Body of European Regulators for Electronic Communications

BEREC

BoR (22) 72

Net Neutrality Regulatory Assessment Methodology

9 June 2022

Table of Contents

1. Executive summary

This document contains BEREC's regulatory assessment methodology intended to provide guidance to National Regulatory Authorities (NRAs) in relation to the monitoring and supervision of the net neutrality provisions of the Open Internet Regulation 2015/2120 [1] ("the Regulation"), and the possible implementation of net neutrality measurement tools on an optional basis. This methodology is also intended to contribute to the harmonisation of net neutrality measurement methodologies. This updated work builds upon previous BEREC guidance on net neutrality, internet access service (IAS) quality monitoring and best practices.

Chapter 3 provides guidance on a harmonised quality of service measurement methodology. It is targeted at maximising measurement accuracy, consistency and enabling the comparison of measurement results between different Member States. The speed measurement methodology is based on multiple transport layer connections by default, and the document describes the subsequent calculation of the measured speed. This document also defines measurement approaches for delay, delay variation and packet loss.

Chapter 4 gives recommendations on methods for detecting traffic management practices that impact individual applications and includes recommendations for detecting traffic management practices that affect the connectivity and ultimately a possibility to use and provide individual applications.

Chapter 5 describes the most important factors that should be considered when assessing the measurement results and gives guidance on information collection. Thus, a number of end-user environment factors may impact the results. These factors include for example Wi-Fi usage, modem and computer performance and radio conditions when measuring speed for mobile subscription.

Chapter 6 provides recommendations for the validation, post-processing and market level analysis of the collected measurement results. The topic of data aggregation for market level assessment purposes is discussed and guidance on monitoring the general IAS quality (IAS as a whole and effect of specialised services on general IAS) as well as individual applications using IAS is provided.

Chapter 7 provides some further guidance on how the speed measurement results should be assessed in comparison to the contractual speed values for end-users.

Finally, Chapter 8 gives guidance on the criteria that NRAs could take into account when providing their own certified monitoring mechanism or certifying a third-party mechanism, while Chapter 9 talks about data protection requirements

2. Introduction

BEREC has developed and subsequently updated this regulatory assessment methodology to help NRAs in the monitoring and supervision of the net neutrality provisions of the Regulation [1] based on various net neutrality measurement tools and harmonised measurement methodology for quality of service indicators.

Furthermore, it is also foreseen that this BEREC net neutrality measurement methodology could contribute to the work of standardisation bodies. The methodology builds upon previous BEREC guidance on net neutrality, internet access service (IAS) quality monitoring and best practices [2].

Under the Regulation, NRAs may have several objectives in measuring IAS:

- Measurement tools can be used by individual end-users to validate that the commitments made to them by the IAS provider (Article 4(1) of the Regulation) are being delivered;
- Measurement tools can be used to detect traffic management practices which may or may not be allowed (Article 3(3) of the Regulation);
- Measurement tools can be used for the establishment of what the 'general quality of IAS' is. This is relevant to assess whether services other than IAS (in the meaning of Article 3(5) of the Regulation) can be provided;
- Measurement tools may be part of a monitoring mechanism certified by the NRA as referred to in Article 4(4) of the Regulation.

NRAs may wish to achieve additional objectives when measuring or detecting certain practices related to the IAS. BEREC notes that it is up to the NRAs to determine the most appropriate measurement tools to serve their objectives. Different objectives may lead to the use of different measurement tools.

In this document, BEREC describes the methodology for the measurement of IAS speed to enable NRAs to assess IAS performance compared to the contractual minimum, normally available and maximum speed values. The methodology also gives guidance on some criteria that NRAs could take into account when providing their own measurement tools as a certified mechanism or certifying a third-party mechanism in accordance with the Regulation and BEREC OI Guidelines [3].

This document aims to describe a measurement methodology that could be combined with a crowdsourcing approach so that it would be possible to provide measurement tools for a large number of end-users. For in-browser or app-based crowdsourcing measurement tools, it is hard or even impossible to have full control over all the factors such as the end-user environment that impact measurement results. This introduces a possibility for error in measurement results that cannot be fully avoided. This methodology provides guidance on how to increase the accuracy and reliability of such measurement results. This is discussed in Chapter 5.

BEREC recognises the standardised measurement approaches from ETSI, ITU and IETF, however BEREC places the greatest emphasis on the practical implementation of the measurement methodology in a crowdsourced scenario, where measurements can be run by any end-user and the metric being measured reflects as closely as possible real end-user experience of internet usage. This is discussed in more detail in Section 3.1.6.

As proposed in the 2012 BEREC NN QoS Guidelines [2], the measurement methods shall encompass

both the IAS as a whole as well as individual applications delivered via IAS. The methodology supports both IPv4 and IPv6 – this topic is further discussed where necessary.

The updates in this document are considered to be broadly compatible with the previous version [8], and as a result BEREC does not anticipate that this update will result in a requirement for significant changes to pre-existing measurement systems. This should be considered on a case-by-case basis however.

3. Measuring internet access service quality

The aim of this chapter is to describe a recommended methodology to measure IAS quality based on the combined goal of maximising measurement accuracy balanced against the need to be able to facilitate easy access to the measurement tool for the public ensuring that the measurement results are comparable between different Member States. The recommended methodology covers a comprehensive number of topics. All are broadly and generically discussed. It is then up to the NRAs to interpret these notions based on their explicit scenarios and adapt them accordingly, if applicable.

Results of these measurements can be also used for the following purposes:

- Empowering the end-user to validate the commitments made to them by their IAS provider;
- Monitoring the general IAS quality to support the confirmation that the performance of IAS is developing sufficiently over time when taking into account technological evolution (see Chapter 6);
- NRAs may also use the data to increase transparency (e.g. interactive maps showing performance in a geographic area);
- To support the detection of traffic prioritisation and/or throttling of selected applications compared to other applications running over IAS (see Section 4.2).

According to paragraph 166 of the BEREC OI Guidelines [3], "*[m]easurements should be performed beyond the ISP leg*" and speed should be calculated "*based on transport layer protocol payload*". In addition, paragraph 140 says that "*[s]peeds should be specified on the basis of the transport layer protocol payload, and not based on a lower layer protocol*".

The methodology described here is targeted at the measurement of IAS quality in both the upload and download directions. It is worth noting that IAS speed (Section 3.1) is just one component of the performance experienced by the end-users, since different applications have different requirements related to IAS delay, delay variation (Section 3.2) and packet loss (Section 3.3).

For measurement tasks involving both the IAS as a whole and individual applications using IAS, the fundamental precondition is that measurements are performed at the edge of the network which provides the IAS (i.e. the modem for fixed access or via the radio access for mobile IAS). It is also worth noting that on some network types, resources need to be allocated before full bandwidth becomes available and that the time period in which resource allocation takes place can be perceived as a phase of reduced bandwidth, thus reduced quality. In some older networks (like 2.5G/3G mobile networks), such a resource-allocation phase can last several-hundred milliseconds and thus have an impact on the measured quality. Such transitional behaviour may be recorded, or it might be discarded depending on the purpose of the measurement.

Where measurements are performed against a test server, this server should be located outside the IAS network[1]. There should be adequate connectivity between the server and the IAS provider to

[1] This is in line with the architectural aspects of a practical measurement system as described by BEREC in previous publications [5].

minimise any influence upon the measurements. Typically, this can be achieved by locating the measurement server at, or close to the national internet exchange point (IXP). Depending on the specific national situation, measurement servers may be located at more than one IXP location[2]. There could also be specific reasons for a measurement server's placement elsewhere, which should be assessed, In the event that multiple server locations are available, the tool should select which one to use based on appropriate criteria.

The hardware running the measurement server(s) should be connected as close to the IXP switch as possible, to minimise latency added due to the communication paths. This means that the number of hops between the main IXP switch and the test server should be kept at a minimum. This is applicable whether the implementation runs inside the network of a hosting provider or directly on hardware under control of the NRA itself.

Since test traffic is internet traffic, BEREC recommends that NRAs ensure that measurement traffic is handled equally to other traffic.

In cases where both the server and client have visibility of test results (e.g. speed measurements), generally the receiver of the data is considered to be the authoritative source of measurement results, but it is recommended that both the client and server measurements are stored for analysis.

In many cases, the primary output of a measurement (e.g. download/upload speed (in bit/s), latency (in ms) or packet loss (as a percentage)) is the one provided to the end-user. However, it is often appropriate for a greater level of detail to be recorded than is displayed to the end-user.

Client-side monitoring mechanisms should mitigate (or at least, to identify), to the extent possible, confounding factors which are internal to the end-user environment. Examples of these factors include existing cross-traffic and the usage of Wi-Fi based interfaces. This topic is discussed separately in Chapter 5.

Server-side monitoring mechanisms should monitor the available capacity of the measurement server such as CPU, memory, OS resources, etc. to detect bottlenecks. Test results which may have been compromised as a result of such bottlenecks should reflect this fact.

The assessment of measurement results is discussed further in Chapters 6 and 7. The certified monitoring mechanism is further discussed in Chapter 8.

3.1. IAS speed measurements

3.1.1. Speed measurement overall methodology

BEREC has based this measurement methodology work on the following NRA requirements for IAS measurements, in particular speed measurements, in a regulatory context:

- Multi-platform - where a speed measurement is initiated by a human end-user, it must be

[2] While the general quality of internet access is best characterised as described, specific endpoints of the internet might have different quality and thus justify additional measurements. Such endpoints might be specific CDNs or some endpoint on the *edge* (implementing the broad concept of edge-computing which is often mentioned in the context of 5G). Installing measurement server components on such specific endpoints might be difficult as it is typically under the control of a third party.

possible to execute it via the equipment that they usually use to access the IAS. No artificial restrictions in the methodology should prevent the measurement from running on other hardware such as game consoles/modem clients/TV-boxes etc.

