v. Test Node Selection

Having a geographically diverse set of test nodes would be of little use if the Whiteboxes running the test did not have a suitable mechanism to determine which node was the "best" to use. "Best" here is used to mean the node with the lowest round trip time between itself and the panelist's Whitebox.

The node actually selected might not always be the geographically closest test node to the panelist; the network route between the panelist's home and the test node will often travel via an indirect route that may take it through one or more cities. This might make another test node that is physically farther away preferable.

To identify nodes with the lowest round trip time, the Whitebox fetched a complete list of test nodes from the SamKnows infrastructure upon first execution of the test batch and performed a simple round trip time measurement to each. It then selected the test node with the lowest round trip time to test against from that point forward.

D. SamKnows Methodology[20]

Each deployed Whitebox performs the following tests:

Test	Primary measure(s)
Download speed	Throughput in Megabits per second (Mbps) utilizing three concurrent TCP connections
Upload speed	Throughput in Mbps utilizing three concurrent TCP connections
Web browsing	The total time taken to fetch a page and all of its resources from a popular website
UDP latency	Average round trip time of a series of randomly transmitted UDP packets distributed over a long timeframe
UDP packet loss	Fraction of UDP packets lost from UDP latency test
Video streaming	The initial time to buffer, the number of buffer under-runs and the total time for buffer delays[21]
Voice over IP	Upstream packet loss, downstream packet loss, upstream jitter, downstream jitter, round trip latency
DNS resolution	The time taken for the ISP's recursive DNS

---

[20] Specific questions on test procedures may be addressed to team@SamKnows.com.
[21] Only the total buffer delay is presented in the tabular results spreadsheet. Results of all tests are in the raw bulk data files.

	resolver to return an A record for a popular website domain name
DNS failures	The percentage of DNS requests performed in the DNS resolution test that failed
ICMP latency	The round trip time of five regularly spaced ICMP packets
ICMP packet loss	The percentage of packets lost in the ICMP latency test
Latency under load	The average round trip time for a series of regularly spaced ICMP packets sent during downstream/upstream sustained tests
Availability[22]	The total time the connection was deemed unavailable for any purpose, which could include a network fault or unavailability of a measurement point
Consumption	A simple record of the total bytes downloaded and uploaded by the router

The following sub-sections detail the methodology used in each of the individual tests.

DOWNLOAD SPEED AND UPLOAD SPEED

These tests measured the download and upload throughput by performing multiple simultaneous HTTP GET and HTTP POST requests to a target test node.

Binary, non-zero content—herein referred to as the payload—was hosted on a web server on the target test node. The test operated for a fixed duration of 30 seconds. It also recorded average throughput at 5 second intervals during the test. The client attempted to download as much of the payload as possible for the duration of the test.

The test used three concurrent TCP connections (and therefore three concurrent HTTP requests) to ensure that the line was saturated. Each connection used in the test counted the numbers of bytes transferred and was sampled periodically by a controlling thread. The sum of these counters (a value in bytes) divided by the time elapsed (in microseconds) and converted to Mbps was taken as the total throughput of the line.

Factors such as TCP slow start and congestion were taken into account by repeatedly transferring small chunks (256 kilobytes, or kB) of the target payload before the real testing began. This 'warm up' period was said to have been completed when three

---

[22] The measurement of availability provided a check on how often tests could not be run and was used as a quality metric overall, but was not used in analysis of broadband performance.

consecutive chunks were transferred at within 10% of the speed of one another. All three connections were required to have completed the warm up period before the timed testing began. The 'warm-up' period was excluded from the measurement results.

Downloaded content was discarded as soon as it was received, and was not written to the file system. Uploaded content was generated and streamed on the fly from a random source.

WEB BROWSING

The test recorded the averaged time taken to sequentially download the HTML and referenced resources for the home page of each of the target websites, the number of bytes transferred, and the calculated rate per second. The primary measure for this test was the total time taken to download the HTML front page for each web site and all associated images, JavaScript, and stylesheet resources. This test did not test against the centralized testing nodes; instead it tested against real websites, ensuring that the effects of content distribution networks and other performance enhancing factors could be taken into account.

