

Marcus, J. Scott; Waldburger, Martin

Conference Paper

Identifying harm to the best efforts Internet

26th European Regional Conference of the International Telecommunications Society (ITS):
"What Next for European Telecommunications?", Madrid, Spain, 24th-27th June, 2015

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: Marcus, J. Scott; Waldburger, Martin (2015) : Identifying harm to the best efforts Internet, 26th European Regional Conference of the International Telecommunications Society (ITS): "What Next for European Telecommunications?", Madrid, Spain, 24th-27th June, 2015, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/127165>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Identifying harm to the best efforts Internet

J. Scott Marcus and Martin Waldburger, WIK¹

Abstract

How are regulators to ensure that public Internet access services (continue to) provide good service to consumers, despite possible incentives on the part of network operators to act otherwise? Many express concerns that Internet traffic management and prioritisation might somehow motivate network operators to degrade non-prioritised traffic, thus turning the best efforts Internet (however defined) into a “dirt road” of poor capacity and quality. National regulators might well be obliged in the near future to assess whether this is in fact the case. How realistic is this fear? How are regulators to make this determination in practice? What pragmatic considerations come into play?

¹ s.marcus@wik.org, and m.waldburger@wik.org, WIK Wissenschaftliches Institut für Infrastruktur und Kommunikationsdienste GmbH, D-53604 Bad Honnef, Germany, +49 2224 9225 12.

Contents

- 1.1 Concerns about a possible “dirt road” effect..... 3
- 1.2 Identifying the “dirt road” effect in practice 3
- 1.3 Structure of this paper 4
- 2.1 The regulatory context 5
- 2.2 Definitions and non-definitions..... 6
- 4.1 Capacity, performance, and quality on the Internet 9
 - 4.1.1 Quality of Service (QoS) on the Internet 9
 - 4.1.2 Quality of Service (QoS), Quality of Experience (QoE), and the needs of the application 10
 - 4.1.3 Instruments of traffic management 11
 - 4.1.4 The mathematics of prioritisation 12
 - 4.1.5 Challenges in measuring evolving QoS/QoE..... 15
- 4.2 Measurement programmes that have been attempted or implemented 19
- 4.3 Complementary measurement approaches 21
- 5.1 Building the evidence base..... 24
- 5.2 A semi-structured problem..... 25
- 5.3 Evaluating changes in end-user performance over time 26
- 5.4 Where should measurements be taken?..... 26

Figures

- Figure 1. Relative sojourn time in an outbound router queue with and without 10% of traffic subject to prioritization as a function of load on the link..... 14
- Figure 2. Relative sojourn time in an outbound router queue with and without 30% of traffic subject to prioritization as a function of load on the link..... 14
- Figure 3. User perceptions of the cause of blocking of online content or applications. 17
- Figure 4. The value chain: Content and application providers, different kinds of ISPs, and the end-user. 18
- Figure 5. Different forms of monitoring are used to fulfill different objectives..... 20
- Figure 6. Many data sources flow into the evidence base, which is assessed using a Decision Support System (DSS) and ultimately a human analyst. 25

1 Introduction

How are regulators to ensure that public Internet access services (continue to) provide good service to consumers, despite possible incentives on the part of network operators to act otherwise? In this introductory chapter, we briefly note ongoing concerns about a possible “dirt road” effect, discuss challenges in identifying whether it is occurring, and note the structure of the rest of the paper.

1.1 Concerns about a possible “dirt road” effect

Our work is inspired by ongoing concerns that Internet traffic management and prioritisation might somehow motivate network operators to degrade non-prioritised traffic, thus turning the best efforts Internet (leaving aside the question of definition to Chapter 2) into a “dirt road” of poor capacity and quality.

The concern is clearly expressed in a BEREC report from 2012.² They note that it is not unusual for an integrated broadband providers that also offers services such as video to positively differentiate in favour of its upstream services, which ‘do not necessarily raise competition problems’; negative differentiation however is characterised as a ‘hypothetical situation’ that ‘when it negatively affects a large number of content providers, is referred to in the net neutrality literature as the “dirt road”’.³

BEREC goes on to observe that ‘[a] vertically integrated [broadband ISP with SMP in a retail internet access market] has incentives to discriminate traffic coming from [content and application providers] which provide contents or applications competing with its subsidiary. Users face negative effects ...’⁴ There is, however, considerable uncertainty about the degree to which retail competition among broadband ISPs serves to constrain harmful behaviour, and thus as to how great a threat this truly is.

1.2 Identifying the “dirt road” effect in practice

There is an active European debate over network neutrality at present, most immediately as a result of the *Telecoms Single Market (TSM)* legislative proposals that the European Commission first put forward on 11 September 2013.⁵ The detailed texts are in a constant state of flux just now, but many of the texts that seem to have the greatest prospects for ultimate adoption call on National Regulatory Authorities (NRAs) to “closely monitor and ensure compliance” developments in order to “promote the continued availability of open internet access services at levels of quality that reflect advances in technology”.⁶

² BEREC (2012), Differentiation practices and related competition issues in the scope of net neutrality, BoR (12) 132.

³ Ibid.

⁴ Ibid.

⁵ European Council (2015), “Proposal for a Regulation of the European Parliament and of the Council laying down measures concerning the European single market for electronic communications and to achieve a Connected Continent, and amending Directives 2002/20/EC, 2002/21/EC and 2002/22/EC and Regulations (EC) No 1211/2009 and (EU) No 531/2012 - Examination of the Presidency’s compromise proposal”, 26 May 2015.

⁶ Ibid.

Even in the absence of the TSM debates, European NRAs would have a duty under Article 8 of the Framework Directive to ensure the ability of all Internet end-users “... to access and distribute information or run applications and services of their choice”.⁷

There is an obvious but complex linkage here between the user’s *ability to access content* and the *quality with which content can be accessed*. *If the quality is degraded enough, the user no longer has effective access, even if the content is nominally still available.*

This, of course, raises numerous questions:

- Informally, what might constitute an operational definition of the “dirt road” effect?
 - More specifically, what constitutes the ability of all Internet end-users “... to access and distribute information or run applications and services of their choice” at “levels of quality that reflect advances in technology”?
 - The various legislative drafts differ greatly in the details, but all permit traffic management under some circumstances. Traffic management inevitably implies that some traffic is impacted. How are NRAs to distinguish permissible from impermissible traffic management?
- Leaving aside the details of technology, what broad approaches are realistically available to enable NRAs to assess what is happening in their respective Member States in regard to the ability of all Internet end-users “... to access and distribute information or run applications and services of their choice” at “levels of quality that reflect advances in technology”? This entails substantial practical and theoretical challenges.

It also implicitly raises questions as to what the NRA should do if it determines that users are unable, or no longer able, “... to access and distribute information or run applications and services of their choice” at “levels of quality that reflect advances in technology”. This is an important question, but it is a discussion for another day.

1.3 Structure of this paper

In this paper, we define the problem to be solved (Chapter 2), briefly consider whether “dirt road” effects are likely (Chapter 3), provide background on measurement of Internet performance (Chapter 4), discuss practical considerations and concerns going forward (Chapter 5), and provide concluding remarks (Chapter 6).

⁷ Directive 2002/21/EC on a common regulatory framework for electronic communications networks and services (Framework Directive) as amended by Directive 2009/140/EC (Better Regulation Directive).

2 Problem definition

In this chapter, we first address the reasons why regulators in Europe might well be called on in the coming years to monitor the possible existence of a “dirt road” effect, and then discuss what terms it is necessary, or not necessary, to define in order to make the task meaningful.

2.1 The regulatory context

This is very visible in the debate (or trialogue) among the European Commission, the European Parliament, and the Council over the Telecoms Single Market (TSM) legislative proposals that the European Commission first put forward on 11 September 2013.⁸ In the Council’s most recent communication⁹ on TSM, we see:

Providers of electronic communications to the public, including providers of internet access services, and providers of content, applications and services shall be free to offer services which are distinct from internet access services and which are optimised for specific content, applications or services, or a combination thereof, in order to meet their requirements for a specific level of quality. ...

There are, however, numerous limitations to these rights. Article 3 paragraph 5 is particularly relevant:

National regulatory authorities shall closely monitor and ensure compliance with Article 3 and with paragraphs 3 and 5 of this Article, and shall *promote the continued availability of open internet access services at levels of quality that reflect advances in technology. For those purposes national regulatory authorities may impose technical characteristics and minimum quality of service requirements in accordance with the second subparagraph of Article 22(3) of the Universal Service Directive. National regulatory authorities shall publish reports on an annual basis regarding their monitoring and findings* [emphasis added], and provide those reports to the Commission and BEREC.

Providers of electronic communications to the public, including providers of internet access services, shall offer such services only if the network capacity is sufficient to provide them in addition to internet access services, they are not usable as substitute for internet access services, and are not to the material detriment of the availability or quality of internet access services for other end-users.

⁸ European Commission (2013), “Proposal for a regulation of the European Parliament and the Council laying down measures concerning the European single market for electronic communications and to achieve a Connected Continent, and amending ...”, 11 September 2013, COM(2013) 627 final.