- End-users measuring their IAS speed should continue to be supported within a web browser or within the restricted sandbox of an on-device app; the methodology must neither require nor prohibit the installation of personal computer client software.
- The resulting speed measurement must objectively reflect the speed available to the end-user, which might be affected by factors such as packet loss or latency.
- The time required to execute an individual speed measurement should be short enough to allow an end-user to complete a measurement without frustration. The duration of the measurement should also consider the volume of data transferred and make it transparent to the end-user.
- While it is recognised that any measurement implementation will have a lower/upper limit of speeds over which acceptable accuracy is returned, the methodology should support the typical speeds available in the relevant markets. In particular, the methodology should be ready to scale to support the speeds delivered over optical and 5G networks.
- The methodology must support both IPv4 and IPv6.

While these requirements primarily focus on speed measurements, they are considered relevant for all IAS measurements in this chapter.

To maximise compatibility in a real-world environment, it is still recommended to measure upload/download speeds based on the time to execute a parallel set of controlled data transfers over HTTP(S). In this way, the speed can be measured based on transport layer protocol payload as referred to in paragraph 140 of the BEREC OI Guidelines [3].

This methodology is supported by the broadest range of relevant platforms, and it conforms to the requirements above. As such it is considered as an adequate balance between the competing demands of accuracy, platform agnosticism, ease of implementation and transparency. This position is also supported by the widespread use of HTTP(S) by normal internet applications and services and therefore it reflects the typical IAS usage by end-users.

To saturate the path being measured, it is generally recommended to use multiple transport layer connections to a single server (or multiple servers at the same physical location) during a speed test. The number of connections may be set according to the characteristics of the path (including the estimated speed and latency). Such an estimation could be based on a pre-test or on the estimated performance of the service based on other information.

The test should be designed to keep the processing burden on the sending and receiving side low to allow the accurate measurement of high-speed connections even on devices with limited computing power.

In the interests of transparency and user friendliness, BEREC recognises the benefits of providing interim test information, such as a display of a graph, illustrating an approximation of the measured throughput over the duration of the test. As such, the test is specified in a way that allows the recording of the progress of data transfer on each connection during the test over short time intervals (e.g. 50-200 ms). This can be achieved by recording the progress of the test on each connection at fixed time-

based intervals or by the organisation of a continuous[3] transfer of blocks of measurement data of a given size. In case a block-based implementation is used, the size of the block (and thus the frequency of block reception) can be set based on a pre-test or on the estimated performance of the service. One consideration of such design is the need to control the resources needed on the endpoints to process the received data.

BEREC recognises that packet losses and consequent packet retransmissions have a negative impact on the throughput of reliable transport layer protocols, in this case TCP (Transmission Control Protocol), and hence the measured IAS speed.

BEREC's recommended methodology follows a few conceptual steps, outlined below. These steps are intended to be agnostic of the exact version of HTTP(S) used (where applicable), and whether the connections are upgraded to use web sockets or not.

Pre-test

A pre-test may be used to preliminarily evaluate the characteristics of the IAS service to ascertain some parameters (e.g. the number of connections, block size) relevant to the main test which follows. In addition, information about the network connection (e.g. cellular technology, physical link speed etc.) could be collected from the environment at this stage when this is possible[4].

Different methodologies could be used to perform the pre-test, for example executing a short copy of the main test. For the determination of the block size, an example strategy consisting of requesting blocks of multiples of a fixed size (e.g. 4096 bytes) until the transfer takes longer than a predefined time (e.g. 200 ms) could be used. The resulting amount of data (4096 byte ^n) might then be used as a reference for the block size in the main test.

Main Test

The main test starts with the receiving[5] side indicating that it is ready to commence measurement. The sending side then begins sending data on each transport layer connection in successive blocks without delay, potentially using the parameters determined during the pre-test phase. The transfer time and size of each block is recorded for subsequent speed calculation. To mitigate any effects of compression, the data transferred should be (pseudo-)random[6] data.

It should be noted that the client should only need to make a single request for test commencement, without needing to make multiple requests for data where possible as this would introduce unnecessary latency and affect the speed measurement.

This approach enables the receiving side to analyse the speed of each connection over the duration

[3] Continuous means that data is sent without (intended) interruptions as a continuous flow of bytes. The fact that internally data is organised in blocks does not imply that each block is requested individually.

[4] On fixed internet access, in order to collect information about the end-user environment (Ethernet vs. Wi-Fi, RSSI, cross-traffic, advertised speed, access technology, etc.), APIs may be used. In this context, Arcep specified an "Access ID API", implemented by ISPs in France in order to better characterise the end-user environment and make QoS measurement tests in fixed lines more reliable.

[5] For upload tests the recipient is the measurement server, whereas in download tests the recipient is the client (end user).

[6] The term pseudo-random means that the data should appear as random to a typical compression algorithm while its cryptographic quality remains irrelevant. Thus, pre-stored "random" data might be reused within a given test, or for subsequent measurements or data from the downlink test might be reused for the uplink test.

of the test rather than as a single total speed for each connection, thereby providing the interim visibility referred to above.

The test completes after a pre-defined interval, however it might be necessary for some connections to keep transferring “trailing blocks” (see diagram below) until all connections have completed transferring measurement blocks.

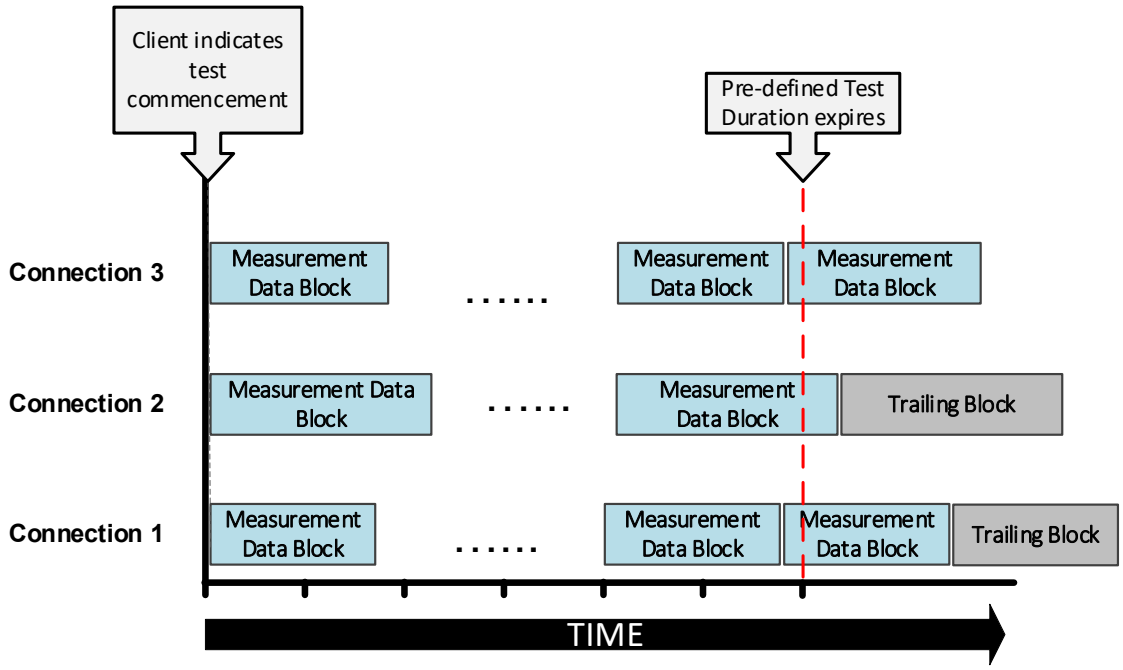

Figure 1 - Conceptual view of speed test

Figure 1 above is intended to illustrate this concept at a high level. Note that after the test duration has ended, “trailing blocks” continue to be sent over Connections 1 and 2 until the completion of the last measurement data block which started before the expiry of the test duration, in this case on Connection 3.

Speed Calculation

The outcome of the measurement will include information about the volume of data received on each connection per intermediate time interval from which the speed is calculated.

The resulting transfer speed is calculated by the recipient, based on the data received on all connections as part of all measurement blocks, including any header data[7]. It should be noted that calculating the data rate based on an average will smooth out the peaks seen in the highly bursty transfer rates seen on mobile networks.

[7] For example, HTTP headers or Web socket frame overhead. While relevant in principle, it should be noted that HTTP headers are comparably small compared with the amount of data transferred using the test. Thus, the error introduced by uncertainty regarding header sizes (e.g. due to compression) may be negligible. In no case should Ethernet or IP headers be included in the calculation.

During an upload test, to provide interim test information to the user, an approximation of the actual speed as measured by the sender could be used. Alternatively, the transfer speed measured by the server could be periodically communicated to the client.

The calculation might include the full period of measurement or might exclude one or more measurement blocks (or sample intervals where applicable) at the start of the test to avoid counting a period of slow-start at the transport layer and any other delay due to the connection setup.

In the event of an error on one of the connections during a test, the effective testing time might be reduced to end at the last correctly received block on that connection, the error should be recorded and taken into account. Depending on the time of the connection error, the entire measurement may be discarded.

Both download and upload speeds should be measured in the same manner and reported in bits/second (e.g., kbit/s or Mbit/s). Note that conversion factors between mega and kilo shall be base-10 rather than base-2 (i.e., 1 Mbit/s = 1000 kbit/s rather than 1024 kbit/s).

3.1.2. Underlying implementation details

HTTP version

While the choice of HTTP version is, in principle, an implementation choice, BEREC believes the following considerations to be relevant.

For HTTP/1.1 based tests, it is recommended that transfers are made using chunked transfer encoding which enables the sending side to send chunks of data of arbitrary size and stop the transfer at the appropriate time. In this case, each chunk would be represented as a measurement data block in Figure 1.

For example, the sender could send the measurement data blocks in fixed chunks of a size potentially determined in the pre-test, stopping when the test duration has fully passed. Irrespective of the implementation chosen, it is important that each chunk is served immediately and without any delay, and consideration given to including any additional data volume transferred as a result of Chunked Transfer Coding (or chunk extensions) in the speed calculation.

In cases where the above implementation is not possible, the HTTP content length should be as large as possible to minimise the number of consecutive back-to-back requests needed to perform the test.

