Body of European Regulators for Electronic Communications

BEREC

BoR (14) 24

Monitoring quality of Internet access services in the context of net neutrality

Draft BEREC report

8 March 2014

Content

1. Executive Summary

Monitoring quality of Internet access services in the context of net neutrality is important to improve NRAs' capacity to perform regulatory assessments of potential degradation of service, as pointed out in the BEREC Framework[1] and Guidelines[2] on Quality of Service in the scope of Net Neutrality.

Furthermore, transparency enables end users to compare Internet access service (IAS) offers and hence strengthen the demand side of the market. It is therefore essential to have appropriate quality monitoring tools to implement the recommendations drawn from the earlier studies in this area.

The main goal of this report is to establish a basis for the creation of Internet access service quality monitoring systems covering two main use cases:

A. Providing transparency on the quality of the Internet access service for end users and
B. Regulatory supervision through monitoring of quality of the Internet access service with regard to potential degradation of service.

Metrics for quality monitoring of IAS

In order to assess the quality of IAS, BEREC recommends measuring actual performance of the service, taking into account as a minimum the following IP layer parameters: Upload and download speed, delay, jitter, and packet loss ratio. In developing measurement methodologies for IP-based communications further development of technical specifications is also needed, primarily by the Internet Engineering Task Force (IETF).

Since connectivity to other networks (autonomous systems) is an essential part of the IAS offer provided by an ISP, this connectivity should also be covered by the measurement methodology. BEREC recommends that measurements beyond the ISP leg, including the interconnection of the ISP, should be used to account for the connectivity of the ISPs towards the Internet.

The recommended IP layer metrics are applicable for fixed as well as wireless/mobile Internet access services. BEREC recommends consideration of the use of additional parameters, e.g. to reflect wireless/mobile network coverage aspects.

BEREC also recommends an emphasis on open source and open data solutions.

Transparency of quality information (Use case A)

Average IAS performance (sub case A1) and *individual* IAS performance (sub case A2)

BEREC recommends implementing end user transparency measurements in a user-friendly manner. A software-based measurement agent downloaded to end user equipment can be sufficient given that measurement results are validated by collecting additional end user information.

[1] BEREC, A framework for quality of service in the scope of net neutrality, Document no. BoR (11) 53, 2011
[2] BEREC, Guidelines for quality of service in the scope of net neutrality, Document no. BoR (12) 131, 2012

Regarding aggregated results, BEREC recommends - for reasons of cost-effectiveness and user-friendliness - that averaging (based on data gathered from all participating users) should be done based on crowd-sourcing.

Regulatory supervision of IAS quality (Use case B)

Degradation of IAS *as a whole* (sub case B1) and *applications* using IAS (sub case B2)

Measurements for supervision of quality of IAS as a whole will typically be conducted in one of two ways. The NRA could either use a controlled system, e.g. with hardware probes, covering a preselected panel, or a less controlled system with software agents and a crowd-sourced user base.

When evaluating potential degradation of IAS as a whole, BEREC recommends that such measurements are conducted over time to allow trend analysis. Measurement results need to be assessed in the light of technical progress and market evolution, with the goal of evaluating potential effects such as the provisioning of specialised services at the expense of IAS.

Regarding monitoring of applications using IAS, BEREC recommends the use of appropriate tools to measure the performance of individual applications (may also be used for the transparency, use case A) and also exploring the use of passive measurements. Leveraging on information from the measurement systems of content and applications providers (CAPs) and other complementary methods could also be considered.

Measurement results obtained by these methods will need to be assessed by experts regarding reasonable and unreasonable traffic management, in order to detect degradation of individual applications using IAS.

Convergence of measurement methodologies

Quality assurance of measurement results and regulatory assessment of the results require deep understanding of the underlying complexities of Internet communications, and of monitoring methodologies. It is expected that this understanding will need to develop over time, and the exchange of experience among NRAs to foster convergence of practices, and participation in and contribution to standardisation activities, are good strategies for harmonisation in this area.[3]

In particular when it comes to gaining experiences with assessment of degradation of service, BEREC recommends that NRAs collaborate to develop a common regulatory practice. Harmonisation of evaluation of potential degradation of IAS as a whole, typically at the expense of specialised service, as well as assessment of degradation of individual applications, should be emphasised.

An evolutionary approach to a potential multi-NRA opt-in monitoring system

To achieve further harmonisation, NRAs may consider the development and adoption of a multi-NRA measurement system on a voluntary basis. For this purpose, BEREC recommends that an evolutionary approach should be taken towards convergence of existing measurement systems.

[3] Some BEREC members (e.g. the Spanish regulator CNMC) face legal constraints for the adoption of measurement systems which implies that they are not able to collaborate in the work at EU level on measurement systems or best practices thereon.

Done on a step-by-step basis, this would allow the identification of best practices, and then the comparison of results before finally enabling cross-border measurements. For example, a set of shared test servers (e.g. placed in major IXPs) could be made available for usage by participating NRAs as a first step. Additional steps could be taken to expand the footprint of the server system. Furthermore, the use of software-based clients could constitute an initial stage.

A collaborative system would further support comparability of measurements, provide for cross-border measurements, and support common regulatory approaches. An opt-in system should allow NRAs to use their system segments for distinct measurement regimes to address national specifics, e.g. separate servers, more parameters, additional technologies, etc. Such a system would typically supplement national systems.

However, a collaborative measurement system would face a number of challenges: cost of cooperation, complexity of system, time constraints related to alignment amongst NRAs, and other factors to be analysed in a future study.

It is recommended that BEREC conducts a feasibility study for a potential future opt-in monitoring system and does not follow up on implementation of a full-blown measurement system. This would draw upon the proposed quality monitoring approach described above, containing recommended measurement parameters and methods. The system should be designed in a way that allows additional measurement scenarios to be integrated smoothly into existing national systems.

Such a study should also consider the effect of dissemination of knowledge among NRAs and further development of best practices. This should facilitate increased harmonisation of measurement methodologies and increased competence in the field of quality monitoring in the context of net neutrality.

2. Scope and background

2.1 Scope and approach of the report

The main goal of this report is to establish a basis for the creation of Internet access service quality monitoring systems capable of enhancing transparency for end users of electronic communication services and prevent degradation of service from Internet service providers (ISPs) in relation to end users and content and application providers (CAPs), as prescribed in the Universal Service Directive (USD)[4].

The report will set out the fundamental aspects of quality monitoring systems of national regulatory authorities (NRAs) in the context of net neutrality[5], at three different levels:

- Using existing systems for this purpose;
- Building recommendations for future systems; and
- Exploring the evolution of a potential multi-NRA opt-in system.

BEREC believes that a quality monitoring system by itself could motivate all parties to take a more net neutrality friendly approach and see an increase in the level of performance experienced by the users. Consumers will have richer information and be able to make better informed purchasing decisions.

Finally, should such an ideal situation not be respected by one or more of the relevant market players, such quality monitoring systems would assist NRAs in intervening with appropriate corrective measures. For example, minimum quality requirements imposed to prevent degradation of service could be based on consolidated analysis of trusted results obtained by such a quality monitoring system.

In summary, two *main use cases* for quality monitoring systems are foreseen; they are not mutually exclusive.

- **Use case A: Transparency about IAS quality.** NRAs ensure the availability of information for end users on the quality of Internet access service. This information may describe the situation at a market level, e.g. the average performance region by region, or ISP by ISP (subcase A1), or it may describe the performance enjoyed by individual users, e.g. by reporting their quality of service at a specific time and location (subcase A2).

  Transparency in the context of net neutrality is described in USD Articles 20 and 21. The transparency use case in this report will only cover how quality monitoring tools can be used to provide measurement results in line with the recommendations of BEREC NN Transparency guidelines. How such results are presented to end users is beyond the scope of this BEREC NN QoS Monitoring report.

[4] DIRECTIVE 2009/136/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws [2009], OJ L 337/11

[5] To BEREC, "net neutrality" describes the principle of equal treatment of network traffic.

- **Use case B: Regulatory supervision of IAS quality.** NRAs have the power to collect information with the specific goal of assessing degradation of Internet access services, either as a whole (subcase B1) or for specific applications (subcase B2). This activity is conducted in the light of regulatory powers and may lead to the imposition of minimum requirements for quality of service and the monitoring of their fulfilment.

  "Degradation of service" is described in USD Article 22(3), and the implementation of this provision is detailed in the BEREC NN QoS guidelines. The regulatory supervision use case will explore how quality monitoring systems can be used to gather measurement results that can be used by the NRA to assess potential "degradation of service".

This document elaborates on the aspects to be considered when creating quality monitoring systems suitable to be applied to the use cases mentioned above and capable of fulfilling the respective measurement objectives. The document first explores how existing quality monitoring systems can be used for this purpose, and whether their methodologies and parameters meet a minimum set of requirements. Then the document draws conclusions on how to build and maintain quality monitoring systems based on these methodologies and parameters. Finally it discusses how to keep existing national practices up-to-date and to adapt them for future measurement tasks in order to develop best practice solutions and keep pace with the technological development. The possibility of a more harmonised approach is explored, which could consist of a common set of measurement parameters and corresponding measurement methods for cross-border uses. The potential of such an addition to national initiatives is discussed.

To that end, the rest of the document is structured as follows:

Chapter 3 describes the regulatory environment in terms of monitoring system governance and stakeholder involvement, and discussed the legal value of measurement results, openness, privacy and security.

Chapter 4 describes implementation aspects, covering measurement metrics and methods, current measurement systems, as well as specific aspects regarding wireless/mobile Internet access and complementary methods.

Chapter 5 describes future perspectives such as measurement system standardisation and cost aspects and discusses a potential multi-NRA opt-in system evolution.

2.2 Background

Legal basis

The Regulatory Framework provides tools which NRAs can use to pursue their regulatory objectives in the context of net neutrality.

The Telecom Package adopted in 2009 gave new powers to regulators and reinforced obligations on operators in relation with quality of service.

- The policy objectives in the Framework Directive provide guidance for how the provisions in the four specific Directives should be understood. They include an overarching objective of guaranteeing access to an open electronic communication

service/network for the interest of the citizens of the European Union: "*promoting the ability of end-users to access and distribute information or run applications and services of their choice*", Art. 8(4)(g) FD.

- The Universal Service Directive introduces a comprehensive set of regulatory tools which aim at promoting a satisfying quality of service and ensuring better information on its actual quality of service which will be received. The tools can be divided into obligations on operators and powers for regulators.

  According to Art 20(1) USD, operators have to provide sufficient information to their current and prospective customers. This is first ensured by an appropriate level of precision of their contracts.

  However, obligations go beyond contractual information, as NRAs can request that operators publish transparent information in other contexts, according to Art 21(3), Art 22(1) and (2) USD. The European legislator therefore acknowledges that contracts might not always be the most appropriate medium to convey complex and changing information, and encourages NRAs to play an active role in increasing the quantity and quality of information available to Internet users.

  Additionally, the 2009 Telecom Package provided a power to NRAs. According to Art 22(3) USD, "*in order to prevent the degradation of service and the hindering or slowing down of traffic over networks, Member States shall ensure that national regulatory authorities are able to set minimum quality of service requirements on an undertaking or undertakings providing public communications networks*." The reach of this regulatory tool has been extensively examined by BEREC, which sees it as the third and final step of the regulatory approach to net neutrality breaches, as explained in its December 2012 NN QoS guidelines.

Articles 20(1), 21(3) and 22(1)-(2) provide the legal basis for main use case A, while Article 22(3) provides the legal basis for main use case B.

BEREC's previous work

Net neutrality has been a prominent area of interest for BEREC in recent years. BEREC has published a number of documents on key topics related to net neutrality, including market findings and guidance for regulators. Among those documents, some are especially relevant to quality of service and explore its interplay with best effort Internet provision, elaborate on its measurement, its publication, and analyse the regulatory process which may lead to minimum quality requirements.

In 2011, BEREC published a framework[6] for quality of service which sets out the foundations for defining QoS in relation to net neutrality, elaborating on quality-related concepts relevant to IP networks and on quality evaluation methods on the Internet. This framework provides a conceptual basis for subsequent works on quality of service.

[6] BEREC, A framework for quality of service in the scope of net neutrality, Document no. BoR (11) 53, 2011

In parallel, BEREC issued guidelines[7] on transparency which explore how the new Regulatory Framework transparency obligations should work in practice. They set out the type of information to be provided and how it should be conveyed. They also elaborate on the requirements for a transparency policy to be effective. They are especially relevant to the current document when considering transparency on actual quality of service.

In 2012, BEREC undertook an investigation[8] into traffic management practices and restrictions currently applied by ISPs. This unprecedented inquiry showed that a majority of ISPs offer Internet access services with no application-specific restrictions; however, some specific practices, such as blocking or throttling of peer-to-peer traffic or VoIP, may create concerns for end users and could sometimes be associated with discriminating actions. These occur more often in mobile networks than in the fixed network sector. The data gathered shows significant differences between countries. As a first exercise, it however did not provide a full and deep view of all practices currently in the markets.