⁹ European Council (2015), “Proposal for a Regulation of the European Parliament and of the Council laying down measures concerning the European single market for electronic communications and to achieve a Connected Continent, and amending Directives 2002/20/EC, 2002/21/EC and 2002/22/EC and Regulations (EC) No 1211/2009 and (EU) No 531/2012 - Examination of the Presidency’s compromise proposal”, 26 May 2015.

2.2 Definitions and non-definitions

The 26 May 2015 Council communication put forward by the Latvian presidency defines an *internet access service* to be a “publicly available electronic communications service that provides access to the internet, and thereby connectivity to virtually all end points of the internet, irrespective of the network technology and terminal equipment used.” That definition suffices for our purposes here.

The concept of the “best effort Internet” goes back to the earliest designs, include Cerf and Kahn (1974) and Cerf (1978);¹⁰ however, there does not appear to be a definitive definition. The key point is that there is no guarantee of (timely) delivery of an IP packet (datagram). Cisco (2009) notes for best effort services that “the network delivers data if it can, without any assurance of reliability, delay bounds, or throughput”.

All Internet traffic is delivered on a best efforts basis, whether prioritised or not, in the sense that there is never a delivery guarantee; however, it is common in the network neutrality literature to use the term best efforts to refer to Internet traffic that does not benefit from prioritisation.

In this report, we refer to *content and application providers (CAPs)*, with the obvious meaning. We avoid the use of the *over the top (OTT)* market players, both because the term has never been well defined and because it carries needless baggage.

Much of the literature on this subject refers to *specialised services*; however, we have considerable discomfort with the term. First, there has never been a clear definition of specialised services. Second, the idea that specialised services are distinct from other Internet traffic was always artificial, and in any case is precisely one of the properties whose correct implementation would have to be verified by “dirt road” traffic measurement in any case. Network operators are likely to find traffic management valuable, with or without admission control (the main distinguishing characteristic of specialised services in most definitions). Prioritised services without admission control are likely to exist unless they are prohibited; and even if prohibited, they are likely to exist unless the prohibition is actively policed.

Consider: A man is lying in the middle of the street, apparently the victim of a traffic accident. To the extent that the regulator is in the role of the medical team making the diagnosis and advising a cure, our concern is with the poor sod lying on the pavement. Our first concern in this sense is with identifying the harm to the victim.

To the police, it will however be important to know which vehicle struck him, and under what circumstances.

We reject the use of the term *specialised services* to the extent that it is confusing and unhelpful for the purpose of identifying whether a “dirt road” effect has emerged, or is at risk of emerging. For this purpose, whether the service that is preferred practices admission controls (correctly) or not is quite irrelevant.

Relative to the forensic and policing function that may also fall to the national regulatory authority, the definition might well be important; however, it is still not necessary to define specialised services at this time. The term may, or may not, appear in the TSM as enacted, *if* enacted, and if so it is that definition that will govern.

¹⁰ V. Cerf and R. Kahn (1974), “A protocol for Packet Network Interconnection”; and V. Cerf (1978), “The Catenet Model for Internetworking,” Information Processing Techniques Office, Defense Advanced Research Projects Agency, IEN 48, July 1978.

3 The likelihood of a “dirt road” effect in practice

In considering how to detect a “dirt road” effect, we would like to highlight the fact that this is a purely theoretical phenomenon that to the best of our knowledge has never been observed in practice. Moreover, there are strong economic reasons to doubt that network operators would find it profitable to allow the network to degrade to a “dirt road” if any alternative were realistically available to them.

In analysing motivations for network operators, it is useful to distinguish between (1) conventional economic analysis, and (2) concerns with purportedly exploding traffic that were making the rounds a few years ago.

As we explain elsewhere,¹¹ there are several different strands of economic thought that are relevant to network neutrality: (1) quality differentiation, (2) two-sided market theory, and (3) vertical foreclosure. The theory of quality differentiation suggests that providers of a service are motivated to benefit from the fact that different consumers of the service have different *willingness to pay (WTP)* for services of different quality.¹² In the absence of market power, price and quality differentiation tend to benefit both consumers and producers – service providers would be motivated to serve all consumers whose willingness to pay exceeds the marginal cost of serving them. In fact, if the cost of serving consumers is high (i.e. close to their WTP), more consumers would tend to be served with differentiation than without.

Consider airline deregulation. Enabling airlines to charge more for passengers who were willing to book in advance and stay over on Saturday night enabled an explosion in air travel on the part of budget vacation travellers with low WTP.

To a first order, one would expect that a network operator that allows the network to degrade is sacrificing revenue from low-WTP consumers who would otherwise pay more than the network operator’s marginal cost to supply.

If a network operator has market power, it is often argued, the network operator would be motivated to reduce quality for low quality services in order to force users to accept high quality, high price services. This begs the question: If the network operator has enough market power to do so, does it not already have enough market power¹³ to simply raise the price of all services, with or without quality discrimination?

Finally, the belief that network operators will be forced to degrade their networks due to exploding traffic volumes reflects a flawed analysis that was popular a few years ago.¹⁴ As we explain elsewhere, Internet traffic growth is healthy, but is by no means exploding – in fact, the percentage growth in both fixed and mobile traffic volumes continues to decline year over year. Unit costs also decline year over year, and in the case of the fixed network this decline exceeds the rate of increase in traffic. Finally, all indications are that prices for both fixed and

¹¹ See J. Scott Marcus (2014), "Network Neutrality Revisited: Challenges and Responses in the EU and in the US", a study on behalf of the European Parliament's IMCO Committee, IP/A/IMCO/2014-02, PE 518.751, at: http://www.europarl.europa.eu/RegData/etudes/STUD/2014/518751/IPOL_STU%282014%29518751_EN.pdf.

¹² See Hotelling (1929).

¹³ The question here is *not about possession of SMP*. The question here deals with the possession of *market power after all remedies have been applied*, i.e. in a modified greenfield sense. This is a very different question.

¹⁴ See for instance BEREC (2012): BEREC Guidelines for quality of service in the scope of net neutrality. BoR (12) 131, at: http://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/1101-berec-guidelines-for-quality-of-service-0.pdf.

mobile broadband services are not “stuck” at any particular level, but rather respond as you would expect to normal forces of supply and demand.¹⁵

Given that most European fixed and mobile broadband markets are reasonably competitive, these factors suggest that network operators will tend to be strongly motivated to provide services that meet the needs not only of consumers who are willing to pay more for relatively better quality, but also the needs of those consumers who want to pay only for best efforts services, however defined. The logic of price and quality discrimination says that firms profit by using different service offers to serve multiple groups of consumers with different WTP.

It is thus unsurprising that we have seen no concrete examples of the “dirt road” effect.

¹⁵ J. Scott Marcus (2014), "The economic impact of Internet traffic growth on network operators", available on SSRN.

4 Measuring the performance of the best efforts Internet

In this chapter, we review capacity, performance, quality of service and quality of experience on the Internet (Section 4.1); provide brief background on measurement programmes that have been implemented or attempted (Section 4.2); and discuss approaches that are complementary to a measurement programme implemented by the NRA (Section 4.3).

4.1 Capacity, performance, and quality on the Internet

Before embarking on a discussion as to how to detect or measure possible adverse changes in the capacity or performance of the best efforts Internet, however defined, it is necessary to review Internet performance in general. The discussion in this chapter draws heavily on our study of network neutrality for the European Parliament.¹⁶

4.1.1 Quality of Service (QoS) on the Internet

As is well known, *Quality of Service (QoS)* in an IP-based packet network is more complex than in traditional networks. IP networks are not designed to be able to carry all of every IP packet (or *datagram*) that every user might attempt to send or receive; rather, they are designed so that excess packets must wait until capacity is available.

If more packets are waiting than can be stored, the excess packets must be discarded – this usually causes no harm, because the network protocols (notably TCP) typically ensure that missing packets are retransmitted. Discarding of packets under these circumstances does not denote a failure – it is a part of the normal functioning of the network.

This approach works extremely well for applications such as e-mail or file sharing, where small delays are perfectly acceptable. It is less well suited to real-time applications such as Voice over IP (VoIP), as we explain shortly.

Contrary to what many believe, differentiated quality of service in the Internet was sought from the outset. The standard 1981 standard for the Internet Protocol already provides for a header field to indicate the desired quality; however, routers are under no obligation to honour the request, nor is there a clear indication of what they are supposed to do with it. “The Type of Service is used to indicate the quality of the service desired. The type of service is an abstract or generalized set of parameters which characterize the service choices provided in the networks that make up the internet. This type of service indication is to be used by [routers] to select the actual transmission parameters for a particular network, the network to be used for the next hop, or the next [router] when routing an internet datagram.”¹⁷

In an IP network, every IP packet (datagram) goes through many point-to-point ‘hops’ from its starting point to its destination. Each hop contributes to the total time (*delay*) that it takes to traverse the network. Key performance parameters for each hop of an IP-based network include:

- **Bandwidth:** the maximum number of bits that a transmission path can carry.
- **Propagation delay:** The time that a packet requires, as a function of the combined length of all transmission paths and the speed of light through the transmission path.

¹⁶ J. Scott Marcus (2014), "Network Neutrality Revisited: Challenges and Responses in the EU and in the US", a study on behalf of the European Parliament's IMCO Committee, IP/A/IMCO/2014-02, PE 518.751, available at: http://www.europarl.europa.eu/RegData/etudes/STUD/2014/518751/IPOL_STU%282014%29518751_EN.pdf.