HTTP/2 would multiplex all requests towards the same server on a single TCP socket by default, thus not using multiple transport layer connections as recommended above. To avoid this issue, multiple server endpoints at the same physical location (or on the same physical machine) could potentially be used, however further details on this are beyond the scope of this document.

At the time of writing, HTTP/3 is still quite new and details of its implementation within popular browsers are not fully known. While it is felt that HTTP/3 and QUIC are promising technologies which are likely to support BEREC’s speed testing use case in the future, BEREC cannot provide specific details of any possible implementation at this time.

Number of connections, BDP & TCP parameters

Tests should generally[8] be executed using multiple transport layer (in practice TCP) connections since the speed of an individual connection is constrained by the default maximum window size.

On links with a low Bandwidth-Delay Product (BDP), a single connection[9] may suffice to saturate the link, however the effect of any packet loss in reducing the measured speed may be more pronounced. Section 3.1.5 contains further information.

In cases where high BDP-value links are being measured, steps should be taken to ensure that the path can still be saturated so that accuracy is maintained. The most common approach is to further increase the number of transport layer/TCP connections, but it is also possible to modify the TCP settings for the server.

Implementations should handle typical BDP-values in scope for measurement.

The detrimental effect of packet loss due to physical reasons (e.g., bit errors because of radio interference in case of mobile or Wi-Fi connections) on the speed measurement depends on the TCP congestion algorithm used on client and server.

Based on the known TCP parameters which have been configured, combined with any path characteristics learned from the pre-test phase (if executed), implementations can decide on the appropriate number of connections to use in each test. However, browser limitations should be considered when increasing the number of connections, as it is possible that not enough connections can be enabled to saturate the path. The number of parallel connections in supported by a browser may also depend on the type of connection and the browser API used.

Web Sockets (RFC 6455)

Within HTTP(S), the WebSocket protocol may be used for the measurement as it provides a duplex connection over TCP with little overhead for sufficiently large WebSocket data frames. In this scenario, each measurement data block depicted in Figure 1 would equate to a WebSocket frame.

When using WebSockets, implementing parties should consider browser-specific performance differences on protocol configuration and connection options available to browsers[10]. When determining the size of WebSocket frames, the resource overhead introduced by the asynchronous nature of the WebSocket API available to browsers should be taken into account, aiming to minimise the number of frames per measurement interval (i.e. one frame per measurement data block, every 50-200 ms). To minimise client resources, WebSocket frames can be discarded immediately after receiving.

It should be noted that, even when using “sockets” via WebAssembly, these connections are, in practice, using the WebSocket protocol by browser implementations.

---

[8] In some cases a single connection test may be appropriate, see Section 3.1.5

[9] For further information on single connection vs. multiple connections tests see https://en.arcep.fr/uploads/tx_gspublication/report-state-internet-2019-eng-270619.pdf#page=16

[10] E.g. differences in performance depending on the “binaryType” specified in the API available to Browsers.

Browser APIs for HTTP

Browser based clients could opt to use APIs such as XMLHTTPRequest or the fetch API. In this case, it would be possible to set the size of measurement data blocks (and therefore the HTTP chunks) to be large enough so as to exceed the test duration and remove any potential for delay between measurement blocks. In this scenario it would be necessary to:

a) utilise "progress" events when using XMLHTTPRequest or to utilise the Stream API when using the fetch API to provide data on interim download progress (see Figure 2), and

b) close the connection to end the test once the pre-defined test duration has expired.

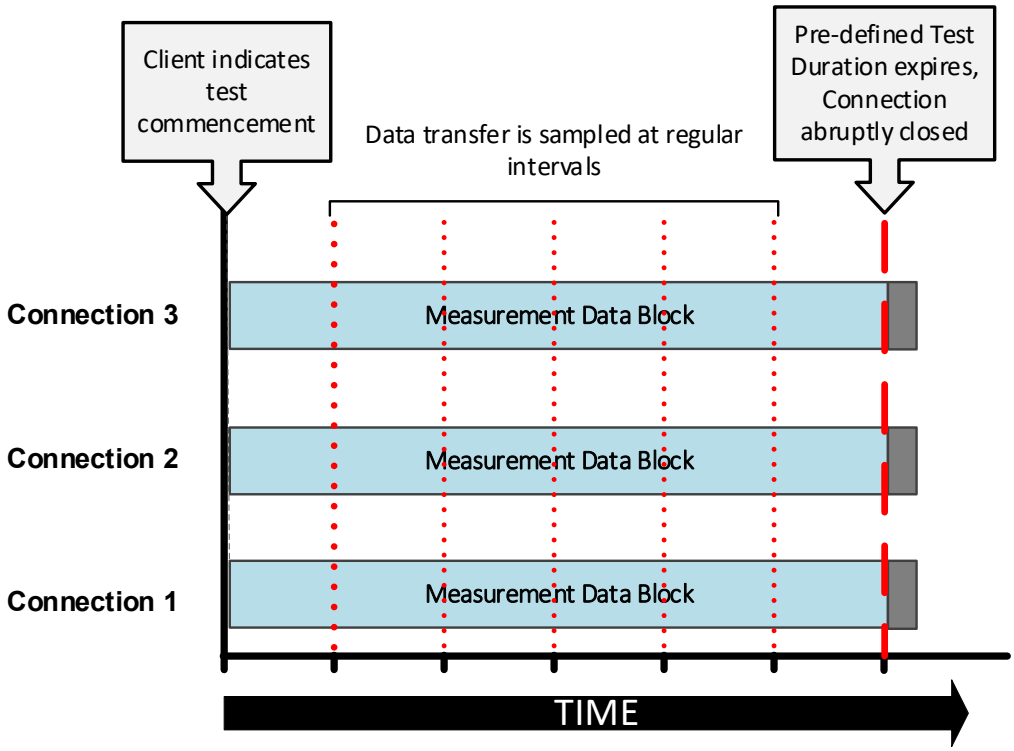

Figure 2 – Sampling test progress via browser based API

Care should be taken for the implementation to ensure the hardware/OS/browser combination is able to accurately report on the progress at regular intervals, that memory exhaustion is not an issue for resource constrained devices, and that connection closure does not result in any unforeseen issues.

Transport Layer Security (TLS)

Given the ever-increasing use of TLS (in the form of HTTPS) by websites on the internet, its use during speed tests is highly recommended except in specific cases where client hardware cannot sustain it without a performance impact

TLS provides the additional advantage of preventing any manipulation from intermediate proxy

servers[11] and does not introduce any significant difference in performance or computational load in end user equipment since encryption is often supported by hardware.

In the event that encryption is not used, unique URLs or the appropriate HTTP headers to prevent caching should be used.

Another reason to recommend the use of HTTPS on standard ports as the TCP payload, is to mitigate any connectivity restrictions (such as might be introduced by intermediate firewall or proxy servers in corporate environments) which could result from the choice of a less commonly used protocol/port.

3.1.3. Miscellaneous details

It must be possible to run measurements both over IPv4 and IPv6, and the version used for a given measurement should be recorded. In case both protocol versions are available to the user, they could be given the choice between these protocols.

3.1.4. Benchmarking the accuracy and stability of internet speed and delay measurements

If benchmarking the accuracy of a measurement system, it is recommended to perform benchmarking measurements over several actual internet connections and under controlled (laboratory) conditions as well to learn the accuracy, and the stability of that system.

These controlled laboratory conditions enable the checking of the accuracy of the speed and delay measurement in a precise way, while the tests performed on actual internet connections can reveal problems that do not occur in the laboratory environment.

One possible principle of the benchmarking of the speed and delay measurement accuracy of the measurement system is to set/generate defined bandwidth and delay values, and to perform repeated tests with the measurement system against these settings.

An additional way to benchmark is to use a “reference” measurement system and compare the measured results of the measurement system being tested with the results of the “reference” measurement system.

Several instruments can be used to set bandwidth values as well as the delay (round-trip time) for such tests. Such instruments are software-based bandwidth limiters (e.g. Linux NetEm), possible shaper functionality of Ethernet routers and of professional grade as well as calibrated instruments.

When checking the accuracy of speed measurements, it is recommended to define BDP-settings, given that the accuracy of speed measurement can be largely affected not only by the bandwidth of the connection but also by the delay of the connection (Bandwidth-Delay Product dependency).

The bandwidth range used for speed verification accuracy tests should cover the specified bandwidth range of the measurement system. Typical values for the delay should be selected for the tests. Multiple test steps (bandwidth-delay pairs) should be specified within the specific speed-latency

[11] While proxies would be prohibited by the open internet regulation for internet access service, they might still be used in office environments where end users may wish to make speed measurement for informational purposes. In this situation, running a test over TLS on port 443 would mitigate any interference from the proxy.

range. The tests should be repeated a minimum of five times for each BDP-pair. If the result is not sufficiently stable, more iterations are needed.

The pause time between each iteration should be chosen so that it does not cause instability in the measurement, as might happen if the shaper buffer is in an inconsistent state at the beginning of each test.

During the measurements, cross-traffic should be avoided, and it is recommended to monitor the CPU and memory usage of the hosts used for the measurement. For the purpose of the measurements, a client computer with a typical performance should be used. Tests should be performed and repeated on multiple combinations of operating systems and browsers.

The bandwidth value measured by the measurement system and the reference value (or conventional value) of the channel bandwidth should be specified at the same OSI layer. In the event that this is not possible, and the value of the bandwidth limit is specified at a lower OSI layer, it should be noted that this calculation is not always completely accurate (due to the TCP header options enabled / disabled by the TCP protocol during the test, for example), this can be considered as an additional uncertainty of the measurement setup.

3.1.5. Single connection vs. multi-connection QoS measurement

While the standard recommendation is to use multiple connections for an IAS speed test, in some cases a single connection test may be appropriate. For example, a single connection speed that is substantially lower than a multi-connection one could potentially reveal an issue worth investigating, for example if an IAS provider or end-user environment were to limit the speed of a single connection.

