INTERNATIONAL TELECOMMUNICATION UNION
TELECOMMUNICATION DEVELOPMENT BUREAU

Document: 2

Global Seminar on Quality of Service and Consumer Protection
Geneva, Switzerland, 31 August-1 September 2006

ICT Quality of Service Regulation:

Practices and Proposals

BACKGROUND PAPER

FINAL VERSION

September 2006

International Telecommunication Union (ITU)

This background paper was prepared by Robert Milne (rem@antelope.org.uk), Antelope Consulting, United Kingdom. The views expressed in this paper are those of the author and do not necessarily reflect the views of ITU or its members.

The author wishes to express his sincere appreciation to Luis Cardozo, Bobby Hannan, Müberra Güngör, Claire Milne, Sebastian Moeller, Enrique Pérez, Nicolas Silva, Nancy Sundberg, Volker Sypli and Eric Yam.

This is one of two background papers to be presented for comments at the ITU Global Seminar on Quality of Service and Consumer Protection, Geneva, Switzerland, 31 August-1 September 2006. Further information can be found on the web site at http://www.itu.int/ITU-D/treg.

International Telecommunication Union

Place des Nations

CH-1211 Geneva, Switzerland

Contents

Summary

For many years some telecommunications regulators have had schemes for quality of service monitoring that recognised its importance in their markets. More recently other regulators have been drafting regulations and, in some cases, publishing results. Yet there have been various apparent setbacks: various regulators have stopped their work on quality of service after publishing consultation papers, regulations or initial quality of service measurements. The problem overall may be that the resources of regulators are insufficient: regulators have many urgent and important tasks and require time to appreciate fully the aims of, and tasks in, quality of service monitoring.

Several regulators now have plans to devise new schemes or revise existing schemes for quality of service monitoring. This paper is intended to help those and other regulators, by starting to study how experience so far can guide quality of service monitoring in the future, particularly in countries where it is not yet well established.

Section 1 looks at the question of when quality of service monitoring by regulators is desirable. Broadly speaking, the section concludes that quality of service measurements can be beneficial in several ways, at least if they are published for all operators, but that quality of service targets can be detrimental unless they are applied only to dominant operators. Though there are exceptions to this conclusion it suggests that fewer targets may be needed than are currently imposed.

Section 2 deals with operational arrangements for quality of service monitoring by regulators. There are various ways in which operators can be involved in the arrangements for defining measurements, setting targets, making measurements, auditing measurements and publishing measurements. There are also ways in which customers can be involved, even in defining measurements and making measurements. The cases studied illustrate wide variations in how quality of service monitoring is done. The more and less successful aspects of the cases are not always what might be expected.

Section 3 turns to the question of what quality of service monitoring by regulators is desirable. Many measurements have been defined and many targets have been set. Regulators and operators need to choose those that match their main aims. The section describes criteria for choosing the measurements to be defined, the targets to be set and the measurements to be published. It also discusses various other choices that regulators and operators must make, such as how measurements are to be made (particularly for voice and related means of communication) and which services are to be monitored in current network and Next Generation Network (NGN) environments. These choices are particularly important as communication services become layered on Internetwork Protocol (IP) networks, when operators will use wholesale services, and bundle retail services, both within and between layers.

Quality of service monitoring inevitably differs in different countries, with their own industry structures and customer preferences. It is unlikely ever to be perfect, as regulators are unlikely ever to have full enough information about the markets. This paper does not pretend to say what regulators always ought to do. However, it does suggest the following overall guidelines for discussion:

- Using operator skills and customer opinions through widespread consultations, working groups and open meetings helps to make monitoring effective.
- The measurements made should be important to customers, practical for operators and comparable between operators. They should concentrate on few aspects of services.
- The measurements published should be accessible to customers, helpful to customers and fair to operators. They should be made available in ways appropriate to the culture of the intended users.
- Any targets set should be useful to customers and realistic for operators. They are most likely to be desirable for wholesale services (and often retail services) of dominant operators.
- Monitoring should entail regularly examining, and understanding the basis for, the measurements.

Some of these guidelines may seem too obvious to discuss, but they are not followed everywhere.

1 The motivation for quality of service monitoring

This section examines the possible aims of quality of service monitoring by regulators and makes inferences about the conditions in which quality of service measurements and quality of service targets are beneficial. It includes discussions of the activities and concepts in quality of service monitoring.

1.1 Aims

The ITU World Telecommunication Regulatory Database indicates that in 2005 regulators had responsibilities for quality of service monitoring in about 84% of countries and for quality of service standards in about 76% of countries[1]. Thus quality of service monitoring could have important effects on the development of telecommunications worldwide.

Quality of service monitoring requires that measurements be made and, sometimes, that targets be attained. As operators often make measurements for themselves already, action by regulators could be unnecessary. However, regulators might monitor quality of service for several reasons. Some are more important when competition is strong, and others are more important when competition is weak. Quality of service monitoring by regulators can have the following aims:

- Helping customers to make informed choices.
- Checking claims by operators.
- Understanding the state of the market.
- Maintaining or improving quality in the presence of competition.
- Maintaining or improving quality in the absence of competition.
- Helping operators to achieve fair competition.
- Making interconnected networks work well together.

These aims are discussed individually below.

1.1.1 Helping customers to make informed choices

The price is an important factor in choosing a service. However, it is not the only one: customers may care about price more than quality, but once they have settled on the price they want the best quality available at that price. In fact quality can be more important than price. Thus in Colombia a survey showed that the most important factor in choosing a mobile telephony operator was coverage for 28% of customers, quality of service for 24% of customers and promotions for 21% of customers[2].

Business customers often care more about quality than price, because problems with quality are more likely to be costly. For example, in the United Kingdom (UK) 59% of small and medium businesses rated the offer of “the best quality and reliability”, not “the cheapest deal possible”, as the most important factor when choosing an operator[3]. Even 55% of residential customers of the main incumbent operator valued quality and reliability above price[4].

Different quality levels can coexist at different prices: in India each long distance operator can have two call-by-call selection codes representing different quality levels, in Australia each fixed telephony operator can offer services without the customer service guarantee as well as ones with it, and in many countries different internet cafes or different calling cards have different quality levels. In such cases customers need to know about quality levels but not to have them constrained.

For services that are bundled together, with one price covering several services, the quality of specific aspects of the services can influence choices greatly: customers are likely to choose bundles that have the best versions of those services that are most important to them.

Measurements that are published can help customers to choose well. Targets that are the same for all operators, however, may limit choices of quality and price.

1.1.2 Checking claims by operators

Operators may make claims in advertisements about their services or the services of their competitors.

An example occurred in the UK, when two mobile telephony operators complained to the advertising regulator that a third mobile telephony operator advertised "the best call success rate ever". The advertisements were based on measurements that were made using a method adopted jointly by the three operators. Earlier measurements that were made using this method had been published. Most of the complaints were upheld, because the third operator had ignored unpublished but audited measurements for the other two operators, had implied that the telecommunications regulator had published the measurements, and had identified the call success rate with network performance[5]. (They also included a discussion of whether call success rates of 99.2% and 99.3% were different for customers in practice.)

Measurements that are published can be used in checking claims by operators about services.

1.1.3 Understanding the state of the market

Figures about rollout may not be enough to show how well policies are succeeding. (For example, numbers of base stations might reach required levels but the base stations might not be maintained well enough to operate properly.) They are therefore of interest to policy makers. They are also of interest to operators, who wish to understand strengths and weaknesses of competitors.

Measurements can show gaps in services that could be filled by market entry or that need new policies for particular groups of people, geographic areas or operating conditions (such as emergencies).

1.1.4 Maintaining or improving quality in the presence of competition

Competition can improve quality. However, cutting prices is often the preferred way of competing, partly because it can be done much faster than improving quality. When cutting prices involves cutting costs, competition can even reduce quality. Quality reductions due to cuts in costs can be difficult to reverse, as new staff may need to be trained or deferred investments may need to be brought forward.

Even in fully competitive markets quality may be poor, because of ferocious competition, rapid expansion or overstretched management. Thus in various countries both customer support and network equipment can become overloaded in aggressive campaigns for internet access customers, for example. Similarly, during rapid expansion of the mobile networks in India quality fell well short of the targets[6].

Measurements that are published, perhaps with targets, can oppose forces that can tend to lower quality levels in some competitive environments.

1.1.5 Maintaining or improving quality in the absence of competition

A dominant operator may be subject to price controls, unwilling to boost supply or insensitive to customer wishes. In such circumstances the operator might try to maintain margins by reducing the quality of a retail service. (For example, public payphone faults and directory enquiry call failures might rise because of cuts in staff.) The operator might even introduce a service at a higher quality level and lower the quality level of the "standard" service to encourage customers to move to the better service[7].

In addition, a dominant operator might suggest that to protect customers any quality of service obligations imposed on it should be imposed on other operators also. Doing this might lead to uniform quality levels but would limit the speed and flexibility with which other operators could offer services.

Measurements that are published, perhaps with targets, for the retail services of dominant operators can counter the adverse effects of other forces on dominant operators. Targets that are the same for all operators, however, may act as barriers to entry.

1.1.6 Helping operators to achieve fair competition

The business of a competitive operator might depend crucially on the facilities of a dominant operator, by interconnecting with, or just reselling, them. The dominant operator is then providing a wholesale service to operators (not a retail service to end users). For retail competition to be fair, a competitive operator buying a wholesale service from a dominant operator should get the same quality level from that service as the retail arm of the dominant operator does[8].

Targets that represent required minimum quality levels for the wholesale services of dominant operators can achieve this (perhaps in conjunction with other actions, such as separating the managements of the retail and wholesale arms).

1.1.7 Making interconnected networks work well together

An end-to-end path may pass through several networks, each controlled by a different operator. No individual operator has responsibility for the end-to-end path, but a dominant operator may have a serious effect on end-to-end paths set up by other operators. (The effect of any other operator on end-to-end paths is on average much less serious.)

Targets that determine required minimum quality levels nationally for end-to-end paths can ensure that one operator (and, in particular, a dominant operator) does not degrade excessively the quality of end-to-end paths set up by others. Of course, targets for measurements of customer support that are not dependent on network equipment are not relevant to the quality of end-to-end paths.

1.2 Ways of achieving the aims

The aims in Section 1.1 mention various activities occurring in quality of service monitoring. Those that contribute directly to achieving the aims of quality of service monitoring are making measurements, publishing measurements and setting targets. Table 1 indicates approximately which activities contribute to which aims.

Table 1 Direct contributions of quality of service monitoring to achieving the aims

activity	monitored operators	aim						
		helping customers to make informed choices	checking claims by operators	under-standing the state of the market	maintaining or improving quality in the presence of competition	maintaining or improving quality in the absence of competition	helping operators to achieve fair competition	making inter-connected networks work well together
making measurements	dominant operators only							
	all operators			+				
publishing measurements	dominant operators only					+		
	all operators	+	+	+	+			
setting targets	dominant operators only					+	+	+
	all operators				+			

Note: A '+' occurring in an entry in Table 1 indicates that the activity contributes to achieving the corresponding aim in Section 1.1. Of course, making measurements is assumed to be required by publishing measurements and setting targets, but making measurements is not awarded '+' in Table 1 just because publishing measurements or setting targets is awarded '+'.
Source: Robert Milne (Antelope Consulting).

The aims in Section 1.1 do not always point to exactly the same conclusions. For instance:

- Different aims can suggest different emphases. In particular, helping customers to make informed choices is important mainly when choices are inconvenient or costly to change. This can happen if customer equipment is tied to a particular operator, contracts last for years or telephone numbers are not portable[9]. It therefore occurs with lifetime subscriptions as introduced in India and the United States (US)[10]. Sometimes, however, customer choices are not inconvenient or costly to change and do not by themselves justify quality of service monitoring. This is so for call-by-call selections of alternative operators without long term commitments.
- Different aims can suggest different measurements. Thus using a single indicator of customer satisfaction for each individual operator, instead of multiple measurements, might help potential customers but not other operators or policy makers[11].
- Different aims can suggest different services for monitoring: For example, helping operators to achieve fair competition might involve imposing targets on wholesale services but not on retail services.
- Different aims can suggest different treatments of operators: resellers might not be obliged to make measurements for the aspects of services that they passed to customers unchanged but might be obliged to publish the measurements for those aspects made by the original operators, just to inform their own customers.

Quality of service monitoring can become complex. It is likely to be made cheaper, easier and more effective by concentrating on particular aims and simple schemes. Doing this involves deciding on the monitoring to be done (making measurements, publishing measurements or setting targets) and on the operators to be monitored (dominant operators only or all operators). Table 2 offers guidance on making these decisions, by summarising roughly how different choices benefit various stakeholders.

Table 2 Direct benefits to stakeholders of quality of service monitoring

monitoring	monitored operators	stakeholders		
		policy makers	customers	other operators
making measurements	dominant operators only			
	all operators	+		
publishing measurements	dominant operators only		+	
	all operators		+++	++
setting targets	dominant operators only	+	+	++
	all operators		+	

Note: A '+' occurring in an entry in Table 2 indicates that the activity benefits the corresponding stakeholders, with the number of occurrences of '+' being loosely the number of aims in Section 1.1 in which the stakeholder benefits from the activity. Of course, making measurements is assumed to be required by publishing measurements and setting targets, but making measurements is not awarded '+' in Table 2 just because publishing measurements or setting targets is awarded '+'.
Source: Robert Milne (Antelope Consulting).