Each Whitebox tested against the following 10 websites (derived from a list generated by Alexa of the top 20 websites in October 2010[23]):

- http://www.cnn.com

- http://www.youtube.com

- http://www.msn.com

- http://www.amazon.com

- http://www.yahoo.com

- http://www.ebay.com

- http://www.wikipedia.org

- http://www.facebook.com

- http://www.google.com

- http://www.netflix.com[24]

The results include the time taken for DNS resolution. The test used up to eight concurrent TCP connections to fetch resources from targets. The test pooled TCP

---

[23] *See* http://www.alexa.com/.

[24] During March 2011 Netflix began enforcing the use of secure web connections to its main web site using TLS security. This was not compatible with the deployed Whitebox software. Consequently, Whitebox web connection requests were rejected, resulting in test failures being recorded in all results after this security feature was implemented. As a result, this website was excluded from this test.

connections and utilized persistent connections where the remote HTTP server supports them.

The client advertised the user agent as Microsoft Internet Explorer 8. Each website was tested in sequence and the results summed and reported across all sites.

UDP LATENCY AND PACKET LOSS

These tests measured 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 was not received back within three seconds of sending, it was treated as lost. The test recorded the number of packets sent each hour, the average round trip time and the total number of packets lost. The test computed the summarized minimum, maximum, and mean from the lowest 99% of results, effectively trimming the top (*i.e.*, slowest) 1% of outliers.

The test operated continuously in the background. It was configured to randomly distribute the sending of the echo requests over a fixed interval of one hour, reporting the summarized results once the interval had elapsed. Approximately 600 packets were sent within a one hour period, with fewer packets sent if the line was not idle.

This test was started when the Whitebox booted and ran permanently as a background test.

VIDEO STREAMING

For the purpose of the video streaming test, the intent was to simulate an end user viewing a streaming video online. This test emulated live video streaming rather than a service such as YouTube that employs a 'progressive download' approach.

The test operated over TCP and used a proprietary client and server side components. The client and server negotiated the test parameters at the start of each test.

A three-second playout buffer was configured and the client attempt to download data from the server at the maximum rate necessary to ensure that this buffer was never empty. A separate client-side thread consumed data from this buffer at a fixed rate, looking for buffer under-runs (which would manifest themselves to users as a pause in video). The Whitebox recorded the time to initial buffer, the total number of buffer under-runs and the total delay in microseconds due to these under-runs.

The test operated at four bit rates: 768 kilobits per second (kbps), 1.25 Mbps, 2.25 Mbps, and 3.75 Mbps.

VOICE OVER IP

The test operated over UDP and, unlike the video streaming test, utilized bi-directional traffic, as is typical for voice calls.

The Whitebox would handshake with the server, and each would initiate a UDP stream with the other. The test used a 64 kbps stream with the same characteristics and properties (*i.e.*, packet sizes, delays, bitrate) as the G.711 codec. The test measured jitter, delay and loss. These metrics were measured by subdividing the stream into blocks, and measuring the time taken to receive each block (as well as the difference between consecutive times).

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

DNS RESOLUTION AND DNS FAILURES

These tests measured the DNS resolution time of an A record[25] query for the domains of the websites used in the web browsing test, and the percentage of DNS requests performed in the DNS resolution test that failed.

The DNS resolution test was targeted directly at the ISP's recursive resolvers. This circumvented any caching introduced by the panelist's home equipment (such as another gateway running in front of the Whitebox) and also accounted for panelists that might have configured the Whitebox (or upstream devices) to use non-ISP provided DNS servers. ISPs provided lists of their recursive DNS servers for the purposes of this study.

ICMP LATENCY AND PACKET LOSS

These tests measured the round trip time (RTT) of ICMP echo requests in microseconds from the Whitebox to a target test node. The client sent 5 ICMP echo requests of 56 bytes to the target test node, waiting up to three seconds for a response to each. Packets that were not received in response were treated as lost. The mean, minimum, maximum, and standard deviation of the successful results were recorded.

LATENCY UNDER LOAD

The latency under load test operated for the duration of the 30 second downstream and upstream speed tests, with results for upstream and downstream recorded separately. While the speed tests ran, the latency under load test sent 20 ICMP echo packets to the target server and measured the round trip time. Packets were spaced 500 milliseconds (ms) apart, and a 3 second timeout was used. The test recorded the mean, minimum and maximum round trip times in microseconds. The number of lost ICMP echo requests was also recorded.

