The ACCC's Pilot Broadband Performance Monitoring & Reporting Program

Technical Appendix

September 2015

Contents

1. Test Methodology provided by SamKnows for Pilot BPMR Report for the ACCC

1.1. Speed Tests

Measures the download and upload speed of the broadband connection in bits per second. The transfer is conducted over one or more concurrent HTTP connections (using the GET verb of download and the POST verb for uploads).

In the download speed test the client will fetch a portion of a 1GB binary (nonzero, randomly generated) payload hosted on an HTTP server on the target test node. The content is discarded as soon as it is received.

In the upload test the client will generate the payload itself (using /dev/urandom as a non-blocking source of random content) to send to the server. The measure of throughput may be optionally carried out on the server side (the receiver) in the upload test.

The speed tests (both download and upload) operate for either a fixed-duration (specified in seconds) or a fixed-volume (specified in MB). Where possible, a fixed-duration test is preferred as it will cater well for all broadband access speeds. However, a fixed-volume test may be necessary where predictability of bandwidth usage is desired.

Four separate variations of the test are supported:

- Single TCP connection download speed test
- Multiple TCP connection download speed test
- Single TCP connection upload speed test
- Multiple TCP connection upload speed test

For multiple TCP connection tests we typically recommend that three concurrent connections are used. In some cases (e.g. where the round-trip time between client and server is very high) it may be necessary to increase this.

Factors such as TCP slow start are accounted for through the use of a “warmup” period. This period begins as soon as the test starts and seeks to establish that the throughput has reached stable rate before starting the real test (which will continue over the same TCP connection(s)). It is important to note that the data transferred in the warm-up period is excluded from the main test results, but it is still recorded separately as a supplementary metric.

The speed test client will record the throughput, bytes transferred and time taken at the end of the test. It may also record these values at multiple intervals during the test. This is commonly used to help characterise the difference between ‘burst’ and ‘sustained’ throughput (where transfer speeds may be inflated at the start of a TCP connection).

1.2. Web Browsing

Measures the time taken to fetch the HTML and referenced resources from a page of a popular website. This test does not test against centralised testing nodes; instead it tests against real websites, allowing for content distribution networks and other performance enhancing factors to be taken into account.

Each Whitebox will test ten common websites on every test run. The time taken to download the resources, the number of bytes transferred and the calculated rate per second will be recorded. The primary measure for this test is the total time taken to download the HTML page and all associated images, Javascript and stylesheet resources.

The results include the time taken for DNS resolution. The test uses up to eight concurrent TCP connections to fetch resources from targets. The test pools TCP connections and utilises persistent connections where the remote HTTP server supports them.

The test may optionally run with or without HTTP headers advertising cache support (through the inclusion or exclusion of the “Cache-Control: no-cache” request header). The client advertises the user agent of Microsoft Internet Explorer 10.

1.3. Voice Over IP (Jitter)

This test uses a fixed-rate stream of UDP traffic, running between client and test node. A bi-directional 64kbps stream is used with the same characteristics and properties (i.e. packet sizes, delays, bitrate) as the G.711 codec.

The client initiates the connection, thus overcoming NAT issues, and informs the server of the rate and characteristics that it would like to receive the return stream with.

The standard configuration uses 500 packets upstream and 500 packets downstream.

The client records the number of packets it sent and received (thus providing a loss rate), and the jitter observed for packets it received from the server. The server does the same, but with the reverse traffic flow, thus providing bidirectional loss and jitter.

Jitter is calculated using the PDV approach described in section 4.2 of RFC5481. The 99th percentile will be recorded and used in all calculations when deriving the PDV.

1.4. UDP Latency and Packet Loss

Measures the round trip time of small UDP packets between the Whitebox and a target test node. Each packet consists of an 8-byte sequence number and an 8- byte timestamp. If a packet is not received back within two seconds of sending, it is treated as lost. The test records the number of packets sent each hour, the average round trip time of these and the total number of packets lost. The test will use the 99th percentile when calculating the summarised minimum, maximum and average results on the Whitebox.