¹⁷ IETF (1981), Internet Protocol: DARPA Internet Program Protocol Specification.

- **Queuing delay:** The time that a packet *waits* before being transmitted. Both the *average delay* and *variability of delay (jitter)* matter, since the two together establish a confidence interval for the time within which a packet can be expected to arrive at its destination.
- **Packet loss:** The probability that a packet never reaches its destination. This could be due to transmission errors, but errors are quite rare in modern fibre-based fixed networks. More often, packets are lost because the number of packets waiting for transmission is greater than the available storage capacity (*buffers*).

Recent BEREC guidance calls for measurements similar to these. “In CEPT’s ECC report 195, the following quality metrics have been selected: upload and download speeds, delay, delay variation, packet loss ratio, and packet error ratio.”¹⁸

In the aftermath of the US FCC’s recent Open Internet Order dealing with network neutrality,¹⁹ many experts complained about references in the Order to packet loss.²⁰ This is somewhat mystifying. In fact, packet loss has always been part of the discussion, and appropriately so in our view; however, it would be fair to say that no single parameter should be viewed in isolation, and that it would be inappropriate to apply a simplistic interpretation to any performance indicators that might be identified. The US experts seem to have come around to the view that the use of packet loss as one of several metrics is reasonable, and that the issue was something of a “tempest in a teapot”.²¹

Not every application is heavily dependent on the QoS. E-mail is, as previously noted, tolerant of high delay or loss, since users do not expect instant delivery. Real-time Voice over IP (VoIP), however, is very sensitive to delay.

4.1.2 Quality of Service (QoS), Quality of Experience (QoE), and the needs of the application

If the application has specific requirements in terms of throughput, speed, or reliability, it can signal those to the Transport Layer (whether UDP or TCP), which could then pass the request through to the Network Layer. This could then be reflected in the Type of Service (ToS) field of each IP packet, which would potentially make the request visible to every router in the path, and as well as to end systems on both sides of the exchange. The notion of distinct Types of Service is thus a basic part of the protocol design; however, no system is obliged to honour the request, and the IP protocol design does not specify what, if anything, should be done with each request.

¹⁸ BEREC (2014a), “Monitoring quality of Internet access services in the context of net neutrality. Update after public consultation”, BoR (14) 117. at:

http://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/4602-monitoring-quality-of-internet-access-se_0.pdf.

¹⁹ US FCC (2015), “In the Matter of Protecting and Promoting the Open Internet: Report and Order on Remand, Declaratory Ruling, and Order”, GN Docket No. 14-28.

²⁰ See for instance Steve Bellovin (2015), “Packet Loss: How the Internet Enforces Speed Limits”, at: <https://www.cs.columbia.edu/~smb/blog/2015-02/2015-02-27.html>.

²¹ Steve Bellovin (2015), “Update on Net Neutrality”, at <https://web5.cs.columbia.edu/~smb/blog/2015-03/2015-03-15.html>; and Nicholas Weaver (2015), “Wait, That Scary Net Neutrality Packet-Loss Clause Isn’t That Bad”, in Forbes, at: <http://www.forbes.com/sites/valleyvoices/2015/03/12/wait-that-scary-net-neutrality-packet-loss-clause-isnt-that-bad/>.

The relationship between QoS at the IP network level and the end user Quality of Experience (QoE) is strongly dependent on the application.^{22, 23} Some important examples are:

- E-mail is, as previously noted, tolerant of high delay or loss, since users do not expect instant delivery.
- The QoE of voice conversations, such as in IP telephony, is dependent on packet delay, delay variation and packet loss. A well-known criterion is that for a proper experience, the one-way delay through the network should not exceed 150 milliseconds.²⁴ Longer delays may cause users on both sides of the connection to begin speaking at once (as with telephone conversations using geosynchronous satellites, where round trip delay is some 270 milliseconds).
- For interactive gaming, delay and delay variation are also important, especially for so-called first person shooter games.
- The parameters that determine the QoE for streaming video include not only those for voice conversation (i.e., delay, delay variation (jitter), and packet loss), but other parameters as well. For users watching video content in a classical TV environment with a defined set of channels to choose from through a remote control, the so-called *zapping time* is important. This is the time that elapses between the selection of a new channel on the remote control and the actual appearance of the new channel on the screen.

4.1.3 Instruments of traffic management

Over the longer term, the network designer should design enough capacity into the network to carry the offered load at reasonable performance. In the short term, however, the network designer has only limited tools at his or her disposal with which to manage either long term shortfalls or short term 'spikes' in traffic.²⁵ The network is in place, and the transmission links have whatever capacity they have. Controlling QoS does not make the transmission links any faster; however, network designers and engineers can control:

- the relative priority with which each router processes the IP packets / datagrams waiting to be sent over each transmission link; and
- during periods where more packets are waiting than a given router is able to store or buffer, *which* packets are to be dropped.

Of the two, prioritisation is the easier to analyse and understand (see Section 4.1.4). Recall that each transmission link has whatever bandwidth capacity it has. Prioritisation does not change this. If some packets are moved forward on the waiting line, others are necessarily moved backward.

²² The QoE for the end user can also depend on factors other than the network QoS, such as the codecs used to code and decode voice and video at the application layer.

²³ Techniques for Measuring Quality of Experience, Kuipers, F.A., R.E. Kooij, D. De Vleeschauwer, and K. Brunnstrom, , Proc. of the 8th International Conference on Wired/Wireless Internet Communications (WWIC 2010), Lulea, Sweden, June 1-3, 2010, LCNS 6074, pp. 216-227, 2010.

²⁴ ITU-T G.114 General Recommendations on the transmission quality for an entire international telephone connection (05/2003).

²⁵ Internet traffic can be highly variable in intensity. See, for instance, the discussion of traffic self-similarity on pages 62-63 of J. Scott Marcus (1999), *Designing Wide Area Networks and Internetworks: A Practical Guide*, Addison Wesley.

It is often forgotten that effects similar to prioritisation can be achieved by *caching* (storing frequently used static data close to the user) and by *replication* (where the same dynamically generated results can be produced in more than one location in the network – cloud services are an example of this kind of distribution or replication of function).

4.1.4 The mathematics of prioritisation

Queuing delay and packet loss are dependent on the amount of traffic that attempts to enter one hop of the network, in comparison to the amount of traffic that it can accommodate. *Queuing theory*, a branch of mathematics that deals with waiting lines, can be used to analyse these characteristics.²⁶

The mathematics of prioritised service delivery have been well established in the literature for decades, but they are not well known to most of those who participate in the network neutrality debate. The mathematics is demanding.

One strand of the literature is scheduling theory, which is relevant not only to the scheduling of industrial processes, but also to the scheduling of activities within a computer system. The classic text is Conway, Maxwell and Miller (1967).²⁷ The second relevant literature strand is queueing theory, the mathematics of waiting lines. Queuing theory is equally relevant to customers waiting to pay for their goods in a retail establishment, and to IP packets / datagrams waiting for service in a router.²⁸

Consider a system with just two priorities, high and low. The magnitude of these effects depends on the relative number of packets in each class.

- If all packets were of high priority, or all of low priority, then prioritisation would have no effect whatsoever.
- If few packets were of high priority, and most were of low priority, then prioritisation would accelerate the small number of high priority packets by moving them to the head of the queue, but the low priority packets would experience only a small additional delay as a result.
- If most packets were of high priority, and few were of low priority, then prioritisation would *only slightly* accelerate the small number of high priority packets by moving them to the head of the queue (although the variability of delay would be reduced), but the low priority packets might experience substantial additional delay as a result.

The second of these cases corresponds to what one might expect if only VoIP datagrams were to be prioritised ahead of other traffic. VoIP is a low bandwidth service, but it is heavily dependent on the average and the variability of delay. Moving VoIP to the head of the waiting queue has little impact on other services, because the number of datagrams is small. This

²⁶ For a brief introduction to queueing theory, see Chapter 16 of J. Scott Marcus (1999), *Designing Wide Area Networks and Internetworks: A Practical Guide*, Addison Wesley Longman. For a comprehensive discussion, see Hisashi Kobayashi (1978), *Modeling and Analysis: An Introduction to System Performance Methodology*, Addison-Wesley.

²⁷ Richard Walter Conway, William L. Maxwell, Louis W. Miller (1967), Addison-Wesley. The text has been out of print for some time, but appears to be available as an e-book.

²⁸ See for instance Hisashi Kobayashi (1978), *Modeling and Analysis: An Introduction to System Performance Methodology*, Addison-Wesley, pages 209-211.

would appear to be an ideal case, where the consumer gains a much more effective and predictable use of VoIP, but with minimal negative impact on other applications.

The third case could be said to correspond to the delivery of linear television-like video (*IPTV*) to consumers over the Internet. A large volume of prioritised IPTV traffic is moved ahead of web traffic in the queue. In principle, this causes a large delay to the web traffic. In practice, this does not appear to generate consumer complaints. Network operators in a competitive market presumably have no incentive to make their customers unhappy, so they apparently limit the total level of IPTV traffic so as to ensure that web traffic also gets through. Given that IPTV traffic is a large but relatively steady stream of traffic, moving it forward on the queue does not introduce substantial incremental variability of delay (jitter) into the web traffic that has been delayed.

