Household bandwidth and the 'need for speed'

Jenny Kennedy [1]
RMIT University
Grenville Armitage [2]
Swinburne University of Technology
Julian Thomas [3]
RMIT University


Abstract

In this paper, we aim to contribute to the policy debate on bandwidth needs by considering more closely what happens in household networks. We draw upon both social and technical studies modelling household applications and their uses to show how queue management protocols impact bandwidth needs. We stress the impact of internet traffic streams interfering with each other, and describe three different categories of internet traffic. We demonstrate how the use of active queue management can reduce bandwidth demands. In doing so we consider how, and to what degree, household internet connections are a constraint on internet use. We show that speed demand predictions are skewed by a perceived need to protect the Quality of Service experienced by latency-sensitive services when using current gateway technologies.

Introduction

One of the more contentious topics in Australian broadband policy is the present and future need for higher bandwidth services by Australian households. Broadband technologies are developing rapidly, with the emergence of a new, more complex ecosystem of domestic connections placing greater strain on household bandwidth requirements. Within the household ecology, mobile devices and Wi-Fi are now an alternative and complementary mode of access for large numbers of Australians. The volume of connected devices and applications have proliferated both within and beyond households. This includes the nascent Internet of Things (IoT). Home broadband traffic is increasingly a mix of both latency-sensitive applications (such as online games and video conferencing services) and latency-tolerant applications (such as content streaming and data sharing services).

Increased bandwidth is attributed to increases in social welfare, GDP and consumption (Centre for Energy-Efficient Telecommunications, 2015 [7]; Deloitte, 2016 [8]). Increased bandwidth is also equated generally with social and economic benefits:

Increased bandwidth is accompanied by increased participation in the digital economy, in online activities, and in the use of entertainment and communication services and technologies; and that increased digital literacy emerges through experience and use of HSB [High Speed Broadband] (Wilken et al, 2011, p. 10 [9]).

In this paper, we aim to contribute to the policy debate on bandwidth needs by considering more closely what happens in household networks. We draw upon both social and technical studies modelling household applications and their uses to show how queue management protocols impact bandwidth needs. We demonstrate how the use of active queue management can reduce the need for bandwidth. In doing so we consider how, and to what degree, household internet connections are a constraint on internet use. We show that speed demand predictions are skewed by a perceived need to protect the Quality of Service (QoS) experienced by latency-sensitive services when using current gateway technologies.

The contemporary networked household

To understand the bandwidth needs of households, we first have to consider the contemporary networked household in more detail. We can point to four critical trends: the proliferation of household connections; a combination of intensive and extensive growth in internet use; the emergence of new connected services; and increasing diversity in Australian households. Each of these four trends are significant in their own right and we consider them briefly below.

The proliferation of household connections

Australian households have an average of nine internet-connected devices; and 9% of households have 20 or more internet-connected devices (Telsyte, 2015 [10]). These
The combination of intensive and extensive growth in internet use

The proliferation of connected devices, combined with mobility, lead to both increased use, and concurrent use, which places greater demand on bandwidth services. Concurrent usage behaviours vary based on a number of contributing factors. For example, users are more likely to multitask if they have access to multiple connected devices (Hassoun, 2014 [13]). Connected general-purpose devices, such as PCs or laptops, are also likely to facilitate multiple connected applications at once. The IoT will vastly increase the number of connected devices in households, further impacting the average peak number of applications concurrently connecting to the home network.

The emergence of new connected services

There is a growing array of internet services: in Australia, one notable example is the emergence of video streaming services such as Netflix, which includes higher resolution content (?4K?, or ?Ultra high definition?). These services, while not necessarily bandwidth intensive in their own right, are examples of an increasing number of applications requiring some level of consistent bandwidth and increasing the aggregate demand. Another example is the growing popularity of online education, which makes use of a wide range of media, but particularly relies on video and audio. Telehealth and home automation are other emerging areas of service that require consistent bandwidth and increase aggregate demand. These examples of new media use require us to rethink common assumptions, such as the idea that demand for ?video? is mainly about entertainment.