As with the availability test, the test operates continuously in the background. It is configured to randomly distribute the sending of the echo requests over a fixed interval, reporting the summarised results once the interval has elapsed.

Typically 600 samples are taken per hour, distributed throughout the hour. If the line is busy then fewer samples will be taken. A higher sampling rate may be used if desired.

1.5. UDP Latency and Loss Under Load

This test seeks to characterise the change in latency and packet loss whilst the line is heavily loaded.

To do this, the latency under load test relies on the existing UDP Latency/Loss test and the download and upload speed tests. It first launches the speed test, and then immediately

after it has launched it begins sending UDP datagrams to the target server for 30 seconds, with packets spaced 500ms apart (and a 3 second timeout). Once the speed test has completed the UDP datagrams cease being sent and the test records the mean, minimum and maximum round trip times in microseconds. The number of lost UDP packets was also recorded.

The test is conducted separately when download and upload speeds are being tested, allowing us to characterise how upstream load affects latency/loss differently to downstream load.

1.6. DNS Resolution

This test measures the DNS resolution time of a selection of common website domain names. These tests will be targeted directly at the ISPs recursive resolvers. A list of appropriate servers will be sought from each ISP in advance of the tests.

1.7. Video Service X Measure

The Video Service X test is an application-specific test, supporting the streaming of video and audio content from Video Service X using their protocols and codecs.

The test begins by seeking out the most popular video in the user's country. This is achieved by fetching a list of the most popular Video Service X videos from a central SamKnows server. The central list of videos is refreshed once every 12 hours using the Video Service X API. We filter for videos that are at least 60 seconds in length and have an HD quality variant. Note that by interacting with the Video Service X API from a central location we can ensure that every probe is delivered the same list of videos.

The test running on the probe will now fetch the Video Service X web page for the most popular video, and parse the Javascript contained within the page. Within this Javascript is held a list of all of the encodings of the video in question and the content server hostname. By making this request from the probe we ensure that the test is receiving the same content server as the user would if they were using a desktop computer on the same connection.

The test will then connect to the content server (using whatever server Video Service X would normally direct a real client on the same connection to) and begins streaming the video and audio. MPEG4, WebM, Dash (adaptive) and Flash video codecs are supported. Although the adaptive codec is supported, the test does not actually adapt its rate; we stream at full rate all of the time, which provides for reproducibility.

The test parses video frames as it goes, capturing the timestamp contained within each video frame. After each frame we sample how much realtime has elapsed versus video time. If video time > realtime at a sample period, then an underrun has not occurred. Otherwise, one has occurred.

The test downloads 10 seconds of audio and video at a time, with a buffer of 40 seconds. So on startup, the test will immediately download (at full speed) 40 seconds of video and audio, and will then download more as required, keeping the 40 second playback buffer full. By default the test will run for a fixed duration of 20 seconds of realtime.

In its default mode of operation the test will capture the 'bitrate that can be reliably streamed' on the user's connection. This is achieved through the following process:

1. Find the fastest recent speedtest result that the probe has completed.

2. As described above, fetch the list of Video Service X videos, find the most popular one, and then select the highest bitrate encoding which is less than the fastest speedtest result found in step 1.
3. Attempt to stream this video, for a fixed duration of 20 seconds of realtime. If successful, then the “bitrate reliably streamed” for this instance is the bitrate that we just fetched.
4. However, if a stall event occurs, then we immediately abort the test and retry at the next lower bitrate.
5. If we find a bitrate that we can stream without a stall event occurring then that bitrate is our “bitrate reliably streamed” for this instance.
6. However, if we encounter stalls for every bitrate, then the “bitrate reliably streamed” is zero.