Consider the behaviour of traffic queueing on an outbound (router) transmission link in a system with just two priorities. The average sojourn time (defined as time spent being served plus time spent waiting in the queue) can be directly computed using queueing theory.²⁹ If only 10% of the capacity of the link is taken up with traffic in the high priority category (see Figure 2), the benefit for prioritised traffic is substantial, even though the impact on low priority traffic is not great (a 2% penalty at 20% overall load, an 8% penalty at 70% overall load); however, the impact is greater at high overall loading of the link than at low. If however 30% of the capacity of the link is taken up with traffic in the high priority category (see Figure 2), the benefit for prioritised traffic is substantial, even though the impact on low priority traffic is somewhat greater: a 17% penalty when the link is 40% loaded, a 30% penalty when the link is 70% loaded.

²⁹ See for instance Ivo Adan and Jacques Resing (2015), “Queueing Systems”, at <http://www.win.tue.nl/~iadan/queueing.pdf>. For ease of exposition, we assume a simple M/M/1 queueing system (which is equivalent to an M/G/1 system with a coefficient of variation of 1.0, but results are available for much more complex systems as well).

Figure 1. Relative sojourn time in an outbound router queue with and without 10% of traffic subject to prioritization as a function of load on the link.

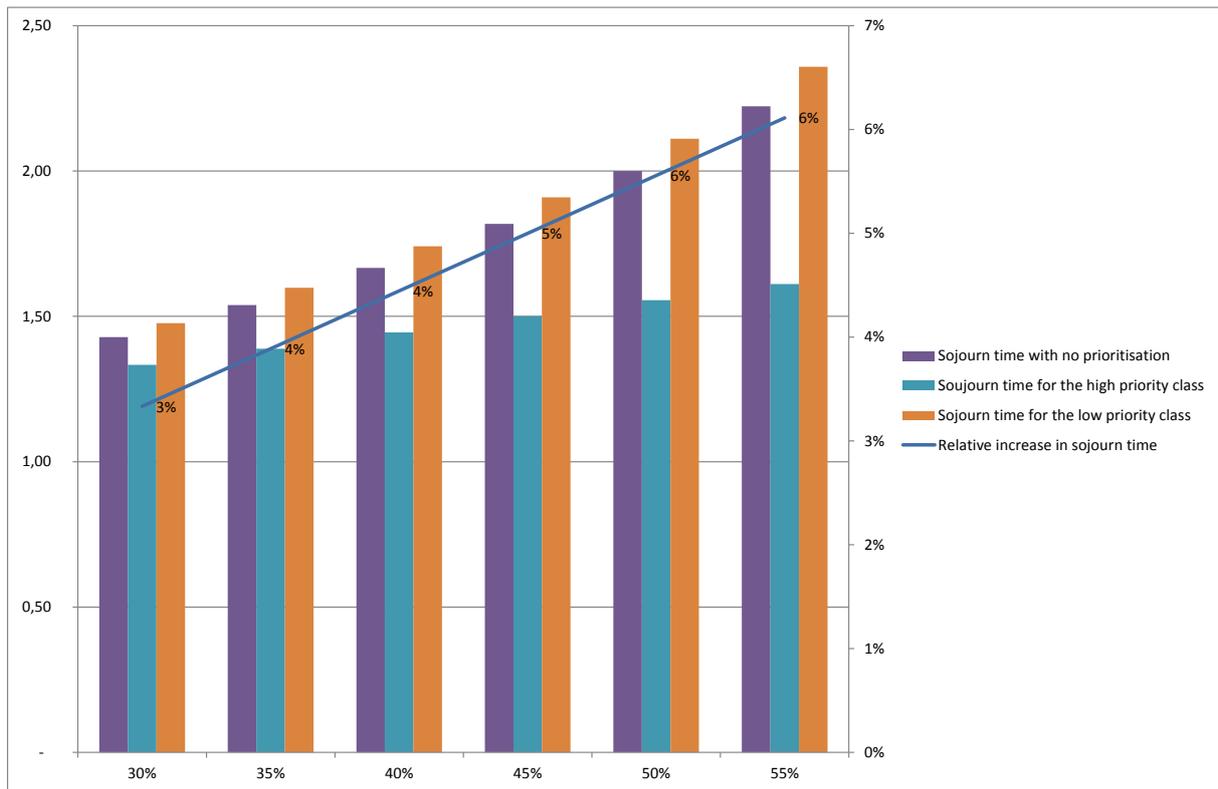
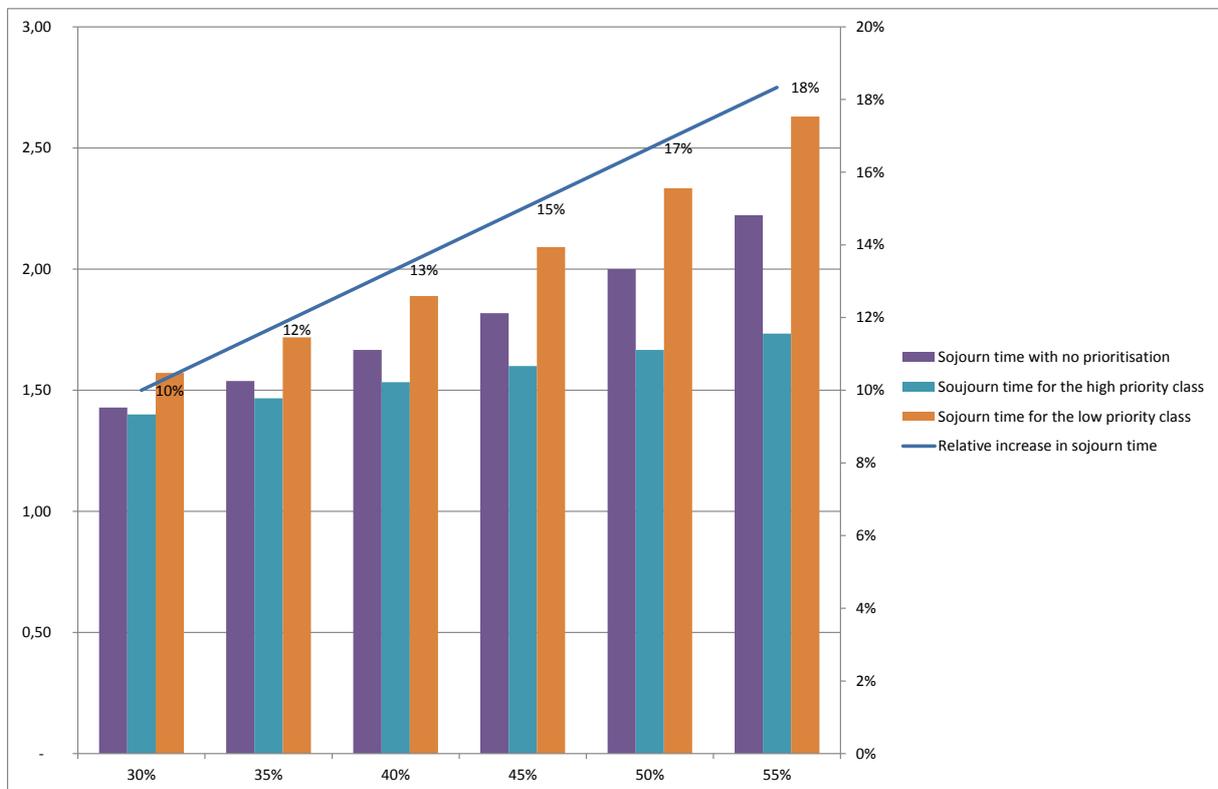


Figure 2. Relative sojourn time in an outbound router queue with and without 30% of traffic subject to prioritization as a function of load on the link.



4.1.5 Challenges in measuring evolving QoS/QoE

A measurement programme of this type faces a great many practical challenges. We note a few of them here.

4.1.5.1 The application-specific nature of QoE

Quality of Experience (QoE) depends not only on objectively measurable QoS parameters, but also on the requirements of the application. Measuring Internet performance is hard enough; measuring application-specific performance poses special challenges, as we explain in Chapter 4.3.

For the particular problem taken up in this paper, this is of less concern than one might think. We are concerned with assessing the performance characteristics of *traffic that is not subject to prioritisation*, in order to ensure that the performance has not declined (but see also Section 4.1.5.3). Since traffic within this class is not supposed to be specifically managed, it *might* not to be necessary to record more than the overall characteristics of this traffic class as a whole. In principle, *the performance of non-prioritised traffic* (call it best efforts traffic for now, but subject to the caveats noted in Section 2.2) *should be roughly the same irrespective of the application for which it is used.*³⁰

In mathematical, these observations can all be viewed as having been drawn from the same underlying distribution. QoS parameters (one way delay, for instance) could be understood, then, in terms of basic statistical measures such as mean and standard deviation, or in terms of quantiles, for the distribution as a whole irrespective of application or application type.