The increasing diversity of Australian households

Considerations of household internet use need to take into account the range of living and working arrangements. The average household in Australia has 2.6 people (ABS, 2015 [14]).

To help illustrate variations in household formations, Telsyte (2015 [10]) proposes four simplified (and heavily urban-centric) household profiles, each with distinctive patterns of use. Though obviously limited, these profiles are productive for demonstrating that households utilise internet services differently at home. Each of the household types represent different models of social activity around internet use, characterised by varying patterns of peak concurrent application use based on household members and consumption patterns.

Table 1 Household profiles adapted from Telsyte (2015 [10]) report

<table>
<thead>
<tr>
<th>Household profile</th>
<th>Peak app stack 2015</th>
<th>Predicted peak app stack in 2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual professional households with children</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td>(?The Hectic Household?)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single or dual-parent households with children</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>(?Suburban dreamers?)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Couples without children</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>(?City Living?)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shared living</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>(?Shared living?)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Contemporary household rhythms and peak app stacks

Rhythms of bandwidth consumption vary based on many factors. Patterns of internet usage cannot be assumed to be uniform given the known impact of socio-economic contexts (Thomas et al, 2016 [15]), infrastructure and geographical factors (Kennedy et al, 2017 [16]; Wilken et al 2013 [17]). Education, employment, social and entertainment practices are increasingly mixed, with activities overlapping both temporally and spatially within the domestic space (Nansen et al, 2009 [18], 2010 [19], 2011 [20]; Gregg, 2011 [21]). Increasing demands are placed on home bandwidth services and these demands intensify at peak times. What constitutes peak times for usage may vary and involve different configurations depending on the household and types of applications in use.

Kenny and Broughton (2014 [22]) identify that the crucial dilemma in future bandwidth use within the home is to do with anticipated peaks of concurrent application use (referred to as an ?app stack?). Kenny and Broughton identify four different types of application: primary (which they categorise as forms of streaming multimedia content, including TV, YouTube, HD video calls, and streamed and interactive games); secondary (content download and uploads, i.e. cloud storage, torrents, software and OS downloads, and non-HD video calls); web surfing; and finally, low-bandwidth traffic, which incorporate all other forms of internet usage. These four categories are characterised by user activity rather than application bandwidth specificities, e.g., Kenny and Broughton state ?primary applications are those apps that are primarily used ?one at a time? by a given individual? (2014, p. 16 [22]). Subsequently, the combinations of concurrent application use depict rhythmic patterns of household activities and presume peak bandwidth use occurs during peak app stack. These categories give a broad overview of the types of applications people use concurrently and show how demands are driven less by the number of devices, and more by the types of applications and their particular bandwidth needs. Still, their categories of application overlook the impact of constrained bandwidth and latency sensitivities that also impacts bandwidth needs.

In the sections that follow we describe our own categorisation of applications based on latency sensitivity, and demonstrate how calculations of bandwidth requirements are less effective when they ignore the latency sensitivity of key applications.

User perceptions of internet Quality of Service, and impact on activities

Bandwidth needs are driven both by application requirements and by user perceptions of internet QoS or performance. A key contributor to QoS is Round Trip Time (RTT), meaning the time it takes for a signal or packet to travel from source to specific destination and back again. Lower RTT usually leads to better QoS.

Table 2 shows some typical RTTs in milliseconds (ms) that might be experienced by home-based applications when accessing remote services over an otherwise idle home gateway (also referred to as RTT base). The RTT base is smallest when remote servers are closer.