It is not uncommon for multi-connection tests to display a faster connection than single connection ones, which can be for several reasons as described in Section 3.1.2. Other potential reasons may include:

- Packet reception order: a connection whose important jitter or link aggregation misconfiguration, for example, makes it impossible to guarantee that packets will arrive in the right order will degrade connection speed considerably;
- Saturation of the terminal’s processor core: a single connection quality of service test may not employ all of a processor’s cores fully, unlike multi-connection tests.

3.1.6. Measurement methodologies specified by other organisations

BEREC notes that other organisations have also published methodologies on throughput measurement which differ from the one outlined here. Notably, the ITU and the Broadband Forum have issued standards based upon UDP-based IP capacity measurement methodologies.

In addition, the IETF has published RFC 9097 *Metrics and Methods for One-way IP Capacity* on the Standards Track. However, current browsers do not provide access to raw UDP sockets, neither over browser APIs nor over toolkits such as WebAssembly. Therefore, utilising a measurement methodology that requires raw UDP socket access is not supported in the end user devices which will be used to perform measurements and is not suitable to fit the NRA requirements described in Section 3.1.1.

3.2. Round-trip delay (ping) and round-trip delay variation measurements

In principle, any kind of short IP packet (e.g. ICMP, UDP or TCP) could be used for round-trip delay measurements. However, the following should be taken into consideration:

- There may be operating system limitations preventing the use of ICMP packets. Further, ICMP packets in particular might be blocked by firewalls and antivirus software and hence their availability to measurement tools cannot be relied upon;
- TCP packets (after connection setup) are subject to flow control. The timing of the TCP session establishment three-way handshake could be used to measure the delay, however this is not routinely measured by the TCP stack and made available to applications;
- In a web browser environment, it is difficult or even impossible to send/receive arbitrary UDP packets which could be used for round-trip delay measurement purposes.

It is recommended that delay is measured by sending and receiving:

- ICMP echo/reply packets if possible, or
- short UDP packets, or
- short TCP packets, or
- WebSocket ping/pong frames sent by the server which are replied-to immediately by the client (such packets/frames should include the minimum data sufficient to enable the sending side to correlate request/response pairs). In case PING/PONG frames cannot be utilised, WebSocket data frames may be used, or
- HTTP requests with no content body response, e.g. utilizing the HEAD or OPTIONS method.

It should be noted that the first measurement in the delay calculation may be high and therefore unrepresentative in case additional delays are imposed by route path discovery or connection setup (e.g. DNS lookup or TLS connection setup).

The client should send as many measurement packets as is feasible in the time allowed by the test, taking care not to send packets so quickly as to compromise the measurement results. All individual latencies should be recorded. The number of measurement packets used should be selected to ensure statistical significance of the test result.

In addition to the latency test (that is done as a separate test), a loaded latency test (delay measured during speed test) could be performed to detect ‘buffer bloat’, as the comparison between latency and loaded latency might indicate issues within the IAS provider’s network.

Irrespective of when the round-trip delay measurement is taken, the resulting number is calculated as the median of round-trip delay measurements, with any timed-out measurements resulting from packet loss being excluded from the calculation. Delay variation (jitter) might also be derived from the delay variation observed during the delay[12].

[12] It should be noted that the number of delay measurements is rather limited, thus the deviation of such data is of limited value

3.2.1. Additional methodology for longer measurements

In order to observe the change of the delay of the connection for a longer period of time, an additional delay measurement with a user-defined duration, could be used. For this purpose, it is recommended to implement the measurement method defined in IETF RFC 2681 *Round-trip Delay Metric for IPPM*. This standard requires random sampling time, which results in a more statistically robust sampling.

3.2.2. One-way delay measurements

RFC 2681 discusses some of the issues in measuring one-way delay on the internet. This would require highly accurate synchronisation of clocks between the client and the server, and the implementation of an application signalling protocol to exchange timing related information. Therefore, BEREC considers that the measurement of one-way delay to be incompatible with the requirements stated above and impractical for regulatory use cases.

3.3. Packet loss measurements

If a packet fails to be delivered through the network within a certain timeout, it is considered as lost for the purpose of packet loss measurements.

On TCP connections a lost packet is automatically and transparently retransmitted by the sender, resulting in performance degradation from the point of view of the user. Unless the platform on which the performance measurement is executed provides access to detailed statistics for each TCP connection, it’s not possible to use this transport protocol to measure the packet loss.

If possible, ICMP echo or UDP packets should then instead be used for this purpose, considering all the related matters already described for the round-trip delay measurement.

Due to the very low packet loss ratio observed in modern networks and the possibly transient nature of their cause (e.g. network congestion, transmission issues), it is recommended to perform this measurement by sending a large number of packets per unit of time, over a long period. Similarly, the rate at which packets are sent should take into account the capacity of the link to avoid influencing the measurement. The number of packets limits the measurement resolution, as such the total number of packets sent/received in addition to the loss ratio should be recorded.

The minimum number of packets used should be based on the required resolution according to the anticipated packet loss for the access method being measured. For example, a packet loss rate of 0.1 % equals a loss rate of 1 in 1000 packets. In this situation a significantly larger number of packets would be required to establish robust results considering the desired margin of error.

However, the principle of running measurements of long duration conflicts with the crowdsourced user-initiated measurement concept: end-users will not accept an extended waiting time for the presentation of results. While long-duration tests are preferred for delay and packet loss measurements, this likely introduces the need for a measurement client running for a longer period in the background.

4. Detecting differentiated traffic management practices

This chapter describes recommendations for detecting traffic management practices that affect the connectivity (Section 4.1) and reachability (Section 4.2) of individual applications.

4.1. Connectivity measurements

This document focuses on measurement methodology and does not give any indications of which traffic management practices are allowed and which are not. Topics include the detection of blocked or partly blocked applications and content (e.g. network-based content filtering, such as ad-blocking and blocked web content) by blocking communication ports, URLs and IP addresses. Therefore, the connectivity measurements described below are an essential part of the net neutrality assessment methodology and should be used according to the need in each national market.

In general, the measurements described here should apply equally to both IPv4 and IPv6. No detection method is provided for protocols other than UDP/TCP (e.g. ICMP), however these are also considered to be relevant in the context of the IAS.

BEREC notes that a crowdsourcing approach may enable the comparison of many results for these measurements from different end-users.

4.1.1. Blocked ports

Whether traffic to certain ports is blocked or not, can be detected by establishing a connection to the port being tested, using the relevant transport protocol. With TCP, a port can normally be considered as being open if it is possible to establish two-way communication to it. Given the connectionless nature of UDP, a measurement system must define a feedback mechanism that tells whether the packet was received.

Measurement tools should be able to test for blocked ports at least over the following protocols:

- IPv4 and IPv6;
- TCP and UDP;
- uplink (connections from the end user to internet host) and downlink (connections from the internet towards the end user); and
- any valid UDP or TCP port number.

It is also worth noting that network address translation (NAT), which might be used by ISPs and modem/routers, affects downlink connectivity such that by default all communication ports are blocked in the downlink direction. This must be considered when assessing the measurement results.

It is important to take into account that the end-user environment (local firewalls or security software) may also affect the results. However, in the case of a crowdsourcing approach, it may be possible to compare large number of results from different end-users. Chapters 5 and 6 provide further information on these topics.

In case a significant share of the measurements under study indicate the same traffic practice, the likelihood that these practices are indeed occurring due to the operator's network setup increases.

4.1.2. Blocked IP addresses

The purpose of this test is to detect on an ad hoc basis if certain IP addresses are blocked. The test is executed by attempting to connect to at least a port on which it is known that the chosen destination address is running a service.

A successful connection to any port (or indeed any response from that address) is not sufficient to detect that the IP address is not blocked, since some ISPs could use middle-boxes to simulate a connection, and even answer on the established connection. Therefore, it is recommended to also send some data and verify the integrity of the received data.

If the connection cannot be established or the received data is not as expected, a new measurement could be performed from another connection point or by using a VPN to access the internet from outside ISP's control so that the ISP does not see the real destination address. If the connection via a proxy is successful, this can be seen as an indication that something in the ISP's network is blocking the IP address.

4.1.3. DNS manipulation

DNS manipulation in the context of the Regulation [1] refers to a situation where a DNS reply is received from the ISP's default resolver[13] (on an A or AAAA request) which falsely indicates that the domain is unknown or where an incorrect IP address is returned. The result of this manipulation is that the client is redirected to a different address.

DNS manipulation can be detected by analysing the responses to DNS requests on known targets (e.g. DNS records of specific domains under the control of the NRA).

4.1.4. Detection of an HTTP proxy

An HTTP proxy is a middlebox that is inserted into the path for end-users' HTTP connections, which may be used to filter or modify traffic. A HTTP proxy could be transparent or otherwise hidden.

A transparent proxy is a middlebox which could be deployed by the IAS provider acting as an intermediary between the client and the target web server. In this context, the ISP routes HTTP traffic via the proxy without user action or knowledge. A transparent HTTP proxy might be detected by checking the HTTP headers for proxy specific content (HTTP_VIA, VIA, FORWARDED, CLIENT-IP).

The HTTP (TRACE) request headers could also be checked for modification between the client and the server and if the intercepting proxy does a DNS lookup on a fake host header. A hidden proxy could be detected by a cache test.

Some HTTP proxies can be detected by connecting to a target domain and checking that the web resource is available and verifying that the content is identical to the content received over a proxy outside the ISPs control.

It may also be possible to detect a HTTP proxy by inspecting properties of the sent traffic such as TTL-flag of the IP packet).

[13] See paragraph 78a of the BEREC OI Guidelines [3]

4.2. Detecting practices that impact QoS of individual applications

The purpose of these measurements is both to suggest other indicators of performance closer to the user experience, and to detect the prioritisation and/or throttling of specific applications. These traffic management practices may be detected by measuring some of the Key Performance Indicators (KPIs) described below and comparing the results based on the following variations:

- Comparison of the same KPIs related to similar applications for the same IAS subscription,
- Comparison of the KPIs for the same application using an equivalent subscription from another ISP, and/or
- Comparison of the KPIs for the same application and the same IAS subscription but using a VPN.