In 2012, BEREC also published guidelines for quality of service[9] which provided recommendations to NRAs on when and how to exercise the new powers to impose minimum QoS requirements on operators. The guidelines for quality of service elaborate on how to assess restrictions, and how to reflect the particular context of national markets.

The 2011 BEREC transparency guidelines provide, among other things, the background for main use case A, while the 2012 BEREC quality of service guidelines provides the background for main use case B.

2.3 Quality monitoring system requirements

Requirements identified for quality monitoring are: **accuracy, comparability, trustworthiness, openness and future-proofness.**

Accuracy: The achieved measurement results should be reproducible and consistent over time. Accuracy requires that results are obtained from a clearly defined population and their statistical treatment is well documented, so that results can be interpreted without bias. Margins of error should be known and published.

Comparability: This includes “plain” comparability of individual sample measurements, but also comparability at higher levels, depending on the goals set by the NRA, such as comparability between IASes, and between countries when possible, so that degradation of certain offers, or degradation caused by specialised services, can be identified with a sufficient level of confidence.

Trustworthiness: The system components must be robust and protected against security attacks, and availability, integrity and confidentiality of the measurement data must be secured during storage and transmission. Privacy is also essential and the consent of end users regarding the treatment of their data must be obtained. The system governance must be designed in a way which mitigates conflicts of interest and ensures credible results. The

[7] BEREC, Guidelines on transparency in the scope of net neutrality: best practices and recommended approaches, Document no. BoR (11) 67, 2011
[8] BEREC, Findings on traffic management and other practices resulting in restrictions to the open Internet in Europe, Document no. BoR (12) 30, 2012
[9] BEREC, Guidelines for quality of service in the scope of net neutrality, Document no. BoR (12) 131, 2012

accountability and legal value of the measurement results has to be taken into account early in the design process.

Openness: Details about the measurement methodology should be made available, and open source code should be considered as an option to achieve this requirement. Furthermore, transparency of collected data (“open data”) should also be sought, with due respect for the limitations of national legislation. As far as possible, a quality measurement system should be based on state-of-the-art specifications, standards and recommendations and best practices, in order to leverage on established know-how and thereby contribute to the best possible measurement results. This will also improve the capability of the system to support comparability (a requirement underlined above).

Future-proofness: The system design should ensure flexibility, extensibility, scalability and adaptability. This implies that cost-effectiveness should also be applied as a general rule-of-thumb to all phases of the measurement system lifecycle, like development, deployment and operation.

These requirements described above contribute to the overall accountability of the quality monitoring and should therefore be taken into account when NRAs design a quality monitoring system.

In the subsequent chapters, quality monitoring solutions will be analysed against these requirements, by considering how well different solutions comply with different requirements. In general, it can be expected that a single monitoring solution will not comply with all the criteria, and therefore trade-offs will need to be made. Also, compliance with some criteria may come at a cost, and the conditions of this compliance should be explored.

3. Regulatory environment

3.1 Introduction

When an NRA establishes and operates a quality monitoring system, the overall regulatory environment plays an essential role, as it provides legal as well as non-legal principles that the NRAs need to consider.

Usually, the applicable national legal framework indicates, or even clearly defines, the objective as well as the constraints of the quality monitoring. On this basis, the regulator decides on the establishment and usability of a monitoring system. As introduced in chapter 2, the two main use cases are typically:

- **ensure transparency** on the quality of the Internet access service for end users (use case A), and
- **regulatory supervision** through monitoring quality of the Internet access service with regard to potential degradation of service, and enabling the NRA, if considered necessary, to impose minimum quality of service requirements (use case B).

When defining the establishment and usability of the quality monitoring system, the regulator should consider the following questions

- which approach should be taken - who should control the monitoring system and how should the stakeholders be involved (the governance issue)?
- what legal consequences and effects can measurement results have (the legal value of measurement results issue)?
- how can information about the methodology of a monitoring system and its results be provided (the openness issue)?
- how can the privacy of the users of such a monitoring system be ensured (the privacy issue)?
- what measures need to be undertaken to ensure the security and integrity of the monitoring system (the security issue)?

3.2 Governance

3.2.1 Regulatory approaches to quality monitoring

There are different regulatory approaches to implementing a quality monitoring system which can be broken down to the following options:

Traditional regulation

The quality monitoring system may be implemented and run by the NRA – either by itself or by using an independent measurement provider chosen by a public procurement. Given a sufficient legal basis, the NRA may also impose the establishment of a quality monitoring system on the ISPs.

The former allows full control over the methodology of the implementation as well as over the generated measurement data. If the implementation of the quality monitoring system is imposed on the ISPs, further steps may be necessary for an effective control.

Since the NRA exercises control over the quality monitoring system, it has to take into account the applicable legal frameworks – not only with regard to the material legitimisation of the quality monitoring system, but also with regard to procedural aspects (see below at 3.2.2).

Co-regulation

Under some circumstances, NRAs may find it appropriate to establish joint regulator-stakeholders systems, rather than imposing implementation merely on ISPs. Under such a scheme, cooperation with stakeholders may be useful to meet specific needs and/or regulatory objectives, such as: (i) system development by independent research institutions; (ii) performing measurement campaigns with the help of consumer organisations; (iii) publishing results on "third party comparison websites".[10]

Therefore, NRAs can choose to involve different stakeholder categories and set up a cooperative monitoring system to share responsibilities and costs with independent research organisations and other third parties. Such cooperative systems using different governance solutions are already in place in some national markets: public-private partnerships based on inter-institutional agreements, establishment of advisory bodies such as steering committees and technical roundtables, or international cooperative forums.

When the set up and operation of such cooperative systems is done with explicit government involvement, the resulting measurement system is based on co-regulatory principles. The USD makes explicit reference to such principles by stating that: "Co-regulation could be an appropriate way of stimulating enhanced quality standards and improved service performance. Co-regulation should be guided by the same principles as formal regulation, i.e. it should be objective, justified, proportional, non-discriminatory and transparent."[11]

Self-regulation

Finally, under some circumstances, NRAs may decide to leave measurement systems to be deployed by the market, and promote *self-regulatory* initiatives for the implementation of relevant measurement methods, as well as the publication of monitoring results, through moral suasion.[12] For instance, NRAs may launch education and information campaigns to increase consumers' awareness of the availability and use of measurement tools, while inviting ISPs to make available user-friendly tools to their customers. Here, the NRA may have some influence, but does not control the methodology of the quality monitoring system, its implementation or the generated data.

The appropriate approach highly depends on the goals of the quality monitoring system, and whether or not the NRA needs to exercise control over the quality monitoring system. The measurement method chosen may also require more or less involvement of the ISPs in the system governance and implementation. It should however be noted that independency of

[10] Third party comparison websites are included by BEREC as one of a number of relevant bodies which can play an important role in providing transparent information to end users regarding the quality of the Internet access service, Ref. BEREC, Guidelines on transparency in the scope of net neutrality: best practices and recommended approaches, Document no. BoR (11) 67, 2011, p. 3

[11] Ref. Recital (48) of the Universal Service Directive, as amended in 2009

[12] See also the EU Better Regulation guidelines and Inter-Institutional Agreement on Better Lawmaking at http://ec.europa.eu/governance/better_regulation/instruments_en.htm

the measuring body, in particular third parties, should be considered as a part of trustworthiness and overall accountability of a monitoring system.

3.2.2 Procedural aspects and stakeholders' involvement

In cases when development of a quality monitoring system is not left to the market (both with regard to use case A and use case B), Art 6 of FD and Art 33 (1) USD require that stakeholders' views are duly considered. The latter provision states: "*Member States shall ensure as far as appropriate that national regulatory authorities take account of the views of end-users, consumers (including, in particular, disabled consumers), manufacturers and undertakings that provide electronic communications networks and/or services on issues related to all end-user and consumer rights concerning publicly available electronic communications services, in particular where they have a significant impact on the market*".

The provision, therefore, requires national legislators to ensure that NRAs can consider the views of interested parties on proposals having an impact on the rights of all end users and consumers of electronic communications networks and services.

The decision about when to inform stakeholders, and to what extent to cooperate with them, depends very much on the respective national situation regarding the NRA, stakeholder relationships, and also on the overall market as well as political situation in the specific country.

Therefore, the NRA should first assess which stakeholders should be involved. As a general assessment, relevant stakeholders may include ISPs, consumers and consumer associations, research centres, measurement system operators, manufacturers and CAPs.

The involvement of these stakeholders could have positive effects on the overall development of the system itself as it:

- ensures that feedback can already be taken into account at an early stage of the development process, enabling improvements and clarifications to be implemented in a much faster, easier and less cost-intensive way;
- minimises or eliminates possible misunderstandings or uncertainties that the diverse stakeholders might have;
- allows for stakeholders to develop confidence and trust in the quality measurement system, and therefore results, in (positive) acceptance of the quality monitoring.

Regarding the form of the stakeholder involvement, the NRA may choose from a number of different options regarding a regulatory proposal, which may include a public consultation.

For instance, in light of Art. 33 (3) USD, NRAs may also decide to promote the creation of cooperative forums including network operators, ISPs and CAPs, in order to ensure comprehensive, comparable, reliable, user-friendly and standardised information regarding quality parameters and traffic management practices which could impact on quality in the scope of net neutrality. Such participatory initiatives would tend to complement public consultations, which are to be held before cooperation between undertakings takes place.

In the same regard, regulators may seek information and support from a number of non-statutory stakeholders, such as research centres and independent measuring organisations, both at the national and international level, by setting up steering groups and/or technical

advisory boards[13]. Enhanced cooperation within such governance structures may be particularly helpful when consumer associations are included. Consumers are key stakeholders because, in most quality monitoring systems, measurements are to be performed inside the customer premises. This implies that NRAs promote proactive participation of consumers and their representative associations.

These participatory methods are not mutually exclusive and may be used to achieve different goals, according to the relevant stages of quality monitoring system development and operation. As a consequence, when keeping stakeholders informed and cooperating with them during the development of a quality monitoring system, it is necessary to determine, from the very beginning, the following:

- Which stakeholders should be involved and why? What could be the risks of involving stakeholders, i.e. do stakeholders have an interest in delaying the process?
- Which and how much information (technical, legal, design, organisational etc.) should they receive? When should they receive it?
- How should they be involved? Should they merely be kept informed or be allowed to participate in a more active way? How much additional time, workload and costs could the involvement of stakeholders mean for the regulator and should it thus be considered in the project schedule?

3.2.3 Funding aspects

The availability of financial resources also impacts on the choice of measurement system and governance structure. Some NRAs may have access to financial resources which allow them to autonomously launch a quality measurement system. In some jurisdictions, the legal framework may also foresee that Internet service providers, or other stakeholders, are required to contribute to funding a system which is designed and supervised by the NRA.

In other circumstances, NRAs may not have sufficient resources to fund a quality measurement system. NRAs could simply rely on an external solution over which they have no influence and which does not require any additional effort or contribution. Alternatively, NRAs could be motivated to look for partners and take part in a multi-stakeholder system. Possible partners are other public institutions at a local, national, European or international level, which may share an interest in recording measurements and could bring funding resources. Research projects and various initiatives coming from universities, end user organizations or private companies may also provide valuable resources. When joining these projects, NRAs can bring their expertise, notably their knowledge of markets.

When NRAs take part in a quality measurement system which they do not fully control, they should pay special attention to the interests of stakeholders and their potential conflicts with regulatory objectives. Even if such systems are likely to decrease the cost for NRAs, they could also decrease the trustworthiness of results and jeopardize the possibility to reuse them in a regulatory context.

[13] The creation of steering committees and technical advisory boards including consumer organisations, manufactures and CAPs is in line with the relevant provisions set out by the Regulatory Framework with the aim of ensuring that relevant stakeholders may actively cooperate in a regulatory environment to achieve the policy goals of quality monitoring.

3.3 Legal value of the measurement results

An important issue for regulators is the question of the legal value of the results of a quality monitoring system. For regulators to set up a measurement system with the overall aim of being objective and provider-independent and enabling users to undertake measurements implies and maybe even intends that an end user will rely on and make further use of the measurement results.

In use case A, an end user having a contractual disagreement with his ISP regarding the minimum guaranteed bandwidth might, for example, present his measurement results as proof. This could be when the end user tries to resolve the problem directly with his provider, when he brings a complaint to an alternative dispute resolution / regulatory conciliation body, or during a court case. Thus, the measurement results could well be presented as evidence at some point.

In use case B, it is equally important to have valid data to decide whether there is a degradation of service and which regulatory measures may have to be taken in response.

This imposes particular responsibilities on the quality measurement system, and the provider of the system. In particular, the provided information about the measurement methods and the measurement results are a major issue, raising questions including:

- can the measurement results be easily reviewed by interested parties (e.g. via open data, a public map)?
- what is the specific measurement methodology used (e.g. is comprehensive information provided about the methodology; is it open source)?
- what needs to be considered by the user before, during and after undertaking a measurement with the quality monitoring system (e.g. is information provided via a comprehensive FAQ; is there the possibility to contact the provider of the system in case of unresolved questions)?