Table 2 can itself be reduced to even simpler guidelines. In brief:

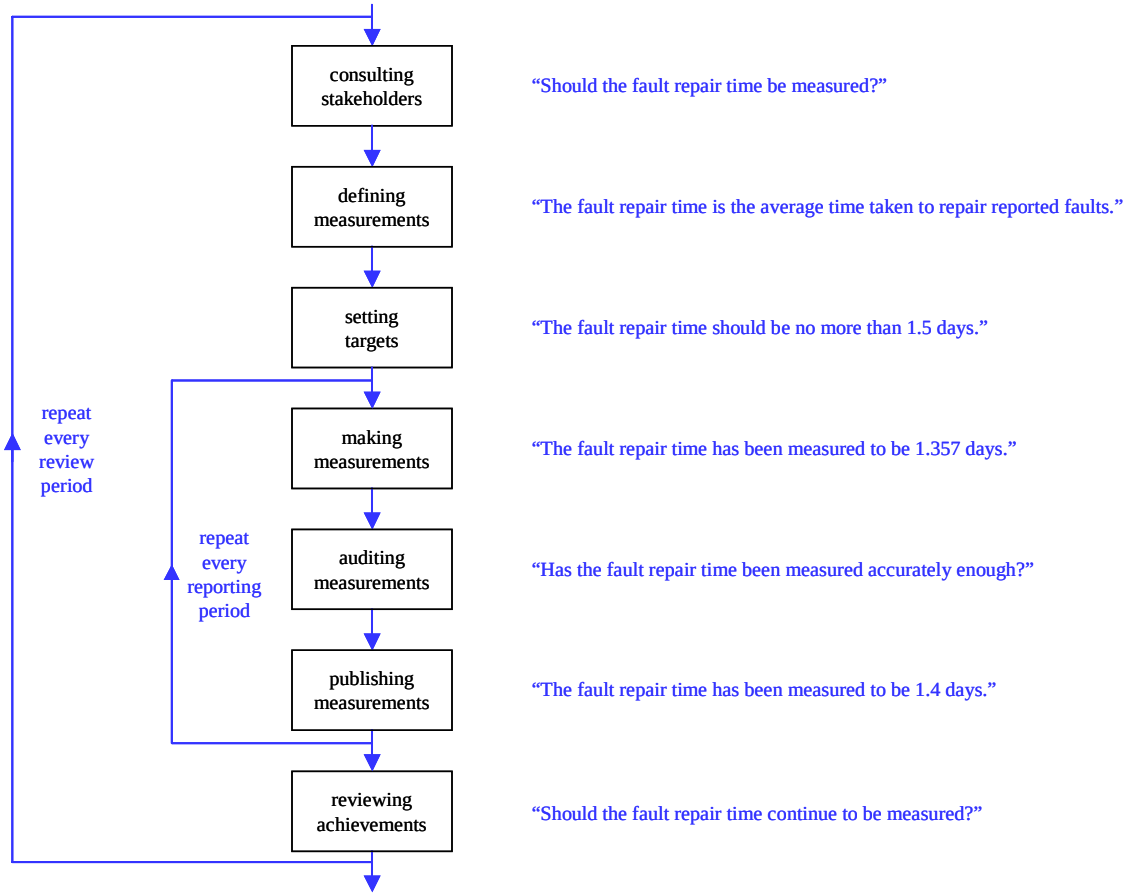
- Making measurements for all operators could benefit policy makers.
- Publishing measurements for all operators could benefit customers.
- Setting targets for dominant operators could benefit other operators.

These guidelines do not mention universal service obligations, which often include minimum quality levels. In reality universal service providers (the operators having universal service obligations) are dominant operators, either through being major forces in the national market or through serving areas where there are few alternatives. In this paper universal service providers are regarded as dominant operators.

1.3 Activities

Figure 1 illustrates a complete process for quality of service monitoring by regulators, in which the activities mentioned in Section 1.2 can recur several times. (The flow chart shows how activities succeed each other and are repeated periodically and gives alongside each activity an example of the questions or statements that form the subject of the activity.)

Figure 1 Activities during quality of service monitoring



Note: The repeat loops in Figure 1 are simplifications, as auditing, for example, may reveal defects in measurement definitions that can be corrected without formal review and consultation processes. The numbers used as examples in Figure 1 are not intended to represent best practice.
Source: Robert Milne (Antelope Consulting).

1.4 Technicalities

There is not one meaning of 'quality of service' that is universally adopted. However, the one chosen for this paper is widely adopted. Box 1 says what it is and identifies various related terms that are not used in this paper.

Box 1 The meaning of quality of service

In ITU-T recommendation E.800, quality of service is "the collective effect of service performances, which determine the degree of satisfaction of a user of the service"[12]. This definition is adopted in this paper. Quality of service therefore concerns aspects of services that users experience directly. It can be contrasted with network performance, which, again according to E.800, is "the ability of a network portion to provide the functions related to communication between users"[12].

Different terminology is used in some documents from ITU and other organisations. In particular:

- The term 'quality of service' may refer to just quality determined by network performance or even just to a way of getting different levels of IP network performance by classifying and controlling traffic. The term 'quality of experience' is then sometimes be used for what is regarded as quality of service in this paper.
- The term 'type of service' was formerly used in IP networks but has now been superseded by other terms related to classifying and controlling traffic.
- The term 'class of service' may describe a level of network performance appropriate to particular services in IP and other data networks.
- The term 'grade of service' has been used for many years. In E.600, grade of service is "a number of traffic engineering variables used to provide a measure of adequacy of a group of resources under specified conditions"[13]. It often refers to a value characterising the quality of one aspect of a service (such as the proportion of successful calls).

There is no firm boundary around what constitutes quality of service: jitter and wander for leased lines could be measured in quality of service monitoring if they were important to users. In practice, however, quality of service monitoring is applied to rather few services.

For the purposed of this paper, when quality of service is monitored there must at least be measurements that are made according to defined measurement methods. These measurements allow discussions of quality of service to be quantitative and to recognise that different aspects of a service may reach different quality levels (as, for example, faults may be repaired fast but the proportion of successful calls may be low).

As indicated in Section 1.2, measurements are not just: some may be published and some may have targets. Box 2 describes how measurements and targets are related.

Box 2 The nature of measurements and targets

Quality of service monitoring requires measurements that characterise the quality of aspects of services. The definitions of the measurements to be made are called 'parameters', 'indicators', 'metrics', 'measures' or 'determinants' in different documents. The term 'parameter' is generally used in standards but is avoided in this paper for simplicity. The term 'indicator' is often used for measures like "the number of main telephone lines per 100 inhabitants" that are outside the scope of this paper.

A target is a potential value (or range of values) for a measurement that must be attained if quality is to be regarded as satisfactory. Targets are called 'objectives', 'benchmarks', 'thresholds', 'standards' or 'reference values' in different documents. In addition, all these terms have other meanings. The term 'target' is sometimes confined to desired or planned quality levels, typically in successive years. Then the term 'threshold' might refer to quality levels that are required immediately, but when used in that way it has the disadvantage that for many measurements high values are bad and low values are good (so 'lintel' would be more appropriate than 'threshold').

Defining measurements involves providing measurement descriptions and measurement methods. A measurement description says what quality is characterised by the measurements. Often a shortened form of the measurement description is used as the measurement name, but the full measurement description is needed for showing when measurements are comparable. Thus the measurement descriptions "the proportion of faults repaired within 2 days" and "the proportion of faults repaired within a time identified by the operator" evidently refer to measurements that are not usually comparable with one another, but the measurement name "the fault repair timeliness" might apply to either of them.

Making measurements involves applying measurement methods. Some measurements may be singled out as publishable measurements, which characterise the quality in a way that is helpful enough to publish. Other measurements may be useful for other purposes, such as analysing trends.

Measurements may have targets. For instance:

- The fault repair time might have "the average time taken to repair reported faults" as a publishable measurement description, with "1.5 days" as a target for it.
- The call setup success ratio might have "the percentage of calls successfully set up after receiving dial tone" as a publishable measurement description, with "97%" as a target for it.
- The packet transmission time might have "the average time taken to transmit 56-byte packets between representative sources and destinations" as a publishable measurement description, with "40 milliseconds" as a target for it.
- The packet loss ratio might have "the percentage of 56-byte packets lost during transmission between representative sources and destinations" as a publishable measurement description, with "1.0%" as a target for it

The measurements made vary widely between regulators. However, they tend to reflect similar concerns and therefore to have related, but not identical, measurement methods. Box 3 illustrates this, by giving names to most of the measurements that are mentioned in this paper. Measurement names are not sufficient for defining measurements in regulations: the call completion ratio, for example, can refer to measurements defined in several different ways (Some of which, incidentally, have little to do with quality of service from operators and much to do with how frequently people answer phones.) However, measurement names are useful in identifying the general intentions of measurements.

Box 3 Conventions for measurement names

Measurement names should be regarded as short explanations of the measurements, not as full measurement descriptions. In particular, the measurement names in this paper may be short for different measurement descriptions in different examples and may differ from the names used by regulators. (For example, what exactly constitutes a fault report varies between regulators.)

Many measurements fit one of the following patterns and, in this paper, have corresponding uniform names:

- The measurements can be times taken to do actions. In this paper they have names of the form "the x time", which signifies the average time taken to do the action "x", from initiating it (by making a request or pressing a button, for example) to completing it. Thus "the service supply time" signifies the average time to supply services, "the fault repair time" signifies the average time to repair faults, "the complaint resolution time" signifies the average time to resolve complaints, "the call setup time" signifies the average time to setup calls, and "the packet transmission time" signifies the average time to transmit packets. (Times other than averages are also sometimes useful, as in "the maximum time taken to repair 95% of faults ", for example.)
- The measurements can be proportions of occurrences of actions done in time. In this paper they have names of the form "the x timeliness", which signifies the result of dividing the number of occurrences of the action "x" that are done on time by the total number of occurrences of the action "x". Thus "the service supply timeliness" signifies the proportion of occasions in which services are supplied on time, "the fault repair timeliness" signifies the proportion of occasions in which faults are repaired on time, and "the complaint resolution timeliness" signifies the proportion of occasions in which complaints are resolved on time. (The term 'timeliness' is specific to this paper, as there is no term in widespread use.)
- The measurements can be frequencies of occurrences of events per customer. In this paper they have names of the form "the x rate", which signifies the result of dividing the number of occurrences of the event "x" by the total number of customers. Thus "the fault report rate" signifies the number of fault reports per customer and "the complaint rate" signifies the number of complaints per customer.
- The measurements can be proportions of occurrences of events having a specific property. In this paper they have names of the form "the x y ratio", which signifies the result of dividing the number of occurrences of the event "x" that have the property "y" by the total number of occurrences of the event "x". Thus "the bill error ratio" signifies the proportion of bills that are erroneous, "the call setup success ratio" signifies the proportion of call setups that are successful, "the call drop ratio" signifies the proportion of calls that are dropped, "the packet loss ratio" signifies the proportion of packets that are lost and "the bit error ratio" signifies the proportion of bits that are erroneous. (The term 'ratio' is often replaced by the term 'rate', bur a rate often implies dependence on time, money or some other quantity, such as the number of customers: a bit error rate and a bit error ratio, for example, could be different.)

2 The operational arrangements for quality of service monitoring

This section considers how quality of service monitoring by regulators operates, particularly with the involvement of operators and customers. Operators can be involved in defining measurements, setting targets, making measurements, auditing measurements and publishing measurements. Customers can be involved, too, in defining measurements and making measurements. The section also deals with techniques for encouraging and enforcing compliance with quality of service obligations.

2.1 Defining measurements

Choices of measurements depend on what customers in the country are most concerned about. Choices of targets depend on what operators in the country can hope to achieve. (How measurements and targets can be chosen is discussed further in Section 3.1 and Section 3.2.) The choices can have unforeseen side effects, as they can focus management attention on some aspects of services to the possible detriment of others. Getting the balance right needs an understanding of internal company dynamics and potential industry performance that many regulators do not expect to achieve. The knowledge and experience of operators can help with this.

Regulators are often unwilling to work closely with dominant operators. However, they may be more willing to work broadly with all operators to define measurements. Also, they may welcome activities by other monitors using magazines and websites that define and make measurements or conduct customer satisfaction surveys. (These "other monitors" form a miscellaneous class, including consumer associations, pressure groups, internet service reviewers and monitoring system vendors.) In general:

- Letting dominant operators continue to use measurement methods that they have used for some time may be expedient if no other operators have to make measurements[14].
- Letting all operators define their own measurements usually leads to measurements that are not are comparable between different operators.
- Encouraging other monitors to define and make measurements assists with monitoring the aspects of services that concern customers. Of course, other monitors might not exist or might not monitor aspects of services mainly of concern to operators and policy makers. They might also make measurements in ways that have low costs but do not represent customer experience adequately.
- Involving an industry group of operators in the definition of measurements uses the skills of operators[15]. The group may take a long time to agree and implement the definitions, partly because individual operators wish to minimise their own disruption and expense and partly because it may need to examine measurement procedures in detail. However, the results may be acceptable to everyone, especially if the group includes customer representatives and external experts. The regulator may still need to check independently that the measurements defined are comparable between operators inside the group. (They are usually not comparable with measurements defined by operators outside the group.)

Examples of regulators making extensive use of industry groups are described in Section 2.5.

2.2 Setting targets

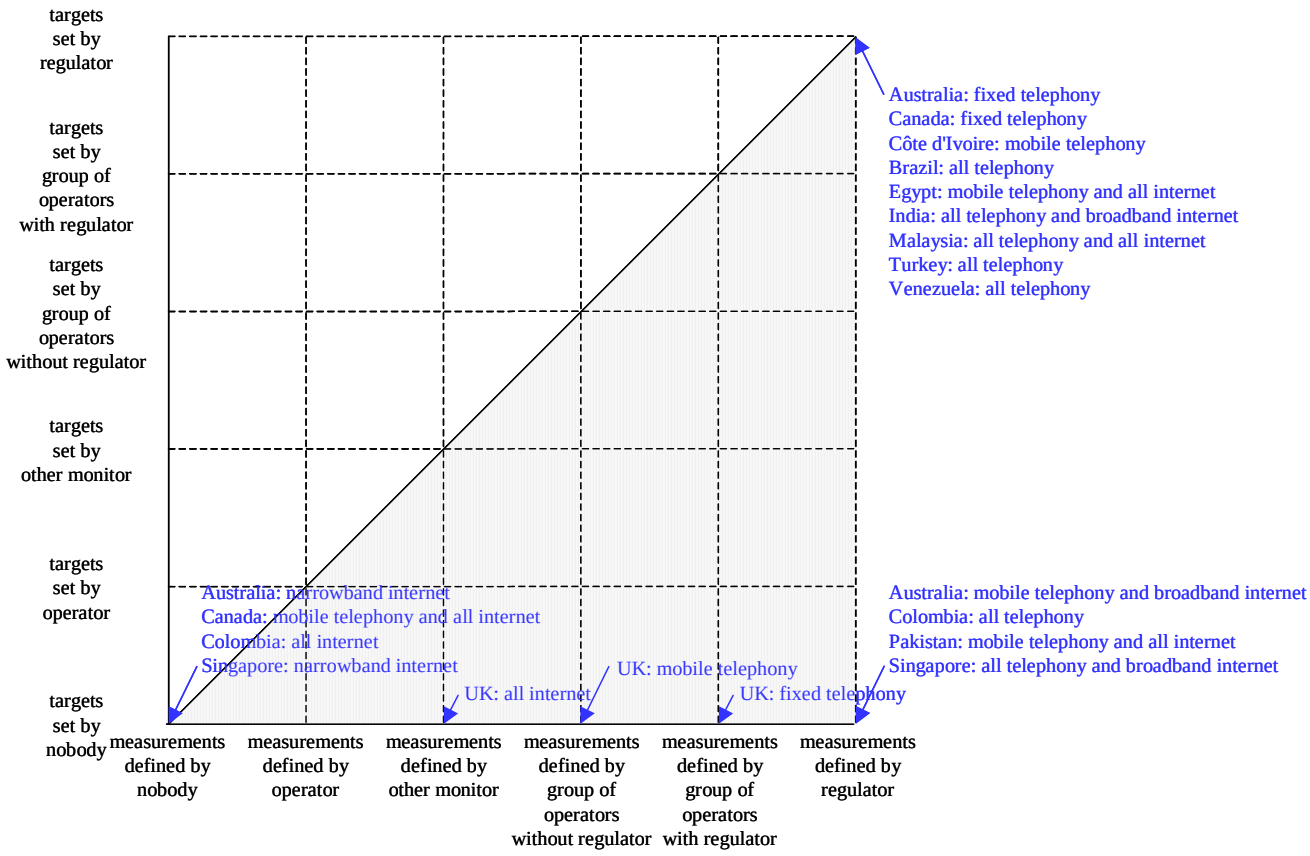
When quality of service is monitored, targets may be set. In principle target setting could be the responsibility of various organisations, just as with measurement definition in Section 2.1. However, in practice, target setting is done by regulators if it is done at all. Yet there are potential applications of target setting by other organisations. In particular, if operators set their own targets they could offer varied choices of quality and price but be obliged to meet commitments: in consultation with regulators they operators would devise, publicise and commit to targets and publish indications of whether these targets were attained.