The key outputs from this metric are:

a) The bitrate reliably streamed
b) The startup delay (the time taken to download two seconds of video)
c) The TCP connection time
d) The number of stalls and their duration (this is only applicable if the test is not running in the ‘bitrate reliably streamed’ mode)

2. Testing Schedules

The volunteer panel was divided into two testing groups because many of the volunteers had limited data plans, while some had very high or unlimited data plans, which enabled higher data intensive testing to be undertaken.

2.1. Testing schedule for regular volunteer panel

Metric	Test target(s)	Test frequency	Test duration	Total Estimated Monthly Volume
Download speed	Melbourne test node	12am-6am: once 6am-12pm: once 12pm-6pm: once 6pm -12am: every two hours	Fixed 5 seconds	1 GB at 10 Mbps 2 GB at 20 Mbps 5 GB at 50 Mbps
Upload speed	Melbourne test node	12am-6am: once 6am-12pm: once 12pm-6pm: once 6pm -12am: every two hours	Fixed 5 seconds	0.1 GB at 1 Mbps 0.2 GB at 2 Mbps 0.5 GB at 5 Mbps
Web browsing	Yahoo, Ebay, ABC, The Age, Commbank, Facebook, Linkedin, Twitter, Amazon, Wikipedia	6 times per day for 10 websites. 12am-6am: once 6am-12pm: once 12pm-6pm: once 6pm -12am: every two hours	Estimated 30 seconds	0.35 GB
Voice over IP (Jitter)	Melbourne test node	Hourly, each day (24x7)	Fixed 10 seconds	1.92 MB
UDP Latency and Packet Loss	Melbourne test node	Hourly, each day (24x7)	Permanent	1 MB
DNS Resolution	Yahoo, Ebay, ABC, The Age, Commbank, Facebook, Linkedin, Twitter, Amazon, Wikipedia	Hourly, each day (24x7)	Estimated 1 second	0.1 MB

Video Service X video streaming	Video Service X's most popular video	12am-6am: once 6am-12pm: once 12pm-6pm: once 6pm -12am: hourly	20 seconds	4.5 GB (25 MB per test)

2.2. Testing schedule for extended volunteer panel

Metric	Test target(s)	Test frequency	Test duration	Total Estimated Monthly Volume
Download speed	Melbourne test node, Hong Kong test node	12am-6am: once 6am-12pm: once 12pm-6pm: once 6pm -12am: every two hours	Fixed 5 seconds	2 GB at 10 Mbps 4 GB at 20 Mbps 10 GB at 50 Mbps
Upload speed	Melbourne test node, Hong Kong test node	12am-6am: once 6am-12pm: once 12pm-6pm: once 6pm -12am: every two hours	Fixed 5 seconds	0.2 GB at 1 Mbps 0.4 GB at 2 Mbps 1.0 GB at 5 Mbps
Web browsing	Yahoo, Ebay, ABC, The Age, Commbank, Facebook, Linkedin, Twitter, Amazon, Wikipedia	12am-6am: once 6am-12pm: once 12pm-6pm: once 6pm -12am: every two hours	Estimated 30 seconds	0.35 GB
Voice over IP (Jitter)	Melbourne test node	Hourly, each day (24x7)	Fixed 10 seconds	1.92 MB
UDP Latency and Packet Loss	Melbourne test node, Hong Kong test node	Hourly, each day (24x7)	Permanent	1 MB
DNS Resolution	Yahoo, Ebay, ABC, The Age, Commbank, Facebook, Linkedin, Twitter, Amazon, Wikipedia	Hourly, each day (24x7)	Estimated 1 second	0.1 MB

Video Service X video streaming	Video Service X's most popular video	12am-6am: once 6am-12pm: once 12pm-6pm: once 6pm -12am: hourly	20 seconds	4.5 GB (25 MB per test)

3. Overview of the testing approach taken in the Pilot BPMR Program against the requirements for an ongoing program as outlined in the Position Paper

This section details the testing methodology adopted in the Pilot Program and explains why certain specifications in the Position Paper were not adopted in the pilot, if applicable.