These measurements can be performed on a regular basis for selected applications, websites or platforms or in targeted situations as needed.

Use cases include web browsing, video streaming, voice over IP (VoIP), video conferencing, audio streaming, cloud services or peer-to-peer file sharing and any other future applications not yet released.

4.2.1. Port throttling

In order to detect if a certain port is prioritised or deprioritised, performance of a tested port should be compared to the performance of a control port. In this context, a speed measurement on port 443 could be considered as a baseline. If there are significant and recurring discrepancies, it could be an indication that there may be a port throttling.

4.2.2. Individual applications using IAS performance measurement

Measurements of the performance of individual applications may show whether blocking or any kind of prioritisation or throttling of specific applications is applied to an IAS offer. Some of these traffic management practices may only be detectable when the network is congested which will require distributed measurements over time in various network segments.

Tools for detecting traffic management practices are likely to provide an indication of the presence of such a practice rather than a clear result. For example, when differences are observed in the calculated weight of a web page (number of bits that are transmitted during the page load) that has been loaded in similar conditions between different ISPs, it could be an indicator to detect blocking of a part of a website (e.g. ads) or data compression.

Another way of detecting traffic management practices could be to compare the measurement results related to a specific ISP, both with and without a VPN. The key idea is to use a VPN proxy located near the ISP's network edge to record and replay the network traffic generated by arbitrary applications, and compare it with the network behaviour when replaying this traffic outside of an encrypted tunnel. If there are significant and recurring discrepancies, it could be a strong indication that there may be impact from traffic management.

Most of the time, a measurement at the application level can only detect the presence of an inadmissible traffic management, but not the cause or responsible network segment.

5. End-user environment

As described in Chapter 3, the IAS quality is measured via an end-to-end connection (client-server). In the fixed environment the client is located at a terminal within the end-user's domain whereas the server is located close to the exit point of the ISP's network towards the internet. The end-user environment represents the end-user's domain which includes the end-user's terminal equipment (TTE) and may include their private network. It consists of different elements, some of which could limit the quality of service as perceived by the end-user when using the IAS.

In case of crowdsourced measurements that rely on end-user provided devices, the TTE is part of the performance assessment and could be a limitation such that the actual performance level provided by the IAS cannot be correctly gauged during the measurement. In principle, if the measurement is performed with specialised hardware probes specific limiting effects caused by end-user provided equipment could be ruled out. However, the measurement approach developed based on NRA requirements (see Section 3.1.1) needs to allow for browser-based measurements and thus the concept and applicability of using hardware probes was not taken into further consideration. A specialised hardware probe with high performance directly connected to the customer premise equipment (CPE) would rule out most of the limitations caused by the client device.

While the end-user environment in this context is just an element of the end-to-end measurement path that may not require specific consideration, for some purposes it is important to have knowledge about the composition of the end-user environment. This is the case for example for a certified monitoring mechanism (see Chapter 8), the objective of which is to assess only the performance that is solely the responsibility of the IAS provider's domain. End-user environment characterisation is also useful regarding post-processing of measurements in order to provide accurate market level aggregation (see Section 6.3.3). This allows irrelevant measurements, potentially biased by end-user environment, to be filtered out and to increase the granularity of the data when assessing general quality of IAS.

In end-user contracts typically the position of the NTP[14] represents the boundary between the user's environment and the IAS provider's domain. Therefore, the position of the NTP can be used as a reference point to identify the elements that may impact how the end-user perceives the IAS performance, beyond the contracted performance.

These limiting factors are listed and described below in two separate sub-sections differentiating between fixed network (see Figure 3) and mobile network environments (see Figure 4).

Whether the measurement results are influenced significantly by these factors depends on the specific configuration of the end-user environment and the technical characteristics of the elements involved. To assess the impact of the individual end-user environment configuration, the local hardware and software needs to be analysed. This can be done either by the measurement client itself by automatic technical means, or by end-user declaration. Technical solutions tend to be more reliable than declarations provided by end-users, but they require the use of installed software on the client

[14]According to Recital (19) of the European Electronic Communications Code (EECC), the NTP represents a boundary, for regulatory purposes, between the regulatory framework for electronic communication networks and services on one side and the regulation of the telecommunications terminal equipment (TTE) on the other.

computer to be able to access the variety of data described below.

A prerequisite of a technical analysis is the informed consent of the user, which should be required by the measurement client before measurement commencement. While processing and collecting end-user and end-user environment related information, please refer to the guidance from Chapter 9.

5.1. Fixed network end-user environment

The following subsections outline the main issues which might prevent an accurate performance measurement of the contracted quality commitment. These issues should be taken into account by certified monitoring mechanisms when presenting measurement results.

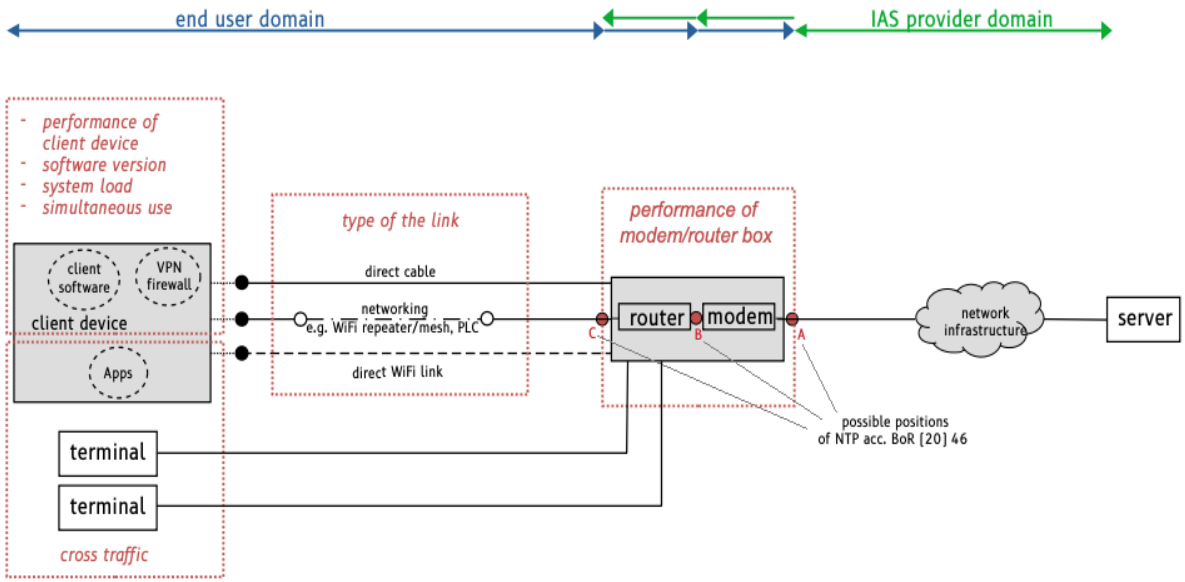

Figure 3 – Illustration of fixed network end user environment

As illustrated in the subsections below, various factors which have their origin in the end-user environment may limit the performance level gauged by the measurement. They can be roughly differentiated between influences caused by the client device, including its connection, and influences caused by additional devices in the local network.

When using crowd-sourced solutions that rely on unknown end-user provided equipment, the end-user should, for example, be requested to

- use a direct link,
- ensure the use of up-to-date hardware and software, and
- avoid the simultaneous use of other applications that might influence the measurements.

Any negative side effects of cross-traffic from additional equipment can be avoided by powering down or disconnecting these devices from the router/modem box during the measurement.

5.1.1. Performance of the modem/router

In cases where the performance of the modem is not sufficient to deliver the bitrate required to support the contractual maximum speed, it is impossible to gauge the true IAS speed.

The performance of the modem/router could also be limited by the (firmware) software version of the operating system such that up-to-date software might be required to deliver adequate performance levels. Given that the modem/router could be provided either by the end-user or the ISP, responsibility for modem/router maintenance like software upgrades, configuration settings etc. depend on the individual situation.

5.1.2. Type of the link

Ideally the link between the client device (end-user terminal where the measurement client software is executed) and the modem/router should be via a direct cable connection (typically Ethernet) that supports at least the IAS speed. If measurements are carried out via a link with lower performance (e.g. Wi-Fi, powerline or wireless repeater), then extra delay, packet loss or a throughput reduction might occur such that the available IAS performance cannot be accurately gauged. In cases where Wi-Fi is used, information on the Wi-Fi version and Wi-Fi signal strength may be relevant.

It should be noted that fixed network measurements can be made using a mobile device via Wi-Fi, in which case the aspects outlined in section 5.2 in addition to 5.1.1 and 5.1.6 would apply.

5.1.3. Performance and load on the client device

The equipment in terms of hardware and software of the client device should be suitable to accurately measure the speeds in principle. Also, if the load upon the client device in terms of RAM and/or CPU utilisation is too high, the measurement client may not be able to generate enough traffic to saturate the IAS and thus the measured performance may not correspond to the actual IAS performance. This could happen when certain software or applications are not closed before starting the measurement procedure.

Also, the client device should not be within energy-saving/low-power mode (e.g. laptops not connected to an external power source) while performing the measurement.

While there should be nothing to prevent consumer electronic devices such as game consoles, set top boxes or TV sets from performing the function of a measurement client, in principle, the hardware/software performance on such devices should be carefully considered, particularly within the context of a certified measurement, where the utilisation of standard computers might be more appropriate..

5.1.4. Version of the client device software

In order to perform correctly, the measurement tool might require up-to-date software such as operating systems, runtime environments and browser versions (in case of browser-based measurements). Outdated software on the client device might not include important performance tuning patches, but also the measurement tool might not be compatible with the latest available version of the software. As a result, information on the versions of software (e.g. client operating system and application software as well as relevant server-side components) in use while performing a measurement should be documented.

5.1.5. Simultaneous usage of other software like antivirus and firewalls

Background software like virtual private network (VPN[15]), anti-virus, content-based filtering (e.g. parental control), firewall and/or any local DNS manipulation that interacts with the traffic of the measurement client, could limit the performance level achievable by the measurement device and potentially affect test results. Such background software could also invalidate the detection of traffic management practices that impact individual applications.