This means that it is possible to measure QoS parameters of the non-prioritised traffic without distinguishing among application types. *In order to understand the influence of QoS upon QoE, we need only to apply the QoE requirements of each application to the known QoS characteristics of the non-prioritised traffic distribution as a whole.*³¹

4.1.5.2 Identifying the cause of a change in QoE

The QoE experienced by the user is the result of performance at every point along the path that the data takes, which is roughly (but not exactly) the same as value chain (see Figure 4). Thus, when the user requests content from a content and application provider (CAP), the speed of response is dependent on:

- The speed of the content and application provider's service, especially the servers;
- The speed of the ISP that serves the CAP (noting that the CAP may self-serve);
- Any number of intermediate ISP transit networks and (at most, in general) one peering interconnection point;
- The speed of the ISP that serves the end user (i.e. the *Internet access provider (IAP)*);

³⁰ To be sure, there are differences: the packet length distribution, inter-arrival time distribution, and more will not be the same. Nonetheless, the assumption to a first order should be correct.

³¹ More subtle impairments are, to be sure, possible. Consider for example ComCast's degradation of BitTorrent traffic. We do not believe that it is cost-effective today to try to detect every possible kind of application-specific degradation through a generalised measurement programme. In some instances, it will be necessary to rely on feedback from end-users, bearing in mind that degradation is not likely to be very interesting to the NRA unless it causes human-perceptible impacts.

- The user's own *customer premises equipment (CPE)*, including routers within the user's home, as well as the consumer's personal computer, tablet, or smart phone.

If present (as is often the case), a *Content Delivery Network (CDN)* will also tend to play a strong role.

The debate over network neutrality has primarily been a debate about degradation caused by the IAP. Degradation could, however, be introduced at any point along this long chain. How should one distinguish between degradation caused by the IAP versus other, less problematic forms of degradation?

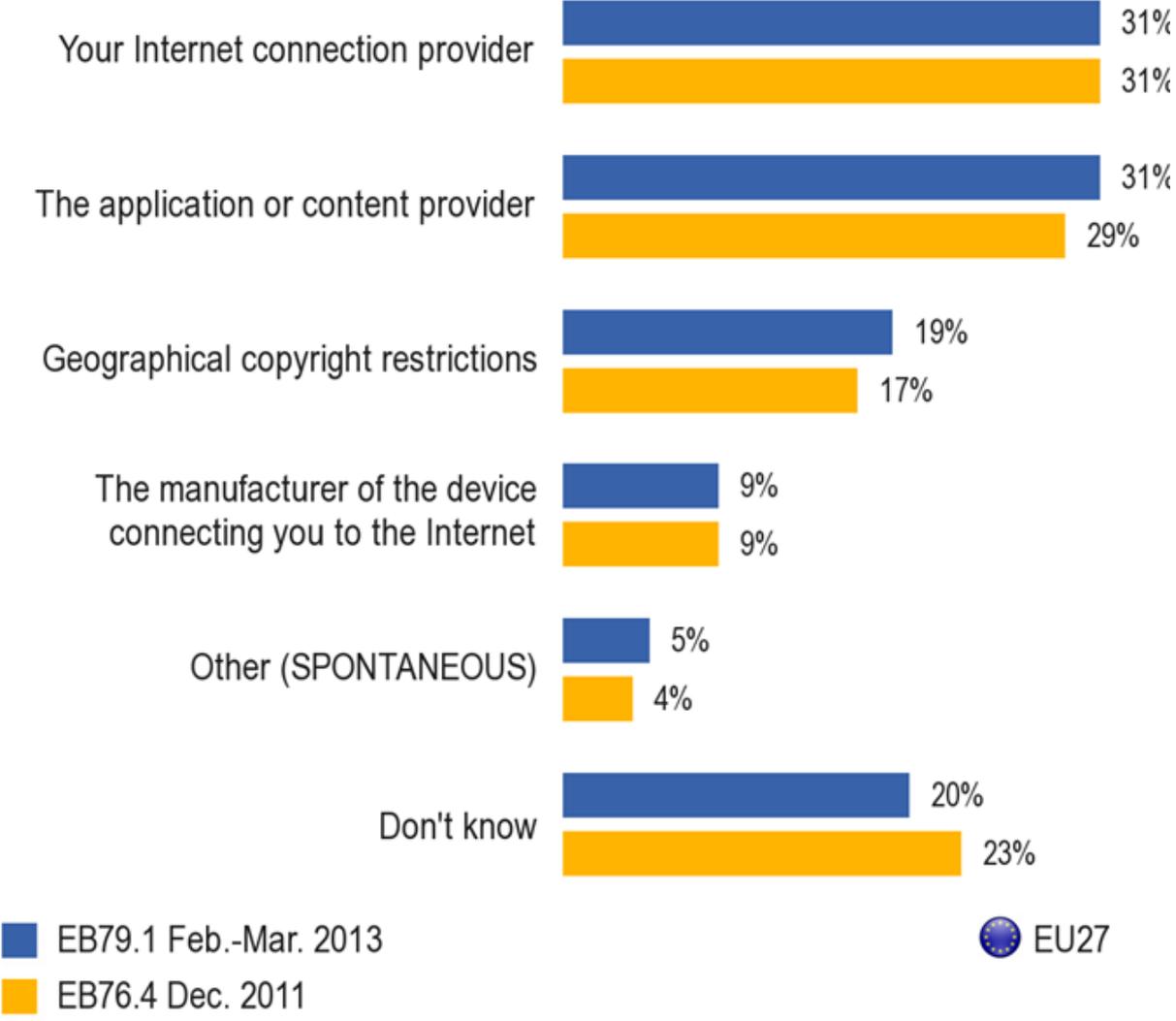
This is a significant issue. A well-known Eurostat consumer survey³² shows that of those European consumers who believe that their service has been blocked at least once, at least as many attribute the blocking, not to the IAP, but rather to the Content and Application Provider (CAP), while others attribute it to geo-blocking, or to the provider of the end device (see Figure 3).³³ These different forms of blocking have quite different public policy implications.

³² Eurobarometer 396 (2013).

³³ See also WIK-Consult, Deloitte, YouGov (2015): *The Value of Network Neutrality to European Consumers*. Summary Report.

Figure 3. User perceptions of the cause of blocking of online content or applications.

QA15. In your opinion, which of the following was responsible for the blocking of online content or applications?

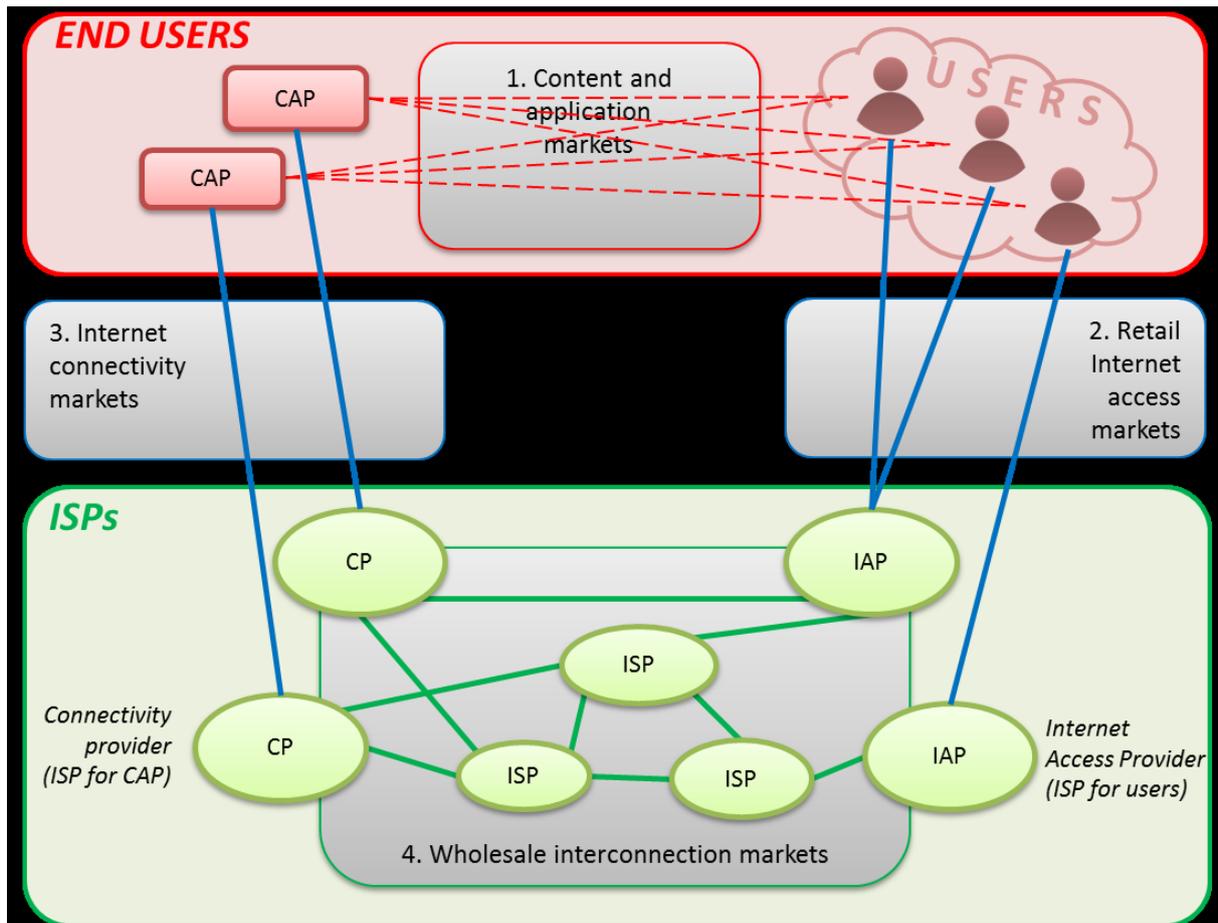


Source: Eurobarometer 396 (2013).