A. Technical approach

Position Paper reference	Position Paper requirements for an ongoing program	Pilot Program adopted these specifications	Why the Pilot Program differed from the Position Paper requirements
A.1	The test setup must be technology neutral i.e. compatible with all forms of fixed broadband including ADSL, VDSL, HFC cable, FTTP, fixed wireless and satellite.	✓	Not applicable
A.2	The test setup must be able to provide accurate information for services with headline data transfer rates in excess of 100Mbps.	✓	Not applicable
A.3	Testing must be automated (requiring no end-user input after initial setup) and able to be performed on a defined schedule including both peak and off-peak times.	✓	Not applicable
A.4	Test results must not be unduly affected by volunteers' in-home network configurations (e.g. Wi-Fi and/or access devices).	✓	Not applicable
A.5	The test setup must generate traffic to simulate end-user behaviour rather than just passively monitoring existing end-user traffic.	✓	Not applicable
A.6	The test setup must be capable of running both simple metrics such as peak/off-peak downstream and upstream data transfer rates, webpage load times and video streaming performance as well as more technical quality of service metrics including TCP/UDP/ICMP latency and	✓	Not applicable

	packet loss, jitter and DNS resolution and failure rates. Other metrics such as specific application testing would be considered optional.		
A.7	The test setup must be able to identify changes to the end-user's service configuration e.g. increases or decreases in subscribed data transfer rate, churn to another RSP etc.	×	Relying on volunteers to advise us of any changes to their service configuration was sufficient for the purposes of an internal Pilot Program. However, in addition to volunteers advising of changes to their services, we would require the test setup in any ongoing program to be able to identify those changes.
A.8	Testing should be conducted primarily between the end-user modem or router and domestic test servers located within each capital city. This is to minimise the effect of network elements outside the control of RSPs and network operators.	×	As the Pilot Program was only conducted in Melbourne, it was not necessary to have test servers located in each capital city. However, any ongoing program would be conducted on a national basis and test servers would be located in each capital city.
A.9	RSPs should be encouraged to host 'on-net' test servers. The data generated by tests on such servers would provide additional validation of broader results but would not be included in public reports.	×	As the pilot was an internal program, we did not consider it necessary to involve RSPs in the pilot. However, we would consider encouraging RSPs to host 'on-net' test servers in any ongoing program to provide additional validation of program results.

B. Impact on volunteers

Position Paper reference	Position Paper requirements for an ongoing program	Pilot Program adopted these specifications	Why the Pilot Program differed from the Position Paper requirements
B.1	The testing tool must be easy for volunteers to install and/or setup and must only require technical support from the program manager in limited circumstances.	✓	Not applicable
B.2	To avoid disruption or degradation to volunteers' broadband services, tests must only be performed when services are not being actively used.	✓	Not applicable

B.3	The testing regime must not consume a large amount of data as this may increase costs for volunteers. The test setup should include a mechanism for tailoring the testing regime based on the volunteer's subscribed data quota e.g. to run a more limited suite of tests or to run tests less frequently where quota is an issue.	✓	Not applicable
B.4	The testing tool must not log volunteers' personal data and the testing company should have appropriate safeguards in place to ensure that security and confidentiality of personal data is maintained.	✓	Not applicable