5.1.6. Cross traffic

Cross traffic might be generated in parallel with the traffic of the measurement client, such as download/upload of data, music streaming, IPTV, video conferencing and software updates running in the background. It consumes capacity of the IAS and thus limits the performance level achievable by the measurement.

Note that cross traffic may be generated by application and/or operating system on the client device or other equipment and also by other network nodes within the end-user environment.

5.2. Mobile network end-user environment

A mobile end-user environment has a more straightforward structure compared to the fixed one. Typically, the end-user just uses one client device that is directly connected to the IAS via a radio link with no other equipment involved. Therefore, the end-user environment consists only of the mobile handset so there is normally no need for a detailed technical analysis of the mobile end-user environment.

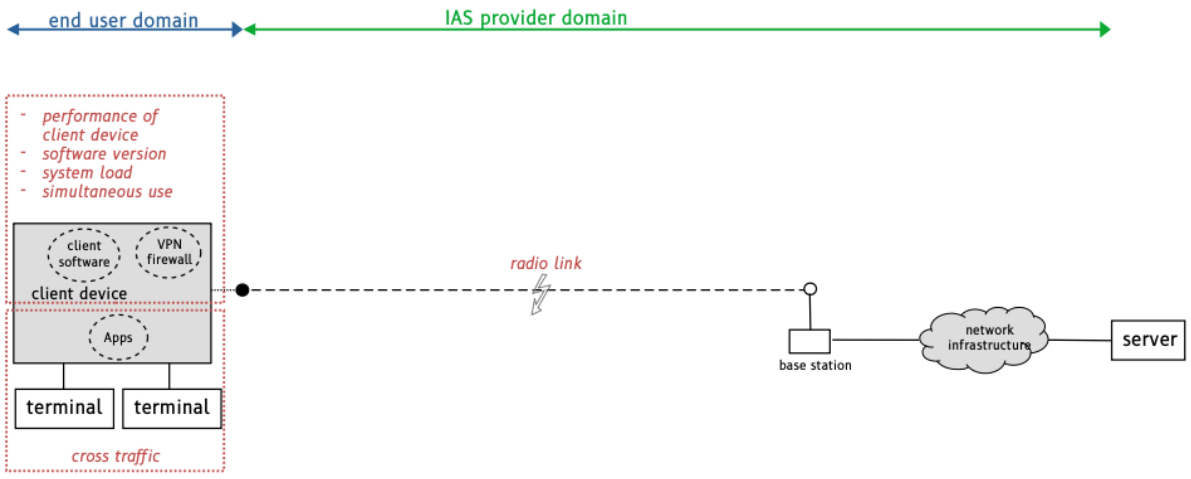

Figure 4 – Illustration of the mobile end-user environment

5.2.1. Performance of client device

The performance of the client device (handset model) involved in the measurements can limit the

[15] VPNs include those provided by as part of the Operating System

performance level gauged by the measurement. Different models perform differently: this could be due by the computational performance and hardware elements of the device but also by the performance of its radio interface in terms of supported standards and transmission speeds.

5.2.2. Simultaneous usage of other software like antivirus and firewalls

This item is covered in the fixed environment section above.

5.2.3. Version of the client device software

This item is covered in the fixed environment section above.

5.2.4. Cross traffic

Cross traffic generated by different applications running in the background on the client device could impact the measurement results. Also, it is possible that the end-user might share the network connection with other devices by using tethering or a mobile router/modem box. In these cases, cross-traffic from additional devices could occur, and the effect of the performance of the local connections should be considered in case the measurement client is not running on the terminal directly connected to the radio link.

5.2.5. Access Method

To avoid any misinterpretation of results, the testing client should verify that it is not using a Wi-Fi connection for the test. If this is not possible, end-users should be advised to turn off Wi-Fi on the device for the duration of the test.

5.3. Useful information to enrich measurement data

Additional information could be collected to increase the overall value of QoS measurement results and to be able to determine whether the cause of poor performance lies within the end-user's or the ISP's domain. Such information can be used to identify (and/or rule out) limiting effects within the end-user environment as listed above (Sections 5.1 and 5.2).

This additional information could be retrieved either by technical means (by the measurement client, or through other tools such an API at the CPE level) or by end-user declaration. This allows the detection of measurements where the end-user environment is a limiting factor. It is recommended to retrieve additional information by technical means as far as possible.

In case of problems perceived by the end-user, this information could potentially be used to explain unsatisfactory measurement results and to identify the root cause. It could also be used to better post-process measurement results for the purposes of market aggregation and potential publication.

Typical information elements useful for the analysis and interpretation of measurement results could include, where available:

- IP address of client (and server used in the measurement);
- time and date of measurement;
- traceroute information of the link client – server and server – client;

- client/server MSS (Maximum Segment Size) and MTU (Maximum Transmission Unit);
- type and speed limits of network interfaces of client device in the fixed environment;
- type and speed limits of the radio link interface in the mobile environment;
- type and firmware version of the modem/router;
- type and software version of the client device's operating system;
- CPU and RAM of the client device;
- Information on client hardware such as handset model;
- level of cross-traffic (either on client device or within the LAN);
- an indication of whether the test traffic could have traversed a VPN.

In fixed networks, while information on IAS offers such as subscription type, speed cap, type of internet connection (FTTH, xDSL, etc.) is not within end-user domain *per se*, it is useful to accurately characterise the fixed line that is being measured.

In mobile networks, the available speed depends, among other factors, to a great extent on the quality of the radio link conditions. Therefore, it is important to retrieve and store the information on the radio conditions which prevailed during the measurement. The available radio parameters vary between different mobile network technologies and operating systems. Therefore, it is recommended to retrieve available parameters provided by the mobile handset. In addition, information about the conditions of the measurement such as indoor/outdoor or in motion is helpful in this context.

The network may for example be technically able to deliver a bit rate higher than the purchased subscription. This could have a significant effect on the performed measurements. Another possibility could be that the speed is throttled to a very low value after reaching the end-user's data cap.

It should be noted that the possibility to retrieve such data by the measurement client itself can be limited when using a browser-based solution, since data from the operating system is not accessible and communication with other equipment within the local network is not possible. In this context, the use of other solutions such as an API at the CPE level to retrieve end-user environment parameters during the test could be relevant.

6. General IAS quality assessment methodology

This chapter provides a framework to assess the general quality of IAS, which consists of several steps summarised in Figure 5 below. The different steps to be considered are related to collecting measurement results (Section 6.1), validating data (Section 6.2), post-processing and aggregating the results while guaranteeing the representativeness of the result database (Section 6.3), analysing the results (Section 6.4), as well as publishing the relevant information while reporting measurement biases (Section 6.5).

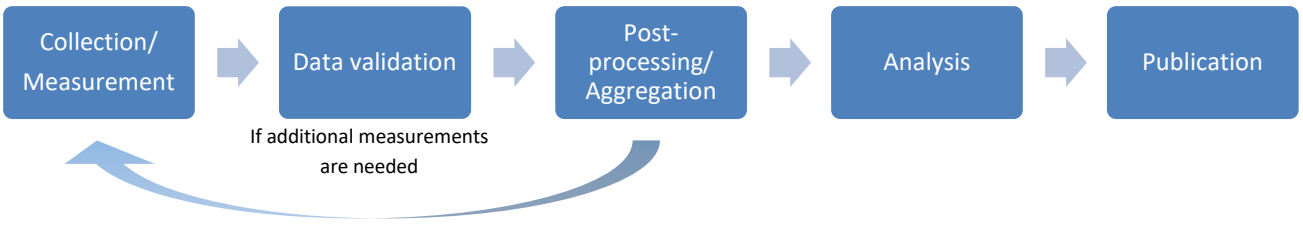

Figure 5 - General IAS quality assessment methodology

When defining the general quality of IAS, the different contexts within which the term is used in the Regulation [1] should be considered. The term is described related to assessment of specialised services (Article 3(5) and recital 17) and related to NRAs' safeguards to promote a general improvement and prevent degradation of IAS (Article 5(1) and recital 19).

The former indicates that the performance of the remaining network capacity when specialised services are introduced, should be assessed. Since it would be difficult or impossible to isolate the two service categories, BEREC proposes to take a "black box" approach to such an assessment. This means that the quality of general IAS must be assessed during external measurements. Given the fluctuating use of capacity by specialised services over time, there is a need to evaluate the performance of the general IAS through more than just isolated snapshots. Longer time series of measurements are therefore needed.

In relation to the latter, the wording of Article 5(1) that NRAs "*shall promote the continued availability of non-discriminatory internet access services at levels of quality that reflect advances in technology*" also requires NRAs to ensure and measure a general improvement in performance over time. The assessment methodology, in addition to considering the general quality at a specific point in time, needs also to consider a reasonable and positive development of the performance indicators that are under observance.

Fixed access versus mobile access

Given that the transparency requirements on ISPs (Article 4(1)(d) of the Regulation) provide a set of specific speed parameters for fixed and mobile internet access services respectively, it is fitting to also consider a separate assessment of the general quality of IAS for fixed and mobile networks.

6.1. Collection/Measurement

The measurement should follow the methodology detailed in Chapter 3 of this document.

The measurement tool generates a data set for each measurement period that is stored in a database. There are two basic measurement approaches:

a) Measurement campaigns using measurement systems with dedicated clients and servers in a controlled environment, or

b) Crowdsourced measurement campaigns relying on end-user-initiated measurements using end-user equipment.

General considerations regarding how to collect measurement data by the use of crowdsourced measurement approaches, and discussion of advantages and disadvantages of this approach, is provided in *BEREC Report on Monitoring QoS of Internet access services in the context of net neutrality* [4], see Section 4.5.2.

The general quality of IAS should be assessed for the whole subscriber base. The measurement results that are collected from assessment of the specific quality for individual subscribers can be reused. In order to enrich the measurement results, supplementary information on the end-user environment can be added either through auto detection or by the end-user, as described in Chapter 5.