Table 2: Typical figures for RTT base from Australia’s east coast (Armitage & Heyde, 2012 [23])

<table>
<thead>
<tr>
<th>RTT base Distance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>10ms</td>
<td>Domestic, intra-ISP servers</td>
</tr>
<tr>
<td>40ms</td>
<td>Domestic, inter-ISP or off-net servers</td>
</tr>
<tr>
<td>180ms</td>
<td>Australia to US West Coast servers</td>
</tr>
<tr>
<td>240ms</td>
<td>Australia to US East Coast servers</td>
</tr>
</tbody>
</table>
RTTs are inflated when there are queuing delays (e.g. due to temporary buffering at congested network routers and switches along the path) and mutual interference between different categories of internet traffic in concurrent use. Such latency issues frustrate users (Ceaparu et al., 2004 [24]). Users are especially likely to report QoS issues when streaming content or conducting voice calls, describing frustration with buffering and call dropout. Thresholds and latency tolerances are somewhat vague and arbitrary. Users are reported to have a threshold of up to 500ms for web page retrieval before becoming frustrated when searching for information (Armitage et al., 2014 [25]) and to prefer sub-100ms for highly interactive first-person shooter (FPS) games (Wac et al., 2011 [27]).

Inflated RTTs impact home internet activities whereby users adapt their practices to accommodate idiosyncratic tolerances of QoS. Strategies involve allowing more time to complete activities, varying and combining tasks, or considering alternate contexts, i.e. completing tasks at a different time, on a different device, or even a different network.

Users perceive that increased bandwidth will reduce RTTs and improve QoS. This perception is both fuelled by and reflected in market attention towards increased bandwidth at a cost of other technological solutions. Increased bandwidth does not directly translate to reduced RTTs in the domestic context. RTTs vary heavily depending on particular household internet traffic streams which mutually interfere with each other. Bandwidth requirements depend on how often mutual interference between internet traffic is likely to occur, and consumer’s tolerance of RTT inflation triggered by such mutual interference.

Categories of internet traffic

Mutual internet traffic interferences are experienced differently depending on the type of traffic. We identify three different categories in the section below.

Latency-sensitive and interactive traffic

Interactive applications involve steady streams of packets between two points on the network at intervals dictated (and limited) by the applications themselves. Some of these applications are continuously interactive in nature, involving a human at one or both ends of the network path, generating and consuming data sent over the network in real-time. Examples include:

- Voice over IP (VoIP)
- Multi-party voice/video conferencing (such as Skype, Facetime, and remote education services)
- Online games (particularly ‘switch’ games like First Person Shooters, or other highly immersive environments)
- Real-time, remote medical monitoring services

Other latency-sensitive applications are sporadically interactive, generating short bursts of traffic infrequently, and benefiting from low RTTs for responses. Examples include cloud-based IoT and home automation services typified by nascent products such as Google’s Home, Apple’s HomeKit, Amazon’s Alexa, and third-party programmable automation platforms such as ‘If This Then That’ (internet.ifttt.com). The timeliness of information transfer is important for all latency sensitive traffic.

There are also behind-the-scenes, latency-sensitive activities. For example, when a user clicks on a new link on a web page their browser does a domain name system (DNS) lookup to determine the target site’s actual IP address before retrieving the remote site’s content. DNS lookups traverse the home’s connection to their ISP, where delays due to interference from other traffic make web browsing more tedious and less responsive. Similarly, services that rely on transmission control protocol (TCP) connections (like web browsing, sending emails, starting new file downloads or uploads, and so forth) can feel far less responsive when a home’s RTT to the outside world is inflated by congestion in the home gateway.

Latency-tolerant (elastic) traffic

Latency-tolerant traffic is elastic, in that the application is flexible in terms of RTT and tolerates slowing down or speeding up as dictated by available bandwidth. Examples include:

- Photo and video sharing applications sync’ing content to/from the Cloud?
- Web browsers retrieving embedded digital objects to render pages
- Sending and receiving emails with large attachments
- Peer to peer file transfer applications
- Remote/offsite backup systems
- Application update systems (such as triggered by Microsoft Windows updates, or iOS Android App Store updates)
- Downloading podcasts, movies, TV show episodes or music tracks in their entirety for later, offline playback
- Instant messaging / notifications with multimedia attachments (Twitter, Apple’s iMessage, and so forth)

Elastic traffic is measured on time to completion (TTC) - how long to upload or download a podcast, photo or app update, and so forth. TTC goes up as the size of objects being transferred grows up and if the bandwidth grows down. Consequently, bandwidth requirements for elastic applications depend on consumer tolerance for TTC, which in turn depend on the social context of application use. For example, one consumer might accept a TTC of 15 minutes when downloading a 60-minute video, while another consumer considers a TTC over 6 minutes to be unacceptable.