In brief:

- Letting dominant operators set their own targets is unlikely to be effective if no other operators have to make measurements, as there will be no ways of comparing quality
- Letting all operators set their own targets could broaden choices of quality and price while requiring operators to fulfil their commitments. However, this does not seem to happen, perhaps because regulators that are more likely to let operators set targets are also less likely to require targets.
- Encouraging other monitors (such as pressure groups and monitoring systems vendors) to set targets is also unusual, possibly because other monitors usually deal only with internet services, for which expectations are less well developed than for telephony services.
- Involving an industry group of operators in the setting of targets uses the skills of operators. It may limit the choices available to customers, just as would setting targets without the involvement of the group. Setting targets needs less detail than defining measurements, so an industry group is less useful for it. In its case the knowledge and experience of operators can be taken into account by due consultation and prior monitoring to determine realistic quality levels.

Figure 2 shows who defines measurements and sets targets in various countries. Different "owners" of the definitions and settings may be suited to different weightings of the aims in Section 1.1: regulators may need to "own" more if they monitor just dominant operators than if they monitor all operators.

Figure 2 Examples of responsibilities for defining measurements and setting targets



Note: The hatched area in Figure 2 indicates possible responsibilities, as no targets can be set if there are no measurements defined.
Source: Robert Milne (Antelope Consulting).

Measurements often have targets associated with them, in the expectation that targets will help to improve quality. However, targets (and the associated measures for ensuring compliance) are not always effective in improving quality. Box 4 describes one such case involving the regulator in India, the Telecom Regulatory Authority of India (TRAI). Targets were set after networks had been monitored for 2 years. They were intended to lead to improvements in successive years. However, the fixed networks could not be modernised fast enough to achieve the targets for fault report rates, which had therefore to be deferred for 3 years. Moreover, the mobile networks achieved their targets in just 75% of situations. (Box 4 also illustrates how measurements may need to change as the market develops, as discussed in Section 3.1.1, and how measurements must be comparable between operators, as discussed in Section 3.1.3.)

Box 4 Monitoring by the regulator in India

Fixed telephony

TRAI published for every 3 months from April 2003 to June 2005 17 measurements with targets for up to 75 fixed telephony operators in 27 geographic areas (treating different operating areas as different operators). Participation in the scheme was compulsory. The targets, set in July 2001, were different for successive years.

TRAI revised the scheme after a review, for the following reasons[16]:

- Quality had remained unsatisfactory (with measurements for fixed wireline networks lagging some years behind the targets), though the targets had been set after 2 years of measurement and consultation.
- Some measurements (such as the delay before getting dial tone) were still required but no longer important and others (such as the delay before refunding deposits) were important but not yet required.
- Fixed wireless networks were more like mobile networks than fixed wireline networks, which probably had more access network faults and fewer dropped calls.

TRAI has published for every 3 months since July 2005 14 measurements with targets for 71 fixed telephony operators in 27 geographic areas[17]. Participation in the scheme is compulsory[18]. The targets are generally the same for successive years. The measurements include the service supply timeliness (separately for initial and various changed services), the fault repair timeliness (separately for various periods of time), the fault repair time, the fault report rate, the bill error ratio, the complaint resolution timeliness, the proportion of rapidly answered attendant calls and the call completion ratio. Not all of the measurements are published: for fixed wireless networks, the measurements published are like those published for fixed wireline networks but the measurements required are like those required for mobile networks.

Mobile telephony

TRAI published for every 3 months from April 2003 to June 2005 10 measurements with targets for up to 122 mobile telephony operators in 23 geographic areas (treating different operating areas as different operators). Participation in the scheme was compulsory. The targets, set in July 2001, were different for successive years.

TRAI revised the scheme after a review, for the following reasons[16].

- Quality had remained unsatisfactory (with measurements failing to reach the targets in 25% of cases), though the targets had been set after 2 years of measurement and consultation.
- Some measurements (such as the fault repair time) were still required but no longer important and others (such as proportions of calls blocked by wireless channel congestion) were important but not yet required.
- Some measurements (such as the proportion of successful calls) were not comparable between operators, as they depended on the operators in ways that were not described precisely enough.

TRAI has published for every 3 months since July 2005 14 measurements with targets for 128 mobile telephony operators in 23 geographic areas[17]. Participation in the scheme is compulsory[18]. The targets are generally the same for successive years. The measurements include the bill correction timeliness, the bill error ratio, the complaint resolution timeliness, the proportion of rapidly answered attendant calls (separately for various periods of time), service coverage (separately for buildings, cars and streets), service availability, the call setup success ratio, the call drop ratio and the conversational voice quality.

2.3 Making measurements

Regulators impose quality of service obligations for telephony services much more often than for internet services. For internet services many rely on customer opinion surveys and on other monitors using magazines and websites. These other monitors may even involve users in making measurements, by distributing test tools and collecting test results. Occasionally these arrangements are made more formal. For example, in Australia the regulator publishes an annual report about tests made by internet users and in Chile the regulator required internet access operators to provide measurement tools to their customers. The future may see much more widespread use of such arrangements: user terminals besides computers are becoming ever more advanced and in principle could report back to operators ample end-to-end quality of service measurements.

Box 5 describes the scheme of the regulator in Chile, the Subsecretaría de Telecomunicaciones (Subtel). In this scheme internet operators make some measurements but also distribute tools so that their customers can make measurements. This scheme is unusual in various other ways, as it does not apply to telephony services but it applies to both access and transit internet services and it obliges operators to publish measurements on their own websites.

Box 5 Monitoring by the regulator in Chile

Internet access

Operators published from February 2002 to December 2004 2 measurements of narrowband (dialup) internet access services[19]. The measurements were expressed as weekly, monthly, quarterly and annual trend graphs on the websites of the individual operators. These websites also provided software for download so that narrowband and broadband internet access users could make other measurements. Participation in the scheme was compulsory[20]. The measurement methods, contents and formats and the software for download were agreed by an industry group working with Subtel. The measurements made by the operators determined the internet session login success ratio and the model pool utilisation. The measurements made by the users determined the internet session login time and the data transfer rates for services including file transfer, web browsing and email.

Subtel suspended the compulsory scheme pending a review, in order to update the telecommunications legislation and introduce quality of service monitoring for new services, such as those using the internet. The internet access operators are no longer publishing measurements.

Internet transit

Operators have published since January 2001 3 measurements of internet transit services separately for each connected internet access operator and each other internet transit operator[19]. The measurements are presented as weekly, monthly, quarterly and annual trend graphs on the websites of the individual operators. Participation in the scheme was compulsory[21]. The measurement methods, contents and formats were agreed by an industry group working with Subtel. The measurements include the packet transmission time (delay), the packet loss and the occupancy of the link with each connected internet access operator and each other internet transit operator.

Subtel suspended the compulsory scheme pending a review, in order to update the telecommunications legislation and introduce quality of service monitoring for new services, such as those using the internet. The internet transit operators are nonetheless still publishing measurements.

Some measurements (such as fault repair times) are best made by operators. Others are best made by external measurement agencies, because doing so makes them more comparable between different operators and, sometimes, less costly. In particular, drive-around tests are more comparable and less costly if they are done at similar times, between similar places and in similar circumstances for all mobile telephony operators. Co-operation between the operators in defining and making such measurements can be valuable in identifying occasional anomalies due to faulty measurement equipment or incorrect applications of measurement procedures.

Operators and regulators often favour the use of customer opinion surveys. Surveys intended for use in quality of service monitoring by regulators could be the responsibility of operators, other monitors or regulators. Savings might be made by obliging operators to include questions due to regulators in their own surveys. However, the surveys might then need to include many questions, because of differences in the information wanted by the operators and the regulators. For example, the operators might want to assess interest in possible new services and the regulators might want to ask about unsolicited sales calls, public payphone faults, directory enquiry responses, and emergency call responses. To regulators customer opinion surveys may be most valuable when customer support from operators is inadequate (so customers do not bother to complain or report faults) and therefore when surveys by operators are likely to be unsatisfactory.

In large countries having several geographic areas, many measurements are likely to be made. Thus in India there can be at least 774 measurements about fixed operators and 1792 measurements about mobile operators every 3 months. Of course the operators may have correspondingly many customers, so the costs of the measurements per customer may be low. In fact the costs of measurements may bear most heavily on small operators, because the number of test calls, for example, needed to make precise enough measurements in a given geographic area is independent of the size of the operator.

Calculating and allocating the costs of measurements can be contentious[22]. Measurements can raise the costs of all operators, including those that already make similar measurements. In principle the measurements made could characterise the quality of many aspects of services for different market segments in different geographic areas. However, the costs are related to the number and variety of the measurements made and can be reduced by concentrating on few particularly useful measurements. Reviewing the choice of measurements from time to time helps the measurements to remain those important to customers and practical for operators, as discussed in Section 3.1.1 and Section 3.1.2.

Where measurements are made by an operator or are audited for an operator the costs are most obviously carried by the individual operator. In other cases the costs may be recovered from licence fees by the regulators or from all the operators in an industry group sharing an external measurement agency, perhaps in proportion to subscriber numbers, coverage or revenue.

2.4 Auditing measurements

Measurements made by external agencies can be submitted to the regulator or even published before the operators see them (as happens for drive-around tests in the UK). Measurements that are seen by the operators before they are submitted to the regulator should be signed off by senior employees in a ‘self-certification’ process. They are therefore probably audited for accuracy by or for the operators.

Nonetheless regulators may want independent auditing of measurements. (For example, in Malaysia the regulator audits all the dominant operators, and other selected operators, every year, even though the law permits jail sentences for failures to fulfil quality of service obligations.) However, independent auditing of all measurements can be inconvenient and costly and can delay publication greatly. To increase confidence in the measurements but limit costs regulators might decide to audit just some measurements (or data used in calculating some measurements), selected at random or through being suspect. Regulators can also check measurements in other ways, such as comparing them with the complaints received by the regulators or performing “mystery shopping” to test complaint handling. All these ways need careful design, as audit procedures can easily become too lax or too burdensome. The possibility of independent auditing can itself be quite effective in ensuring the accuracy of measurements: quality of service measurements may not produce major newspaper articles, but legal proceedings for falsifying them will do so.

The definitions of measurements themselves, as well as the subsequent measurements, require auditing both before the definitions are implemented and thereafter: though the definitions may be intended to be precise enough to achieve comparability, they may remain open to slightly different interpretations. If the auditors for different operators and regulators meet together, they can find and resolve minor problems of comparability, agree the main parts of each measurement procedure needing investigations in audits, and validate the auditing techniques of each other.

Box 6 shows that the regulator in Ireland, the Commission for communications Regulation (ComReg), and its predecessor, the Office of the Director of Telecommunications Regulation (ODTR), decided that fairly simple quality of service monitoring was not helpful enough to continue. The consequences of the audits (rejected measurements, delayed publications and extra costs) were factors in reaching this decision. The decision contrasts with that in the UK described in Section 2.5, perhaps because the influence of quality on customer choices was thought to differ in the two countries.

Box 6 Monitoring by the regulator in Ireland

Fixed telephony

ODTR and subsequently ComReg published for every 6 months from January 2001 to December 2003 3 measurements for up to 9 fixed telephony operators of residential or business services (treated separately)[23]. Participation in the scheme was voluntary (except for the universal service provider). Operators did not always need to make all of the measurements. Mergers and withdrawals from participation left only 6 operators late in the life of the scheme. The measurements included the service supply timeliness, the fault repair timeliness and the complaint resolution timeliness.

ComReg suspended the voluntary scheme after a review, for the following reasons:

- Lack of comprehensive coverage reduced the usefulness of the scheme to customers.
- Measurements were often rejected after audits for not being comparable with others.
- Costs were disproportionate to the margins of small operators.
- Choices of operators by residential customers (but not necessarily business customers) were based mainly on pricing.

Comreg is developing new quality of service obligations that will apply to the universal service provider.

Mobile telephony

ODTR and subsequently ComReg have not imposed quality of service obligations on mobile telephony operators.

2.5 Publishing measurements

Publishing measurements is central to achieving some of the aims described in Section 1.1. In particular, it is essential if customers are to base choices of services on general conclusions, not just personal anecdotes.

Publishing measurements usually involves distributing written words through websites. However, written words may not be the most suitable for the general public, at least in rural and remote districts or in oral societies where spoken words predominate. Publishing some measurements through radio broadcasts or freephone calls could reach more people. Of course journalists and other intermediaries might still prefer written words that they could summarise for the general public. Where written words are suitable there are other ways of providing them besides using websites or even newspapers. In particular, quality of service deficiencies could be publicised in SMS messages and bills.

In the countries discussed in this paper, publishing measurements depends on written words, often in tables of numbers. However, tables of numbers are not always used exclusively: bar charts are used in Singapore, trend graphs are used in Chile, and bar charts and coverage maps are used in the UK, for example.

Table 3 gives a simple example of a table of numbers for a particular measurement period and geographical area, displaying three measurements (the fault repair time, the packet transmission time and the packet loss) for four internet services from three operators (X, Y and Z). The table is designed according to some of the criteria listed in Section 3.3 so that it can be accessible to customers, helpful to customers and fair to operators: the measurements are presented for different operators in the same table, published without extra irrelevant numbers, rounded to at most two figures, expressed

consistently (in this case with higher numbers always meaning worse quality), separated for different services from the same operator, and accompanied by explanations of unusual quality levels.