Depending on what is measured, and where, and how, it might well be the case that an adverse change is observed, but that it is difficult to determine the cause of the performance loss. In the authors’ experience, it is often difficult to pinpoint the source of a performance loss, even with good staff, good tools, and good will all around.

All of this has important implications for the design of any measurement system. It is important to be clear as to what is to be measured, and from what vantage point.

Figure 4. The value chain: Content and application providers, different kinds of ISPs, and the end-user.



Source: BEREC (2012).³⁴

4.1.5.3 Changing expectations over time

The price-performance of nearly everything that has to do with information and communication technologies (ICTs) has progressively improved thanks to Moore’s Law. In a famous 1965 paper, Gordon Moore recognised that the number of components that could cost-effectively be implemented with a single integrated circuit was doubling per unit time.³⁵ Today, it is widely understood that, with present technology, the number of components per integrated circuit approximately double every two years.³⁶ We are all familiar with these effects. If we buy a personal computer today, it will cost no more than a personal computer that we could have bought two years ago, but it will be twice as fast, will have twice as much memory, and will likely have a hard disk drive (or semiconductor equivalent) that is twice as large.³⁷

³⁴ BEREC (2012), “Differentiation practices and related competition issues in the scope of net neutrality”, BoR (12) 132.

³⁵ Gordon Moore, “Cramming more components onto integrated circuits”, *Electronics*, Volume 38, Number 8, April 19, 1965. The original paper suggested a doubling per year, but subsequent work found the rate to be a doubling every eighteen to twenty-four months.

³⁶ Fundamental physical limits will eventually put a limit to trend, but it is likely to continue for a few more cycles at least.

³⁷ Disk storage is also benefiting from rapid improvements in technology.

For this reason, merely requiring that the performance of Internet applications become no worse over time might well miss the point. A user receiving 2015 performance in 2020 might well be subject to intentional degradation, even in the absence of a decline in measured performance.

It is for this reason that the current proposal to the Council from the Latvian presidency call on NRAs to “promote the continued availability of open internet access services at levels of quality that reflect advances in technology”.³⁸

A measurement programme needs to take this into account, but it need not pose a fundamental difficulty. The *thresholds* used to indicate a potential problem will need to “ratchet” upwards from year to year. As long as the underlying measurement data remains intact, doing so should not preclude longitudinal studies over time.

4.2 Measurement programmes that have been attempted or implemented

It is clear that NRAs implement a wide range of programmes to measure Internet Quality of Service (QoS) today. An extensive list appears in BEREC (2014b),³⁹ the BEREC annex to the report on “Monitoring quality of Internet access services in the context of net neutrality”.

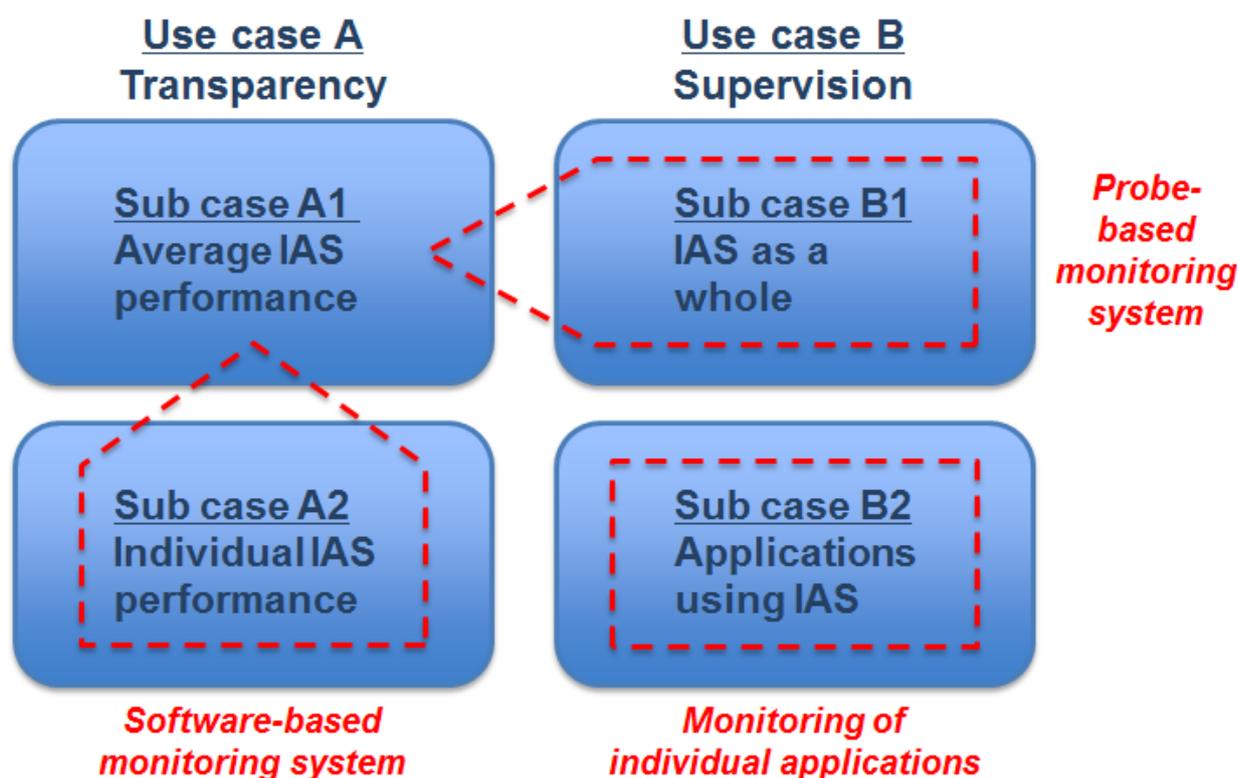
As BEREC notes, different programmes are used for different purposes (see Figure 5). Some are seen as being preferred for regulatory supervision, while others are more useful for facilitating transparency; and most address the network as a whole, with only a few measuring the performance of individual applications.

Active programmes are more widely used, but some passive programmes are in place. Measurements from fixed locations tend to be preferred for purposes of regulatory supervision, but crowd-sourced measurements from (potentially large numbers of) consumers also have their place.

³⁸ European Council (2015), “Proposal for a Regulation of the European Parliament and of the Council laying down measures concerning the European single market for electronic communications and to achieve a Connected Continent, and amending Directives 2002/20/EC, 2002/21/EC and 2002/22/EC and Regulations (EC) No 1211/2009 and (EU) No 531/2012 - Examination of the Presidency’s compromise proposal”, 26 May 2015.

³⁹ BEREC (2014b): Annex of: Monitoring quality of Internet access services in the context of net neutrality. Update after public consultation. Annex to BoR (14) 117.

Figure 5. Different forms of monitoring are used to fulfill different objectives.



Source: BEREC (2014a).⁴⁰

Measurement programmes also vary as to whether they measure the fixed network or the mobile; and whether they use dedicated probes versus measurements taken from the end user's own device.

The network quality initiative undertaken by zafaco on behalf of the German NRA BNetzA and the Misura⁴¹ network quality initiative undertaken by the Fondazione Ugo Bordoni (FUB) on behalf of the Italian NRA AGCOM are both noteworthy. They are elaborate representatives of measurement programmes focusing on speeds. The key question is not in relation to securing the best effort Internet; rather, the key focus is whether consumers get speeds that are close to the maximum speeds that their contract indicates.

SamKnows is one of the few cases that use dedicated hardware to measure.⁴² It is also quite unique in terms of the range of different tests it covers: Speeds (single QoS parameters), application classes (such as web browsing), and specific applications (BitTorrent, YouTube, Netflix).⁴³

⁴⁰ BEREC (2014a): Monitoring quality of Internet access services in the context of net neutrality. Update after public consultation. BoR (14) 117, op. cit.

⁴¹ See <https://www.misurainternet.it/architettura.php>.

⁴² See <https://www.samknows.com/meet-the-whitebox>.

⁴³ See <https://www.samknows.com/broadband/uploads/methodology/SQ301-005-EN-Test-Suite-Whitepaper-4.pdf>.

Looking (to some extent) beyond Europe, nearly every OECD member state engages in speed tests.⁴⁴

4.3 Complementary measurement approaches

While the paper thus far focused on measurement programmes, this section identifies and potential complementary solutions, such as consumer surveys, and measurements undertaken by Content and Application Providers (CAPs) or other third parties.

Many NRAs worldwide conduct (or previously conducted) consumer satisfaction surveys on a regular basis and published reports about the quality as perceived and reported by the participants. For instance, the Malaysian Communications and Multimedia Commission (MCMC) lists on a dedicated webpage⁴⁵ annual⁴⁶ consumer satisfaction survey reports since 2001. MCMC stopped conducting the surveys after 2011. We picked MCMC's case as it is exemplary for many other NRAs. It appears as if consumer satisfaction surveys were mostly a thing of the past. They are a rare breed nowadays.

