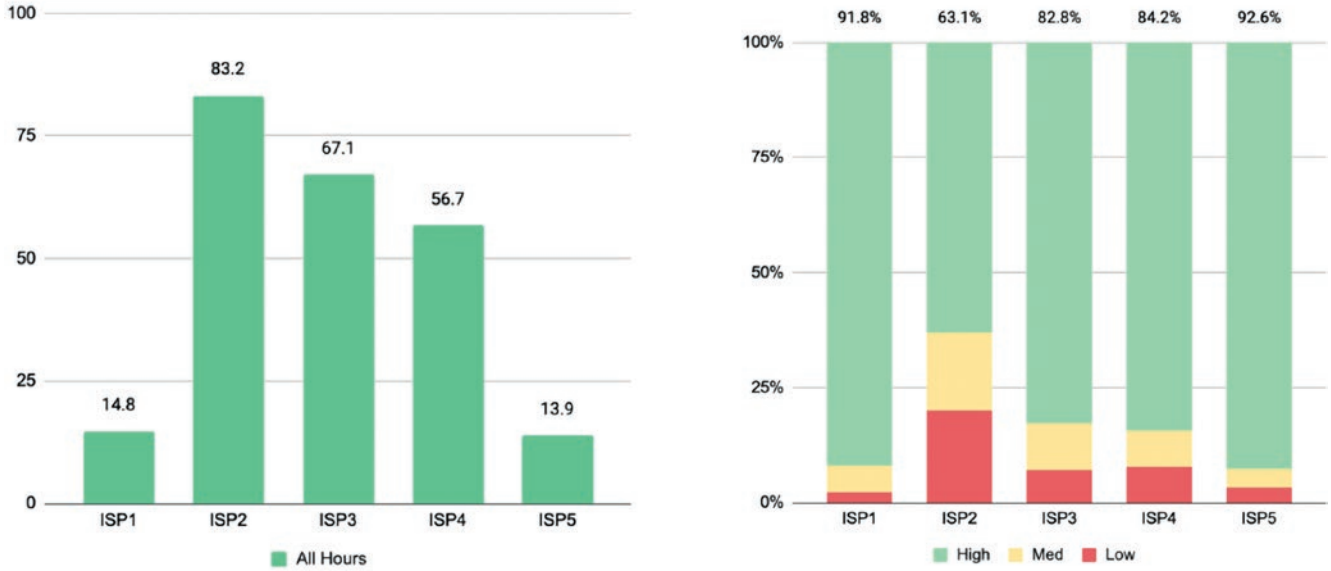

Fig. 3: (a) ISP comparison of 95%-ile jitter (ms) for CS:GO shooter game, ISP5 offers the most stable latency; (b) Gamer experience for shooter genre games. ISP5 offers the highest percentage of time delivering "high" gaming experience.

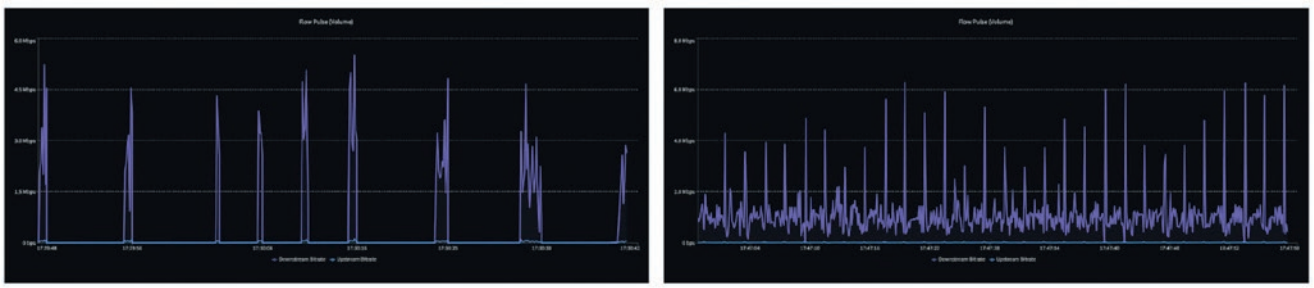

Fig 4: (a) "Pulse" of a Netflix on-demand video stream; (b) "pulse" of a Twitch live video stream

idle otherwise. In contrast, a Twitch live-stream (Fig. 4b) also fetches chunks, but typically at 2-second intervals, without any idle periods. We have trained AI engines [3] on these stochastic behavioural patterns to not only distinguish various traffic types (on-demand video, live video, gaming, conferencing, etc.), but also the various providers (e.g. Netflix from Disney+, Zoom from Teams, etc.) with high confidence. Further, by examining the pulse, such as the height, width, and spacing of the spikes, the AI engines can also deduce the quality of the on-demand [4] or live [5] video stream in terms of resolution and client buffer health. The methods have been further extended to measure latencies and lag spikes for online games [6], and stutters for conferencing session, even when all traffic is encrypted.

The ability for AI engines to make inferences about the traffic type and user experience as described above depend on the ability to extract the pulse for millions of flows every 100ms at Terabit speeds. Traditional fixed-function switches do not have this capability, and custom-building it would be prohibitively expensive. Fortunately, a new breed of programmable switching silicon is emerging, exemplified by the Tofino series of chips from Intel [7], capable of 12.8Tbps and available in 32 x 400G port switches today. Such switches, low in cost and commoditised by the hypescalers, can be programmed in a language called P4 to extract fine-grained per-flow telemetry at scale. Indeed, the system described above has been implemented and deployed commercially in multiple Tier-1 networks globally.

[1] Broadband Forum, 2021. An economic argument for moving away from Mbps. https://www.broadband-forum.org/an-economic-argument-for-moving-away-from-mbps

[2] J. Han, M. Lyu and V. Sivaraman, “On the Validity of Internet Speed Test Tools and Broadband Measurement Programs”, Asian Internet Engineering Conference (AINTEC), Hiroshima, Japan, Dec 2022, https://www2.ee.unsw.edu.au/~vijay/pubs/conf/22aintec.pdf

[3] R. Babaria, S. Madanapalli, H. Kumar, V. Sivaraman, “FlowFormers: Transformer-based Models for Real-time Network Flow Classification”, Intl. Conf. Mobility, Sensing, and Networking (MSN), UK, Dec 2021, https://ieeexplore.ieee.org/document/9751578

[4] S. Madanapalli, H. Habibi Gharakheili and V. Sivaraman, “Inferring Netflix User Experience from Broadband Network Measurement”, IFIP Traffic Measurement and Analysis (TMA) Conference, Paris, France, Jun 2019.

[5] S. Madanapalli, A. Mathai, H. Habibi Gharakheili, V. Sivaraman, “ReCLive: Real-Time Classification and QoE Inference of Live Video Streaming Services”, IEEE IWQoS, Tokyo, Japan, Jun 2021.

[6] S. Madanapalli, H. Habibi Gharakheili, V. Sivaraman, “Know Thy Lag: In-Network Game Detection and Latency Measurement”, Passive and Active Measurement (PAM) Conference, Mar 2022, https://dl.acm.org/doi/abs/10.1007/978-3-030-98785-5_17

[7] Intel Intelligent Fabric Processors, https://www.intel.com/content/www/us/en/products/network-io/programmable-ethernet-switch.html

For more information, see https://canopusnetworks.com/