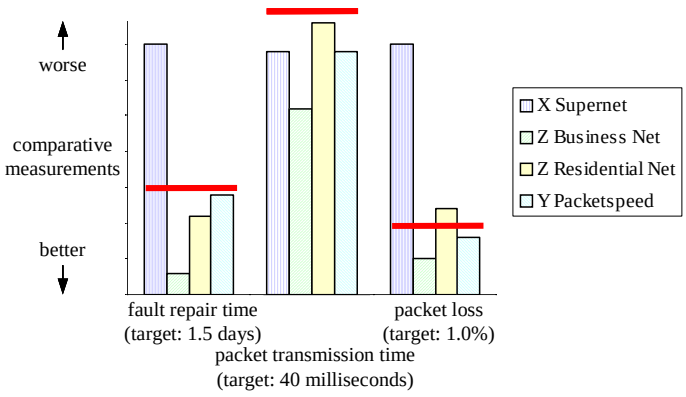
Table 3 Examples of publishable measurements arranged in a table

service name	fault repair time (target: 1.5 days)	packet transmission time (target: 40 milliseconds)	packet loss (target: 1.0%)	explanatory remark
X Supernet	3.5 days	34 milliseconds	3.5%	The quality was affected by slow fault repair by the backhaul operator.
Z Business Net	0.3 days	26 milliseconds	0.5%	The measurements were made when the service had very few customers.
Z Residential Net	1.1 days	38 milliseconds	1.2%	The measurements were made for this area jointly with others, not for this area separately.
Y Packetspeed	1.4 days	34 milliseconds	0.8%	

Note: The coloured numbers in Table 3 exceed the targets. The numbers and the uses of targets for internet services in Table 3 are not intended to represent best practice.
Source: Robert Milne (Antelope Consulting).

Figure 3 represents much of the same numerical information as a bar chart (with different scales for the three different measurements). The explanations of unusual quality levels (and indeed the numbers themselves) could be added to the bar charts as comments.

Figure 3 Examples of publishable measurements arranged in a bar chart



Note: The horizontal lines in Figure 3 represent the targets. The numbers and the uses of targets for internet services in Figure 3 are not intended to represent best practice.
Source: Robert Milne (Antelope Consulting).

Table 3 represents how measurements might be tabulated when they are intended to be compared by customers (as opposed to journalists, for example). In many countries the tables have other arrangements: in Brazil and Colombia, for example, the measurements for fixed telephony operators are presented for different operators in different tables and do not provide explanatory remarks.

In some countries many measurements from many operators are published at many times. Most of these measurements are unlikely to interest many potential customers. However, they may help to demonstrate that the regulator is fair and open. Box 7 outlines the detailed monthly publishing done by the regulator in Brazil, the Agência nacional de telecomunicações (Anatel). This can result in the publication of at least 1190 measurements about fixed operators and 1034 measurements about mobile operators every month. (In this example, as in the other examples in this paper, the counts of measurements do not include numbers that are published but that are essentially irrelevant, in that they are simply used in deriving the measurements.) In both India and Brazil the numbers of measurements are extremely high partly because measurements are published for many separate geographic areas.

Box 7 Monitoring by the regulator in Brazil

Fixed telephony

Anatel published for every month from July 1998 to December 1999 14 measurements for 35 fixed telephony operators in 34 geographic areas (treating different operating areas as different operators). Anatel then replaced the 14 measurements by 34 measurements, as further detailed measurements were wanted.

Anatel has published for every month since January 2000 34 measurements with targets for 35 fixed telephony operators in 34 geographic areas[24]. The measurements are presented as tables separately for each operator. Participation in the scheme is compulsory[25]. The measurements include the service supply timeliness for address changes, the fault repair timeliness (separately for various periods of time), the fault report rate, the fault repair timeliness for public payphones, the fault report rate for public payphones, the bill error ratio (separately for local, national and international bills), the complaint resolution timeliness, the proportion of rapidly answered attendant calls (separately for day and night), the call setup success ratio (separately for day and night and for local, national and international calls) and the call completion ratio (separately for day and night and for local, national and international calls).

Anatel has also published for every month since March 2005 the number of complaints per 1000 subscribers for each of these operators.

Mobile telephony

Anatel published for every month from January 2000 to February 2004 9 measurements with targets for 35 mobile telephony operators (treating different operating areas as different operators). Anatel then replaced the 9 measurements by 22 measurements, as the mobile telephony operators changed to different technologies and operating areas.

Anatel has published for every month since September 2003 22 measurements with targets for 47 mobile telephony operators[26]. Participation in the scheme is compulsory[27]. The measurements include the fault repair timeliness, the bill error ratio, the complaint rate (separately for general and coverage or congestion complaints), the complaint resolution timeliness, the proportion of rapidly answered attendant calls (separately for morning, evening and night), the call setup success ratio (separately for morning, evening and night), the call drop ratio (separately for morning, evening and night) and the call completion ratio (separately for morning, evening and night).

Anatel has also published for every month since March 2005 the number of complaints per 1000 subscribers for each of these operators.

If only one, dominant, operator has quality of service obligations it may publish its own measurements. However, if several operators are to be compared easily, their measurements are best presented near each other in the same publication. Measurements of multiple operators are published on single websites for every 3 months in Canada and Singapore and for every 12 months in France, for example, as well as in almost all the countries discussed in detail in this paper. (However, only some

of these websites can present measurements of multiple operators together on a single page, and most do not cover all classes of operators.)

Collecting measurements from multiple operators introduces delays and administrative overheads. Slightly less co-ordination is needed if measurements are published by individual operators separately, as is done in Belgium and Spain for telephony services and in Chile for internet services. Doing this may be satisfactory if comparisons are expected to be made only by enthusiasts and experts such as specialist journalists, not by typical members of the general public. However, it still leaves regulators with responsibilities for specifying uniform contents and formats for the measurements of different operators, vetting the measurements and telling customers about the existence of the measurements.

Publishing many measurements from many operators at many times requires skills that several regulators and operators do not have[28]. However, the task can be simplified in various ways. For instance:

- The numbers of measurements can be reduced. Customers are unlikely to be interested in many measurements, especially when the measured differences between operators are small.
- The numbers of operators can be reduced. Operators that have too few customers or too few revenues (or that are not dominant, of course) could be exempted from quality of service obligations.
- The numbers of times of publication can be reduced. In a market that is developing quickly doing this may not be helpful to customers or fair to operators, but in a market that is fairly mature many actions to improve quality will not have noticeable effects for several months.
- The numbers of geographic areas with separate measurements can be reduced. However, in many countries geographic differences (including those between urban and rural districts) can be important.

2.6 Ways of sharing responsibilities

Responsibilities for several of the activities outlined in Section 1.3 can be shared between the regulator, the operators, other monitors and the customers. The factors involved in allocating responsibilities to different stakeholders are discussed in Section 2.1, Section 2.2, Section 2.3, Section 2.4 and Section 2.5. Table 4 summarises the discussions by expressing the advantages as disadvantages of alternative allocations of responsibilities. As with any such summary the statements are simplifications and not substitutes for proper examinations of the factors.

Table 4 Disadvantages of allocations of responsibility in quality of service monitoring

activity	monitored operators	responsible organisation						
		regulator	group of operators with regulator	group of operators without regulator	other monitor	operator	customer	nobody
defining measurements	dominant operators only	does not use skills of operators.	may be slow to conclude.	may not look at all aspects. may be slow to conclude.	may not look at all aspects. does not use skills of operators.	may be self-serving.	is not feasible.	does not achieve any quality aims.
	all operators					does not provide fair comparisons.		may lead to poor quality.
setting targets	dominant operators only		may be slow to conclude. limits service offerings.	may be slow to conclude. limits service offerings.	may not look at all aspects. does not get commitment by operators.	may lead to poor quality.	is not feasible.	may lead to poor quality.
	all operators	limits service offerings.						
making measurements	dominant operators only	does not use findings by operators. is not realistic always.	is not realistic for current telephony. needs audits.	is not realistic for fixed telephony. needs audits.	is not realistic for current telephony. may not test enough.	does not get economies of scale. needs audits.	is not realistic for current telephony. is not represent-ative.	does not achieve any quality aims.
	all operators							may lead to poor quality.
auditing measurements	dominant operators only			needs spot checks by regulator.	is not feasible.	needs spot checks by regulator.	is not feasible.	is not prudent.
	all operators							
publishing measurements	dominant operators only			does not get endorsement by regulator.	does not get endorsement by regulator.	does not provide easy comparisons.	is not feasible.	may lead to poor quality.
	all operators							

Source: Robert Milne (Antelope Consulting).

As illustrated in Section 2.2 responsibilities are allocated differently in different countries. Table 5 provides further examples of allocations.

Table 5 Examples of allocations of responsibility in quality of service monitoring

activity	monitored operators	responsible organisation						
		regulator	group of operators with regulator	group of operators without regulator	other monitor	operator	customer	nobody
defining measurements	dominant operators only	Canada: fixed telephony						Canada: mobile telephony
	all operators	India: all telephony Spain: fixed telephony	Chile (suspended): all internet France: mobile telephony UK: fixed telephony	UK: mobile telephony	UK: all internet			
setting targets	dominant operators only	Canada: fixed telephony Spain: fixed telephony						Canada: mobile telephony
	all operators	India: all telephony	France: mobile telephony				UK: all telephony	
making measurements	dominant operators only					Canada: fixed telephony		Canada: mobile telephony
	all operators	France: mobile telephony India: mobile telephony	UK: fixed telephony	UK: mobile telephony	UK: all internet	India: all telephony Chile (suspended): all internet	Chile (suspended): all internet	
auditing measurements	dominant operators only							Canada: mobile telephony
	all operators	India: all telephony	UK: fixed telephony	UK: mobile telephony				
publishing measurements	dominant operators only	Canada: fixed telephony						Canada: mobile telephony
	all operators	India: all telephony France: mobile telephony	UK: fixed telephony	UK: mobile telephony	UK: all internet	Chile (suspended): all internet Spain: fixed telephony		

Source: Robert Milne (Antelope Consulting).

2.7 Ensuring compliance

Quality of service monitoring by regulators imposes on operators obligations to make measurements and, sometimes, to attain targets. The obligations may be formalised in licences for separate operators (as in Algeria, Morocco and Oman, for example) or in regulations for multiple operators (as in Côte d'Ivoire, Malaysia and the United Arab Emirates, for example).

Quality of service obligations are only likely to be fulfilled if legitimate concerns are addressed. A requirement for successful quality of service monitoring by regulators, therefore, is to take into account the opinions of operators and consumers fully and openly. This can be achieved through a consultation process, for obligations in licences as well as for obligations in regulations.

Typically at the start and end of a consultation process the regulator circulates documents that discuss the policy options and proposals. The process then both brings extra knowledge and experience to bear on complex problems and shares "ownership" of the solutions: in Jordan recently the consultation process resulted in a greatly reduced burden for the regulator as well as the operators, by halving the number of measurements and the frequency of auditing[29].

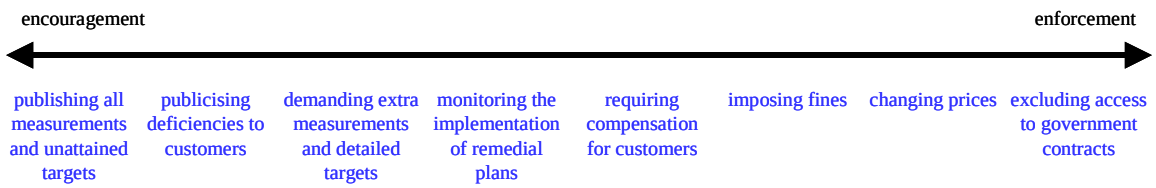
Imposing obligations without due consultation about something as complex as quality of service is likely to diminish the authority of regulators, not enhance it. Yet even after due consultation regulators need techniques for ensuring that operators fulfil their obligations. These techniques may rely on encouragement (and, in particular, on the power of publicity) or on enforcement (and, in particular, on immediate effects on profits). In general:

- The power of publicity is greatest when there is competition and the published measurements distinguish between different operators: condemning operators collectively may achieve little. Major differences in quality levels can be singled out for special attention while minor ones are left for investigation by experts. Regulators in countries with many years of competition sometimes tend towards using encouragement, arguing that different quality levels could emerge following market forces and that detailed regulation needs too many people.
- Immediate effects on profits attract attention from senior management in operators but may distort measurement reporting within operators and to regulators. Usually enforcement entails comparing measurements with targets (not just comparing measurements between operators), though the targets might be set to match the prevailing quality levels. Regulators in countries with few years of competition sometimes emphasise enforcement more than encouragement when they introduce quality of service obligations, but they may find that it needs more effort than they can spare.

In some countries, encouragement and enforcement are applied more forcefully to dominant operators than to other operators, especially where targets are set only for dominant operators.

Figure 4 highlights techniques for ensuring that operators fulfil quality of service obligations. Serious and persistent failures to fulfil many obligations might be handled by using even more drastic enforcement techniques that are not related directly to quality of service, such as withdrawing licences or transferring franchises.

Figure 4 Techniques for ensuring fulfilment of quality of service obligations



Source: Robert Milne (Antelope Consulting).

These techniques are not always practicable (or available under the law, of course). In particular:

- Publishing all measurements and unattained targets can be laborious. However, it can help to show that the regulator is fair and open. Publishing at least some measurements is central to helping customers make informed choices: as discussed in Section 2.5, it is often the main technique for encouraging compliance with quality of service obligations.
- Publicising deficiencies to customers (by putting remarks in bills, messages or advertisements) needs firm comparisons with other operators or against targets.
- Demanding extra measurements and detailed targets could lead to an emphasis on measurement procedures instead of problem solutions. Nonetheless, it can be appropriate when the actions needed to improve quality can be effective fast. Thus when quality fell well short of the targets at points of interconnection in India, all operators were directed to provide interconnections within 90 days of being paid[30]. They were also required to make measurements of congestion at points of interconnection every month, instead of every three months, in the expectation that reductions in the times taken to provide interconnections would have observable effects rapidly[31].
- Monitoring the implementation of remedial plans may require external agencies skilled in network design and operation to assist the regulators. Accompanying it by direct intervention in the activities of the operators may cause confusion about responsibilities and duties.
- Requiring compensation to customers may not be feasible when customers do not need accounts (so credits and rebates cannot be given, as with certain calling cards) or bills are often wrong. It is most useful when customers have better information than regulators about the quality that they receive, can request compensation directly without recourse to regulators or arbitrators, and notice different quality levels very easily (so the compensation can vary with the severity of the deficiency). Compensation is therefore often paid mainly for long fault repair times in fixed wireline networks, as in Chile and India (and also in the UK and some states of the US, for example)[32]. (However, in Mexico it is paid more generally.) For similar reasons compensation is suited to wholesale services, so in Canada it is paid to both wholesale and retail customers of incumbent operators[33].
- Imposing fines can involve extensive legal processes and may take a long time (as may various other techniques for enforcing compliance with quality of service obligations). For example, the regulator in Brazil imposed fines on several operators in 2005, as well as in earlier years, for quality problems that arose in 1999, 2000 and 2001 (though if fewer cases need to be investigated the delay in imposing a fine can be just 1 year.),[34].
- Changing prices (by introducing quality factors into price controls, effectively with rewards for good quality as well as penalties for bad quality) needs careful design if it is to act as an incentive to improving quality[35]. As indicated in Section 2.8 the relation between quality of service and price controls is not always clear.
- Excluding access to government contracts can be difficult to make proportionate to failures by operators and may not be applicable to several operators at once.

The techniques require effort and expertise, of various sorts, with processes that operate efficiently and regularly whenever operators report measurements. Nonetheless they deserve consideration if the aims in Section 1.1 are to be achieved. In many countries several techniques are available, to provide graduated penalties. Thus in Turkey the penalties may involve monitoring the implementation of remedial plans, requiring compensation to customers, or imposing fines.