Therefore, together with the measurement results, sufficient information must be provided to the public about the methodology used and its potential limitations and biases.

3.4 Openness about methods and results

Measurement results only have a meaning when they are supported by documentation on how - and under what conditions - the results were established. This is especially true when there is no single generally accepted methodology on how measurements should be done. Ideally the description of the methodology is useful for both consumers and technical experts. In other words, the detailed technical description should be accompanied by a high level summary of the most significant parameters/conditions. In the case that measurements are performed by third-party contractors, it needs to be made clear that all details of the measurement study should be made publicly available and these details should not be restricted due to a non-disclosure agreement.

Quality measurements systems are built upon software. Knowledge of source code is therefore the ultimate tool to make the measurement methodology transparent. While manuals and documentation describe the intended behaviour of software, only the source-code can reveal the actual implementation. Software which allows users to review the source-code is usually referred to as “open-source software”. There are also other forms of

giving insight into source code, e.g. some vendors of commercial software sometimes allow specific groups (e.g. government representatives) to review the source code of their software.

BEREC believes that open source is a good practice which supports transparency of methodology and the effective proliferation of measurement systems, but it is not the only option, and in many circumstances regulators will not have the choice to make the source code available.

A measurement campaign creates a data set for every single measurement. That dataset contributes to the collected data. In the case of an end user measurement, the data will be of special interest to that individual user (e.g. for the users' individual statistics and history). The total set of collected data is often referred to as "raw data". Further processing is required before a report can be generated. These steps might include removal of invalid measurements, statistical analysis and presentation. The final report includes tables, diagrams and figures of the measurement results.

If it is decided to publish results, typically the report is published. But that report is a specific interpretation of the raw data, someone might be interested in some specific statistics which were of no interest for report or were left out for other reasons. To do so, raw measurement data would be required. Such public data is often referred to as "open data".

Making (raw) collected data available might be challenging:

- Privacy has to be taken into account when data are collected (see next section). This implies that the user must be informed which data are collected, for what purpose and what information shall be included, how data will be used, and specifically that some information may be made available to the public.
- Raw data requires a lot of interpretation; therefore it is necessary to make the methodology of data collection transparent as well.
- It is necessary to define how data can be used. It is possible to define an individual license or to use existing licenses, e.g. those prepared by the Creative Commons initiative[14]. Some countries' open data initiatives propose some specific license requirements[15].

The decision about whether to make data available as open data, will also be guided by general policy decisions such as freedom of information legislation.

3.5 Privacy

Another crucial point that needs to be taken into account when setting up a quality monitoring system is the issue of privacy. It is of utmost importance that the provider of a quality measurement system strictly abides by the respective European and national legal requirements, and also provides users and the public with adequate information about this. If

[14] "Creative Commons", accessed January 24th, 2014, http://creativecommons.org

[15] e.g. "Open Government License", UK Government, accessed January 24th, 2014, http://www.nationalarchives.gov.uk/doc/open-government-licence/ , or Eibl, G. et al, "Rahmenbedingungen für Open Government Data Plattformen", e-Government Bund-Länder-Gemeinden White Paper, Version 1.1.0 (2012). Online: http://reference.e-government.gv.at/uploads/media/OGD-1-1-0_20120730.pdf (accessed January 24th, 2014).

legal requirements are followed, critical questions which might arise on publication of the data can be responded to very quickly. Additionally, trust and confidence in the system, as well as the reputation of its provider, i.e. the regulator, will be assured.

As the development of the measurement system could be considerably affected by any legal requirements, it is necessary to determine the following aspects from the beginning:

- what are the respective legal requirements at the European, and especially national, levels regarding the treatment of data when a measurement with a quality monitoring system is undertaken;
- which specific data is actually needed to provide qualitative measurement results;
- could the usage of any data, especially *personal data* and *sensitive data*[16], raise a concern or even cause criticism publicly and what can be done to best handle and alleviate such situations?

It should be noted that the respective legal requirements can be rather diverse and comprehensive, in formal as well as material content, by country[17]. Examples would be:

- the concrete definition of which data are or might be considered as *personal data* (e.g. in some European countries the IP addresses are classified by the highest courts as personal data, while in others they are not);
- the definition of what is precisely understood by the legal term *sensitive data*;
- the usage and treatment of *data*, *personal data* and *sensitive data* by the provider of the quality measurement system (this can differ depending if the data is merely *processed* or if all data or part of the data is *transmitted* within or even outside of the EU; also specific data, such as *personal data*, must often be completely deleted after a certain amount of time, independent of whether these data are available to the public or not)[18];
- measures that need to be fulfilled in order to secure the data (see next section);
- steps that need to be taken to assure the consent of the user of the measurement system regarding the treatment of his data (providing a privacy policy; the (active) acceptance of the privacy policy by the user before being able to undertake a measurement for the first time; and the logging of this consent in order to have a proof; renewing the acceptance in case of updates or new releases of the measurement system; possibilities for the user to withdraw his acceptance);

[16] The EU defines *personal data* (which also encompasses the legal term *sensitive data*) as *"any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity"*.
A special category of data is *sensitive data*: *"Member States shall prohibit the processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life."*.
Viz. Art. 2, 8 of the reference document: Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995], OJ L 281

[17] The main national sources in this regard are the data protection act, the telecommunication act and decisions of the respective highest national courts.

[18] Data that is *transmitted* is here understood as data that third parties, thus parties other than the provider of the measurement system and the user who undertakes a measurement, have access to. Such data may be publicly available or not.

- the necessity of a mere notification or also the permission of the respective data protection body or data protection commissioner;

In addition to complying with the law, it is highly recommendable to be informative and transparent towards the user of the quality measurement system as well as the interested public.

3.6 Security

The main goal of a quality monitoring system is to measure quality parameters of the Internet access service. As a result, in any type of implementation both measurement agent endpoints may be exposed to different kinds of attacks.

It is of high importance for the provider of a quality monitoring system to ensure and uphold the security and integrity of the system and to fulfil respective national and European legal requirements. At the European level, emphasis in this regard should be put on the *Data Protection Directive.* In a wider context, also Art. 13a FD as well as the *Cybersecurity Strategy* of the EU and the current discussions regarding this topic ought to be considered.[19]

The client side of the quality measurement system is located at the premises of the end user, usually in uncontrolled software, hardware and network environment that makes the client susceptible to attacks. A compromised computer with the measurement client can pose a threat to the entire quality monitoring system and the end users participating. In the central site, the main attack threat is the various kinds of denial-of-service attacks against the measurement server, because its network interface is exposed to the Internet.

A security incident can lead to serious consequences for the quality monitoring system; service outage, loss of measurement data, breach of user data, and even the conversion of a measurement client network into a botnet. Any of these security incidents could weaken the trust between the voluntary users and the regulator, which may threaten the feasibility of the entire measurement system (refer also to section 4.2.3).

Regarding these attack vectors and the possible consequences, the cybersecurity issues have to be dealt with the same – or maybe even more – seriousness as other issues, e.g. measurement technical issues.

[19] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995], OJ L 281, or the reference document: European Commission.
"Cybersecurity Strategy of the European Union: An Open, Safe and Secure Cyberspace", 2013, online: http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=1667 (accessed January 24th, 2014).

3.7 Conclusions and recommendations

Regulatory approaches

When choosing between the different regulatory approaches, the appropriate solution will typically depend on the use case: In particular, when the main objective for quality monitoring is regulatory supervision, a regulator-controlled monitoring system will, in most cases, be the best.

The technology chosen and the tools available on the market should also be taken into account when choosing the governance approach. For instance, measurement tools from independent organisations (e.g. research institutions and other non-profit entities) may also contribute to provide consumers with reliable and comparable information regarding quality of IAS.

On the other hand, it may also be appropriate to set up complementary measurement systems to address the goals of the NRA and to limit risks. For example, a stakeholder-controlled, hardware-based system could effectively be complemented by a regulator-controlled, software-based tool, as each one has different risks, biases and benefits.

Procedural steps and stakeholders' involvement

The level of stakeholders' involvement according to Art 6 of FD and Art 33(1) USD depends on the overall market conditions, regulator-stakeholder relationships and also on the relative cost-effectiveness of such involvement in a specific regulatory environment. Therefore the need for information and cooperation from different stakeholder categories should be carefully assessed, and the different kind of support (technical, financial, operational) they may seek at the different stages of the monitoring system lifecycle. To the extent that information and cooperation from stakeholders is deemed to be advantageous, BEREC recommends assessing whether it is useful to:

- take into account the views of stakeholders at the very early stages of the process;
- encourage proactive involvement by relevant stakeholders in line with Art. 33(2) USD;
- promote the creation of cooperative forums.

A useful step for planning a participatory process is the preliminary identification of relevant stakeholders, which may differ according to the quality monitoring system objectives and architecture. In general terms, BEREC recommends involving the following categories of stakeholders:

(i) research organisations, consumer and business users' associations, CAPs and measurement providers are especially relevant in order to address quality monitoring to net neutrality objectives;
(ii) ISPs, consumer organisations, individual end-users and measurement providers are especially relevant for measurement of performance.

Possible forms of stakeholder involvement may include public consultation, cooperative forums including network operators, ISPs and CAPs, as well as advisory bodies or steering groups. These forms are non-exclusive participatory methods which may be used to achieve different goals, or to include different stakeholder categories, according to the relevant stages of quality monitoring system development and operation.

On the legal value of measurements results

- The provider of a quality measurement system should keep in mind that measurements undertaken with their system may imply legal usage of these measurement results (e.g. before a court).
- This imposes particular characteristics on the quality measurement system: it should be objective, robust and legally grounded, and information should be provided about the measurement method and results.

Being transparent about methods and respecting privacy at the same time

Trust and confidence in the quality monitoring system should be encouraged notably by considering the following:

- providing information on how, and under what conditions, measurement results are established (e.g. relying on the usage of open-source software),
- making “raw data” publicly available (open data principle),
- respecting and transparently ensuring compliance with European and national legal requirements regarding the issue of (data) privacy as well as the security and integrity of the system,
- limiting the processing of such data to cover only that which is actually needed to provide qualitative measurement results.

4. Implementation aspects

4.1 Measurement metrics

Many of the metrics currently used to evaluate quality of the Internet Access Service (IAS) help end users to choose an appropriate IAS offer. The metrics also allow NRAs to evaluate the performance of IAS offers in their respective markets. Standards developing organizations (SDOs) have specified various parameters which can be used to evaluate quality of the IAS as a whole. Nevertheless, this information is often too technical to allow end users choose an IAS offer suitable for their usage of the Internet.

Popular applications impose different demands on the underlying Internet access service. For example, applications such as web access and streaming media need high data transmission throughput. In contrast, VoIP and gaming are sensitive mainly to delay and delay variation. Whilst sensitivity to packet loss or packet error is in many cases mitigated by the use of TCP data retransmission, this may not be suitable to real-time applications.

In this section we consider international/European standards and recommendations from CEPT, ETSI, ITU and IETF.

In CEPT's ECC report 195,[20] the following quality metrics have been selected: upload and download speeds, delay, delay variation, packet loss ratio, and packet error ratio. The criteria CEPT used to choose the relevant standard was primarily based on the ETSI Guide EG 202 057.[21] The first two definitions rely on the ETSI Guide; the last three are also referenced in the guide, but the main sources are ITU-T Rec. Y.1541 and G.1010.

IETF also defines a set of quality metrics in a similar way to those defined by ITU and ETSI. However, some variations exist. For example, RFC 2679 and RFC 2681 define one-way and round-trip delay, RFC 3393 defines delay variation, and RFC 2680 and RFC 6673 define one-way and round-trip packet loss metrics. RFCs 3148 and 6349 define frameworks for measurement of transmission speed using the TCP protocol.

When measurement samples of quality metrics such as speed or delay are collected, they will have temporal and spatial variation. To account for temporal variation, statistically derived metrics such as average values, percentiles (e.g. highest 95% and lowest 5%) and standard deviation are used. Spatial variation is described in section 4.5 below.

In some cases, theoretical maximum values are used to describe the performance of a service. However, from the end user's point of view, the actual value of the performance is more relevant for evaluating real experiences when using the service. Deriving average actual values requires thorough statistical analysis, and usually it will be preferable to distinguish between peak-time and off-peak-time during the analysis.

[20] CEPT Electronic Communications Committee, "Minimum Set of Quality of Service Parameters and Measurement Methods for Retail Internet Access Services", ECC Report 195, 2013, online: http://www.erodocdb.dk/Docs/doc98/official/pdf/ECCREP195.PDF (accessed January 24th, 2014)

[21] EG 202 057 is relatively old and was written for dial-up Internet access. It is only applicable for the network section from NTP to the RADIUS server. This is a different scope than the IAS discussed in this report.

As described in BEREC's 2012 NN QoS Guidelines, an important quality indicator is achieved by monitoring the effects of congestion in the network. The difference between peak and off-peak values is a way of quantifying this effect, as well as the duration of peak-time on a daily and weekly basis.

