



Measuring Broadband New Zealand

Data dictionary

In 2018, the Commerce Commission appointed SamKnows to measure New Zealand's internet performance. The programme, called Measuring Broadband New Zealand, gives internet users in New Zealand access to SamKnows Whiteboxes to measure the quality of their fixed-line internet. The aim of the programme is to increase transparency about actual in-home broadband performance and provide consumers with independent information about internet performance across different providers, plans, and technologies, to help them choose the best broadband for their homes. It will also encourage providers to improve and compete on their performance.



Data dictionary

This data dictionary details how to interpret the fields contained in the raw data files. The raw data files contain the test results and environmental data used to produce the Autumn 2019 Report.

Raw measurement data

Measurements are presented in their raw, unaggregated form. However, only measurements that were used within the report have been included in this raw data package. This means that tests other than download, upload, latency, loss, Netflix and YouTube have been excluded. Additionally, metadata fields that were not used in the report have been excluded.

File listing

| File | Description |
|---------------------------------------|--|
| ./raw_data | Contains all of the raw measurement data files |
| ./raw_data/curr_httpget.csv | Download speed test data |
| ./raw_data/curr_httppost.csv | Upload speed test data |
| ./raw_data/curr_udplacency.csv | Latency and packet loss data |
| ./raw_data/curr_youtube.csv | YouTube video streaming |
| ./raw_data/curr_nflx.csv | Netflix video streaming |

Data dictionary

curr_httpget.csv (Download speed)

| Field Name | Description |
|--------------------|--|
| unit_id | Unique identifier for an individual unit |
| dtime | The time of the test (local time) |
| target | Hostname of the test server |
| address | IP address of the test server |
| fetch_time | Time the test ran for in microseconds |
| bytes_total | Total bytes downloaded across all connections |
| bytes_sec | Running total of throughput, which is sum of speeds measured for each stream (in bytes/sec), from the start of the test to the current interval. Multiply by 0.000008 to convert to Mbit/s |
| bytes_sec_interval | Throughput at this interval. Will match bytes_sec when only a single interval is reported on. Multiply by 0.000008 to convert to Mbit/s |
| warmup_time | Time consumed for all the TCP streams to reach a warmed-up state (Units: microseconds) |
| warmup_bytes | Bytes transferred for all the TCP streams during the warm-up phase |
| sequence | The interval that this row refers to. Will always be 0 when only reporting on a single interval |
| threads | The number of concurrent TCP connections used in the test |
| successes | Number of successes (always 1 or 0 for this test) |
| failures | Number of failures (always 1 or 0 for this test) |
| location_id | Please ignore (this is an internal key mapping to unit profile data) |

Data dictionary

curr_httppost.csv (Upload speed)

| Field Name | Description |
|--------------------|--|
| unit_id | Unique identifier for an individual unit |
| dtime | The time of the test (local time) |
| target | Hostname of the test server |
| address | IP address of the test server |
| fetch_time | Time the test ran for in microseconds |
| bytes_total | Total bytes downloaded across all connections |
| bytes_sec | Running total of throughput, which is sum of speeds measured for each stream (in bytes/sec), from the start of the test to the current interval. Multiply by 0.000008 to convert to Mbit/s |
| bytes_sec_interval | Throughput at this interval. Will match bytes_sec when only a single interval is reported on. Multiply by 0.000008 to convert to Mbit/s |
| warmup_time | Time consumed for all the TCP streams to reach a warmed-up state (Units: microseconds) |
| warmup_bytes | Bytes transferred for all the TCP streams during the warm-up phase |
| sequence | The interval that this row refers to. Will always be 0 when only reporting on a single interval |
| threads | The number of concurrent TCP connections used in the test |
| successes | Number of successes (always 1 or 0 for this test) |
| failures | Number of failures (always 1 or 0 for this test) |
| location_id | Please ignore (this is an internal key mapping to unit profile data) |