Using several techniques instead of one can lead to less legalistic procedures and more proportionate remedies. Box 8 outlines a scheme used by the regulator in Australia, the Australian Communications and Media Authority (ACMA), and its predecessor, the Australian Communications Authority (ACA). This scheme itself depends on several techniques (publishing measurements, monitoring the implementation of remedial plans and requiring compensation to customers).

Box 8 Monitoring by the regulator in Australia

Fixed telephony

ACA and subsequently ACMA have published for every 3 months since January 1998 4 measurements for up to 8 fixed telephony operators in 7 geographic areas (subdivided further into "urban", "major rural", "minor rural" and "remote" districts)[36]. The measurements are presented as tables separately for each operator. Participation in the scheme is compulsory for fixed telephony operators having large enough customer numbers[37]. The measurements include the service supply timeliness (separately for initial and various changed services), the fault repair timeliness and the proportion of fulfilled appointments. Reviews have changed some scheme details, most recently by removing the obligation on resellers to make and report measurements[38].

The scheme was developed to get operators to provide and maintain telephony services for residential and small business customers (having at most 5 connections) speedily. It specifies times in which operators should supply services, repair faults, fulfil appointments and pay compensation. It thereby defines a customer service guarantee from participating operators that is enforced using the following rules[37]:

- An operator who exceeds the time on any occasion must pay compensation of AUD12 (USD9.5) to residential customers and AUD20 (USD16) to business customers (roughly half the monthly rental in both cases), for each of the first 5 days of lateness, and AUD40 (USD32) for each subsequent day of lateness.
- An operator who exceeds the times on at least 10% of occasions in at least 2 successive quarters or who shows significant performance decline is asked about possible causes and remedial plans.
- An operator can be exempted if it exceeds the times for reasons beyond its control.
- An operator must offer services with the customer service guarantee but may also offer services with better or worse guarantees.

In 2005, the dominant operator exceeded 4% of service supply times, 9% of fault repair times and 6% of appointment times, with average compensation payments of AUD58 (USD46), AUD25 (USD20) and AUD14 (USD11) respectively[39].

Mobile telephony

ACA and subsequently ACMA have other quality of service monitoring schemes. Among them are ones assessing the performance of operators offering mobile telephony, the reliability of the universal service provider, and the responsiveness of operators giving priority to customers with life-threatening medical conditions. Some of these give force to work by an industry group of operators, the Australian Communications Industry Forum (ACIF), which has drawn up several codes of practice for telephony service customer protection and network operation[40]. (ACIF is now part of a larger group, the Communications Alliance.) Another industry group, the Internet Industry Association (IIA), has drawn up codes of practice for internet service customer protection[41].

Some operators will not fulfil quality of service obligations voluntarily, so encouragement needs to have the force of law behind it. This problem of compliance occurs with other obligations (such as adherence to codes of practice) and makes many operators prefer regulation, which shares burdens fairly, to persuasion, which imposes burdens only on compliant operators.

Box 9 describes one case, that of the regulator in the UK, the Office of communications (Ofcom), and its predecessor, the Office of telecommunications (Oftel), where encouragement to fulfil quality of service obligations was insufficient. Schemes involving industry groups were used for both fixed and mobile telephony. Originally operators were not obliged to participate. However, the voluntary scheme for fixed telephony operators was found to be insufficient and was replaced by a compulsory one, though the corresponding scheme for mobile telephony operators was retained with the proviso that it was developed further.

Box 9 Monitoring by the regulator in the United Kingdom

Fixed telephony

An industry group published for every 3 months from July 1995 to December 2004 5 measurements and 4 customer satisfaction survey results for up to 13 fixed telephony operators of residential or business services (treated separately). Participation in the scheme was voluntary (except for the universal service providers). Operators did not always need to make all of the measurements. Mergers and withdrawals from participation left only 5 operators late in the life of the scheme. Nonetheless in 2003 the publication website received an average of 77000 visits per month, with a maximum of 140000 visits per month.

Ofcom replaced the voluntary scheme with a compulsory one after a review, for the following reasons[3]:

- Lack of comprehensive coverage reduced the usefulness of the scheme to customers.
- Falling participation in the group made measurement and audit procedures difficult to improve.
- Costs could be decreased by fuller participation in the group and improvements in procedures (except perhaps for customer satisfaction surveys).
- Choices of operators were influenced heavily by quality.
- Existing ways of comparing between operators were not regarded as satisfactory by residential customers.

A new industry group is publishing for every 6 months from January 2006 5 measurements for 20 fixed telephony operators of residential or business services (treated separately)[42]. The measurements are presented as bar charts comparing operators. Participation in the scheme is compulsory for fixed telephony operators having large enough revenues, call volumes and life times, even if they do not own networks[43]. The measurements were defined by the group in a framework laid down by Ofcom. The measurements include the service supply timeliness, the fault repair timeliness, the fault report rate, the complaint resolution timeliness and the bill error ratio.

Oftel also published for every 6 months from October 1999 to September 2003 the number of complaints per 1000 subscribers for each of these operators.

Mobile telephony

Oftel published for every 6 months from April 1999 to May 2003 3 measurements for 4 mobile telephony operators in 9 geographic areas (covering 70 towns as well as main roads). Participation in the scheme was voluntary. Operators sometimes blocked publication and took a long time to develop improved measurement and audit procedures. Very few people knew about the publication website, which in 2003 received an average of 4000 visits per month.

Ofcom replaced the voluntary scheme with another voluntary one (after a review, for the following reasons[3]:

- The group was expecting to consider improvements in measurement and audit procedures.
- Comparability would be increased and costs would be decreased by co-ordinating tests already done separately by operators in the group.
- Choices of operators were influenced heavily by quality.

Another new industry group is publishing from September 2006 3 measurements for 4 mobile telephony operators. The measurements er updated twice monthly and are intended to cover the country in 6 months[44]. The measurements are presented as tables and coverage maps comparing operators. Participation in the scheme is voluntary, but there may be a further review if the information published does not help customers or does not cover almost all service provision[43]. The measurements were defined by the group. The measurements include the call setup success ratio, the call drop ratio and the conversational voice quality.

Oftel also published for every 6 months from October 1999 to September 2003 the number of complaints per 1000 subscribers for each of these operators.

2.8 Relating quality to profitability

As indicated in Section 2.7 regulators may try to enforce quality of service obligations by changing the prices charged by operators. In particular they may introduce quality factors into price caps, as has been done in some states of the US, for example. Box 10 describes what is done by the regulator in Colombia, the Comisión de Regulación de Telecomunicaciones (CRT).

Box 10 Monitoring by the regulator in Colombia

Fixed telephony

A government agency has published for every 3 months since January 2003 9 measurements with targets for up to 42 fixed telephony operators in up to 31 geographic areas (treating different operating areas as different operators)[45]. The measurements are presented as tables separately for each operator. Participation in the scheme is compulsory[46]. The measurements include the service supply time (separately for initial and changed services), the fault repair time (separately for internal and external plant) and the fault report rate (separately for internal and external plant). The regulations prescribe details of a customer satisfaction survey as well as of these measurements[47].

The price cap rules allow variations in local tariffs between different operators and geographic areas and are not imposed on some operators in some geographic areas. The variability extends to the treatment of quality through the quality factor, which is different for each operator that is subject to the price cap rules. Each year, for each such operator the quality-dependent price cap is the product of the quality factor and the quality-independent price cap (which is formed by subtracting a productivity factor from the retail price index).

For the start of the scheme in 2000 the quality factor was set to 100% or 84% (depending on when the operators reported their measurements to CRT). Since then it has been a weighted sum of figures derived from measurements of the customer satisfaction level, the service supply time, the fault repair time and the fault report rate. The figures are derived by raising to a certain power the results of mapping the measurements on to truncated linear scales between 100% (for more satisfactory quality) and 50% (for less satisfactory quality). The "certain power" changed in 2005 from 0.25 to 0.20 (so the minimum possible quality factor changed from 84% to 87%)[46]. In addition the weightings, "more satisfactory" values and "less satisfactory" values change from year to year, to reflect changes in what is realistic for the networks. For instance[48]:

- The customer satisfaction level had a weighting of 25% in 2001 and 40% in 2006, a "more satisfactory" value of 68.1 in 2001 and 80.0 in 2006, and a "less satisfactory" value of 53.7 in 2001 and 76.0 in 2006.
- The service supply time had a weighting of 25% in 2001 and 10% in 2006, a "more satisfactory" value of 100.8 days in 2001 and 10 days in 2006 and a "less satisfactory" value of 410.6 days in 2001 and 15 days in 2006.
- The fault repair time had a weighting of 25% in 2001 and 20% in 2006, a "more satisfactory" value of 2.5 days in 2001 and 1 day in 2006 and a "less satisfactory" value of 5.6 days in 2001 and 2 days in 2006.
- The fault report rate had a weighting of 25% in 2001 and 30% in 2006, a "more satisfactory" value of 39.2% in 2001 and 23% in 2006 and a "less satisfactory" value of 83.6% in 2001 and 33% in 2006.

The calculated quality factor is often more than 90% but is sometimes the minimum possible (84% or 87%).

Mobile telephony

Another government agency has published for every 3 months since January 2003 3 measurements for 6 mobile telephony operators[49]. The measurements include the proportion of successful customer appeals against decisions of operators.

Even when regulators do not introduce quality factors into price caps, they wish to ensure that regulating the profitability of operators makes quality levels more appropriate (or at least does not make them less appropriate). Box 11 summarises the evidence from the US about the relation between quality levels and ways of regulating the profitability of dominant fixed telephony operators. At this stage it offers no final clear conclusions.

Box 11 Relations between quality of service and regulation of profitability in the United States

The states of the US have been regulating incumbent fixed telephony operators for many years. They do so in a framework determined by the national regulator, the Federal Communications Commission (FCC), that allows many variations between states but that largely determines the measurements made for quality of service. FCC published the measurements for every 3 months from January 1991 to December 1995. FCC has also published the measurements for every 12 months since January 1996[50]. The measurements include the service supply time, the fault repair time, the fault report rate, the complaint rate and service availability.

Each operator serves several states. In principle the measurements can therefore be analysed to check whether differences in quality levels are correlated with differences in how states regulate operators.

The states have typically shifted from 'rate of return regulation' to various forms of 'incentive regulation'. Rate of return regulation limits the profits kept by operators by assessing them in such a way that operators recover their costs. Incentive regulation encourages operators to reduce costs by letting them keep some of the profits due to cost reductions. The most widespread kind of incentive regulation involves price caps to limit the prices charged to customers, with the price cap being defined by subtracting a 'productivity factor' from a retail or wholesale price index. Other kinds of incentive regulation involve sharing profits with customers and deferring reductions in prices to 'rate of return' levels.

Rate of return regulation does not encourage cost reductions greatly. Indeed it might be expected to improve quality above optimum levels, because operators would suffer no penalties for capital investment that would make life easier for their engineers. Incentive regulation, by contrast, might be expected to lower quality levels because operators would be constrained in how they maintained margins. Some states in the US have introduced a 'quality factor' alongside the 'productivity factor' in price caps to overcome this potential problem in incentive regulation: the profits kept by operators would be increased or decreased according to whether quality levels had risen above or fallen below targets.

Several studies of retail services in the US have used the FCC publications to consider whether differences in quality levels are correlated with choices between rate of return regulation and incentive regulation[51]. However, at this stage the evidence is inconclusive and might even seem contradictory on initial inspection: different studies use different ways of defining incentive regulation, counting quality deficiencies and examining other influences on quality levels, so their results are not comparable[52].

Fewer studies have examined wholesale services. Their quality levels do seem to have fallen while there have been price caps[53]. However the fall does not seem to be correlated with the use of incentive regulation instead of rate of return regulation[54].

2.9 Resolving quality problems directly

Quality of service obligations make the operators take the responsibility for solving problems. This may not be appropriate to all problems in all circumstances: in countries where incumbent operators have few funds and alternative operators cover limited geographic areas, investments in quality may need to be centralised. An approach that aims to eradicate the root causes of the problems may achieve more than encouragement or enforcement.

In particular, if the problems can be traced to deficiencies in staff skills, there may be requirements for national training programmes available to the staff of all operators. Poor customer support is sometimes a symptom of such deficiencies.

High fault report rates and low proportions of successful calls generally indicate a need for improved network equipment. In fact network modernisation can improve quality of service greatly[55]. As network modernisation is motivated largely by wishing to reduce running costs, regulators can stimulate it by relating the profits kept by operators to prices instead of costs[56].

Quality of service improvements that depend on staff development may fall away without sustained effort, but those that result from capital investment are likely to last until the next major change in technology. Suitable staff recruitment and process simplification can sometimes produce improvements quickly, but many actions to improve quality are slow to have their full effects.

3 The selection of areas of focus for quality of service monitoring

This section examines the likely constraints on quality of service monitoring by regulators. It describes criteria for choosing the measurements to be defined, the targets to be set, the measurements to be published and the measurements to be made. The section then discusses which services are to be monitored in current network and NGN environments.

3.1 Measurements to be defined

The numbers of measurements required by regulators vary widely between countries: the examples in Section 2.4 and Section 2.5 show variations by factors of 10. There are many measurement definitions to choose between: ITU has currently at least 150 recommendations related to quality of service, and several other organisations, such as ETSI and IETF, have also developed related standards[57]. Regulators can narrow their choices to measurements that are relevant to one or more of the aims in Section 1.1. They can then restrict attention further to measurements that as far as possible are:

- Important to customers.
- Practical for operators.
- Comparable between operators.

These points are discussed in some detail below.

3.1.1 Importance to customers

The measurements defined for quality of service monitoring by the regulator need to be important to customers (though operators will often choose to have extra measurements for their own purpose). This paper suggests, in particular, that:

- The measurements should be the ones needed by those current or potential customers who would be most seriously affected by unsatisfactory quality. (In particular, measurements that are published should be helpful to customers in the ways listed in Section 3.3.1.) In some countries the customers most seriously affected might be people with disabilities. In other countries they might be the users of the most popular service. (For example, in the Philippines SMS messages, often from poor users, contribute as much revenue as mobile voice calls, so at some stage SMS transmissions might require quality of service monitoring even if voice calls did not.)
- The measurements should be made over short enough periods of the year and for small enough geographic areas to point to problems[58]. Average measurements are not always adequate: for small operators quality may fluctuate greatly over some months and for large operators quality may vary between districts. (For example, unsatisfactory pair gain systems in some districts of Australia prevent broadband internet delivery, have poor voice quality, and get congested, but there are too few of them have much effect on aggregate quality of service measurements.)
- The measurements should be reviewed, to see whether they need to be changed, as the market changes and different aspects of services become most important. In a developing country, the number of years taken to supply fixed network connections may be important. As a network expands, the proportion of successful calls may need attention. Once a network matures, customer support attracts most complaints. The details of individual measurement definitions may also evolve as quality improves: service supply times for initial network connections may come to be measured in days rather then in years and eventually in terms of adherence to times specifically requested by customers. (In fact developments in technology doe not always lead to improvements in quality: conversions to and from voice over IP, together with mobile access, could make end-to-end delay unsatisfactory when it was not unsatisfactory before.)
- The measurements should be consistent between services, particularly if the services are seen as substitutes for one another. Several measurements are likely to deal with aspects such as customer support that are common to many services, as discussed in Section 3.5.3.