It should be noted that consumer satisfaction surveys were typically focused on perceived voice quality. They thus naturally adopt an application class-specific, or even an application-specific perspective. When used today and with an eye to applications that rely on a best effort service model, consumer satisfaction surveys would obviously have to adapt to the current portfolio of relevant online applications, but we do not see any fundamental technical reason that would hinder such transposition of the concept in principle. The more important question might however be whether there is sufficient value for an NRA to use the instrument. Given that most NRAs profit from a complaints office for consumers, and since many NRAs are invested in some form of measurement programme, the need for consumer satisfaction surveys might not be evident. If a physician had the choice between conducting a series of tests with a patient and hearing from patients how well they think they were treated, the physician would probably go for the former. Likewise, an NRA would probably rate the ability to measure over the ability to ask consumers about their impression.

Consumer satisfaction surveys may nonetheless determine an interesting instrument, albeit more to CAPs than to NRAs. For instance, Skype presents end-users with the question how they rate the voice or video quality after completing the call. End-users can rate perceived quality on a scale from one to five, ranging from very bad to excellent call quality. Skype does not publish any statistics on call ratings. Other CAPs, however, have initiated measurement programmes on their own – and use the respective reports actively in the public debate on network performance of different ISPs.

Notable examples for CAP-driven measurement programmes include the ISP speed index⁴⁷ and Google's video quality report⁴⁸. Both are application-specific; they reflect two of the largest video streaming applications world-wide, namely Netflix and YouTube. Both are speed-oriented, meaning they report mainly on throughput. The ISP speed index from Netflix

⁴⁴ OECD (2014), "Access Network Speed Tests", OECD Digital Economy Papers, No. 237, OECD Publishing. <http://dx.doi.org/10.1787/5jz2m5mr66f5-en>.

⁴⁵ See <http://www.skmm.gov.my/Resources/Statistics/Consumer-Satisfaction-Survey.aspx> (Accessed: May 2015).

⁴⁶ There was however no report for the year 2005.

⁴⁷ See <http://ispspeedindex.netflix.com/> (Accessed: May 2015).

⁴⁸ See <https://www.google.com/get/videoqualityreport/> (Accessed: May 2015). Please note that the website only shows results if it is accessed from a location which Google covers in the measurement programme. The website otherwise provides information on the programme, e.g. its methodology.

calculates for each country observed the maximum, minimum and average speed. Within a country, the index ranks observed ISPs by average speed, and it indicates whether individual ISPs moved up or down in the ranking in comparison to previous periods.

Various scenarios are conceivable in which CAPs may see a benefit in building a measurement programme on their own and to go public about it. The motivation for Netflix and Google apparently is to use the reports to put pressure on those ISPs that do not provide sufficient bandwidth to the video streaming application in question, possibly even throttle the respective traffic. Pointing fingers to underperforming ISPs helps them make consumers aware that the reason for inferior experience may be with the party that transports data, not with the video platform itself. It also helps them create incentive for ISPs to invest into network capacity (or avoid throttling traffic) available to the best effort Internet.

CAP-driven measurement programmes therefore may be seen as a complementary solution to NRA-driven measurement programmes; however, they have limitations. Most importantly, they are funded and operated by a non-neutral actor. While the overall methodology in use is documented for a public audience, the measurement programmes have not been audited by a neutral party. These programmes are not fully transparent. Whether the pressure they place on ISPs to enter into agreements with CAPs⁴⁹ ultimately serves the public remains to be seen.

There are also open Internet performance measurement platforms and annual network quality studies by independent parties. When the Measurement Lab (M-Lab) issued a report, which concluded based on data collected by the M-Lab platform that “ISP interconnection has a substantial impact on consumer internet performance”⁵⁰, it sparked a complex debate⁵¹ on degraded experience induced by the behaviour of transit ISPs like Cogent. The key conclusion from this case is that open platforms such as M-Lab may constitute invaluable additional sources to provide insight on select topics that otherwise might have stayed hidden.

Periodic network quality studies conducted by independent parties may serve as similarly helpful complementary instruments. Prominent examples include the semi-annual J.D. Power U.S. wireless network quality study⁵² and Connect’s annual wireless network study⁵³ for Germany, Austria, and Switzerland. J.D. Power is a market research firm; Connect is a telecommunications-oriented media outlet. Both studies use automated drive testing. While on the road or in a train, they replicate the scenario of a mobile end-user making phone calls or using the mobile Internet. They focus on metrics like call drop rates, call setup time, website loading times, or failed website request rates. They address a wider audience including consumers by answering primarily the question, which wireless network is overall best suited for a certain mobile user profile. These network quality studies focus on non-data applications (e.g. mobile telephony). They therefore cannot serve as a replacement for the measurement programmes discussed elsewhere in this report. They do not deliver nearly enough fine-grained information on the capacity of the best effort Internet. Moreover, automated test drives are very costly, so that network quality studies collect data only during a relatively

⁴⁹ Such agreements might be legal in some jurisdictions; in others they might not.

⁵⁰ See Measurement Lab (2014), p. 4.

⁵¹ See for instance <http://blog.streamingmedia.com/2014/11/cogent-now-admits-slowed-netflixs-traffic-creating-fast-lane-slow-lane.html> and <http://arstechnica.com/information-technology/2014/11/during-netflix-money-fight-cogents-other-big-customers-suffered-too/> (Accessed: May 2015).

⁵² For a summary of the latest issue (2015 Vol. 1) as regards study description and its primary findings, see http://www.jdpower.com/sites/default/files/2015024%20Wireless%20NQ%20Release%20V1_Press%20Release_Final.pdf (Accessed: May 2015).

⁵³ See <http://www.connect.de/netztest/> (in German; Accessed: May 2015).

short period in a year. They therefore represent snapshots rather than continuous quality measurements.

Voluntary industry codes of conduct, together with self-regulatory or co-regulatory policing mechanisms, represent an overall alternative or complement to monitoring by the NRA. Voluntary industry initiatives such as the Broadband Stakeholder Group's traffic management⁵⁴ and open Internet⁵⁵ codes have led many UK-based ISPs to release relatively easy to understand and compare information on their traffic management practices according to a common set of indicators,⁵⁶ thus arguably addressing BEREC's transparency objectives. Many ISPs have committed themselves to avoid blocking legitimate traffic and content. If these codes were conceptually adapted to cover in addition specific commitments as regards the current and future capacity made available for the best effort Internet, they could perhaps evolve into a model for a voluntary industry initiative with the potential to mitigate "dirt road" concerns overall.

⁵⁴ See Broadband Stakeholders Group (2011).

⁵⁵ See Broadband Stakeholders Group (2014).

⁵⁶ A regularly updated online repository of the signatory ISPs' traffic management practices is available at: <http://www.broadbanduk.org/bsg-openinternettrafficmanagement/trafficmanagementkfis/> (Accessed: May 2015).

5 Pragmatic considerations

There are many design choices that impact what a programme to measure whether Internet access services are being provided “at levels of quality that reflect advances in technology”⁵⁷ is able to deliver. These choices will determine the limits of such a programme.

BEREC is already seeking to exploit commonalities among the requirements of various NRAs; however, it is fairly clear that a wide range of active and passive measurement tools are already in place,⁵⁸ and we assume that this will continue to be the case.

With this in mind, we discuss in this chapter the distinction between data capture versus data analysis, together with the need to systematically build an evidence base for subsequent analysis (Section 5.1); the rationale for viewing this as a semi-structure problem amenable to analysis by a Decision Support System (DSS), but still requiring human judgment (Section 5.2); statistical approaches to assessing changes in overall network performance over time (Section 5.3); and a few words about the placement of measurement probes (Section 5.4).

5.1 Building the evidence base

NRA are already using a wide range of tools and methodologies (both active and passive) to capture data relevant to network neutrality.⁵⁹ It does not seem to be either helpful or practical to identify a single, preferred way to capture the data.

It is, however, helpful to draw a clear distinction between *data capture* and *data analysis*.

However it is captured, data flows into the evidence base used for subsequent analysis.

As we have seen in Section 4.2, NRAs are already using a wide variety of tools, including both active and passive measurement, managed versus crowd-sourced approaches, to monitor QoS today. We assume that this will continue to be the case, and see no reason why things should be otherwise. NRAs will have to accommodate local circumstances.

We nonetheless feel that it is possible to conceptualise the supervisory process in an integrated way. Data flows from a plethora of sources into the national evidence base. From there, a collection of tools (collectively termed a *Decision Support System*, see Section 5.2) will be used to help a human analyst to assess whether there are indicia of an emerging “dirt road” effect.

As previously noted, BEREC is seeking commonalities among Member State NRA requirements.⁶⁰ The “front end” of the evidence base will tend to differ from NRA to NRA to the extent that they are using different collection tools; however, there is a real possibility that back end analysis tools could enjoy a degree of commonality.