Elastic applications often use TCP as their underlying transport. Unless limited by the application itself, elastic traffic will consume as much bandwidth as TCP can extract from the network at any given instant. This causes home gateway congestion that increases RTT for all internet traffic sharing that gateway. TTC can degrade as RTT increases; new TCP connections take longer to start-up, and active TCP connections take longer to recover from regularly occurring packet losses (from brief slow-downs to multi-second stalls as competing TCP connections step in and briefly consume all of a path’s capacity).

Streaming content traffic

Examples of streaming content include:

- internet TV and movie services (such as Netflix)
- internet radio services

Streaming services are initially interactive (and latency sensitive) when consumers are using online menus to select content. Once the selected content has begun playing (streaming), the service can adapt to variations in network latency.

Average bandwidth requirements can be estimated from the audio/video encoding rates of the content being streamed. However, the short-term behaviour of streaming traffic can have distinctly aggressive characteristics akin to repeated bursts of short-lived elastic traffic.

Services built on technologies such as DASH (Dynamic Adaptive Streaming over HTTP) will usually begin a stream by pulling down tens of seconds of content as fast as possible, then settling into a regular pattern of short bursts of data traffic as the client retrieves chunks of content piece-meal from the server over time. In addition, DASH-like services will usually adapt the content quality (and hence size in bytes) of newly requested chunks depending on the speeds achieved while retrieving previous chunks. Commonly deployed on top of conventional TCP, such traffic causes periodic short bursts of congestion on the home broadband link as each new content packet is retrieved.
Combining the bandwidth needs of each category

Today's home internet users are likely to be dissatisfied by an internet service offering (a) high latencies for latency-sensitive applications, (b) long TTC for important elastic applications, and (c) poor streaming quality. They require sufficient downstream and upstream bandwidth to ensure satisfactory service during periods of mutual interference where applications from all three categories are simultaneously active.

Streaming traffic is an attractive category on which to base rough bandwidth estimates, as the average requirements can be estimated from the consumer's desired video and audio quality. For example, it is reasonably simple to estimate the number of MBs it takes to stream TV or DVD-quality movie content per minute using common encoding and compression schemes. However, rather than flow smoothly, most streaming traffic hits the home's access link with short bursts of elastic-like traffic, consuming as much spare bandwidth as is available at the time during each burst.

Elastic applications are more complicated. They only require enough bandwidth to achieve an acceptable TTC. But by using TCP for transport, elastic applications will typically consume whatever extra bandwidth happens to be available at the time, leading to even shorter TTC than the user may need. What constitutes an acceptable TTC (and hence minimum bandwidth requirement) depends greatly on the application itself, usage-patterns and user tolerance thresholds.

Significance of queue management in home gateways

In the home, there is typically RTP inflation during heavy traffic loads due to conventional home gateways using the first-in-first-out (FIFO) queueing protocol.

The internet requires certain amounts of buffering (queue storage) in routers and gateways to absorb transient bursts of traffic. During periods of low (or no) congestion, buffers are mostly empty and packets pass through with minimal additional delay. However, during periods of congestion late arriving packets can experience additional queuing delays. When bulk data transfers cause long-lived or cyclical queue build up, the conventional FIFO queue architecture means all traffic gets backed-up inside the queue, and everyone experiences worst case RTTs.

RTT inflation is higher through gateways with more buffer space, and lower as access bandwidth goes up. This is a key reason why many users perceive higher bandwidth service to be the critical factor for a positive internet experience. Unfortunately, many gateway vendors provide excessive amounts of buffering in an attempt to minimise packet losses by TCP connections. Often referred to as "bufferbloat" (Gettys & Nichols, 2011 [28]), the effect is to exacerbate RTP inflation during congestion. Dropped packets cause the higher layer TCP connection to slow down and retransmit the lost packet, which increases TTC not network layer RTT. RTP is inflated by having more queuing delay before the loss occurs.