Box 12 outlines the experience of the European Union (EU) and the Organisation for Economic Co-operation and Development (OECD), in which various measurements have gradually lost their importance.

Box 12 The evolution of quality of service monitoring in the EU and OECD countries

The European Commission set up a study in 1989 of possible quality of service measurements[59]. Some measurements (such as the delay before getting dial tone) were dismissed as being unnecessary in digital networks, but others seemed relevant according to the surveys reported in the study. The European Commission chose 9 measurements for incorporation in law in 1998[60]. These measurements underlie quality of service monitoring for universal service providers (and sometimes other fixed telephony operators) in countries that have joined or are joining the EU. In the new regulatory framework of 2002 the European Commission made two of them (the proportion of unsuccessful calls and the call setup time) optional as their values were good enough[61].

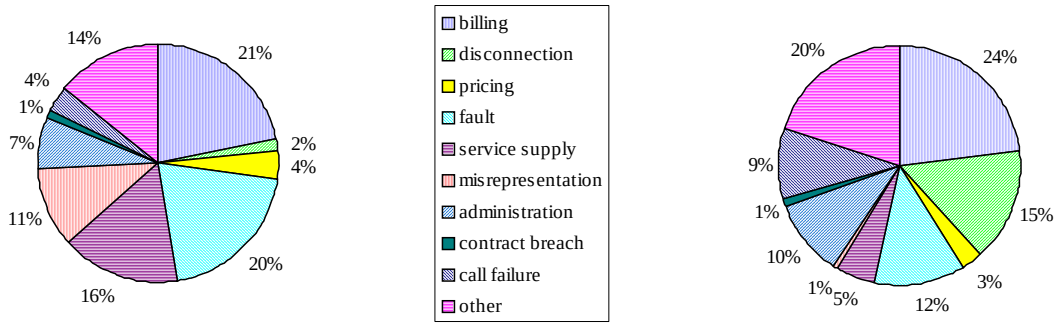
The OECD chose 6 measurements like those of the European Commission in 1990[62]. These measurements are discussed regularly in the OECD biannual reports on telecommunications. They have evolved with the development of the market. In particular, for many OECD countries now the service supply time for initial network connections is too small to be reported and the response times for operator services are less important than the prices of the services[63].

Ideally the aspects that concern customers are found, and periodically reviewed, through customer opinion surveys. Such surveys can be expensive: one regulator rejected a pilot survey covering only two geographic areas because of the expense. However, they may reduce the costs of operators by showing where operators are paying too much attention to aspects of services that customers are not concerned about[64].

In the absence of surveys, records of complaints can help to guide the choice of measurements. However, often these records are open to interpretation in various ways, as described in Section 3.4.1, and do not represent the concerns of potential customers fully: people give up trying to complain. (Supply times for connections are particularly hard to complain about when there are no connections.)

Often many complaints are about prepaid or postpaid billing. Figure 5 illustrates this by summarising analyses of complaints made to two regulators in Africa in 2002-2003 and in 2004-2005[65]. (The analyses resemble each other in several respects but not in all respects: the proportions of complaints allocated to the classes “billing”, “pricing”, “administration” and “contract breach” are rather similar for the two countries, but the proportions of complaints allocated to the classes “disconnection”, “service supply” and “misrepresentation” are very different for the two countries.)

Figure 5 Analyses of complaints to regulators in two countries in Africa



Source: Robert Milne (Antelope Consulting).

3.1.2 Practicality for operators

The measurements defined for quality of service monitoring by the regulator need to be implementable by operators for reasonable costs in reasonable times using consistent measurement and audit procedures. This paper suggests, in particular, that:

- The measurements should not require more tests and calculations than are likely to be needed to characterise differences in quality that are perceptible to customers, now or in the future. Most services may need few measurements each and several of those measurements may be common to many services (as in Malaysia).
- The measurements should be the same as or similar to ones that operators already make (or will benefit from making) for their own purposes, if possible.
- The measurements should deal with matters that operators can control: operators make choices (about numbers of base stations, for example) that weigh up the costs and benefits of reducing problems and that may affect levels of quality and pricing.
- The measurements that need to be made for separate periods of the year and for separate geographic areas should be aligned with those already used by the operators for customer support and network operation, if possible.

Operators can benefit from sharing in quality of service monitoring. They may find that it helps them to work out where to direct their investment: high fault repair times may point to improving customer support, whilst high fault report rates may point to improving network equipment[66].

The "probable causes" of particular deficiencies might be identified and quantified as far as possible in treatments of complaints and fault reports. They could be useful to operators and regulators but not to customers. They would be omitted from documents provided to customers or other operators, for reasons of confidentiality and comprehensibility. (For example, to operators disconnections caused by non-payment of bills, equipment failure, and equipment theft are all different but to customers, one disconnection looks much like another.)

3.1.3 Comparability between operators

For comparability to be achieved, measurements must allow comparisons between operators offering similar services. (Similar services can substitute for each other to some extent, so services are certainly similar if they differ only in implementation features that customers do not experience directly.) This paper adopts this view in proposing that:

- The measurement descriptions (explained in Box 2) should expose any significant dependencies on choices by individual operators. This has particular importance for measurements that are counts of

events occurring in certain intervals of time: the intervals might be defined once for all operators or separately by individual operators. (For example, the proportion of faults repaired within 2 days might be described as characterising "the ability of the operator to repair faults", and the proportion of faults repaired within a time identified by the operator might be described as characterising "the ability of the operator to fulfil fault repair commitments" but should not be described as characterising "the ability of the operator to repair faults".) Also, if there are fewer significant dependencies on individual operators the measurements will generally be more helpful to customers.

- The measurement methods (explained in Box 2) should be precise enough that differences in interpretation and implementation should not lead to systematic (non-random) differences in measurements. (As measurement methods are statistical their results are subject to some randomness.) This may be hardest to achieve for measurements of complaint handling, but it can even be difficult for measurements of traffic handling. (For example, measurements of call set up time may need to take account of whether calls end on user terminals or on network devices such as central mail boxes, when calls are forwarded, and how number portability and number translation are implemented, as well as the terminal types and application implementations for user terminals.)

The details of measurement methods may need to be discussed in groups of operators before they can be settled. Table 6 provides an example of what might need to be decided even for counting fault reports, by listing some apparent impairments that when reported by a user might or might not generate new fault reports. Some of these, such as the repeated reports of events, could affect the fault report count greatly. Others might be given different treatments by operators without affecting the count significantly[67].

Table 6 Examples of apparent impairments possibly not generating fault reports

reported by that user already
reported by a different user already
on more than one connection using the same route
on more than one connection not using the same route
in customer premises equipment maintained by the operator
in customer premises equipment not maintained by the operator
in additional application software supplied by the operator
in additional application software not supplied by the operator
due to normal operation of the service
cured by the user before testing
not detected during testing
not present after testing
due to maintenance at times notified to customers in advance
due to maintenance at times not notified to customers in advance
due to problems of another operator notified to customers in advance
due to problems of another operator not notified to customers in advance
due to disasters caused by human activity
due to disasters not caused by human activity

Source: Robert Milne (Antelope Consulting).

After the measurement descriptions and measurement methods have been refined, comparability may be achieved. However, it may be comparability between operators in one group in one country only. Often even the measurements for fixed telephony and for mobile telephony are not comparable (though this may change as fixed-mobile convergence occurs). Not only the number of measurements, but also the precision of measurement definitions varies within and between countries[68]. Where measurements are the same in outline they may be different in detail, so comparisons can be very approximate.

3.2 Targets to be set

As discussed below, if targets are set they should as far as possible be:

- Useful to customers.
- Realistic for operators.

3.2.1 Usefulness to customers

Targets can be put to various uses: as indicated in Section 1.1, they can be useful in maintaining quality or in improving quality. Different kinds of target are needed for quality maintenance and for quality improvement. More fully:

- Some targets represent required minimum quality levels[69]. (For example, the proportion of successful calls might be required to be at least 95%.) They are intended for maintaining quality above levels that have already been achieved in the country. To set them regulators have to know what quality levels have been reached.
- Other targets represent desired or planned quality levels, typically set for successive years. (For example, the proportion of successful calls might be desired or planned to be at least 96% after 1 year, 97% after 2 years and 98% after 3 years.) They are intended for improving quality from levels that have already been achieved in the country. To set them regulators have to know what quality levels could be reached, in what times and at what costs.

The choice between these kinds of target depends on the main aims of quality of service monitoring. Required minimum service quality levels are more widely chosen, and more generally applicable, than desired or planned quality levels.

Targets of either kind may be more appropriate to wholesale services than to retail services and to dominant operators than to other operators, for the following reasons:

- Policy makers are likely to be concerned with end-to-end quality levels determined by access and core networks. The incumbent operator is likely to be dominant in the markets for making and receiving calls using fixed access and core facilities[70].
- Customers are sometimes more interested in differences between quality levels than in absolute quality levels, at least for retail services[71]. In such cases comparisons between operators may be more useful than comparisons with targets. Of course customers may still welcome constraints on the quality levels of a dominant operator.
- Operators are likely to be interested in the quality levels that they can expect to be available from those operators whose facilities they use for interconnection or resale. Often the main facilities available are those of a dominant operator.

3.2.2 Realism for operators

Though targets can be useful they must be realistic if they are to be effective. Making them realistic can be difficult, for the following reasons:

- The targets should relate to the quality levels that customers want. However, these levels might change as quality improves and might vary for different market segments in different geographic areas.
- The targets should be set after observing what is possible for the operators, so the regulator appreciates what quality levels have been reached and what quality levels could be reached. However, making the observations delays the introduction of the targets
- The targets might be set by adapting values determined through earlier measurements by the operators, used in other countries or proposed in international standards. However, these values may be for measurements that are not comparable with the measurements chosen. In particular, targets often differ greatly between countries. For example, in India and the Philippines the target is to have no more than 1 error per 1000 bills and in Malaysia and Romania the target is to have no more than 20[72].
- The targets should avoid limiting the available choices of quality and price, provided that operators and regulators publish enough information for customers about the choices. For internet services, for example, differences in quality due to differences in actual bandwidth may justify different prices for connections having the same nominal bandwidth.

There are many international standards relating to quality of service measurements and targets: several organisations have defined measurements and some have set targets. Box 13 mentions briefly those that are most relevant to this paper. They are ETSI, which often looks at topics from a broader perspective than that of the EU, and ITU, which carries out standardisation work through ITU-T and ITU-R study groups. The study groups in ITU covering quality of service are ITU-T Study Group 2 ("Operational aspects of service provision, networks and performance"), ITU-T Study Group 12 ("Performance and quality of service") and ITU-T Study Group 19 ("Mobile telecommunications networks")[73].

Box 13 Some sources of measurements and targets from ETSI and ITU standards

ETSI has been responsible for standardising the quality of service measurements required by the European Commission. It has therefore had definitions of measurements for fixed and mobile telephony services for many years and definitions of measurements for internet services for rather fewer years[74]. The proposals for measurements of internet services follow user surveys to ascertain the aspects of services that customers were most concerned about[75]. These definitions leave open many choices. Though many of them are suitable as inputs to national definitions, few of them can provide comparability without further work.

ETSI has also been devising quality of service measurements for mobile services with 3GPP. Besides voice telephony and messaging the services include video telephony, video streaming, file transfer, web browsing and email[76]. Measurements are defined both including and excluding application interfaces (thereby taking into account more or less of the user experience). The measurement methods are accompanied by information on test procedures, profiles and calculations leading towards comparability. Targets are not considered, though some experimental investigations of targets have been done[77].

Many ITU-T recommendations consider quality of service. They are useful in several areas (though they do not provide consistent sets of measurement methods covering all aspects of services). Those most immediately relevant to fixed and mobile telephony are E.721 and E.771, which include targets based on typical reference connections for various calls (local, national, international, fixed and mobile)[78].

Some targets in ITU-T recommendations relate to network performance, rather than quality of service, as they are calculated "bottom up", from network equipment, not "top down", from user experience. However, there are exceptions, such as the upper bounds on end-to-end delays that prevent conversations from deteriorating and that are recommended in G.114[79]. Also targets for voice, video and data applications suggested by some empirical work are summarised in G.1010[80].

Some such voice, video and data applications are put in the context of IP network performance by Y.1541[81]. However, several targets for IP networks are still rather tentative, as some applications have not been assessed widely by users and other applications may never actually be implemented on IP networks. (In particular some of the packet loss targets in Y.1541, but not those in G.1010, are unnecessarily demanding for most applications that are expected to use IP networks: they are intended to deal with the transport of tightly synchronized bit streams over IP and do not take account of the packet loss concealment that is implemented with many voice encodings or the reliable delivery that is used in file transfers and similar applications.)

3.3 Measurements to be published

As discussed below, if measurements are published, they should as far as possible be:

- Accessible to customers.
- Helpful to customers.
- Fair to operators.

3.3.1 Accessibility to customers

Measurements intended for publication should be published at locations or times that let their intended users access them. This paper suggests, in particular, that:

- The measurements should be published in a form directed either to customers or to others (such as journalists) active for customers. These are the "intended users".
- The measurements should be presented appropriately to reach customers. In some countries publishing through radio broadcasts or freephone calls would reach many more people than publishing through newspapers or websites. In other countries readers would expect written words and prefer tables of numbers to graphs or bar charts (contrary to widespread belief).
- The measurements should have their locations or times of publication widely publicised and easily found.

- The measurements should be published at the same locations or times for all operators that are intended to be compared.

3.3.2 Helpfulness to customers

Measurements intended for publication should be expressed in ways that their intended users find helpful. (This may seem obvious, but it is overlooked surprisingly often.) This paper suggests, in particular, that:

- The measurements should deal with the aspects of services that customers are most concerned about, as discussed in Section 3.1.1. In particular, as explained in Section 3.4.3, measurements of traffic handling should be end-to-end, not network-by-network, if possible.
- The measurements should relate to the aspects of services that customers can experience directly. (For example, customers experience dropped calls directly but do not experience wireless network traffic channel congestion directly, though it is one cause of dropped calls.)
- The measurements should be published without being accompanied by extra, irrelevant, numbers. (For example, the call setup success ratio may be of interest to users, but the number of test calls is not of interest to users: auditors or others should have checked whether the number of test calls is large enough to make the measurements precise enough.)
- The measurements should be described in terms that are comprehensible to customers. This should be possible if customers can experience the implications of the measurements directly, though keeping the descriptions precise enough not to mislead may be difficult.
- The measurements should assist with comparisons between operators. For this purpose they are better presented in tables with entries for all the operators in one measurement period, not, for example, in tables with entries for one operator over all the measurement periods.
- The measurements should be expressed no more and no less precisely than is needed to characterise differences in quality that are perceptible to customers and to suppress differences in measurements that are likely to be random. Rounding the measurements to at most 2 significant figures usually helps to achieve this[82]. It also corresponds very approximately with some experimental results about human perception (though differences in perception vary with circumstances). An alternative replaces measurements by a scale having perhaps five points (as in France), where, to minimise subjective ranking, each point must represent a range of possible numbers[83].
- The measurements should be presented consistently with each other. (For example, whether calls can be made and whether messages can be sent should be presented in terms of either successful calls and successful messages or unsuccessful calls and unsuccessful messages, not in terms of successful calls and unsuccessful messages.)