In the context of net neutrality, performance of individual applications is also important, because it can be used to detect potential degradation of individual applications. The following table, based on ECC report 195, illustrates popular applications by non-professional users and the relevance of the quality parameters on the performance of those applications. In the table below, the relevance goes from '–' (not relevant) to '+++' (very relevant). When evaluating quality aspects of IAS in the context of net neutrality, it is essential to evaluate potential degradation of individual applications based on such considerations.

Table 4-1

Application	Data transmission speed		Delay	Delay variation	Packet loss	Packet error
	Downstream	Upstream				
Browse (text)	++	-	++	-	+++	+++
Browse (media)	+++	-	++	+	+++	+++
Download file	+++	-	+	-	+++	+++
Transactions	-	-	++	-	+++	+++
Streaming media	+++	-	+	-	+	+
VoIP	+	+	+++	+++	+	+
Gaming	+	+	+++	++	+++	+++

Glasnost is a well-known measurement tool that can be used to detect degradation of individual applications.

Technical standardisation is a continuously ongoing process. Within the area of Internet communication, IETF has a central role, being the organization that has developed, and continues to develop, the IP technology. Therefore, IETF has a particularly important position regarding IP quality measurement.

It is important to standardise and harmonise definitions, scenarios and methodology for how to measure the quality of the Internet access service. This harmonisation should be applied across all ISPs in a country and ideally in the pan-European space.

4.2 Measurements using injected test traffic

4.2.1 Test measurement architectures

Two different typical measurement architectures are used by the ISP and/or by other entities, such as NRAs. The first approach is used to measure the quality of the ISP leg. The second approach extends beyond the ISP including connectivity to the Internet. This normally extends to a central IXP where the ISPs under consideration interconnect with each other. It may also extend to other prominent IXPs, to which ISPs are not directly connected, or it could be cloud-based. In the following subsection both of these scenarios are analysed.

Measurement of the ISP leg

Here, the path to the measurement server stays within the ISP's network and will cover a segment of the network infrastructure used for the provision of IAS, depending on the server's location. The control of the ISP is related to the part of the network operated by the ISP, i.e. the infrastructure between the network termination point (end user side) and where the interconnection to other ISPs takes place. The following figure gives a schematic representation of this scenario.

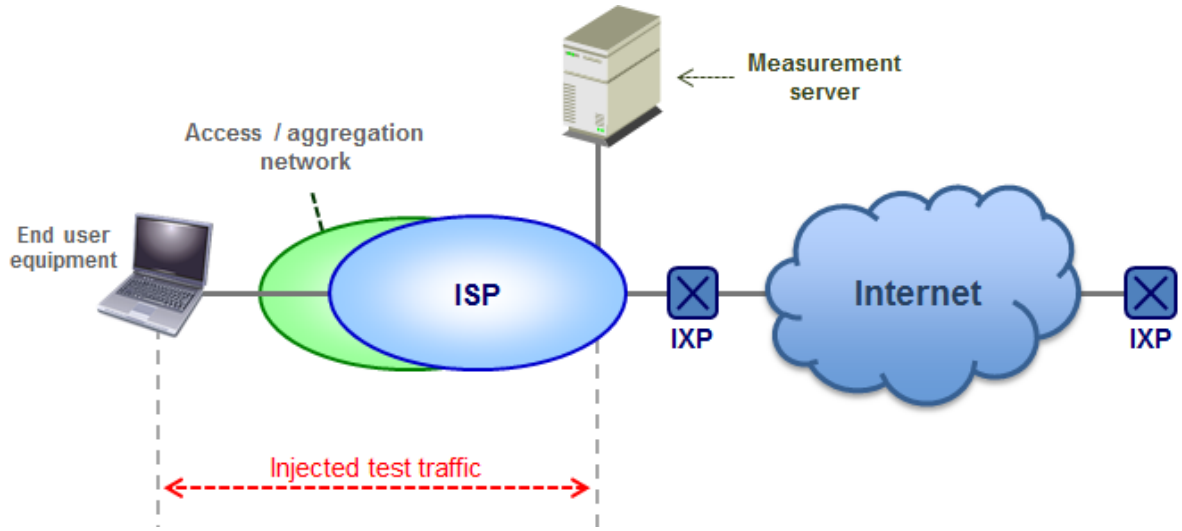

Figure 4-1 – Measurement of the ISP leg

An important point related to this scenario is that the measurement server should be located as close as possible to the edge of the ISP's network (e.g. the peering router) in order to measure all potential bottlenecks of this network. By assuring that all ISPs are measuring metrics based on the same methodology and comparable reference points, the comparison between network performance of ISPs is in theory possible, although real comparability is unlikely to be achieved since separate systems may have differences in implementation.

Nevertheless, the end user is not able to obtain a complete view of the quality of the IAS itself. This is because the metrics refer to the ISP's internal network and do not take into account the connectivity to the Internet. This connectivity is important to evaluate because the quality of IAS is also determined by connectivity to the Internet, i.e. the higher the interface capacity and interconnectivity of the interconnection agreements are, the better the quality of the IAS becomes.

Measurement beyond the ISP leg

In this scenario, the measurement path includes a complete Internet connection from client to measurement server beyond the ISP and may thus be carried out without involvement of the ISP, e.g. by an NRA or an independent measuring organization. This scenario reflects all aspects of the quality of the IAS as opposed to the scenario with the ISP leg described above, since interconnectivity to the rest of the Internet is included in the measurement.

In some cases there may not be larger differences between measuring over the ISP leg and measuring towards a more distant location, like an IXP. However, if an ISP provides better performance to some interconnections than to others and thereby introduces different levels of performance to different destinations, it is valuable for the NRA and the end users to have transparent quality information about this effect.

Furthermore, to the extent that the measurement results from different scenarios give the same values; centralised servers could replace individual servers to save cost, in particular the costs of operation and maintenance. In general, the comparability of the IASs can be better achieved when using a central measuring point. On the other hand, since Internet traffic can take different paths, and most ISPs have several interconnections to the Internet, a *statistical average over several central servers* will give a better estimate of the quality of IAS under real-life use conditions.

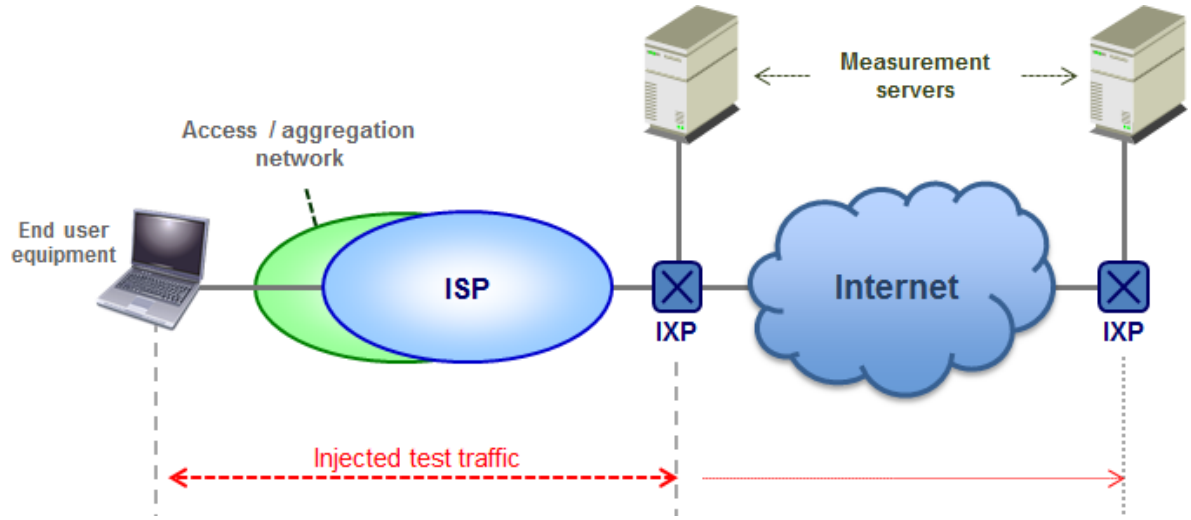

Figure 4-2 – Measurement to the IXP

The disadvantage with a “test server connected to *one* interconnected IXP scenario” is that it may not accurately reflect the actual end-to-end Internet experience which end users receive. Ordinary traffic from end users may be routed via different paths which would influence the quality. A risk also exists that the ISP may “groom” the route to the measuring point to ensure better performance than ordinary Internet traffic.

4.2.2 Hardware based versus software based methods

Quality measurements can be performed with or without full control of the measurement termination unit at the end user side. The following subsections consider both alternatives. (Passive measurements are covered in a section 4.3.2 below.)

Hardware-based methods

Here, a dedicated hardware probe is used for the sole purpose of measurement. There are two options to consider:

- The probe completely replaces the end user's equipment. No other equipment can be connected to the Internet access while the probe is performing measurements. This is applicable also in the case of mobile Internet access.
- The probe shares the Internet access with ordinary traffic, e.g. by connecting a probe to a residential gateway. In this case, there may be bottlenecks when specialised services (e.g. IPTV) use the access, since they may throttle the Internet traffic, creating problems if the probe is not able to check all traffic that flows through it.

Hardware-based methods may allow the possibility of executing measurements during periods when the users are not actively using their IAS, i.e. when there is a low level of end user generated traffic. For that reason, periodic tests are possible and measurement process do not need to disturb the end user's usage of the IAS.

On the other hand, these methods imply higher costs because of the costs of the probes themselves, as well as their installation. The cost also includes panel selection and guidance of each panellist separately. Thus, this method may be more suited to measuring a limited number of end users (such use case B1).

Software-based methods

In this category, there are two options to consider:

- A web-based tool, where the download and execution of the measurement software is initiated via the end user's web browser by accessing a specific web page.
- A dedicated software client, where the measurement software is permanently installed on the end user's terminal equipment. In this case, different versions of the software are needed to support different operating systems and terminal equipment.

The web-based tool generally has a low cost, and installation within the operating system is not needed. Nevertheless, it still requires collaboration from users, who must initiate the measurements. Measurements made by this approach are likely to be influenced by different browser software, its interaction with the operating system and the presence of other activities on the access.

The dedicated software client method is also a relatively low cost solution, but will require installation of the software by the user. The software may be activated by the end user for each test sequence. Alternatively, it may be made more intelligent by detecting the level of activity on the end user equipment and its Internet interface in order to perform the measurements autonomously.

Other kinds of verifications are normally necessary to implement in the software client in order to make the software-based measurements more controlled, e.g. excluding measurements over wireless LAN.

Note that with both options the end user may be required to give some information regarding their IAS, such as contract information. Whilst these methods facilitate a wide range of users and a quicker deployment time compared to hardware based methods, they can only perform measurements and data collection when the users are online.

4.2.3 Controlling the measurements

When designing a measurement system, it is essential to guarantee that there will be enough server capacity so that the measurement service does not itself act as a bottleneck and distort the results. The capacity can be pre-allocated and controlled in a way that clients can perform measurements only according to a pre-defined schedule. With this mechanism it is possible to make sure that there is always enough capacity and resources for all active measurements.

The other option is to implement an access control scheme, where the system queues clients and only allows a given number to measure at the same time; i.e. buffering measurement requests when the system becomes overloaded. This type of load control is essential for a system that allows clients to measure at their own initiative. Otherwise the reliability of the measurement results will be compromised.

4.3 Measurements using ordinary user traffic

4.3.1 Application measurements

Application measurements cover end-to-end performance of specific applications. All the parts of the Internet effecting user perception influence application measurements. The drawback is that when low performance is detected, it is difficult to determine which section of the end-to-end communication path causes the impairment. To determine this, it is necessary to compare the results on different applications and from different ISPs and / or to perform specific measurements on some sections of the relevant parts of the networks used.

Collecting information from CAPs allows values of technical quality parameters to be obtained which represent the experience of the end user using the corresponding application. Some CAPs provide such information today.[22] The figure below shows this measurement scenario.

[22] See for example “Netflix ISP Speed Index”, Netflix Inc., accessed January 24th, 2014, http://ispspeedindex.netflix.com/ .