Data dictionary

curr_udplacency.csv (Latency and packet loss)

| Field Name | Description |
|-------------|--|
| unit_id | Unique identifier for an individual unit |
| dtime | The time of the test (local time) |
| target | Hostname of the test server |
| rtt_avg | Average round-trip time in microseconds |
| rtt_min | Minimum round-trip time in microseconds |
| rtt_max | Maximum round-trip time in microseconds |
| rtt_std | Standard deviation round-trip time in microseconds |
| successes | Number of successful packets (note: use failures/(successes + failures)) for packet loss |
| failures | Number of packets lost |
| location_id | Please ignore (this is an internal key mapping to unit profile data) |

Data dictionary

curr_youtube.csv (YouTube video streaming)

| Field Name | Description |
|--------------------------|--|
| unit_id | Unique identifier for an individual unit |
| dtime | Time test finished in UTC |
| video_id | The YouTube video ID (will change frequently, as the most popular video changes) |
| codec | The codec in use for the video |
| download_duration | The test duration in microseconds |
| stall_events | The number of stall events encountered |
| stall_duration_avg | The average stall duration in microseconds |
| stall_duration_total | The total stall duration in microseconds |
| video_duration | The duration of the video we fetched in microseconds |
| prebuffering_duration | From the DNS request until 2 seconds of video have been downloaded, in |
| bytes_sec | The download speed in bytes per second. Note that YouTube rate-limit server side |
| video_itag | The YouTube format identifier of the video |
| video_download_bytes_sec | The video download speed in bytes per second. Note that YouTube rate-limit server side |
| video_download_bytes | The amount of bytes of video download |
| video_download_duration | The duration spent downloading video in microseconds |
| video_hostname | Hostname of server |
| video_address | IP address of the server |
| video_connect_time | How many microseconds it took to establish a TCP connection to the video server |
| video_bitrate | The bitrate of the video we chose to fetch (in bytes per second) |
| audio_itag | The YouTube format identifier of the audio |
| audio_download_bytes_sec | The audio download speed in bytes per second. Note that YouTube rate-limit server |
| audio_download_bytes | The amount of bytes of audio download |
| audio_download_duration | The duration spent downloading audio in microseconds |
| audio_hostname | Hostname of server |
| audio_address | IP address of the server |
| audio_connect_time | How many microseconds it took to establish a TCP connection to the audio server |
| audio_bitrate | How many microseconds it took to establish a TCP connection to the audio server |
| max_bitrate | The maximum available bitrate for this video (so we can calculate video_bitrate/ |
| web_connect_time | How long in microseconds it took to connect to the YouTube web server |
| startup_delay | The time in usec to download the HTML page from "www.youtube.com" and then |
| range | The size of the read-ahead buffer, will always be 40 seconds by default |
| error_code | An internal error code from the test. 600 and 604 mean "no error" |
| error_msg | A detailed error message if something went wrong |
| successes | Number of successes (always 1 or 0 for this test) |
| failures | Number of failures (always 1 or 0 for this test) |
| location_id | Please ignore (this is an internal key mapping to unit profile data) |

Data dictionary

curr_nflx.csv (Netflix video streaming)

| Field Name | Description |
|-----------------------|--|
| unit_id | Unique identifier for an individual unit |
| dtime | The time of the test (local time) |
| target | Netflix OCA hostname |
| address | Netflix OCA IP address |
| bitrate | Bitrate of emulated video in bytes/sec |
| max_bitrate | The maximum bitrate supported by Netflix (currently always 15.6Mbps) in bytes/sec |
| stall_events | How many times has it stalled at this bitrate |
| stall_duration_total | Sum off the durations (usec) of all the stalls |
| connect_time | Time it took to establish the TCP connection with the content server in usec |
| download_duration | Duration of the test in usec |
| prebuffering_duration | How long it took to fetch 2 seconds of video at the specified bitrate in usec |
| bytes_sec | Download speed in bytes/sec after the prebuffering finished |
| error | Detailed error code: NO_ERROR / |
| error_msg | Detailed error message if a problem occurred fetching the content from Netflix |
| successes | 1 if the test runs for the full duration (may have stalls though) i.e. it was not aborted and stepped down |
| failures | 1 if the test was aborted for some reason |
| location_id | Please ignore (this is an internal key mapping to unit profile data) |