6.2. Data validation

Depending on the kind of measurement approach chosen, the data validation could be complex and extensive. Before aggregation, measurement results should be anonymised, and records should be normalised with unsuitable ones filtered out.

For the measurement approach based on a pre-validated setup using dedicated clients and servers, basic plausibility checks like timestamps matching the measurement schedule, correct client identification etc. may be sufficient.

For crowdsourced measurement approaches, more extensive steps should be taken since the conditions at the client side are not predetermined, i.e. it is unknown whether the end-user environment fulfils the requirements for an accurate measurement. To some extent this can be cross-checked by the use of supplemental information (see Section 5.3) gathered at the time of measurement which should be validated where possible.

The validation process of end-user provided information is a multi-step process. Such a process starts with the removal of implausible data and could include verifying internet service provider identification, potentially discarding data relating to providers which are not relevant for the measurement campaign.

Cross-checking for the correct measurement set-up is done by the use of measurement result metadata as described above. Depending on the end-user environment requirements, certain metadata should be collected together with each measurement result. Such records could include type of connection (e.g. Ethernet, Wi-Fi), type of terminal used, status of terminal equipment (e.g. processor load, cross-traffic, parallel active applications), network environment (firewall) or the kind

of access technology of IAS (e.g. identifying modem type) etc. as described in Section 5.3[16].

6.3. Measurement results post-processing and aggregation

6.3.1. Measurements post-processing

Post-processing of the collected data is a crucial stage for eliminating false, manipulated or irrelevant measurements. It creates the ability to ensure that the results are representative and as widely comparable as possible. It also helps to protect against attempted fraud.

NRAs should consider implementing efficient data processing algorithms to deliver the most reliable results possible. For example, it is particularly important that measurements are excluded when obtained from a target server that has proved to be a limiting factor (notably when the server's connection capacity is below or equal to that of the line being tested). Other examples could be:

- merging multiple results from the same end-user during the same time frame (1 hour for example) in fixed access;
- detecting and removing automated tests that are made in bad faith in order to influence measurement results of a particular ISP.

6.3.2. Statistical representativeness

The general quality of IAS should be assessed for the whole subscriber base. To ensure a representative results data set, further actions may be required to supplement the existing results.

Thus, before analysing measurement results, the need for additional steps, to ensure that the data set statistically represents reality, should be assessed. In case of insufficient data, representativeness cannot be guaranteed. Any identified bias due to the testing method that is likely to distort the representativeness or create comparability issues, should be mitigated (through additional measurement for example) and/or disclosed.

6.3.3. Market level aggregation

After the measurement data is processed, the results could be aggregated. At the market level, the measurement results could be summarised into aggregated values for different categories such as IAS offers (subscription type, data cap, etc. where available), ISPs, geographical area, technology type, mobile technology generation, mobile device type, etc.

Over time, the user-base tend to migrate from one speed class to another, and from one access technology to another (for example from DSL to fibre). Therefore, it could also be relevant to aggregate measurement results independent of these categories. Such aggregation could provide valuable information about the development of the general quality of IAS at the market level and assist the NRA to focus its efforts on the promotion of the continued availability of non-discriminatory internet access services at levels of quality that reflect advances in technology.

[16] When collecting and processing information, the applicable privacy rules must be taken into account.

Aggregated results of IAS performance at the market level may be used for regulatory supervision, including monitoring of the general IAS quality. Ideally, monitoring general IAS quality would be based on the ongoing and continuous collection of measurement results, however it is also possible to run specific measurement campaigns as required.

The market level aggregated measurement data could be used to monitor that the overall[17] available quality of IAS (e.g. speed, delay and packet loss) improves over time. In addition, it is important to assess whether an ISP treats individual applications equally (see section 6.4).

Furthermore, to reflect the fact that the performance of the network may differ significantly over the geographical coverage of the network, the aggregation should include the measurement location at a granularity which balances the anonymity of the end-user against the level of detail required for the assessment to be broken down to smaller geographical areas.

6.4. Analysis

Following measurement aggregation, results could be plotted against selected dimension(s) and time (e.g. download speed for a given 4G network over time) which could enable NRAs to extrapolate a realistic improvement in the general quality of internet access for later analysis. This approach could be used to assess the impact of specialised services on the general quality of IAS.

6.4.1. Measuring the improvement in general IAS quality

Based on the measurement results database, for a given subset of the measurement data (e.g. fixed or mobile access for a certain region), the average download and upload speeds are calculated for the preceding years (for example a five or more years' perspective). For those speeds, a predictive function could be applied that forecasts the following year's speed, considering factors such as the limitations of the underlying access technologies. Then in future years, the predicted values can be compared with the measured values for that year.

Based on the aggregation dimensions, this assessment could also be applied at a granular level (e.g. to each ISP). This would give a better indication of whether specific ISPs show particularly low performance (development) compared to the others.

If the observed average measured values of download and upload are significantly lower than the predicted values[18] for the corresponding period, it might indicate that the general quality of IAS has not improved adequately over time.

Furthermore, other QoS parameters, such as latency, jitter and packet loss, could supplement the assessment. In case one or more of these parameters are significantly worse compared to preceding years, this might be caused by increasing congestion in the network.

Finally, to assess whether particular geographical areas experience degradation of IAS general quality, the measurement results which are aggregated for smaller areas could be analysed per area (e.g. per county or municipality).

[17] This could be the mean or median (or, even, certain k-th percentile)

[18] Considering the appropriate confidence intervals

Any prediction made should be accompanied by information on the parameters used to make it, considering the data listed in section 6.5 in the first instance.

6.4.2. Illustrations of the predictions

The following illustration provides an example on how the predictions could be used for visualising trends found in the data set under study. NRAs will decide on which parameters, and combinations of such, are to be employed.

A well supported time-series prediction requires a substantive dataset, preferably covering a historical period longer than the one being predicted. Moreover, possible seasonal variation (e.g. different levels of home broadband usage in the winter vs. summer) should also be factored into the model, given its potential impact on the time series.

Figure 6 illustrates predictions on how some KPIs in mobile networks could be expected to develop in the next few years. Note that any given performance indicator will have an upper limit in terms of improvement, as the development will also be technology-dependent. NRAs need to take such realities into account when analysing their findings.

An illustration of a time-series based prediction of average mobile download speed could be done like in Figure 6. In this example, a speed decrease in year N could be caused by the increased number of active users in the network which was not followed by an increase in the network capacity.

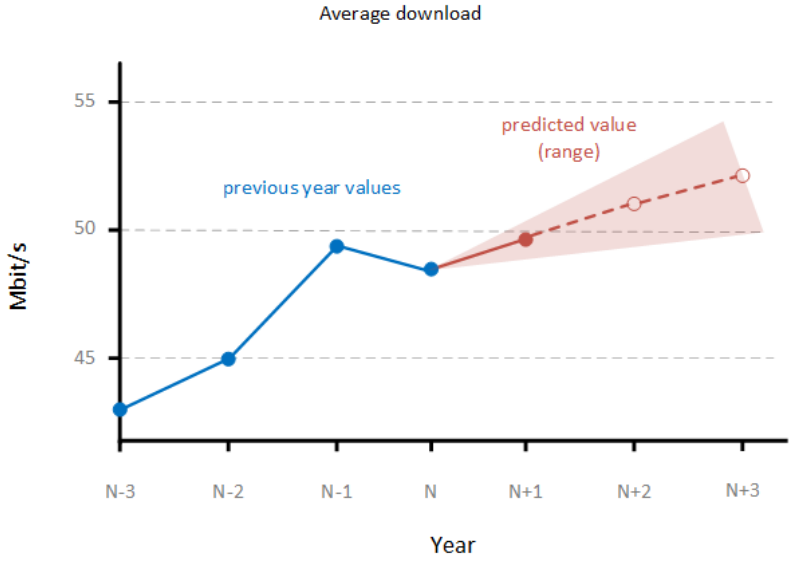

Figure 6 - Example on visualising predicted trends of average download speed

6.4.3. Further Analysis: effect of specialised services on IAS

According to Article 3(5) second subparagraph of the Regulation [1], specialised services (SpS) shall not be provided "*to the detriment of the availability or general quality of internet access services for end-users*". Therefore, the task for NRAs is to check that SpS are not provided at the expense of IAS.

The BEREC OI Guidelines, paragraphs 106-115, describe multiple approaches as to how NRAs can supervise this including information requests from ISPs followed by an assessment of the IAS provider's capacity expansion plans.

In addition, an NRA could use the aggregated IAS QoS measurement results to analyse the relevant network performance in the relevant area, comparing the results for both before and after the introduction of a certain specialised service. If the measured speed values are in general lower after the SpS introduction, this could be an indication that the SpS is provided at the expense of IAS. The NRAs may monitor this e.g. by following the trend of how the average speed measurement results for each ISP are evolving. When the introduction of a SpS affects the general quality of IAS, this could be visible also from general IAS performance results.

6.5. Publication

The market level data could also be used for transparency purposes by publishing statistics, as well as interactive maps showing fixed IAS performance or mobile IAS performance in a geographic area. Such data would thus provide end-users with an overview of the general quality of IAS.

Any publication of market level data should disclose any identified bias related to measurement methodology likely to create comparability issues, especially when comparing ISPs. The publication could also provide transparency over a number of other parameters, for example:

- Indicating the exact period covered by the publication;
- Indicating the number of tests for fixed or mobile access technology and per ISP;
- Providing as many details as possible on the data processing performed (methods to adjust the results);
- Indicating test location, for example in the case of mobile network, it would be relevant to indicate the percentage of customers per region, if publishing findings by region; the percentage of tests conducted while mobile (significant distance travelled between the beginning and end of the test, if this is captured);
- Specifying the operating system: percentage of operating systems per operator taken into account in the results for fixed network and the percentage of Android and iOS devices used for mobile network;
- Indicating the percentage of tests conducted in IPv4 and in IPv6;
- Providing the distribution of default servers in case the tool offers several test servers: the publication should indicate the breakdown of tests by operator, with respect to the choice of test servers;
- Communicating on the other factors that are likely to introduce significant bias on the analysis of market level aggregation publications. For example, in mobile networks, where there are significant differences tied to the devices, the results could be broken down: e.g. by type of device for each operator, or by indicating the percentage of tests per smartphone model, on the most widely used devices. In fixed networks, if a publication is comparing different ISPs without differentiating access technology, it must clearly indicate that this combination of technologies used by ISPs introduces significant biases in the results. The publication should also indicate, when possible, any other possible bias (test server's limitations, user device, etc.).