C. Sample size and selection

Position Paper reference	Position Paper requirements for an ongoing program	Pilot Program adopted these specifications	Why the Pilot Program differed from the Position Paper requirements
C.1	**Geographical:** defined segments such as Large City (Sydney, Melbourne, Brisbane, Perth, Adelaide), Medium City (Gold Coast-Tweed, Newcastle-Maitland, Canberra-Queanbeyan, Sunshine Coast, Wollongong), Small City (Hobart, Geelong, Townsville, Cairns, Darwin). Additionally, an aggregated national regional/rural segment measuring results across the range of RSPs, technologies and speed tiers for volunteers in areas with limited broadband infrastructure.	×	As the Pilot Program was an internal program limited to Melbourne participants, it did not collect data about service performance in other locations. However, any ongoing program would require statistically significant samples incorporating defined geographic segments across Australia.
C.2	**Service types:** NBN-based services including FTTP, FTTN, HFC, fixed wireless and satellite, as well as non-NBN services such as ADSL and potentially FTTP. While included in the program, ADSL services would be tested at a	×	As the Pilot Program was limited to approximately 90 volunteers, not all fixed broadband technologies could be covered in the sample size. The Pilot Program tested ADSL, HFC, FTTP and FTTN services but there were no volunteers on fixed wireless and satellite services.

	more aggregated level with results across all RSPs being combined to provide general guidance on performance rather than more specific comparative guidance.		However, any ongoing program would test all fixed broadband technologies.
C.3	**RSPs:** at least the top five RSPs by subscribers within each defined geographic segment, as well as the top five RSPs by subscribers in the broadly defined regional/rural segment. To ensure a representative sample, any other RSP for which a sufficient number of volunteers opted in would also be included.	×	As per the reasons outlined in C.1 above, data was not collected for defined geographic segments and therefore we did not need to look at the top RSPs by subscribers in each location. However, any ongoing program would ensure a representative sample across the top RSPs by subscribers in each geographic segment.
C.4	**Speed tiers:** defined segments including 'up to 24Mbps', '25 to 50Mbps' and 'greater than 50Mbps'. These segments would only inform how many samples are required for each RSP and how retail plans are grouped when presenting results.	×	As a proof of concept, the Pilot Program did not require a representative sample and therefore it was not necessary to work out how many samples were required for each RSP. However, any ongoing program would look at defined speed tiers to ensure a representative sample.
Section C	The overall sample would need to be representative across the geographical, service type and speed tier dimensions, and ensuring this would be part of the sample selection exercise. There would also be a threshold minimum number of valid observations per sub-segment (e.g. Large City end-users with RSP #1 on HFC at 50+ Mbps); this threshold would be set based on the accuracy and margin for error of the testing solution ultimately deployed.	×	As noted in this report, we did not consider that a statistically robust sample was required for the purposes of the pilot. However, any ongoing program would require statistically significant samples incorporating each of the geographical, service types, RSPs and speed tier dimensions as outlined in the Position Paper.

D. Volunteer recruitment and management

Position Paper reference	Position Paper requirements for an ongoing program	Pilot Program adopted these specifications	Why the Pilot Program differed from the Position Paper requirements
D.1	Volunteers would be recruited through a public campaign,	×	As the pilot was an internal program, a public recruitment campaign was not

	including social media, and would be strictly opt-in.		required. However, volunteers would be recruited through a public campaign in any ongoing program, and participation would be strictly opt-in.
D.2	The testing company must collect and validate critical consumer data from volunteers including their street address, RSP, broadband technology and subscribed 'speed' tier.	×	As the Pilot Program was a proof of concept, we did not consider it necessary for the testing company to validate data (e.g. address and broadband plan details) received from volunteers. Furthermore, the Pilot Program only tested a small sample of services and so it would not have been possible to draw conclusions based on geographic location. However, any ongoing program would require validation of volunteer details to ensure the accuracy of test results. We would also consider involving RSPs in any ongoing program to improve the accuracy of the validation process.
D.3	For ADSL and VDSL services, the length of the copper line from the relevant DSLAM (whether in a node or exchange) to the volunteer's premises must be determined. Samples for ADSL and VDSL services would be narrowly defined such that only volunteers whose copper line lengths fell within a specified range would be accepted.	×	For the reasons outlined above for D.2, volunteer address details were not collected in the Pilot Program and therefore line length data was not available. However, any ongoing program would collect address details and this would be used to determine copper line lengths.
D.4	Volunteers must not receive monetary compensation in return for their participation in the program and they may exit the program as they wish. However it would be open to the testing company to provide non-monetary incentives such as access to real-time and historical data on the volunteer's own connection to encourage participation.	✓	Not applicable
D.5	The testing company must not disclose a volunteer's identity to their RSP.	✓	Not applicable
D.6	The testing company must manage the distribution of the test solution (whether hardware or software) and monitor the volunteer base to	×	As the pilot was an internal program, we considered it appropriate for us to distribute hardware probes to volunteers directly. However, we would require the testing