We argue that traffic queues management is the most crucial factor when determining bandwidth needs. Downstream traffic queuing is managed at the internet service provider (ISP) end, whereas upstream traffic queuing is managed at the home gateway. Active Queue Management (AQM) at home gateways can reduce the upstream bandwidth required to support satisfying user experiences.

Motivated by concerns over "bufferbloat", recent Internet Engineering Task Force (IETF) interest has focused on AQM schemes such as Proportional Integral controller Enhanced (PIE) (Pan et al, 2013 [29]), Controlled Delay (CoDel) (Nichols & Jacobson, 2016 [30]) and FlowQueue-CoDel (FQ-CoDel) (Hoeiland-Joergensen et al, 2016 [31]). These AQM schemes provide burst-tolerant congestion signalling at far lower levels of queuing delay than is typical of classical FIFO (or tail drop) queue discipline. A summary of work studying the performance of different AQMs for a variety of traditional Internet applications can be found in Hoeiland and Joergensen (2015 [32]).

In particular, FQ-CoDel assigns different traffic flows to sub-queues, uses CoDel to manage each sub-queue, assigns relatively even bandwidth share between sub-queues, and briefly gives priority to newly-populated sub-queues. As a result, an FQ-CoDel bottleneck achieves latency reductions, capacity sharing, and priority for low-rate or transactional traffic (such as DNS, TCP connection establishments and VoIP).

We use a controlled experimental testbed to demonstrate the positive impact of switching from FIFO to FQ-CoDel buffer management (Armitage et al, 2017 [33]). Figure 1 illustrates the RTT of a VoIP flow competing with an upstream elastic TCP flow with 1, 2, 3, 4 or 5Mbps upstream link speed. With FIFO, the RTT inflation is severe: over 1 second at 1Mbps and still over 250ms at 5Mbps. Using FQ-CoDel over the same range sees the VoIP flow experience RTTs well under 100ms.

![Figure 1: Significant RTT reduction when using FQ-CoDel instead of FIFO buffer management](image)

Predictions of bandwidth need

To further demonstrate the significance of home gateway queue management for bandwidth requirements, we can emulate the probable internet usage scenario of a typical household. In doing so, we can give an estimation of the bandwidth demands for using FIFO queue management versus using AQM techniques.

The Hectic Household, as described by Telsyte (2015 [10]) averages 12 apps during peak time use. Drawing on ethnographic fieldwork on household rhythms of technology uses we can describe in some detail a realistic profile of such a household (Kennedy et al, 2015 [39], 2017 [16]; Wilken et al, 2011 [8], 2013 [17]).

Many applications start and stop throughout the day. We estimate that the peak app stack in this type of household of two parents and two children occurs in the early evening, at which time the following devices could be expected to be actively accessing the internet:

- Laptop #1
- Laptop #2
- Tablet #1
- Tablet #2
- Smart phone #1
- Smart phone #2
- FitBit
- Sonos speakers x3
- XBox
- Smart TV (or TV with set-top box)

<table>
<thead>
<tr>
<th>Traffic Category</th>
<th>Application</th>
<th>Bandwidth requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency-sensitive applications</td>
<td>Online game</td>
<td>100 kbps down/up</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If audio-only call</td>
</tr>
</tbody>
</table>

Table 3: Application bandwidth requirements
Video call

100 kbps down/up

Video & audio

Latency-tolerant applications

200-500 kbps down/up

Retrieving photos from ‘the cloud’ / file downloads / receiving emails with attachments / software/firmware updates

Bulk downloads

5Mbps down

Sync’ing photos or documents ‘to the cloud’, sending emails with attachments

Bulk uploads

0.5Mbps up

Social media / instant messages with attachments

Short-lived TCP notifications

0.2Mbps up/down

In Nov/Dec 2016 the average web page was 2.4MBytes of content spread over 106 HTTP requests across 35 different TCP connections.