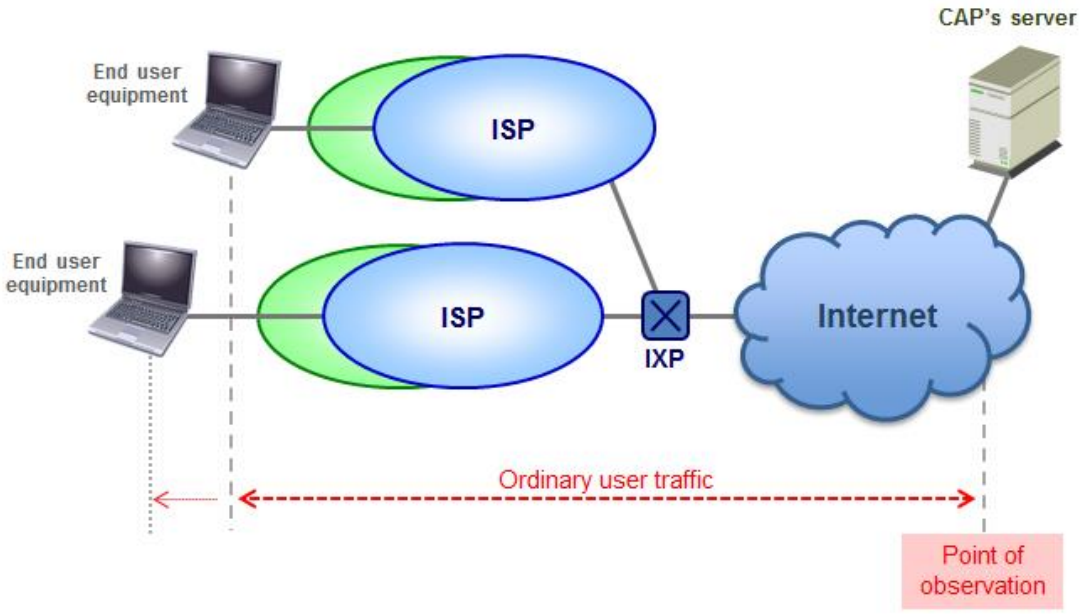

Figure 4.3 – Application measurements

CAPs are in a good position to observe the performance of end user communication as it can be done automatically every time their users access the content. However CAPs do not normally have any control over the other applications used by the end user or visibility of the traffic generated by them. Also, CAPs' own service, connectivity and server capacity may limit the achieved results. Furthermore, it is important to perform an objective evaluation of results from CAPs' measurements if such results are to be used by NRAs.

Additionally, quality measurements based on user-initiated communication with CAPs can be handled from the client side. This can, for example, be used for monitoring the download time of webpages and performance metrics for streaming video.

4.3.2 Passive measurements

While the most commonly used measurement methods today rely on active measurements using a client/server configuration, passive monitoring is an alternative to such approach. Unlike active monitoring, passive monitoring does not inject test traffic into the network and does not rely on dedicated test servers. Instead, passive monitoring tools analyse ordinary traffic in order to infer quality-related parameters. Both approaches can be used to meet the objective of monitoring quality of IAS in the context of net neutrality, each one with its pros and cons, and they should be regarded as complementary. The following figure gives a schematic representation of the passive scenario:

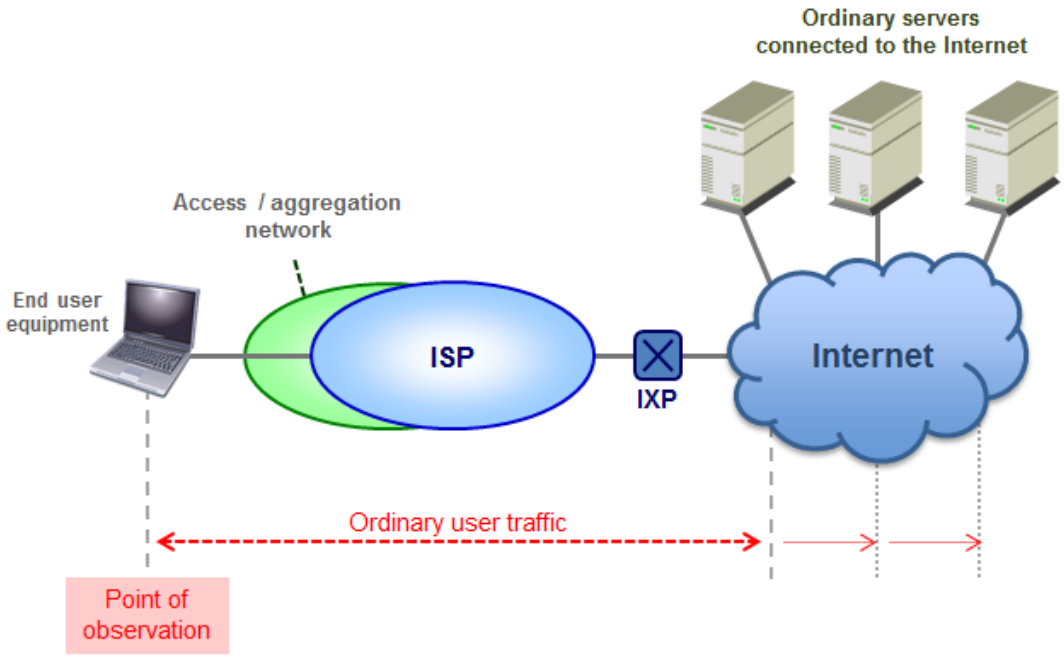

Figure 4-4 – Passive traffic measurements.

In relation to degradation of IAS as a whole, active measurements typically inject artificial traffic emulating web communications using the HTTP application layer protocol. Here, passive monitoring can play an important role, because it measures traffic originated from ordinary user applications and the reported quality metrics will be the ones that the user is currently experiencing end-to-end. Passive monitoring can measure performance of IAS as a whole by monitoring aggregated traffic from a mixture of applications sent to, and received from, the Internet.

By doing so, it is possible to verify whether the reported quality values provide sufficient performance for a best effort IAS service as a whole, by making comparisons with a minimum level of performance for a mixture of different types of applications. In this case, the results obtained by active monitoring cannot generally be extrapolated to the user experience, because server location, traffic in the network, and the test applications used will not meet the same conditions as for ordinary Internet communication.

In relation to degradation of individual applications using IAS, passive monitoring also provides advantages compared to active monitoring, since it is particularly helpful in identifying net neutrality violations. Existing systems for detecting differential treatment that rely on active monitoring are typically specific to an application or to a particular differentiation mechanism. If the ISP changes differentiation practices, the active monitoring tool may not be able to detect it. However, the users need a method to detect differentiation for any application that might be subject to differentiation and for any mechanism used to achieve it, and passive monitoring provides a better alternative.

Passive monitoring can partially solve these problems because its aim is to measure network performance for user-originated traffic. Passive monitoring does not try to identify the policies implemented by an ISP but only measures the end-to-end performance. From here it is possible to use statistical methods to identify if any traffic management practices

are being implemented, since it is possible to establish causal relationships between the ISP's practice and performance degradation.

Another advantage of passive monitoring is that it does not change the traffic load imposed on the network in any way. This is particularly important in mobile environments. Whilst most of the fixed IAS offers provide a flat rate tariff, on a mobile IAS offer this might not be the case. Data rates and data caps are offered on most of the mobile offers and this might limit the usage of active tests since they reduce the remaining capacity for the user as a result of traffic injection.

However, there are some drawbacks with passive monitoring too. The main inconvenience for passive monitoring is the lack of control on the server side and on the network conditions. Server and network capacity, connectivity and load will influence the resulting quality, and the conclusions reached need to take into account those aspects. Nonetheless, statistical data post-processing or the use of complementary active monitoring tools, allow for the isolating of traffic management practices from other causes of degradation such as overload, incorrect configuration, technical failure, etc. It is also important to note the lack of commercial experiences with passive monitoring due to its complex implementation.

NANO is a passive measurement tool that can inspect traffic to and from end users' hosts and detect potential degradation of individual applications and thereby be used to investigate net neutrality violations.

4.4 Current measurement systems

A number of European NRAs have been engaged in IAS quality measurement initiatives for many years and in varying levels of details and focus. The measurement parameters and methods used are basically the same, independent of the approach chosen. Differences can be observed with respect to the measurements for detection of potential degradation of individual applications. The majority of measurements, however, are focused on the measurement of the quality of IAS as a whole.

4.4.1 General system aspects

The clear trend for measurement approaches covering both use case A and B is a client/server architecture using injected test traffic. Two basic architectures can be observed:

1. **Provision of a software-based monitoring system with measurement agents that can be downloaded to any end-user IAS**. The measuring agent may, in principle, be both hardware- or software-based. However, a software-based solution is usually preferred due to lower costs, easier distribution and achieving full coverage. The software-based solution relies on the end users' own terminal equipment and measurements are user-initiated.

2. **Use of a probe-based monitoring system built for a pre-defined (limited) number of measurement end points (clients)**. The probes are located at network termination points, i.e. end user-like IAS accesses. They are either operated by the measuring organisation or by distributing hardware probes at end user premises. The number and distribution of probes follows a specific measurement system design depending on the scope and the desired statistical significance of the measurement

campaign. In both cases, server-side measurement end points are typically located at peering points.

The former approach is typically used for regulatory supervision where the focus lies on providing tools for precise and controlled measurements of specific network scenarios (use case B). The latter one is typically used for transparency issues (use case A) where the focus lies on achieving a high coverage of the whole market, ideally all end users.

When *individual end user information* is the goal, software-based measurements are preferred due to the coverage achieved, ease of implementation and lower cost. The intention is to allow end users to cross-check whether the contracted IAS quality is fulfilled (use case A2).

This architecture can also be used to provide *average results* (based on statistical analysis of collected results from all measurements) to inform end users about the quality to be expected at a specific location prior to signing a contract (use case A1). Measurement results are then collected based on a crowd-sourced approach.

Probe-based measurements are used where full control of the clients is needed, typically for scheduled measurements over a longer period. Prescheduled measurements are performed for a predefined set of Internet access services with respect to type, number and distribution. This approach allows for regulatory supervision of potential degradation of *IAS as a whole* (sub case B1).

NRAs that have a probe-based monitoring system can also use this to cover transparency regarding average IAS performance (sub case A1) when distributing a statistically relevant number of probes. Regulatory supervision of potential degradation of *individual applications run over IAS* (sub case B2) will typically be performed with dedicated measurement tools.

Based on the main use cases with associated sub cases, the overall monitoring methodology can be summarised like this:

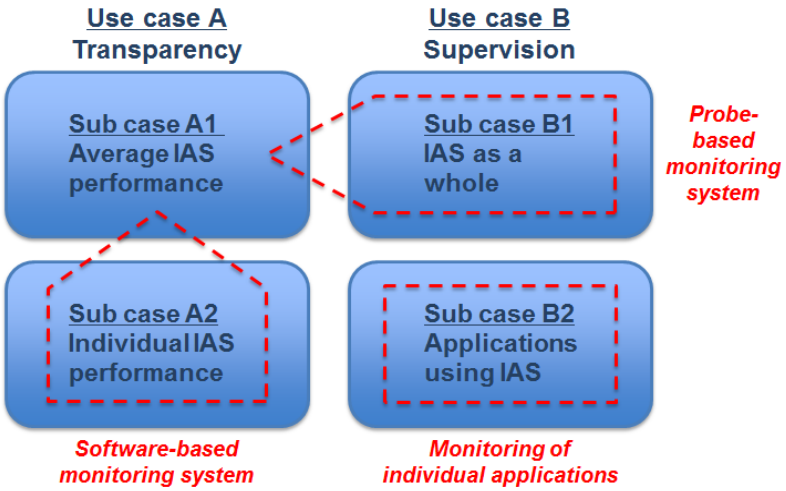

Figure 4-5 – Measurement systems vs. use cases.

4.4.2 General cost aspects

An internal BEREC survey conducted in 2013 identified at least 18 NRAs involved in quality monitoring systems (and two NRAs currently developing their own tools). These NRAs are either in charge of measurements or they cooperate with different stakeholders. Furthermore, in 21 European countries there are other quality monitoring tools made available to end users by public (e.g. university/research centers) or private institutions (e.g. ISPs, consumer associations, test houses).

To assist their decision process when setting up a quality monitoring system, NRAs can make a preliminary assessment of the actual costs they will incur by comparing costing structures and trends of different existing systems. In this respect, and in light of providing NRAs with useful information for such economic assessments, this section summarises findings from the BEREC survey regarding costing aspects of quality monitoring.

Different technical solutions (hardware-based vs. software-based), regulatory goals (transparency, regulatory supervision or a mix), steering entities (NRAs vs. stakeholders) are some of the reasons why there is a high variability both in the total cost and in the costing structure of the different national quality monitoring systems developed in a regulatory environment.

Costing structure

The survey confirmed that there are some common trends concerning the costing structure (that is the ratio between set-up costs and running costs).

Set-up costs are characterised by a high variability**,** making a precise assessment impossible. There are, however, a minimum number of cost elements which are common to all technical solutions: manpower, minimum hardware equipment and software licences.

Set-up costs appear less relevant for software-based solutions, although cost savings compared to hardware-based solutions do not seem large because software development and professional resources (IT specialists) represent the most significant cost component. Validation activities and preparation of reporting systems have little effect on set-up costs (or indeed on total costs). Adding additional metrics, while increasing overall costs, is expected to only imply relatively low marginal costs.

One common trend in cost structures is the fact that set-up costs are higher than running costs in most cases, and in some cases it may account for a very large part of the total cost. This is especially true for solutions like hardware probes, whereas web-based solutions, for example, have a more balanced costing structure.

Some governance and technical solutions allow for a certain degree of savings. For instance, in public-private international systems, nodes and equipment are hosted at the premises of associated measuring organisations at zero cost to participating NRAs. Other international partners may cover day-to-day operations and other running costs (e.g. remote support for maintenance).

In any case, as a whole, set-up costs (costs incurred to develop and install the measuring tool), seems relatively larger than running costs (annual costs sustained to run the quality monitoring system).