	ensure that samples remain active and representative.		company to both distribute and monitor the probes in any ongoing program to ensure that samples remain active and representative.
D.7	The testing company must provide complaint handling and technical support services for volunteers as part of the day-to-day management of the program.	×	As the pilot was an internal program, we considered it appropriate for us to handle feedback and provide technical support to volunteers directly. However, we would require the testing company provide complaint handling and technical support services in any ongoing program as part of the day-to-day management of the program.

E. Data analysis

Position Paper reference	Position Paper requirements for an ongoing program	Pilot Program adopted these specifications	Why the Pilot Program differed from the Position Paper requirements
E.1	Results must be collated and 'washed' through statistical analysis to ensure that outliers and/or errors are removed. For example, tests run after shaping or throttling due to excess data usage would need to be removed to avoid skewing the overall results.	×	As the Pilot Program was used to demonstrate what data we could obtain from a broadband monitoring program, results were not required to be 'washed' through statistical analysis to ensure that outliers and/or errors are removed. However, any ongoing program would require that statistical analysis be applied to results.
E.2	In the case of ADSL and VDSL services, statistical methods would need to be applied to address potential variance based on copper line quality. This requirement is additional to the sample selection requirement described in section D.3.	×	As outlined in D.3 above, the Pilot Program did not collect information on copper line lengths. However, any ongoing program would require application of statistical methods to address potential variance based on copper line quality.
E.3	Ownership of the raw data generated by the testing program would reside with the ACCC, and this would be used for other purposes and/or disclosed to other parties as the ACCC considered appropriate.	✓	Not applicable

F. Reporting approach

Position Paper reference	Position Paper requirements for an ongoing program	Pilot Program adopted these specifications	Why the Pilot Program differed from the Position Paper requirements
F.1	The reporting of results would be managed by the ACCC rather than the company retained to conduct the testing and preliminary data analysis.	✓	Not applicable
F.2	The testing and data analysis setup would need to support a 'hybrid' reporting approach whereby the ACCC would publish summary results on a limited set of metrics monthly, and a more fulsome report once or twice per year.	×	As the pilot was a once-off internal program, it was not necessary to publish any ongoing reporting in the format prescribed in the Position Paper. However, any ongoing program would be required to address all the reporting requirements prescribed in the Position Paper.
F.3	Monthly summary data would be presented primarily in graphical form (potentially via an interactive website) with some standing commentary, explanatory statements and disclaimers but no detailed commentary. At a minimum the metrics covered would include average peak-hour and 24-hour throughput both numerically and as a proportion of advertised speed. Broader 'quality of service' metrics would also be included, but a decision on which metrics to include would be made once the overall 'shape' of the testing program had been finalised.	×	
F.4	Detailed reports would include trend analysis, full explanation of more technical metrics, commentary from RSPs represented in the reports and supplementary information on the testing methodology and the sample characteristics.	×	
F.5	The ACCC would not seek to incorporate pricing or plan configuration data into either the monthly or detailed reports due to the frequency with which these change.	✓	Not applicable

F.6	The ACCC would seek the cooperation of other organisations such as ACCAN and the TIO to help promote the program to consumers.	×	As the pilot was an internal program, it was not necessary to seek assistance to help promote the program to consumers. However, the ACCC would seek the cooperation of other organisations in any ongoing program.