Measurements can be weighted and summed to give single figures of merit. As indicated in Section 2.8 this happens in Colombia. Weightings need to result from conducting customer opinion surveys if important information is not to be lost by using them[84].

3.3.3 Fairness to operators

Measurements should be expressed in ways that are fair to operators. For instance:

- The measurements for two services from the same operator should be published separately when the services are intended to have different levels of quality (for business and residential customers, say).
- The measurements in two geographic areas might be published separately for all operators if services in the two geographic areas are likely to have different levels of quality. However, separate publication in this case might not always be realistic given the networks and organisations of the operators.

- The measurements might be annotated with explanations of deficiencies by an operator when the operator needs to use the facilities of another operator in a way affecting quality. (For example, an alternative operator may have faults on lines that are leased from a dominant operator that does not offer fast protection switching.) Customers can then know both how satisfactory the service is and why the service may have problems.
- The measurements for a service might be published with annotations by the regulator when they may not reflect a true or lasting state of affairs. This might happen if the service is too new to have enough customers to provide statistically significant (precise enough) measurements.

3.4 Measurements to be made

When measurements are being defined, the measurement methods must be specified carefully if comparability is to be achieved. Many fine details need to be clarified, as illustrated in Section 3.1.3. Also, choices need to be made in each of the following broad classes of method:

- Objective and subjective methods.
- Passive and active methods.
- Network-by-network and end-to-end methods.

3.4.1 Objective and subjective methods

Measurement methods may be objective (when they involve counting or timing events) or subjective (when they involve asking for the opinions of customers). Measurement methods that involve counting records of complaints are somewhere in between: the counting is objective, but there may be subjectivity in decisions to complain or in interpreting complaints. In particular, measurements of complaint handling are related to how operators perform customer support. For instance:

- Whether calls from customers ever reach call centre agents, and are ever counted, can depend on how many agents operators employ.
- Whether calls from customers are classified as complaints or as queries can depend on how articulate and assertive customers are.
- Whether complaints about faults can be checked immediately can depend on which parts of networks operators control.

Though quality of service is concerned with user satisfaction, objective measurement methods are often preferable to subjective ones, for the following reasons:

- Subjective measurements may place too much weight on the opinions of people who express themselves strongly.
- Subjective measurements result in distributions of opinions that can be difficult to summarise helpfully and fairly. (For example, a mean opinion score has little value if users are split equally between those who are very satisfied and those who are very dissatisfied.)
- Subjective measurements make comparability within cultures difficult to achieve. (For example, surveys by different operators would need to use the same questions, be made at the same times, be focussed on the same populations and use the same mail or other techniques for getting responses.)
- Subjective measurements make comparability between cultures difficult to achieve. The notion of satisfaction may be quite different for different ethnic and religious groups.
- Subjective measurements are complicated and costly.

As far as possible objective measurements methods are correlated with, or at least intuitively related to, user perceptions.

Over time there may be some movement from objective measurements of mechanical service (such as the proportion of calls that are successful) to subjective measurements of human service (such as the proportion of customers that are satisfied with customer support). Measurements made by objective measurement methods in different countries are rarely comparable in detail even if they are in outline, but they might be made comparable by refining the measurement descriptions and measurement methods. Measurements made by subjective measurement methods, might never be comparable between countries, because of cultural differences.

Services have satisfactory quality only if users can communicate accurately enough. What "accurately enough" means may vary between users and between uses. It is an obvious candidate for subjective measurements but objective measurements are often favoured. Box 14 discusses this briefly, especially for voice.

Box 14 Measurements for voice, video and text

Voice quality can be assessed by various methods described in ITU-T recommendations. In particular:

- Subjective assessments use surveys of customer opinions about particular calls to calculate a Mean Opinion Score (MOS), described in P.800[85]. In practice the opinions must be obtained by asking about test calls, as otherwise they are difficult to distinguish from general feelings about the services. Even when test calls are used the results can be affected strongly by the attitudes and environments of the survey participants and may not be reproduced in other similar surveys.
- Objective assessments for listening use estimators such as Perceptual Evaluation of Speech Quality (PESQ), described in P.862[86]. The results can predict the quality of single messages (as in voice mail) but not the quality of interactive conversations: they estimate distortion but not some other factors that influence voice quality for conversation (such as absolute delay, listening level, loudness loss, sidetone and echo) and that can be particularly important for voice over IP. The results can be converted into numbers on the MOS scale in ways that predict the MOS for listening to some extent.
- Objective assessments for conversation use the E-Model, described in G.107[87]. The results take into account many factors that influence voice quality for conversation but are not yet established experimentally in some cases (such as when low bit rate voice encodings are used together or with room noise). For network planning the results are calculated using reference connections devised by operators, but for quality of service monitoring they are calculated using typical connections experienced by customers, preferably with many of the E-model factors derived from In-service Non-intrusive Monitoring Device (INMD) measurements as in P.562[88]. The results can be converted into numbers on the MOS scale in ways that predict the MOS for conversation to some extent.

Measurements of perceived or predicted quality made by different methods are unlikely to be comparable with each other, so to achieve comparability all operators should use the same method.

Quality assessments for video are less well developed than those for voice, though work has been done on subjective and objective assessments for viewing digital television, for example, as in J.144[89]. Even the audio requirements can range from those for telephony to those for "surround sound" in interactive gaming. Just as with voice there is a distinction between single messages (as in video streaming) and interactive conversations. At this stage video telephony services, for example, are likely to be treated in quality of service monitoring as if they were voice telephony services (so that aspects like lip synchronisation are ignored) but some targets can become inappropriate: the call setup time, for example, may be greater for video telephony than for voice telephony.

Quality assessments for text can usually have simple tests of matching between the sent and received texts.

3.4.2 Passive and active methods

Measurements of traffic handling (such as those for the call setup success ratio) can use real traffic or test traffic. More fully:

- Measurements using real traffic require network equipment to count the traffic. They are "passive" or "non-intrusive", in that they do not add extra traffic to the network. They may not be

immediately comparable even within individual networks, because equipment from different suppliers can have different counting mechanisms. (Also, they do not typically match the experiences of users because the traffic is counted only if it gets into the networks along the wireline or wireless links from their edges.)

- Measurements using test traffic require customer equipment from which test calls are made at the edge of the network, so the experiences of users can be represented. They are "active" or "intrusive", in that they add extra traffic to the network and could lead to extra disruption and expense, especially for a poorly performing network. They are comparable between different networks only if different operators make them at similar times, between similar places and in similar circumstances. In particular, drive-around tests for mobile networks that are co-ordinated between the operators in the main geographic areas lead to measurements that are more comparable and less costly[90].

If measurements of traffic handling are to be comparable, they have to use real traffic for all operators or test traffic for all operators. However, operators may well wish to supplement measurements of test traffic with their own measurements of real traffic (such as off-line analyses of signalling records), to understand network performance better, especially at critical points.

Measurements using test traffic need enough tests to give the required confidence in the measurements[91]. The traffic would pass from traffic-weighted locations to traffic-weighted locations during a part of the day chosen to balance the wish to find how networks perform during the busy hour (when the traffic is most intense) with the wish to economise by spreading tests through the day.

3.4.3 Network-by-network and end-to-end methods

The quality of service provided for a network path depends on all the networks that the path goes through. (For example, as illustrated in Figure 6, the delay between an end user on network X and an end user on network Z that passes through a network Y is the sum of three delays, which are from the end user in X to a point of interconnection of X with Y, from the point of interconnection of Y with X to a point of interconnection of Y with Z, and from the point of interconnection of Z with Y to the end user in Z.)

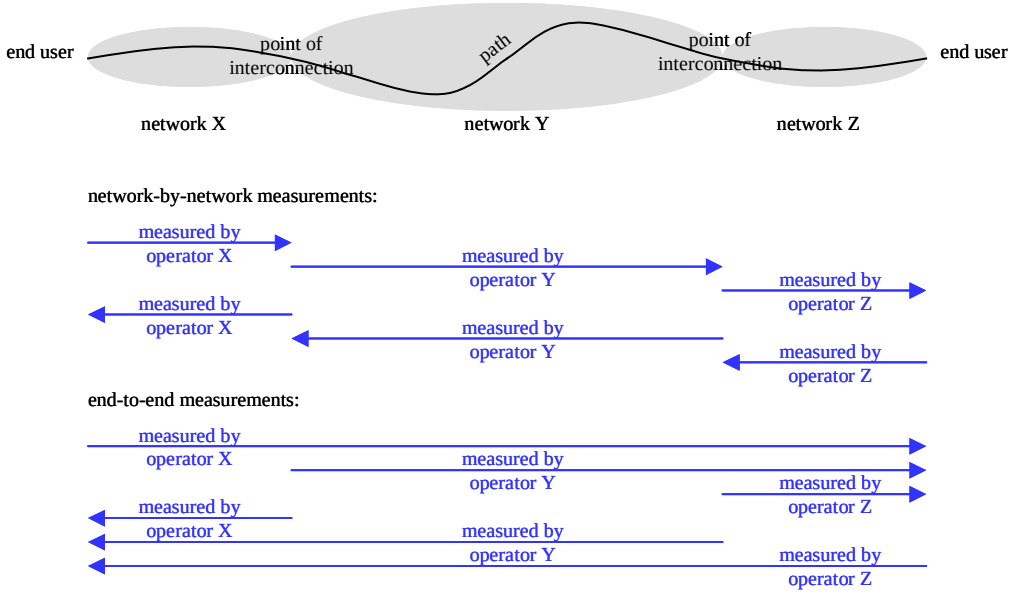
Measurements of traffic handling (such as those for delay and packet loss) can be network-by-network or end-to-end. More fully:

- In network-by-network measurements, the individual networks are considered in isolation, so an end-to-end delay along a path is calculated from the delays across the individual networks used by the path. The measurements for a network therefore do not represent what customers experience, at least for paths ending in other networks, but they do represent what the operator controls.
- In end-to-end measurements, the individual networks are considered in combination, so an end-to-end delay along a path is measured directly (and the delay averaged over all paths through a network takes account of paths ending outside the network). The measurements for a network therefore represent what customers experience on average, but they do not represent what the operator controls.

End-to-end measurements are more helpful to customers than are network-by-network ones, as they concentrate on aspects of services that customers experience directly. They can be fair to operators if published measurements are annotated to explain service deficiencies due to other operators. (For resale, too, measurements should concentrate on aspects of services that customers experience directly, by ensuring that, for example, fault repair times include both the times taken by resellers to process fault reports and the time taken by the original operators of the resold networks to repair faults.)

Figure 6 shows the distinction between network-by-network and end-to-end measurements when there are three networks (X, Y and Z), one of which is providing transit services to the other two. (It is applicable to other measurements, such as packet loss, as well as delay.)

Figure 6 Examples of measurements across three networks



Source: Robert Milne (Antelope Consulting).

3.5 Services to be monitored

Services arise from complex combinations of customer wishes and network implementations. For quality of service monitoring, particularly in an NGN environment, they can be analysed in terms of the following notions:

- Layers in services.
- Network segments used by services.
- Aspects of services.

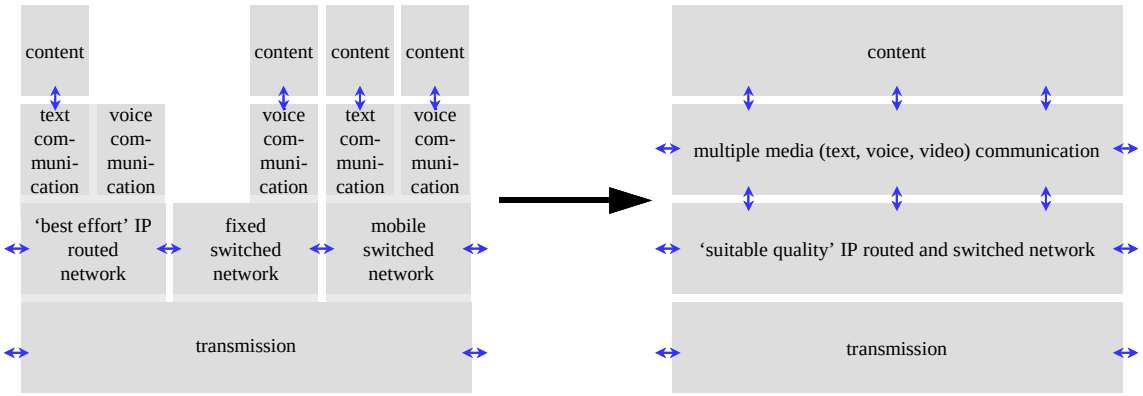
3.5.1 Layers in services

Telephony originally combined transmission, switching (over access and core networks) and voice communication. One operator would supply all these in combination, perhaps along with some content services such as directory enquiries. Quality of service monitoring covered the combined service. Customers never noticed that the implementation had transmission and switching layers.

Then other services were introduced, such as narrowband (dialup) internet access. These other services could be regarded as either as applications on a par with directory enquiries or as different forms of communication. Often they had poor quality levels during their early years but were nonetheless exempt from quality of service monitoring to avoid inhibiting their growth.

Now, however, there are more new services. Some of these are parts of former combined services that are now sold separately. (For example, carrier preselection separates access and core networks and voice over IP splits out voice communication.) Splitting services into separately sold parts may be even more common in an NGN that lets many services use one network. Figure 7 outlines roughly how the current network environment, in which different services are supported by different networks, is giving way to the NGN environment, in which different services are supported by shared networks. (Of course, there are more services and networks than are shown in Figure 7: large operators may have 50 separate networks.)

Figure 7 The transformation from current networks to next generation networks



Source: Robert Milne (Antelope Consulting).