Total cost

Total costs differ significantly with regard to the system architecture chosen, the costs of which are of a different order of magnitude. Furthermore, total costs depend on several other characteristics, ranging from the market involved (fixed and/or mobile Internet access services) and including aspects related to the requirements of the quality monitoring system, such as equipment, technical personnel, housing and common costs.

- System architecture

In aggregate, total costs for existing quality monitoring systems vary over a wide range. Because system architectures and governance solutions are very different among countries, comparison based on total costs could be misleading or, at least, very difficult.

That said, it is worth noting that minimum resources needed to set up a quality monitoring system and associated implementing activities correspond, roughly, to the steps needed in order to implement a software-based system. Therefore, the software-based approach can be considered as a technical baseline model for quality monitoring systems. A hardware-based system includes more elements on top of the baseline system. It is therefore logical that hardware-based solutions would tend to be more expensive.

- Location of measurement servers

The server location impacts not only on the measurement costs but also on the total cost. A larger number of measurement points obviously imply higher costs than one single measurement point, in terms of both hardware and software requirements. There are indeed some economies of scale that should be taken into account when planning the measurement server location.

Placing the measurement servers within a national IXP could provide a significant saving. According to several academics and policy makers, once an IXP is established, it becomes a natural location for hosting a variety of other services dealing with bandwidth, speed and reliability of Internet. (In case of collaboration within an opt-in quality measurement system, several NRAs can share the cost of deploying a larger number of measurement servers, something explored in chapter 5.)

4.5 Composition of test traffic

The measurement architectures and methodologies for conducting active measurements by using injected test traffic are provided above. In addition to these considerations as a part of the design of a measurement regime, it is important to decide on the amount and distribution of measurement end points from which the test traffic is injected in order to achieve meaningful results.

As presented in section 4.4 there are two basic measurement architectures used that are related to the use cases A and B. Use case B – the regulatory supervision of specific network scenarios – requires a dedicated measurement set-up and a pre-defined set of IAS offers with respect to type, number and distribution of measurement end points. Use case A – transparency measurements – is based on user-initiated measurements and the aggregation of available measurement results (samples).

For each use case a different methodology for generating respective test traffic is followed. Use case B, which is by itself a specifically designed measurement scenario, calls for the design of a specific distribution of measurement end points (Internet access services) prior to the start of the measurements. The Internet access services selected to inject this traffic is referred to as the panel.

The motivation of use case A is to allow end users to perform measurements and to ideally collect measurement samples from all end users to create a database. So rather than designing specific measurement scenarios, the focus is first to allow for individual test measurements (samples) and secondly to achieve as large a database as possible. This is referred to as a crowd-sourced approach.

4.5.1 Preselected panel approach

The basic idea of this approach is to identify a set of IAS access points providing a valid representation of whole the population of existing access points. This set is referred to as a (preselected) panel. Each end user access point serves as a measurement end point (client). In combination with the measurement server(s), the clients constitute a "measurement network". The end user access points are equipped with hardware probes that control the client-related measurement aspects. By injecting test traffic at the end points of this network, measurement samples are generated and collected for subsequent analysis.

Regarding the panel selection, there is a trade-off between panel size, statistical accuracy and ultimately cost. The first step towards determining the participants of the panel is to specify the particular market aspects which are of interest. Measurement points should be selected based on statistical sampling. IAS packages may be first identified based on the popularity as well as the market share of the respective ISPs and/or the similarity of the packages to other ISPs. Assuming that technology and geography are of interest, there should also be a consideration of how the selected packages are technologically delivered and geographically distributed.

During the selection, the necessary number of participants per-technology-type per-package are chosen in view of the market aspect required to be studied. For example, if a particular package market (population) is to be studied, a group of participants subscribing to the package will be needed. The group size must be drawn from all the technologies over which the package is delivered. The number of participants for the package must be sufficient in size to take into consideration the confidence interval for the population. There could also be other relevant criteria, such as geography (urban/suburban/rural).

In order to achieve the target number of probes for each particular market aspect, a recruitment campaign is conducted with the aim of establishing an initially large number of potential volunteers which may eventually be selected for the evaluation process. Given the unpredictability of the standard deviation of the chosen group of participants, it is possible that the number of participants will have to be increased in an incremental manner until a satisfactory margin of error is achieved.

If panel participants are not directly selected by the measuring organisation based on their IAS offers, but instead by a call for volunteers, then a pre-screening and preliminary speed measurements should be undertaken along with checks on IP addresses. This is in order to reduce the impact of respondent misconceptions regarding which package they were using.

For example, during the pre-screening of suitable panellists, an average speed reading estimate may be undertaken using a software tool to give a rough estimate. If this reading is consistent with the panellist's package and their IP range matches that assigned to that IAS package, they are accepted onto the panel.

In general, once the pre-screening checks are completed and successfully passed, the particular volunteer becomes a participant of the representative panel and subsequently issued with all the necessary equipment. The actual panel is constructed by meeting the required number of probes for each particular market from the identified potential volunteers. A safety margin should be incorporated to account for future changes in the recruited panel.

4.5.2 Crowd-sourced approach

With this approach, any IAS access can act as a measurement end point (client). The access is typically provided with a software-based measurement agent that allows for user-initiated measurements executed on the end user's own equipment. The client measurement agent exchanges test traffic with one or more measurement servers. Each user-initiated test measurement generates a measurement sample that is stored in a central database for subsequent statistical analysis.

The measurement samples continuously generated by the crowd-sourced approach are post-processed and published to provide average IAS performance information. Ideally, all members of the population, i.e. all IAS access points, should be present in the collected data sets. Based on this database, regulatory supervision and quality analysis can be made, e.g. the performance of specific IAS packages. However, under real-life conditions only a subset of the population will participate.

In order to provide robust quality measurement results, the crowd-sourced approach must aim to collect a large number of participants. The larger the number of participants is, the less likely it will be that the statistical analysis of the collected measurement samples becomes biased. The success of the crowd-sourced approach relies on a low threshold for participation. Therefore crowd-sourced campaigns normally rely on the distribution of measurement software that can be executed on any end user equipment.

Since crowd-sourcing relies on volunteers that can participate without any admission control, the number and distribution of the measurement samples is unknown. Also, the participants volunteering may originate from a population that was not targeted by the quality evaluation (e.g. participation by business Internet access services with particularly high-speed access). Therefore, a thorough statistical post-processing of the raw data collected has to be performed.

As a result, crowd-sourced measurement software will not only measure the actual quality metrics, but will also collect additional information on the measurement environment (hardware, operating system, location etc.) for validation purposes. This will typically be accompanied by a short questionnaire to be answered by the participants on, for example, the type of contract, tariff and location.

It is advisable to perform a statistical monitoring throughout the entire measurement campaign in order to keep track of the composition of the measurement samples. This

allows for countermeasures, such as advertisements or promotion/recruitment campaigns, if it is appears likely that the required number or distribution of participants will not be met.

4.6 Wireless/mobile aspects

Wireless and mobile Internet access services occupy a significant market share of the overall IAS market. This section investigates wireless/mobile IAS specifics compared to fixed IAS and presents different approaches of evaluation of quality. Wireless/mobile IAS specifics are due to the dependency on radio signal propagation conditions in open air and the end user's high level of mobility.

4.6.1 Mobile/wireless quality parameters

Despite the diversity of IAS physical layer technologies, they all use the IP protocol at the network layer. In terms of technological neutrality, this feature can be used to evaluate IAS quality in the same way for all types of access technologies, including wireless/mobile IAS. General technical quality parameters commonly used to evaluate IAS quality, such as data speed, delay, and jitter, can be used for wireless/mobile IAS too.

On the other hand, for wireless/mobile IAS few additional quality parameters are standardised to achieve a complete picture of overall quality of an individual wireless/mobile IAS. Quality parameters defining *wireless/mobile network availability and accessibility* as well as *IP service accessibility and integrity* could be used.

4.6.2 Radio coverage

Wireless/mobile IAS technologies rely on radio signal propagation through open air. To achieve permanently reliable communications it is necessary to ensure that the transmitted radio signal is strong enough to be received and decoded at the receiving end.

In the case of wireless IAS provided between antennas at fixed locations on both the end user and the ISP's sides, quality measurements can be carried out in a similar fashion to that for wireline IAS. In addition, wireless/mobile IAS quality is affected by variability of signal propagation conditions caused by nature such as rain, snow, fog, or sun storms. This is in addition to disruption or degradation due to interference from other radio signals.

In the case of mobile IAS, base stations have fixed locations and the radio signal is adjusted to optimise mobile network performance. In areas where the base station's signal strength is too low, the mobile IAS will become unavailable. More complicated cases occur when an end user is located at the edge of a cell. In such circumstances mobile IAS quality depends strongly on terminal position or potential radio interference. At such locations, mobile IAS tends to be highly unreliable.

Many mobile operators provide coverage maps on their websites defining areas where different levels of transmission speed could be achieved. When using such maps it is important to note that they are calculated using models and will in some cases not be able to represent the situation in practice. For example, coverage maps describe signal strength and/or IAS availability outdoor, and therefore indoor coverage of mobile IAS cannot be inferred directly from outdoor coverage map.

4.6.3 Mobility aspects

An important feature of mobile IAS is support for end user mobility. However, taking into account the cellular nature of mobile IAS and the fact that a moving end user crosses the boundary between neighbouring cells, disruption or degradation of mobile IAS frequently occurs because of interference from two (or more) cells in border areas or because of handover procedures not happening smoothly.

Cases of IAS degradation of quality or network congestion due to a high number of end users or due to an end-user's application's demand of high data throughput could easily happen for wireless/mobile IAS. In cases when traffic load exceeds the installed capacity, congestion will occur.

4.6.4 Tools/systems to evaluate mobile/wireless IAS quality

Measurement of mobile IAS is a specific application of the different measurement approaches presented in section 4.4. At the IP layer, mobile/wireless IAS performance measurements (e.g. speed and latency) are the same as with fixed. The challenge with mobile/wireless IAS is to address the variability of performance in the radio access section and to achieve geographical coverage.

For some test scenarios, a controlled measurement system approach is used and can be set up in the form of a drive test. Note that a full coverage of networks with drive tests is impractical. If the aim is to provide information relevant to end users which allows for pre-contractual comparison, a crowd-sourced solution with software clients on end user equipment may be preferred.

In order to provide representative results, the design of mobile/wireless measurement probes has to take into consideration specifics of the relevant radio access technologies, conform to minimum technical requirements, and function in the same way as ordinary end users' handsets. Furthermore, the measurements performed should not lead to network overload over longer time periods.

Crowd-sourced quality measurements may be achieved using specifically designed mobile apps which are installed and executed on end users' smartphones (e.g. Android or iOS). The app then performs measurements, either on a regular basis or when an end user initiates a measurement.

The following points should be taken into consideration when designing a measurement system for mobile/wireless IAS:

- Measurement methodology should take into account the above mentioned aspects
- Data volume produced by measurements (mobile data tariffs tend to not be flat-rate)
- The effect of terminals which might have limited capabilities
- How to perform sampling to achieve statistical reliable results

4.7 Complementary methods

4.7.1 Introduction

Analysis of stakeholders' opinions may help in the analysis of results achieved from objective measurements, by providing complementary methods. For the purpose of transparency, technical measurements may be enough for specific objectives but, for a thorough net neutrality investigation, complementary methods are often an advantage.

Complementary evaluation methods are important since the information obtained through them is helpful to understand the relation between objective measurement results and the quality perceived by the users. In this case, instead of accuracy, the observations offer a picture of how well the results can be trusted. Important indicators could include be collection and analysis of user's complaints, traffic management investigations, mean opinion scores on user's levels of satisfactions, public opinions and debates.

4.7.2 User's complaints

Collecting and classifying end user's complaints is a helpful exercise to provide an early indication of a degradation of quality of a service. A series of a specific type of complaints is already an indicator, normally resulting from contact with the service provider to clarify the reason; a well-structured collection of complaints may help to draw the quality profile of a specific provider.

Clause 4 of the EC recommendation C(2010)3021[23] *on the use of a harmonised methodology for classifying and reporting consumer complaints and enquiries* may help to establish a main structure for the collection and analysis of the complaints. It is likely to need a more detailed structure in order to adapt to particular countries' needs.

Some NRAs have systems of collecting user's complaints but often the detailed relation between user and supplier is not covered by the electronic communication act and the data available by the NRA is insufficient. The most appropriate partners for NRAs may be consumer protection organisations or Alternative Dispute Resolution bodies to solve disputes between users and service providers.

4.7.3 Traffic management investigations

BEREC organised and may repeat a questionnaire addressed to ISPs in Europe. This exercise is particularly useful to understand the evolving traffic management practices and relate them to results obtained from users' satisfaction surveys and / or objective quality measurement campaigns.

[23] See further information on:
European Commission, "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a harmonised methodology for classifying and reporting consumer complaints and enquiries", 2009.
Online: http://europa.eu/legislation_summaries/consumers/consumer_information/ co0014_en.htm (accessed January 24th, 2014) and
European Commission, "Communication from the Commission: Monitoring consumer outcomes in the single market: the Consumer Markets Scoreboard", 2008.
Online: http://europa.eu/legislation_summaries/consumers/consumer_information/ l10140_en.htm (accessed January 24th, 2014).