Web browsing (with multiple windows open)

Requires both high speed (to download content) and low RTT (minimising TCP connection set up times)

3Mbps down

ABC iView @ high quality

1.5Mbps down

Netflix

SD@ 480p: 3Mbps down

HD@720p: 5Mbps down

Streaming applications

Ultra HD (4K): 25Mbps down

Stan

SD: 3Mbps down

HD (720p): 4.5Mbps down

HD (1080p): 7.5Mbps down

Sonos

Audio streaming

256 - 320 kbps down

In Table 3 we estimate the applications that could conceivably be in use simultaneously during the peak app stack, together with the typical bandwidth requirements of each application. Using these approximations, we can estimate the bandwidth required to provide uninterrupted and timely service for the household during this period. Adding the speeds required by each category in each direction separately gives us a downstream of 16.9Mbps (see Table 4) and upstream of 2.0Mbps (see Table 5).

Table 4: Downstream bandwidth requirements

<table>
<thead>
<tr>
<th>Traffic Category</th>
<th>Total bandwidth requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency-sensitive applications</td>
<td>0.6Mbps</td>
</tr>
<tr>
<td>Latency-tolerant applications</td>
<td>8.8Mbps</td>
</tr>
<tr>
<td>Streaming applications</td>
<td>7.46Mbps</td>
</tr>
</tbody>
</table>

Table 5: Upstream bandwidth requirements

<table>
<thead>
<tr>
<th>Traffic Category</th>
<th>Total bandwidth requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency-sensitive applications</td>
<td>0.3Mbps</td>
</tr>
<tr>
<td>Latency-tolerant applications</td>
<td>1.3Mbps</td>
</tr>
<tr>
<td>Streaming applications</td>
<td>0.4Mbps</td>
</tr>
</tbody>
</table>

A superficial conclusion would argue this household’s needs could be met by a 18/2 Mbps service. However, this would fail to account for the RTT inflation of bulk transfer and streaming services hitting standard FIFO bottlenecks in the upstream and downstream directions. The household would notice significant degradation of latency-sensitive interactive activities during peak periods if the home were serviced at 18/2 Mbps with FIFO queue management.
In order to keep RTTs moderately bounded during bursts of elastic upstream traffic, this household requires at least 5Mbps in the upstream. Assuming the FIFO buffers aren't too long in the downstream direction, a hypothetical 18/5Mbps service might service this household with tolerable degradation from time to time. In the current market, they would need to purchase a plan of at least 25/5Mbps speeds.

Depending on personal thresholds, this household's needs could be catered for with 12/1Mbps speeds. But during peak app stack periods, while streaming services may remain arguably acceptable (albeit degraded), web browsing, bulk uploads and notifications would experience noticeably longer TTCs, and interactive applications would be regularly disrupted to the point of uselessness. The unsatisfactory solution for households today is to adapt their usage patterns and eliminate simultaneous use of latency sensitive, latency-tolerant and streaming applications. However, this solution is increasingly ineffective as more unattended applications on tablets, phones and IoT devices launch bursts of elastic traffic at unpredictable times throughout the day.

Household requirements with AQM

When considering the household's overall bandwidth requirements if their broadband service used FQ-CoDel for active queue management in both directions, the speeds required by each category of traffic in each direction are the same as above. What changes is the overall bandwidth required from the ISP to minimise mutual interference.

FQ-CoDel isolates the different traffic flows from each other and keeps RTTs low. The household's latency-sensitive traffic is protected and low RTTs are experienced, regardless of elastic and streaming traffic. In this scenario, a hypothetical 18/2 Mbps service becomes conceivable.

FQ-CoDel is also beneficial to elastic and streaming applications, as the underlying round-robin scheduling gives even sharing of available capacity during peak app stack periods. Compared to FIFO scenarios, with FQ-CoDel it is far less likely for one elastic traffic flow to impact capacity of other flows for periods of time. Consequently, the household would perceive the TTCs of various activities to be more consistent throughout the day.