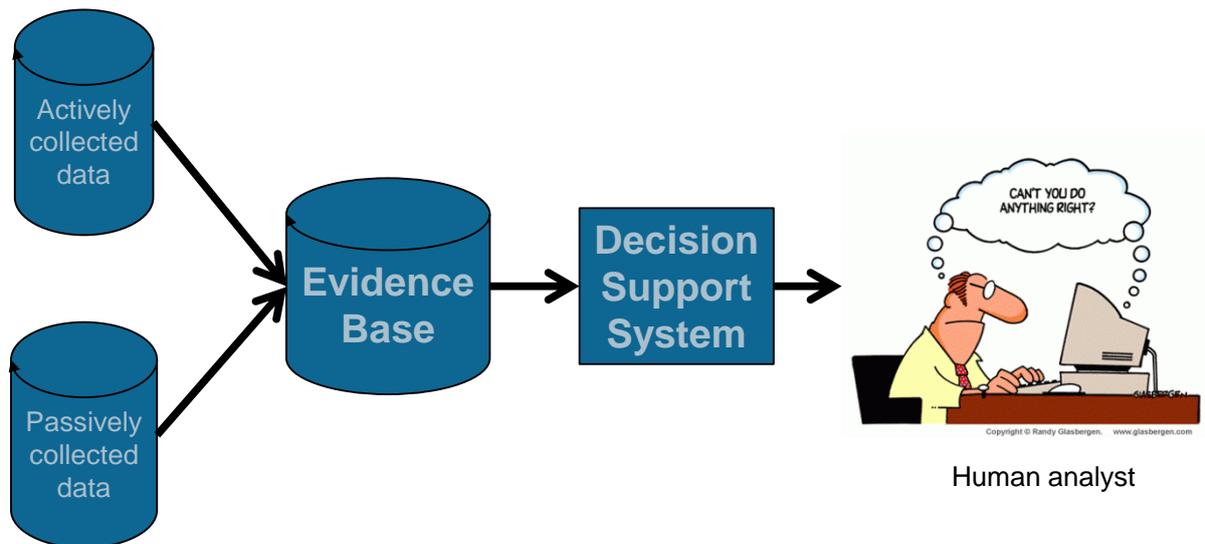
⁵⁷ This is the formulation from the European Council’s memorandum of 26 May 2015, op. cit.

⁵⁸ BEREC (2014b): Annex of: Monitoring quality of Internet access services in the context of net neutrality. Update after public consultation. Annex to BoR (14) 117.

⁵⁹ Ibid.

⁶⁰ BEREC (2014a): Monitoring quality of Internet access services in the context of net neutrality. Update after public consultation. BoR (14) 117, at: http://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/4602-monitoring-quality-of-internet-access-se_0.pdf.

Figure 6. Many data sources flow into the evidence base, which is assessed using a Decision Support System (DSS) and ultimately a human analyst.



5.2 A semi-structured problem

Detecting changes in the quality of the best effort Internet over time is to a great extent non-trivial since it is not a *structured problem*.

- A *structured problem* is a routine problem that is amenable to a fully automated solution using technologies available at present. Both current and desired states as well as the relations that lead from current to desired state of a structure problem are known.
- An *unstructured problem* is not at all amenable to automated solution. Human judgment is needed. Automated tools are of little or no relevance.
- A *semi-structured problem* is a problem where human judgment is required, but automated tools can help or support the human analyst. *Decision Support Systems (DSS)* are often used to support the human analyst.

Over time, as a problem becomes better understood, it is not unusual for an unstructured problem to evolve into a semi-structured problem, or for a semi-structured problem to evolve into a fully structured problem.

Our belief is that determining whether Internet access services are being provided “at levels of quality that reflect advances in technology”⁶¹ is a semi-structured problem today. Over time, it might perhaps evolve into a fully structured problem, but we are a long way from that point.

Semi-structured problems show some structure, i.e. some elements to decision making are known, while other elements remain uncertain a priori. For the specific problem envisioned, we find structure when it comes to network performance measurements. Monitoring network performance according to a set of well-established set of quality parameters such as throughput, delay, jitter, packet loss rate and error rate has been done for quite a while. It is a routine network management task. We describe the wide range of measurement programmes that have been implemented in Section 4.2.

⁶¹ This is the formulation from the European Council’s memorandum of 26 May 2015, op. cit.

5.3 Evaluating changes in end-user performance over time

Within the Decision Support System (DSS), we anticipate that a range of statistical tools will be needed, as well as data management and data visualisation tools. Various tools to analyse time series and detect trends will clearly be required, as well as tools based on analysis of variance techniques that try to distinguish real trends from unrelated movements that in this context represent random “noise”. In effect, we are seeking to establish confidence intervals.⁶²

In principle, the problem bears similarities to panel data analysis; however, the large data volumes would appear to imply that simpler, heuristic techniques will need to be employed.

In assessing changes and tendencies over time, it is natural to think of using many of the same techniques that industrial engineers utilise to detect shifts in an industrial process. (The same could perhaps be said of market analysts.)

One of the most common approaches in industrial engineering is the *exponentially smoothed average*. In it, new data is given a weight α , while the previous average is given a weight of $(1-\alpha)$. For instance, new data has a weight of 20%, old data a weight of 80%.

The use of an exponentially smoothed average reduces the risk of over-reaction to a brief “spike” in some parameter, but converges fairly quickly to the new value when there is a real change. The choice of α controls the speed of this convergence. A low value of α minimises the risk of over-reaction to “spikes”, but also risks failing to respond promptly enough to a real change; conversely, a high value of α enables quick convergence, but risks over-reaction. It is thus clear that α must be chosen with care, and that the choice entails at least as much art as science.

5.4 Where should measurements be taken?

As BEREC itself has noted, the nature and location measurement probes is important, but does not seem to be an altogether settled matter among the NRAs.⁶³

In this paper, we are considering detection of a possible “dirt road” effect by the NRA, which is a form of regulatory supervision in BEREC’s terminology. As BEREC notes, where regulatory supervision is the goal, there is often a need for probes at one of more fixed locations in order in order to have repeatability.⁶⁴ “Crowd sourced” solutions where consumer’s run text programmes at times and from locations of their own choosing can potentially generate a far greater number of samples, but without an assurance that these samples are representative in times of time of day, geographic distribution, or characteristics of the user base (since those who generate the tests are potentially subject to *self-selection bias*), and with the risk of bias or variation due to other concurrent activities in the end user’s device.

At the same time, a complementary role for crowd-sourced measurements might well be possible and appropriate.

⁶² How to achieve this will be a significant focus in our ongoing study on behalf of the German BNetzA.

⁶³ BEREC (2014a): Monitoring quality of Internet access services in the context of net neutrality. Update after public consultation. BoR (14) 117, at: http://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/4602-monitoring-quality-of-internet-access-se_0.pdf.

⁶⁴ Ibid. “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 IASes with respect to type, number and distribution. This approach allows for regulatory supervision of potential degradation of IAS as a whole (sub case B1).”

6 Concluding observations

It is clear that overall operational monitoring of network neutrality can be very challenging; operational monitoring at network level in order to identify “dirt road” effects, however, appears to be a relatively straightforward problem.

Obtaining useful results appears to depend on:

- A clear specification of the problem;
- Ruthless paring down to avoid needless distractions; and
- Good engineering discipline and methodology.

Nonetheless, substantial uncertainties will remain until there is sufficient implementation experience.

References

BEREC (2012): BEREC Guidelines for quality of service in the scope of net neutrality. BoR (12) 131, pp. 1-67. Available at:

http://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/1101-berec-guidelines-for-quality-of-service-_0.pdf.

BEREC (2014a): Monitoring quality of Internet access services in the context of net neutrality. Update after public consultation. BoR (14) 117. Available at:

http://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/4602-monitoring-quality-of-internet-access-se_0.pdf.

BEREC (2014b): Annex of: Monitoring quality of Internet access services in the context of net neutrality. Update after public consultation. Annex to BoR (14) 117.

Broadband Stakeholders Group (2011): Voluntary industry code of practice on traffic management transparency for broadband services. Pp. 1-10. Available at:

<http://www.broadbanduk.org/wp-content/uploads/2013/08/Voluntary-industry-code-of-practice-on-traffic-management-transparency-on-broadband-services-updated-version-May-2013.pdf>.

Broadband Stakeholders Group (2014): Open Internet Code of Practice: Voluntary Code of Practice Supporting Access to Legal Services and Safeguarding Against Negative Discrimination on the Open Internet. Pp. 1-11. Available at: <http://www.broadbanduk.org/wp-content/uploads/2015/01/BSG-Open-Internet-Code-of-Practice-amended-November-2014.pdf>.

Cisco Systems (2009): Cisco IOS Quality of Service Solutions Configuration Guide. Release 12.2SR, pp. 1-778. Available at:

http://www.cisco.com/c/en/us/td/docs/ios/qos/configuration/guide/12_2sr/qos_12_2sr_book.pdf.

European Parliament and Council (2009a): 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. OJ L 337, pp. 11-36. Available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009L0136&from=EN>.

European Parliament and Council (2009b): Directive 2009/140/EC of the European Parliament and of the Council of 25 November 2009 amending Directives 2002/21/EC on a common regulatory framework for electronic communications networks and services, 2002/19/EC on access to, and interconnection of, electronic communications networks and associated facilities, and 2002/20/EC on the authorisation of electronic communications networks and services. OJ L 337, pp. 37-69. Available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009L0140&from=en>.

Measurement Lab (2014): ISP Connection and its Impact on Consumer Internet Performance. Release 1.0, pp. 1-37. Available at: http://www.measurementlab.net/static/observatory/M-Lab_Interconnection_Study_US.pdf.

WIK-Consult, Deloitte, YouGov (2015): The Value of Network Neutrality to European Consumers. Summary Report.