Formulating questions and analysing answers helps NRA to better specify a clear boundary between discriminating and differentiating traffic management methods, and helps ISPs to understand it and enhance their practice. Such investigations should be based on clear and unambiguous questions and be made repeatedly, evolving to stay relevant to the traffic management practices met in the market.

4.7.4 Quality perceived by end users (QoE)

The quality perceived by end users (QoE) is essentially the relationship between the performance expected from a specific service and the subjective perception obtained after the use of the service which is largely depending on the QoS.

Some CAPs collect QoE feedback, e.g. each time a user has used a service, and they may have an interest in collaborating with NRAs on a quality monitoring system. CAPs could be invited to perform similar studies, and be encouraged to provide NRAs with such results. (Ref. also section 4.3.1 above on "Application measurements")

Direct consultation of the opinions of end users is another possibility. Experts on consumer surveys and mean opinion scores may determine with acceptable levels of confidence the general opinion of the end users within a scale from 1 to 5.[24] This may be a more expensive exercise but has the advantage that it can be adapted to the population under study and to relevant applications.

4.7.5 Public opinions and debates

Tracking commercial campaigns of ISPs and CAPs, public discussions or parliamentary debates may also help to understand the behaviour of the market players and provide guidance on how to adapt the quality monitoring system to the evolution or needs of the market. Major complaints from CAPs and business users are often discussed in the media.

4.8 Conclusions and recommendations

The evaluation of quality of Internet access services in the context of net neutrality can be divided into two different aspects; IAS as a whole vs. individual applications using IAS.

Regarding IAS as a whole, BEREC recommends that, as a minimum, the following metrics measured at the IP layer should be used to evaluate quality of the Internet access service:

- Upload and download speed
- Delay and jitter
- Packet loss ratio

BEREC puts emphasis on the end users' experiences when using the IAS, and therefore information about the *actual performance* of the service is essential. By using statistical methods, the recommended metrics should be reliably sampled and analysed in a way that

[24] ITU-T recommendation G.107 is an example of a well-established model for voice telephony, but others exist and some are being developed; ITU-T recommendation G.1011 offers an overview of objective and subjective evaluation methods as well as models (like G.107) for different types of services.

provides actual values (statistics of measurement), both at *peak and off-peak times*, describing variation of performance over time.

Regarding variation of performance towards different destinations, measurement servers covering different geographical / topological conditions should be considered.

As indicators for *level of congestion*, BEREC recommends measuring at different times of the day and week in order to obtain the difference between peak and off-peak values, and also the duration of congestion. Furthermore, test measurements towards different server locations are recommended in order to monitor whether congestion is particularly prone at specific interconnections from the ISP to the rest of the Internet.

When measuring IP layer metrics, the transport layer protocol (typically TCP or UDP) and application layer protocol (e.g. HTTP) are relevant for the measurement methodology. Furthermore, the network scenario used regarding measurement clients and servers is essential, and this is further discussed below. Each metric needs an associated complete measurement methodology clarifying such aspects.

The recommended IP layer metrics are applicable for fixed as well as wireless/mobile Internet access services. BEREC recommends considering the use of additional parameters, for example, in order to reflect wireless/mobile network coverage aspects.

An important part of the evaluation of the quality of IAS in the context of net neutrality is to detect potential *degradation* of IAS as a whole (use case B1|). To achieve this, comparable periodic measurements are needed over time, to distinguish long-term evolution from short-term variations. These measurement results should then be assessed in the light of technical progress and market evolution to draw conclusions about the quality level of IAS.

Although several standards exist for quality metrics, BEREC believes that there is a need to further specify and clarify metrics and corresponding measurement methods to achieve a consistent IAS quality evaluation toolkit. BEREC therefore encourage standards developing organisations (SDOs) to continue their effort in this regard. Furthermore, NRAs should continue, and also consider enhancing, their participation and contribution to SDOs, notably IETF regarding IP-based quality measurements.

Regarding *individual applications using IAS* (use case B2), BEREC recommends that assorted applications (i.e. content *and* applications) are measured to check whether degradation occurs. The assortment of applications should contain the most used applications, but other relevant applications should also be included to account for the diversity of applications used among end users. Traffic management investigations will typically give an indication about applications that are particularly important to check.

These recommendations about what to measure are supported by requirements set out in chapter 2. The emphasis on statistical reliability is particularly important to achieve accurate measurement results. Openness is sought through use of standardised measurement methods, and this will also allow for comparability between measurement results.

How to measure?

BEREC recommends using measurements *beyond the ISP leg* to allow for the inclusion of the interconnections of the ISPs in the measured results. It is recommended to increase comparability between measurement results of different IASes, and even different countries. Comparability between IASes is an important element of transparency. Regulatory supervision for the detection of degradation of service may, among other things, be supported by comparison between IASes and between regions or countries. Centralised measurement servers may also be a cost-effective solution if servers are shared between NRAs.

Considerations regarding spatial distribution at the end user side are closely linked to the objective of the measurements performed (ref. main use cases A and B).

In the case where the purpose is to give transparent information about the IAS offer subscribed to by individual end users (use case A), BEREC suggests that a less controlled measurement system using software measurement agents can be sufficient. Web-based software or apps will usually be preferred due to low cost, ease of distribution and installation. Software agents allow end users to initiate test measurements at their own initiative, e.g. performing speed meter tests.

When this quality measurement approach is used by a large number of end users, it can constitute a crowd-sourced measurement campaign. Achieving a sufficiently broad coverage of the crowd-sourced user base is challenging compared to a preselected panel using a controlled system. Such a crowd-sourced approach can be preferred for cost-effectiveness reasons.

In the case where the purpose is to perform in-depth and/or long-term quality supervision and evaluation of the IAS offers in the market (use case B1), the use of a controlled measurement system, e.g. with hardware probes, usually provides a higher level of reliability. For this approach, the panel should be preselected based on the required geographical coverage and statistical reliability.

These recommendations about how to measure are supported by requirements set out in chapter 2, depending on the use case considered. The NRA will need to strike a balance between covering all (as many as possible) IASes (i.e. use case A) in a cost-effective way, and achieving high level of accuracy for measurement results from a limited number of IASes (use case B1).

BEREC recommends that NRAs increasingly put emphasis on evaluating performance of IAS as a whole, to assess potential degradation due to specialised services.

When measurement campaigns are conducted under use cases A and B, the focus has so far mainly been on evaluation of the quality of IAS as a whole. *BEREC also recommends that NRAs increasingly puts emphasis on evaluation of performance of individual applications using IAS, to assess potential application-specific degradation.*

Detection of *degradation of individual applications* (use case B2) is the most challenging part of the regulatory supervision of net neutrality. Even though blocking can easily be detected, it is likely to be complicated to verify the throttling of applications with a high degree of certainty. A few aspects are however identified.

In general, detecting degradation of individual applications is a relatively new area which is still under research. A few tools exist, and BEREC recommends gaining further experience with tools measuring performance of individual applications.

Building on information from CAPs' own measurement systems is in principle a method suitable for this purpose, but this would require extensive evaluation before results could be used and NRAs would need to be vigilant if/when engaging in this kind of measurement. For the time being, BEREC is unsure about the usability of this method.

Finally in this category, targeting detection of degradation of individual applications through passive monitoring also has promising characteristics. As for general application-specific test tools, BEREC recommends looking more deeply into passive measurement methods. It is important that privacy is respected related to these kind of measurements.

When performing measurements where it is important to reflect the actual usage of wireless/mobile end users, BEREC recommends performing measurements through the use of mobile apps for crowd-sourcing or using drive tests, including indoor and outdoor usage.

It is expected that further experiences about technical measurement methods and architectures will be gained over time, and this should be fed into harmonisation activities in organizations like BEREC and CEPT, and standardisation work, notably in IETF, but also ETSI and ITU. BEREC therefore recommends that national and European institutions, including NRAs, contribute to the exchange of experiences and development of specifications, standards and recommendations, within these organizations.

BEREC considers complementary methods - such as analysis of user's complaints, traffic management investigations, opinion scores, and monitoring public opinions and debates - are valuable sources for information which assist the technical quality monitoring. For example, user's complaints can be used for increasing scrutiny of ISPs' traffic management practices, such as checking for potential throttling of traffic from specific applications.

BEREC recommends that consumer organisations and associations of CAPs' share information with NRAs when unreasonable traffic management is perceived. NRAs should follow up such information and other incidents observed, by performing targeted inquiries.

5. Future perspectives

5.1 Introduction

Converged quality monitoring solutions

As shown in the previous chapters, the present stage of development of quality monitoring solutions is characterised by national initiatives with varying regulatory goals and different measurement methodologies applied. Most NRAs are either engaged in or are planning to start quality measurement systems. The national initiatives differ with respect to their design and stage of development due to different national goals, structure of broadband markets, timescales of NRAs and geographical scope of NRAs' jurisdiction.

In the future, a more harmonised approach could be explored. Even if the NRA's focus remains on national quality monitoring, harmonisation can be beneficial since replication of development, deployment and operation efforts can be avoided and costs, e.g. for software development, could be shared.

Furthermore, NRAs may want to support a stronger convergence due to the wish to reflect the transnational topology and usage of the Internet, which makes it relevant to measure quality parameters across borders (based on clients and servers in different countries) and obtain comparable results between countries (based on common metrics and harmonised methodologies),

Also, the emergence of a European single market, where supply and demand sides, competition and regulation are expected to become less heterogeneous on a continental scale will lead to more harmonised quality monitoring solutions.

The decision to engage in such a process of convergence will depend on NRAs and other stakeholders involved in the management of measurement systems. NRAs which already have an existing system in operation may find no need to adapt their measurement tools compared to those who are still in the designing process. However, whichever system is used, it is anticipated that obsolescence will be reached at some point in time due to changing NRA requirements (e.g. reinforced legal framework and duties); future developments of Internet technology; ageing of monitoring system hardware/software etc. When obsolescence is reached, NRAs may choose to implement harmonised measurement tools and features in their monitoring systems.

Nevertheless, it should be noted that in some Member States (e.g. in Spain), the competences to establish minimum quality requirements or to monitor consumer rights are not within the remit of the independent NRA, and there are legal constraints for the development of any measurement system. This institutional set-up would hinder the process of convergence even in the case that a broad majority of NRAs choose to implement a harmonised measurement approach.

NRAs using monitoring systems supporting harmonised tools and features could use each other's systems for cross-border measurements since the measurement software of clients

and servers will be compatible. NRAs could cooperate and combine their national systems to form a distributed measurement system going beyond the coverage of the original disparate systems. It could then be used for specific cross-border measurement tasks, but would require coordination among participating NRAs in order to grant access to each other's resources (e.g. temporary usage of servers at various IXPs), or new shared servers could be established.

Such a system could in principle be extended to form a multi-NRA monitoring system by introducing a central management plane that administers and operates clients and servers of the system. System resources would be supervised by the central management plane and could not necessarily be accessed by a NRA directly. On the other hand, the system could be equipped with multiple management interfaces, giving each NRA control over resources dedicated to individual NRAs. Such functionality depends on system requirements to be decided during early phases of system development.

Two approaches to designing a multi-NRA system

Based on this, two main approaches to designing a multi-NRA system can be identified:

Gradual efforts to converge existing measurement systems

This approach would be characterised by the adoption of convergent measurement methodologies and the sharing of measurement resources among NRAs in order to perform cross-border measurements. It may simplify administrative procedures and reduce costs of administrative compliance. This gradual approach may be particularly preferable for NRAs that already have a national measurement system in place. Convergence following three stages would allow for recognising best practices, comparing results and finally allow for cross-border measurement set-ups:

- *Convergence of metrics and methods:* Existing measurement initiatives would be continued, and their methodologies would be clearly disclosed and compared at a European level in order to recognise best practices and identify areas where results are comparable. Improved technical specifications would be developed and existing systems adapted to these.
- *Sharing and comparison of results:* The results of existing national systems would be analysed by experts who identify areas where comparisons could reasonably be made. Comparable results would be published on a common medium. Caveats would be clearly disclosed and non-comparable results would be explained.
- *Partly shared system:* In order to allow for cross-border measurements, existing measurement systems could be adapted to share a set of features: some shared measurement tools and servers. At the same time, NRAs continue to maintain specific features and tools for national purposes. For instance, particular servers or test metrics could be of special interest in one country, while being less relevant in others.

Implementation of a full-blown common measurement system

This approach aims at establishing a full-blown common measurement system. NRAs could rely on this common system, which is managed by a central entity (e.g. an association of NRAs or another body). Such an approach may potentially be considered as stage 4. This approach is easier to implement for an NRA which does not yet have any measurement system in operation than for NRAs, which might supplement their existing national systems

with a common one, or replace/modify their system to integrate it in a common system. Moreover, it needs to be considered that the deployment of a full-blown common system would be a time-consuming activity that would need many resources in order to coordinate several stakeholders from many countries.