Even 12/1Mbps speeds could be a realistic option if the household was willing to accept a modest degradation in TTCs for elastic applications and reduced streaming quality during peak app stack periods. During peak periods, the interactive traffic is protected (due to its low speed demands) and the remaining upstream and downstream capacity is divided equally between all other competing traffic.

The bandwidth requirements of most current latency-sensitive applications are dwarfed by the requirements of streaming media or elastic applications with demands for low TTC. However, it only takes one latency-sensitive application to be impacted by RTT inflation for the household to perceive their broadband service to be inadequate. Consequently, demand for significant (e.g. greater than 5Mbps) upstream bandwidth will exist whether we envisage one VoIP call during peak app stack periods, or multiple VoIP calls and multiple online games. Alternatively, we deploy an AQM (like FQ-CoDel or similar) at least in the upstream direction of home gateways to mitigate RTT inflation.

While the Telustry report (2015) provides some alternative models for household populations (and hence peak app stack), there remains a need for greater insight into TTC tolerances and expectations of typical consumers for activities such as sending multi-media notifications, syncing content to and from portable devices, and so forth. It is also important to constrain overly optimistic estimates by recognising limits imposed by the physical context of typical households. For example, the number of living areas, bedrooms, and so forth puts practical upper bounds on the number of streaming or internet access devices operating concurrently at any given time. The physical sizes of viewing areas also limits the size of practical TV screens (and hence, the degree to which a household may be satisfied with combinations of SD, HD and/or ultraHD streaming).

AQM deployment challenges

A number of practical issues mean that AQMs will take time to deploy. In the long term, equipment on either side of the broadband link between homes and their ISPs need to be upgraded. In the short term, significant benefits will accrue simply from having home gateways incorporate something like FQ-CoDel to manage buffers on the upstream side of their broadband service port(s).

ISPs typically use equipment from experienced data networking device vendors to control their end of the access link. However, upgrade schedules are guided by existing multi-year equipment or software contracts, and their vendors' willingness to incorporate AQM technologies in the medium to long term.

In contrast, home gateways may be owned and supplied by the ISP, or independently purchased by the consumer from a range of low-margin vendors. This raises several challenges.

ISP-owned gateways might be upgraded to support FQ-CoDel through automated, remote firmware update and reconfiguration. But this presumes the gateway's vendor plans to release new firmware with FQ-CoDel or similar AQM, and that the existing firmware allows remotely-controlled firmware updates. Well-known brands have been inconsistent in supplying updates to devices sold more than a few years earlier. This may well continue in regard to adding AQM.

Technically-savvy consumers might be motivated to try upgrading their own gateway's firmware. But only a small subset of consumers are likely to explore this option, constrained by reliance on new vendor-supported firmware or third-party open-source firmware that supports AQM.

A significant fraction of basic consumer gateways retail for less than $100 and are based around embedded versions of the Linux 2.x series kernel that do not support FQ-CoDel. If late 3.x or current 4.x series Linux kernels are on a vendor's product development roadmap, supporting FQ-CoDel in future models is easy. If a firmware upgrade of their existing FIFO-based gateway is not an option, consumers face a choice between replacing their existing gateway with a new sub-$100 AQM-capable unit, or paying monthly for a higher downstream/upstream speed tier from their ISP. The former is likely to be a very attractive option for people whose downstream speed requirements are already met by one of the lower speed tiers.

Conclusions

Users and ISPs are in a position to leverage AQM technologies in the upstream direction of home gateways, independent of the political and market forces driving country-wide upgrades to broadband access services. This benefit is currently under-tapped and represents a niche market opportunity. Deployment of FQ-CoDel (or similar) would provide significant improvement to consumer experience of interactive services for those on a 12/1Mbps speed tier or ADSL2+ plans. Even with 25/5Mbps and higher tiers available, many households may find their needs for VoIP, games and casual internet use met by a 12/1Mbps plan coupled with FQ-CoDel.

Acknowledgements

This work was supported in part by a 2014-2016 Swinburne University of Technology / Cisco Australia Innovation Fund project titled ?An evaluation of household broadband service requirements for educational innovation and Internet of Things?.

References