In addition to market level aggregation, NRAs should consider publishing measurement results as Open Data to increase transparency. Further recommendations for publication of information are available on section 5 of [6].

7. Individual results assessment

This chapter is intended to provide guidance to NRAs in assessing measurement results related to an individual subscriber.

7.1. Speed measurement assessment

Broadband speed is a commonly cited metric when characterising the quality of a broadband offering. Speed measurements may be made to assess both the generic as well as the individual qualities of a network. The interpretation of the results of these measurements needs to be as realistic as possible as the information can be relevant for NRA's supervision and enforcement policy.

According to Article 4(1)(d) of the Regulation, ISPs shall declare the minimum, normally available, maximum and advertised download and upload speed in their fixed network contracts. For mobile network subscriptions, ISPs must declare estimated maximum and advertised download and upload speeds. These notions are further specified in the BEREC OI Guidelines [3].

Minimum speed

According to paragraph 143 of the BEREC OI Guidelines regarding fixed networks, "*The minimum speed is the lowest speed that the ISP undertakes to deliver to the end-user, according to the contract (…). In principle, the actual speed should not be lower than the minimum speed, except in cases of interruption of the IAS. (…)*".

To verify the minimum speed, each valid individual speed measurement result should be compared against the minimum speed value defined in the contract checking if any measurement is below it.

Maximum speed and estimated maximum speed

According to paragraph 145 of [3], regarding fixed networks "*The maximum speed is the speed that an end-user could expect to receive at least some of the time (e.g. at least once a day). (…)*".

As described in paragraph 153 of [3] regarding mobile networks "*The estimated maximum speed for a mobile IAS should be specified so that the end-user can understand the realistically achievable maximum speed for their subscription in different locations in realistic usage conditions. (…)*".

To verify the maximum and estimated maximum speed based on a sufficiently large set of results over the relevant period, each individual speed measurement should be compared against the (estimated) maximum speed value defined in the contract.

Normally available speed

According to the recitals of the Regulation and the BEREC OI Guidelines [3] paragraph 147, regarding fixed networks "*The normally available speed is understood to be the speed that an end-user could expect to receive most of the time when accessing the service*", and "*BEREC considers that the normally available speed has two dimensions: the numerical value of the speed and the availability (as a percentage) of the speed during a specified period, such as peak hours or the whole day*".

Further, according to paragraph 148 of the BEREC OI Guidelines, "*The normally available speed should be available during the specified daily period.*" For example, an NRA may set a requirement "*that normally available speeds should be available at least during off-peak hours and 90 % of time*

over peak hours, or 95 % over the whole day; (…)".

The normally available speed should be calculated based on the results of a series of speed measurements.

Advertised speed

For fixed IAS, according to paragraph 150 of the BEREC OI Guidelines [3], NRAs should "*In the event that speeds are included in an ISP's marketing*" ensure that the advertised speed is included "*in the contract for each IAS offer*".

For mobile IAS, according to paragraph 156 of [3], "*The advertised speed for a mobile IAS offer should reflect the speed that the ISP is realistically able to deliver to end-users.*"

Paragraph 151 of [3] provide that "*NRAs could set requirements in accordance with Article 5(1) on how speeds defined in the contract relate to advertised speeds, for example that the advertised speed should not exceed the maximum speed defined in the contract*" related to a fixed IAS. In the same vein, Paragraph 157 of [3] provides "*that the advertised speed for an IAS as specified in a contract should not exceed the estimated maximum speed as defined in the same contract*" related to mobile IAS.

Measurements can be used to compare overall ISP performance with advertised speeds.

7.2. Other QoS parameters and traffic management assessment

Other parameters, in tandem with speed, contribute to presenting a more holistic picture in terms of quality and acceptable level of service for a particular application or service. These are packet loss, delay and delay variation, the measurement methods for which are covered in Chapter 3.

The acceptable margins for such values will vary from service to service. For example, a video streaming service based on progressive download might require a certain bitrate with low packet loss, however it might not require stringent latency or jitter performance. Whereas a video calling/conferencing service might not have a stringent packet loss requirement but could be highly sensitive to latency and jitter.

In this context, an NRA may wish to define thresholds for acceptable performance for different services. Such thresholds could be applied to the results of a QoS measurement, with an indication of which services would perform acceptably on the current line. Such an indication could take a user-friendly form such as a red/yellow/green traffic light notation

8. Certified monitoring mechanism

As already set out in section 7.1, ISPs shall describe the minimum, normally available, maximum and advertised download and upload speed in their fixed network contracts, according to Article 4(1)(d) of the Regulation. For mobile network subscriptions, ISPs must describe estimated maximum and advertised download and upload speeds.

Article 4(4) of the Regulation [1] defines that an end-user may use "*a monitoring mechanism certified by the national regulatory authority*" to check that the actual performance meets what has been specified in the contract. This measurement information can be used for triggering the remedies available to the consumer in accordance with national law.

This entails a decision on whether the subscription meets the different speed values defined in the contract and whether there is a "*significant discrepancy, continuous or regularly recurring*". Note that in some Member States, the NRA may not be competent to resolve disputes between consumers and undertakings providing electronic communications services, including deciding on whether there is significant discrepancy, and such decisions may be made by a different authority or body.

To be able to issue a declaration either that there is no significant discrepancy between actual and indicated performance or that there is such a discrepancy empowering the user with the right to trigger "*the remedies available to the consumer in accordance with national law*", a number of conditions should be satisfied from a regulatory point of view for giving legal value to this "evidence". The final ruling over which "evidence" is sufficient for triggering legal consequences however is still subject to court rulings. Therefore, decisions of NRAs should be made transparently; all measurement data should be available for further legal considerations of the respective court.

The Regulation [1] does not require Member States or an NRA to establish or certify a monitoring mechanism. Therefore, it is worth noting that a certified monitoring mechanism may be available only in some Member States. The Regulation does not define how the certification should be done, so this is a national matter. If the NRA provides a monitoring mechanism for this purpose, it should be considered as a certified monitoring mechanism according to Article 4(4) of the Regulation. As the Regulation talks about "*a monitoring mechanism certified by the national regulatory authority*", the question of when to certify a monitoring system and how to certify can be considered to be up to an NRA according to the national legislation and circumstances.

8.1. Guidance on criteria regarding certified monitoring mechanism

This section gives guidance on criteria NRAs could take into account when providing their own certified monitoring mechanism or certifying a third party mechanism in accordance with the Regulation [1] and BEREC OI Guidelines [3].

a) The certified monitoring mechanism should fulfil the requirements specified in Chapter 3 and take the considerations of Chapter 5 into account.

b) The certified monitoring mechanism should be in compliance with the applicable legislation such as privacy rules.

c) End-users should be able to make a straightforward comparison between measurement results and the contractual speed values.

d) The NRA is recommended to give guidance on in which cases a significant and continuous or regularly recurring difference is established by the certified monitoring mechanism. Non-compliance on a single indicator is sufficient to give the user the right to use "*the remedies available to the consumer in accordance with national law*".

e) The NRA is recommended to ensure the integrity of the operation of the certified monitoring mechanism in case the mechanism is provided by a third party. It is also recommended to take into account the independence and business model of the entity providing the monitoring mechanism where it is not provided by the NRA itself.

9. Privacy

When an NRA considers offering a measurement tool, privacy questions should be considered from the very beginning. The legal framework for this can be found, amongst other, in the General Data Protection Regulation ("GDPR", Regulation (EU) 2016/679). Specifically, the kind of data that falls within the scope of "personal data", and grounds on which this personal data can be processed need to be considered. Also, other data protection requirements, such as data minimisation, rights of the data subject and requirements for a privacy policy have to be taken into account.

A more detailed overview on what aspects to cover in this assessment can be found in Chapter E of the *Net neutrality measurement tool specification* document [7].

10. References

[1] Regulation (EU) 2015/2120 of the European Parliament and of the Council of 25 November 2015 laying down measures concerning open internet access and retail charges for regulated intra-EU communications and amending Directive 2002/22/EC and Regulation (EU) No 531/2012, https://eurlex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02015R2120-20181220

[2] BEREC Guidelines for quality of service in the scope of net neutrality (BoR (12) 131), November 2012, http://berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/guidelines/1101-berec-guidelines-for-quality-of-service-in-the-scope-of-net-neutrality

[3] BEREC Guidelines on the Implementation of the Open Internet Regulation (BoR (22) 81), June 2022, https://berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/guidelines/10280-berec-guidelines-on-the-implementation-of-the-open-internet-regulation

[4] BEREC Report on Monitoring QoS of Internet access services in the context of net neutrality, BoR (14) 117, September 2014, http://berec.europa.eu/eng/document_register/subject_matter/berec/reports/4602-monitoring-quality-of-internet-access-services-in-the-context-of-net-neutrality-berec-report

[5] BEREC Report on Feasibility study of quality monitoring in the context of net neutrality (BoR (15) 207), November 2015, http://berec.europa.eu/eng/document_register/subject_matter/berec/reports/5576-feasibility-study-of-quality-monitoring-in-the-context-of-net-neutrality

[6] BEREC Guidelines detailing Quality of Service Parameters (BoR (20) 53), March 2020, https://berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/guidelines/9043-berec-guidelines-detailing-quality-of-service-parameters

[7] Net neutrality measurement tool specification document (BoR (17) 179), October 2017, https://berec.europa.eu/eng/document_register/subject_matter/berec/reports/7296-net-neutrality-measurement-tool-specification

[8] BEREC Net Neutrality Regulatory Assessment Methodology (BoR (17) 178), October 2017, https://www.berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/methodologies/7295-berec-net-neutrality-regulatory-assessment-methodology