Two types of full-blown system design can be distinguished:

- *Existing systems*: Collective agreement of NRAs to rely on an existing open or proprietary measurement system which could be relatively quickly adopted (or is even already deployed). Examples are open systems like RIPE Atlas, M-Lab, but commercial systems also exist which already provide measurement agents and a database of measurement results.
- *New systems:* NRAs and other stakeholders join forces to develop a new common system from scratch, which aims to address the collective needs of NRAs, notably building a trusted source of measurement data which serves the objectives of informing both end users and regulators, and provides a reliable resource which could be used to support a regulatory intervention, if needed. There is an array of credible technical solutions (ref. previous chapters) available which could be combined. The funding and management of such a system poses a challenge to NRAs. By its nature, this approach will require a collaborative approach and agreement amongst participating NRAs and will result in a system with a large geographical footprint but with effort to be shared amongst participating NRAs and potentially other stakeholders

.

5.2 Measurement system standardisation

When building a centrally-administrated quality measurement system where different measurement methodologies and components from different systems (existing or new) are integrated, a common architecture and management framework would be required. It would preferably be based on standardised solutions, thus enabling a larger set of available components to pick from, including off-the-shelf products.

In IETF the *Large-Scale Measurement of Broadband Performance* (LMAP) working group[25] has been launched with the goal to standardise a measurement system *architecture* for performance measurements of broadband Internet access services, usable for, for example, network diagnostics by ISPs and collecting information by NRAs, including end user initiated measurements.

The LMAP charter sets out that the working group will develop *a framework* that contains common terminology and architecture elements, *use cases* clarifying the basis of their work, *an information model* for metrics, schedules and results, and *management protocols* to use between elements of the architecture.

LMAP is developing general measurement architecture, and will not standardise quality metrics, since this is left to other relevant bodies. Furthermore, deciding the set of measurements to run is beyond the scope of LMAP, and has been left to the organisation

[25] "Large-Scale Measurement of Broadband Performance (LMAP) - Charter", IETF, accessed January 24th, 2014, http://datatracker.ietf.org/wg/lmap/charter/

which will manage the measurement system once deployed. In this way, the LMAP deliverables will develop an open and flexible architecture, with a likelihood of supporting the needs of the NRAs.

The IETF *IP Performance Metrics* (IPPM) working group[26], which has existed for many years already, is developing standardised quality measurement parameters and methods as referred to earlier in this report (ref. section 4.1). IPPM has recently been re-chartered to accommodate the development of, among other things, a *registry* of commonly used metrics. Such a registry is foreseen to become a valuable supplement to the general information model of the LMAP architecture, thus facilitating its population with concrete measurement metrics and methods.

Both LMAP and IPPM seek to cooperate with other relevant standards development organizations, such as ETSI and the Broadband Forum. It is particularly interesting that the Broadband Forum has also launched a project aimed at developing a broadband measurement framework called *Broadband Access Service Attributes and Performance Metrics*.[27]

It is important that NRAs participate in standards development organizations' activities related to quality measurements. This will allow NRAs the opportunity to both influence their work and leverage more directly on these organizations' findings. This is in addition to the general increase NRAs' own know-how within this complex area of Internet-related measurements.

Enhanced emphasis on initiatives from the Internet community (as opposed to the traditional telecom community) should be emphasised. The IP technology is standardised within IETF, and following closely the emerging standards related to quality metrics and measurement systems within the same organization, would probably provide NRAs with first-hand knowledge about upcoming Internet measurement technologies.

5.3 Cost aspects

Considering that any decision depends on the use case, and taking into account the requirements listed in section 2.4, the following cost-related aspects may become relevant for evaluating whether a measurement system is future-proof. However, the following preliminary considerations must be taken with great care as they are looking at the costs only from a very generic viewpoint.

Irrespective of which of the two options - gradual or full-blown - is favoured, the implementation of such systems implies fixed costs. System development, as previously analysed, requires investments in terms of system design and procurement of assets, both specific (such as hardware solutions, servers and possibly probes for test processing) and generic inputs (i.e. housing). Consequently, set-up costs are significant compared to running costs. Therefore, once the quality measuring system is built-up, the larger the number of users, the lower the unit cost of quality measurements.

[26] "IP Performance Metrics (IPPM) – Charter", IETF, accessed January 24th, 2014, http://datatracker.ietf.org/wg/ippm/charter/

[27] "Broadband Forum - Technical Work in Progress", The Broadband Forum, accessed January 24th, 2014, http://www.broadband-forum.org/technical/technicalwip.php

At a very general level, the occurrence of fixed costs, as part of the overall total expenditures needed for the measurement system, favours the implementation of a multi-NRA system. Whether economies of scale - associated to the presence of fixed costs in both the setting-up and the activities carried on during the recurrent monitoring activities - could actually be exploited is however difficult to assess at a general level. For instance, an NRA already having a national system in place has already incurred fixed costs which need to be considered as sunk, and thus it may be less beneficial for them to switch to a multi-NRA system. Therefore, the starting position of an NRA, i.e. already having a system or starting from the scratch, plays an important role.

From an economic point of view, potential activities eligible for “centralisation” include long-run investments for the deployment of the quality measuring system and fixed costs behind the implementation of the measurement system. Short-run investments and variable costs necessary to the day-to-day operations are qualified, on the contrary, to a decentralised system management.

Considering the gradual effort to converge existing measurements systems, the three main phases involved in prototyping the quality measurement system (specification phase, development phase and deployment phase) might be coordinated across countries. National institutions and stakeholders (typically the NRA) would be in charge of implementing and further adapting the system to national specific circumstances and local market conditions, then deploying and operating it at the national level.

When assessing the economic viability of a full-blown common measurement system, significant additional coordination costs need to be considered too. In case of national systems, the number of actors involved in the entire monitoring process is lower and the participation of local and small stakeholders could be favoured. The NRA and stakeholders have the chance to interact quite often, even during periodic, formal and informal, meetings occurring on general regulatory aspects.

When assessing the merits of a national system, other regulatory, technical and economic reasons need to be considered. The following reasons may provide a rationale for using a national system:

1. particular electronic communications network characteristics adopted by network operators in a specific geographic area;
2. specific supply-side conditions in terms of pricing, range and quality of electronic communication services commercialised in the concerned area;
3. demand-side specifics in terms of consumers' habits, i.e. bundles of electronic communication services, actual download/upload speed etc.;
4. specific monitoring and supervision reasons to detect infringements of net neutrality, in the event that national regulation is in place.

Within this framework, relying only on either national systems or international systems might not be the most cost-efficient solution, particularly in the short run. Relying on a multi-NRA opt-in system alone may produce a crowding-out effect as investments have already been funded and not yet amortised; its development from scratch would imply higher costs and the deployment could not be immediate. Many national systems have proved to deliver valuable information to end users and policy makers for both transparency goals and

regulatory supervision needs. Advancement in the quality monitoring systems has to reconcile these aspects.

Full-blown systems are not considered realistic at this point in time and will therefore not be followed up in the feasibility study considered as the next step.

Overall, the choice of quality measurement systems depends on the use case foreseen. Such systems have to fulfil the requirements of accuracy, comparability, trustworthiness, openness and future-proofness. The latter criterion implies that system design should ensure flexibility, extensibility, scalability and adaptability. This includes applying cost-effectiveness as a general rule-of-thumb to all phases of the measurement system lifecycle, including development, deployment and operation. If two systems equally fulfil the requirements listed by the NRA the more cost effective system should be chosen.

5.4 Evolution of a potential multi-NRA opt-in system

Whilst the use of existing measurement systems can undoubtedly provide NRAs with the means to achieve their national objectives, this cannot be said for current or future NRAs' objectives that may include cross-border aspects. When considering the two use cases, it emerges that a cross-border capable measurement system is likely to be needed for use case B when certain issues arise.

The focus of the present study has been on net neutrality related issues which are leading to common ways of evaluating quality of IAS. Hence, this is already putting NRAs on a path towards convergence of methodologies and tools. Once this degree of convergence is achieved, NRAs could easily share their systems for cross-border measurements (see *Gradual effort to converge existing measurement systems in section 5.1*). Any NRA may opt in to such a partly shared system.

The measurement system design would allow NRAs to collaborate on cross-border quality measurements and thus benefit from dividing effort amongst NRAs to provide a richer set of information without a massive increase in effort, by sharing resources of existing systems. Furthermore, it should also allow NRAs to opt into the system at a later date after implementing a set of methodologies and tools in their measurement system. It would function at a multi-country level but yet can cater for NRAs' needs to parallel measurement to local specifics and maintain jurisdictional boundaries.

For transparency purposes (use case A) both approaches could be used. When applying the *gradual* approach a multi-NRA cross-border solution is simply a distributed set of measurement servers. Such a system is relatively easy to set up. It implies that participating NRAs agree that their measurement servers can be accessed by other NRAs as well, or dedicated shared servers can be established. Each NRA would tailor their own software client and distributes it to their end users.

5.5 Conclusions and recommendations

This section discusses the evolution of a potential future multi-NRA opt-in quality monitoring system dedicated to regulatory purposes. If developed, it should meet current and future needs of NRAs and provide the basis for harmonised measurements across the jurisdictions of the NRAs, i.e. cross-border measurements.

BEREC recommends that NRAs collaborate on a voluntary basis on the development of such a multi-NRA opt-in system in order to benefit from dividing the efforts of developments amongst the NRAs in order to provide a richer set of information for consumer, citizen, policy and potential intervention.

BEREC recommends that a distributed system philosophy is adopted in order to allow the system to extend over the European Internet infrastructure, while also catering for NRAs' needs to adapt to local specifics and maintain jurisdictional boundaries. This approach will also allow the integration of existing systems or other NRAs to opt in at a later date.

This study has drawn recommendations for NRAs to adopt common ways of evaluating quality of Internet access services in the scope of net neutrality, underpinned by the common belief amongst participating NRAs in the benefits of achieving a greater level of harmonisation.

For the purpose of harmonisation, BEREC recommends that an evolutionary strategy is emphasised, where harmonisation itself is viewed as multi-stage process along the following:

- Stage 1: Convergence of metrics and methods
- Stage 2: Sharing and comparison of measurement results
- Stage 3: Harmonisation in terms of cross-border measurements

Here, BEREC recommends that NRAs should aim at using a convergent set of measurement parameters and corresponding measurement methods and participating in the activities of Standards Developing Organizations (SDOs), with a focus on IETF. NRAs should exchange methodologies and experiences when implementing the standards and provide feedback in order to improve measurement methods. Furthermore, sharing of measurement results (open data) should be considered. This allows for the first two stages of harmonisation.

To go beyond the first two stages, BEREC believes a closer collaboration amongst NRAs would be preferable. In the first instance, collaborative effort will be needed between participating NRAs to agree on the governance, design and development of tools and implementation of a system covering different European states. Once agreement is established and servers deployed, NRAs would be free to adopt the national systems they will use to interface with the servers, preferably using standards (e.g. LMAP).

During stage 3, a set of shared test servers (e.g. placed within major IXPs in Europe) and software components could be made available for usage by participating NRAs as an early step. Additional steps could later be taken to expand the footprint of the system.

The above leads BEREC to believe that seeking convergence of measurement methodologies would not undermine current NRA investment in measurement systems or

vice-versa when the long term view is considered, which must take into account aspects such as obsolescence of current systems. A multi-NRA opt-in system following the *gradual* approach would not prevent NRAs using their existing solutions to address specific national needs, e.g. separate servers, more parameters, complementary technologies, etc.

By its nature, effective governance and careful planning would be necessary in order to build such an opt-in system. Indeed, this activity would require the setting up of a project to oversee the development of the system. BEREC recommends that the development of an opt-in quality measurement system following the gradual approach should adopt a formalised system engineering approach in order to provide the necessary system development assurance which is vital for harmonisation.

The above leads BEREC to recommend that an initial feasibility study is conducted to investigate how an opt-in approach (convergence of methods and sharing of infrastructure) could be realised in practice, including NRAs with no existing systems. Such a study should also consider the effect of dissemination of knowledge among NRAs as well as increased convergence of measurement methodologies and increased competence in the field of quality monitoring in the context of net neutrality.

Glossary

AS Autonomous System
BEREC Body of European Regulators of Electronic Communications
CAP Content and Application Provider
CEPT European Conference of Postal and Telecommunications Administrations
ETSI European Telecommunications Standards Institute
FD Framework Directive
HTTP Hypertext Transfer Protocol
IPPM IP Performance Metrics
IAS Internet access service
IETF Internet Engineering Task Force
IP Internet Protocol
IPTV IP Television
ISP Internet service provider
IT Information technology
ITU International Telecommunication Union
IXP Internet eXchange Point
LAN Local Area Network
LMAP Large-Scale Measurement of Broadband Performance
NRA National Regulatory Authorities
NN Net Neutrality
QoE Quality of Experience
QoS Quality of Service
SDO Standards development organization
TCP Transmission Control Protocol
UDP User Datagram Protocol
USD Universal Services Directive