This was an early version of the latency under load test and was incorporated following input from MIT. Enhancements to the test (such as making it use UDP datagrams rather than ICMP packets) will be incorporated into future versions.

---

[25] An "A record" is the numeric IP address associated with a domain address such as www.fcc.gov.

AVAILABILITY TEST

This test measured the availability of the network connection from the Whitebox to multiple target test nodes by sending and receiving TCP segments to a receiving server located on each test node.

The Whitebox established long-lived TCP connections to the server on each test node, periodically sending TCP packets containing a timestamp in microseconds.

The server echoed back the same data to the Whitebox and if it failed to respond or the connection was reset via TCP RST or FIN then the Whitebox would attempt to re-establish the connection. If the Whitebox was unable to re-establish the connection to all three servers simultaneously, it was inferred that Internet connectivity was at fault, and the test recorded a failure locally, along with a timestamp to record the time of failure.

To aid in diagnosing the point in the route to the target test nodes where connectivity failed, a traceroute was launched to all target test nodes, the results of which were stored locally until connectivity was resumed and the results could be submitted.

This test was started when the Whitebox booted, and ran permanently as a background test.

Table 5: Estimated Total Traffic Volume Generated by Test

Test	Target(s)	Duration	Total Est. Daily Volume
Web browsing	10 popular US websites	Est. 30 seconds	80 MB
Video streaming[1]	1 off-net test node	Fixed 10 seconds at 768kbps, 1.25Mbps, 2.25Mbps, 3.75Mbps	60 MB
	1 on-net test node	Fixed 10 seconds at 768kbps, 1.25Mbps, 2.25Mbps, 3.75Mbps	60 MB
Voice over IP	1 on-net test node	Fixed 30 seconds at 64k	1 MB
	1 off-net test node	Fixed 30 seconds at 64k	1 MB
Download speed[2, 3]	1 off-net test node	Fixed 30 seconds***	4.5 GB at 50Mbps / 1.72 GB at 20Mbps / 858 MB at 10Mbps / 357 MB at 3Mbps / 129 MB at 1.5Mbps
	1 on-net test node	Fixed 30 seconds***	4.5 GB at 50Mbps / 1.72 GB at 20Mbps / 858 MB at 10Mbps / 357 MB at 3Mbps / 129 MB at 1.5Mbps
Upload speed[2, 3]	1 off-net test node	Fixed 30 seconds***	174 MB at 2Mbps / 87 MB at 1Mbps / 44 MB at 0.5Mbps
	1 on-net test node	Fixed 30 seconds***	174 MB at 2Mbps / 87 MB at 1Mbps / 44 MB at 0.5Mbps
UDP latency	1 off-net test node	Permanent	1 MB
	1 on-net test node	Permanent	1 MB
UDP packet loss	1 off-net test node	Permanent	N/A (uses above)
	1 on-net test node	Permanent	N/A (uses above)
Consumption	N/A	N/A	N/A
Availability	3 off-net test nodes	Permanent	1 MB
DNS resolution	10 popular US websites	Est. 3 seconds	0.3 MB
DNS failures	10 popular US websites	N/A (As DNS resolution)	
ICMP latency	1 off-net test node / 1 on-net test node	Est. 5 seconds	.3MB
Latency under Load	1 off-net test node / 1 on-net test node	Est. 5 seconds	.3MB
ICMP packet loss	1 off-net test node / 1 on-net test node	N/A (As ICMP latency)	N/A (uses above)

[1] Video streaming rates: Lines will only stream the rates they are capable of, according to the latest speed test results. If a rate is deemed unreachable (e.g. a 3.75Mbps rate on a 1Mbps line), then it will be skipped.

[2] Download/upload daily volumes are estimates based upon likely line speeds. All tests operated at maximum line rate so actual consumption may vary.

[3] Frequency: twice every hour. All other tests report results once per hour.

In addition to the tests described above, for 60 seconds prior to and during testing, a 'threshold manager' service on the Whitebox monitored the inbound and outbound

traffic across the WAN interface to calculate if a panelist was actively using the Internet connection. The threshold for traffic was set to 64 kbps downstream and 32 kbps upstream. Statistics were sampled and computed every 10 seconds. If either of these thresholds was breached, the test was delayed for a minute and the process repeated. If the connection was being actively used for an extended period of time, this pause and retry process would continue for up to 5 times before the entire test cycle was abandoned.