The layers in the implementation are changing to separate more sharply content, communication, network and transmission. The changes in the layers come with changes in businesses. For instance:

- Operators that operate in single layers need both wholesale services to interconnect networks (the horizontal arrows in Figure 7) and wholesale services to use adjacent layers (the vertical arrows in Figure 7). For the regulators, layer-by-layer regulation becomes important alongside, or sometimes instead of, service-by-service regulation. Quality of service measurements of traffic handling for interconnecting networks (such as those for delay and packet loss) can be used both in assessing the end-to-end quality of that layer and in designing the layer above. Quality of service measurements for customer support may be needed both within and between layers.
- Operators that operate in several layers create new bundles of services straddling those layers, especially by getting exclusive rights to content: they introduce "walled gardens", in which their customers access particular content with particular quality of service, instead of other content over the public internet. For their customers, making informed choices becomes harder, because many services need to be considered at once and tariff packages become more complicated. For their competitors, offering equivalent services becomes harder if the new bundles exploit exclusive rights. Quality of service measurements for the individual services are important for comparing disparate bundles from different operators.
- Operators that control access networks and supply voice communication services want to charge for the use of 'suitable quality' IP, which can support real time services such as voice communication better than can current 'best effort' IP, at least when the traffic is fairly heavy[92]. Their attitudes are discussed briefly in Section 3.5.2. Quality of service measurements for 'best effort' IP may be needed to confirm that charging for 'suitable quality' IP is not a strategem for lowering the quality level of 'best effort' IP to encourage customers to buy 'suitable quality' IP.
- Operators that formerly bundled user terminals in their services need to support many differently perfoming terminal types and application implementations. Quality of service measurements of traffic handling must either identify or exclude the terminal types and application implementations.
- Operators may offer new real time services. The quality requirements and expectations of users (including people with disabilities) need to be understood. Quality of service measurements need to be defined, even if they are not used in quality of service monitoring by regulators.
- Operators may exploit new terminals. Quality of service measurements of traffic handling could be made end-to-end by providing suitable terminal capabilities.

Though these changes in the layers are significant, for the regulator they may mainly require the application of existing techniques in slightly new situations.

The scope of quality of service monitoring differs between different layers. In particular:

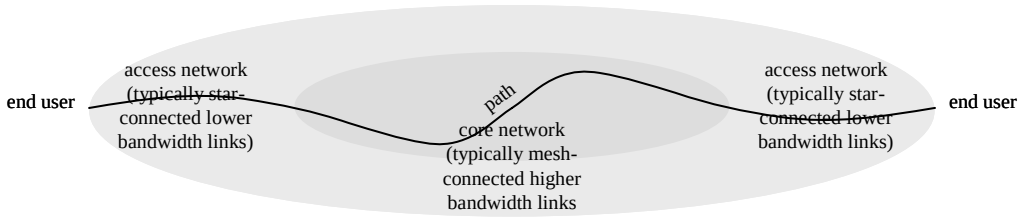
- Services in the content layer have little quality of service monitoring. The main exceptions are operator assistance calls, directory assistance calls and emergency calls, which may be monitored for the frequency and speed with which they are answered. (Other calls to operator call centres may also be conveniently monitored in this way, even though their subjects are not specific to the telecommunications industry.) For other services in the content layer, the content itself may be monitored for other reasons, such as ensuring legality and conformance with rules about the contents and lengths of calls with particular prices. However, this monitoring is not usually regarded as a form of quality of service monitoring and may be performed by organisations quite different from telecommunications regulators.
- Services in the communication layer, such as telephony and messaging, are currently the main focus of quality of service monitoring for end users. The monitoring is discussed in Section 3.5.3.
- Services in the network layer providing IP routing and switching have increasing importance for quality of service monitoring when they can be bought separately from communication services. Currently internet services, and managed business IP network services such as IP virtual private networks, can be bought in this way. In the future IP network services with different quality levels are likely to become available more widely. At least when they are provided by dominant operators, monitoring their quality can help both end users and alternative operators in the same layer or the layer above.
- Services in the transmission layer are bought by end users (corporate customers constructing their own private networks) and by other operators. The quality of service characteristics of the transmission, such as the bit error ratio of a digital leased line, are likely to be monitored by corporate customers that are strong enough to need only occasional regulatory intervention in retail services, but other operators buying wholesale services may need regulatory support.

Effective regulation of wholesale services can stimulate the growth of new operators, in the same layer or the layer above. Ultimately there might be enough operators that customers have adequate choice. There could still be a need for quality of service measurements for retail services to inform customers, but there would be no need for quality of service targets for retail services, just as there would ultimately be no need for retail price control[93].

3.5.2 Network segments used by services

Regulatory requirements such as carrier preselection and local loop unbundling draw attention to the distinction between access and core networks. Figure 8 illustrates this distinction in a simplified form. (It ignores, for example, the possibility that aggregation or distribution networks may lie between access and core networks.)

Figure 8 Access and core networks



Source: Robert Milne (Antelope Consulting).

Even in the current network environment, quality problems are more usual in access networks than in core networks, because the network equipment has more vulnerable locations and is less easily duplicated for protection. However, any shortages of bandwidth for voice communication are due to core networks and mobile networks, not to fixed networks: in fixed networks the availability of bandwidth for a call is usually guaranteed if the call setup is successful.

In an IP network the availability of bandwidth is not guaranteed in this way. (Part of the work on creating standards for the NGN environment concerns letting IP networks have properties like this.) Typically the bandwidth is more severely limited in IP access networks than in IP core networks. IP access networks are likely to differentiate between packets associated with different services, so that they can offer shorter delays and fewer losses to some services. IP core networks can sometimes be implemented without doing this (though they need to maintain the markings that differentiate between packets traversing them, to ensure similar behaviours in the originating and terminating IP access networks). Measurements of delay and packet loss for IP access networks, and the corresponding charges for interconnection, might therefore need to relate to specific services, even if those for IP core networks do not do so.

Thus in IP access networks there might be a need to provide extra mechanisms so that real time services such as voice communication can attain high quality levels. (These mechanisms might not be needed to accommodate just one voice call per IP access line, but they could be needed for multiple video calls, for example.) In particular, operators controlling access networks wish to be compensated by other operators for providing these mechanisms and for losing revenue from their own voice communication services. They might also wish to increase the appeal of their own services, ranging from voice communication to video content, by bundling them together with the IP access network service.

This is why some operators in the US oppose regulation that supports “network neutrality”. (The term loosely describes avoiding some forms of discrimination between services and between operators in access and core networks.) Opponents of regulation to support network neutrality argue that operators should be able to charge other operators extra for supporting services that use extra network resources. Proponents of regulation to support network neutrality argue that without it operators might (and in some cases do already) discriminate in favour of their own services, even to the extent of blocking voice over IP. Both opponents and proponents of regulation can claim to be favouring investment (particularly in new networks, in the case of opponents, and in new applications, in the case of proponents). Whether any such regulation is required can vary between countries: sharp separations between wholesale and retail services of dominant operators, and competition through local loop unbundling and other forms of resale could be enough.

3.5.3 Aspects of services

Usually services have many aspects, each of which may need its quality to be characterised. There can therefore be many measurements to be defined. Several models have been devised to help with defining measurements systematically[94]. This paper mainly draws a distinction between measurements for customer support and measurements dealing with other aspects of services.

In many respects customer support is independent of services: all services require contact with customers, for supplying them, resolving complaints, and repairing faults. Measurements of quality related to customer support are therefore essentially independent of services: they concern the numbers of times that current or potential customers make requests to operators (as measured by the proportion of bills queried, for example) and the responsiveness of operators to requests (as measured by the time taken to satisfy customers querying bills, for example). The most widely used of them relate to supply (the service supply time or timeliness), faults (the fault repair time or timeliness and the fault report rate), complaints (the complaint resolution time or timeliness and the complaint rate) and billing (the bill error ratio).

To customers, measurements of quality related to other aspects of services also seem largely independent of services. Even measurements that might be needed for future services are fairly predictable (though the importance of those measurements to customers may change over time and the targets may be very different for different services). Box 15 provides a model that describes these measurements. (It is expressed in terms of services in the communication layer but is more generally applicable, especially in conjunction with the naming conventions of Box 3.)

Though customer support is likely to receive attention for many services in many layers, in an NGN environment quality of service monitoring by regulators is otherwise likely to be confined to those services in the communication and network layers that most concern customers.

Box 15 Measurements related to communication

The user requirements for services often follow a pattern. In particular:

- Users want to start uses of services successfully (dependably and speedily). For example, they want to set up calls, transmit SMS messages, log into internet sessions or send emails. The quality is characterised by the probability of starting uses successfully and the time taken by successful starts. Defining measurements can be systematic and straightforward (for calls, for example, leading to the call setup success ratio and the call setup time). However, the measurements are not necessarily important to users: the probability of SMS messages transmitted successfully may be too low, but the time taken by successful SMS transmissions may be satisfactory enough not to be measured. For services that users do regard as having starts (such as GPRS and broadband internet sessions) measurements are usually unnecessary. Here "successful" needs to be interpreted relative to user understanding of the service: for store-and-forward services like SMS, MMS and email "successful" can mean "getting to or from the central point of storage" and sending and receiving messages can be treated as separate aspects of the service that need separate measurements.
- Users want to continue uses of services for long enough to complete their communications. For example, they want to use the connections made by successful call setups and internet session logins without perceiving any breaks in them until breaking them deliberately. The quality is characterised by the probability of continuing uses for long enough. Defining measurements can be systematic and straightforward (for calls, for example, leading to the call retention ratio). For services that users do not perceive as having connections (such as SMS) measurements are usually unnecessary. When measurements are made using test traffic (not real traffic) "long enough" needs to be related to the distribution of lengths wanted by customers, such as the durations of real calls.
- Users want to use services for communication accurately. For example, they want the connections made by successful call setups to allow voice conversations. The quality is characterised by user perceptions. These may depend on the intended application: network delays, for example, are much more relevant to interactive conversations than to single messages. Defining measurements is straightforward for services that just provide only text messages but can be problematic for other services, especially if they provide voice or video conversations.

These aspects are essentially generalisations of those referred to as 'accessibility', 'retainability' and 'integrity' in ITU-T recommendation E.800 (though the terms 'accessibility', 'continuity' and 'fidelity' might be more appropriate)[12]. Accessibility itself essentially comprises attributes referred to as 'reliability' and 'speed' in I.350[95]. The ETSI and 3GPP work on mobile services follows approximately similar lines[96]. Other models are discussed in G.1000[94].

Conclusion

This paper starts to assemble and analyse information about current quality of service monitoring by regulators that can be used to guide future schemes. It describes the motivation (in Section 1), operational arrangements (in Section 2) and selection of areas of focus (in Section 3) for quality of service monitoring. On the basis of experience so far it puts forward detailed guidelines for discussion for each of these topics, as well as overall guidelines.

Clearly there is still much to be done to extend and examine the information. However, the paper does at least point to possible ways forward.

Abbreviations

3GPP	Third Generation Partnership Project
ACA	Australian Communications Authority
ACIF	Australian Communications Industry Forum
ACMA	Australian Communications and Media Authority
Anatel	Agência nacional de telecomunicações
Cintel	Centro de investigación de telecommunicaciones en Colombia
ComReg	Commission for communications Regulation
CRT	Comisión de Regulación de Telecomunicaciones
CRTC	Canadian Radio-television and Telecommunications Commission
ECTEL	Eastern Caribbean TELecommunications authority
ETSI	European Telecommunications Standards Institute
EU	European Union
FCC	Federal Communications Commission
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications
ICT	Information and Communication Technology
IETF	Internet Engineering Task Force
IIA	Internet Industry Association
INMD	In-service Non-intrusive Measurement Device
IP	Internetwork Protocol
ITU	International Telecommunication Union
ITU-D	ITU – Telecommunication Development Sector
ITU-R	ITU – Radiocommunication Sector
ITU-T	ITU – Telecommunication Standardisation Sector
MMS	Multimedia Messaging Service
MOS	Mean Opinion Score
NGN	Next Generation Network
ODTR	Office of the Director of Telecommunications Regulation
OECD	Organisation for Economic Co-operation and Development
Ofcom	Office of communications
Oftel	Office of telecommunications
PESQ	Perceptual Evaluation of Speech Quality
SMS	Short Messaging Service
Subtel	Subsecretaría de telecomunicaciones
TCP	Transmission Control Protocol
TRAI	Telecom Regulatory Authority of India
TRC	Telecommunications Regulatory Commission
UDP	User Datagram Protocol
UK	United Kingdom
US	United States

Endnotes

[1] In this paper the term 'regulator' covers government ministries, not just separate regulators, and the term 'operator' covers all service providers, not just network operators.

[2] See Cintel, *Brief Estudio de Mercado 2005*, February 2006, at http://www.cintel.org.co/noticia.php3?nt=4716.

[3] See Oftel, *Reporting Quality of Service Information to Consumers: A consultation on the regulatory options for acquiring and publishing quality of service measures for telecommunication services*, November 2003, at http://www.ofcom.org.uk/consult/condocs/qualitystate/qosoriginal/qos.pdf.

[4] See Oftel, *Consumers' use of fixed telephony: Q14 August 2003*, October 2003, at http://www.ofcom.org.uk/static/archive/oftel/publications/research/2003/q14fixres1003.pdf.

[5] For the adjudication by the industry group that regulates advertising see Advertising Standards Authority, "Non-broadcast Adjudication," November 2005, at http://www.asa.org.uk/asa/adjudications/non_broadcast/Adjudication+Details.htm?Adjudication_id=40467.

[6] See TRAI, *Report on QoS Parameters related to Congestion on Point of Interconnections (POIs) of CMSPs*, January 2006, at http://www.trai.gov.in/trai/upload/Reports/2/report17jan06.pdf.

[7] For a discussion in economic terms of such problems with dominant operators see Sappington, D. E. M., "Regulating Service Quality: A Survey," *Journal of Regulatory Economics*, November 2005, Volume 27, Issue 2, Pages 123–154.

[8] For interconnection, wholesale service targets for measurements of traffic handling are more demanding than the corresponding retail service targets (because they consider partial paths, between a point of interconnection and an end user, not complete paths, between one end user and another). For resale, wholesale service targets for measurements of traffic handling are the same as the corresponding retail service targets, but wholesale service targets such as that for the fault repair time may need to be more demanding than the corresponding retail service targets (to allow time for customer care procedures, for example).

[9] Residential broadband internet service exemplifies this in several countries. The operators may well require contracts lasting one or two years, to cover the costs of equipment in the customer premises and in the access network. They may well also provide email addresses, such as customer_name@isp_name.com, that are not portable from their networks, and storage that is not accessible only from outside their networks. Nonetheless there may be no co-ordinated published quality of service monitoring.

[10] Reducing churn might be a reason for introducing lifetime subscriptions in India, where monthly churn rates vary between 3.5% and 6.0% approximately.

[11] Also, customer satisfaction surveys can be misleading: customers that have changed their operators may not want to admit that doing so was unwise, so satisfaction with the alternative operators to whom they have changed may appear higher than satisfaction with the dominant operators from whom they have changed. This effect can be reduced if the survey sample size is made large enough to distinguish each individual operator, but doing this can be expensive. Thus customer satisfaction surveys that produce results about individual operators are usually insufficient on their own: in Australia and Colombia, for example, such surveys are performed every 12 months and every 6 months respectively as accompaniments to the quality of service measurements.

[12] See ITU-T Recommendation E.800, *Terms and definitions related to quality of service and network performance including dependability*, August 1994, at http://www.itu.int/rec/T-REC-E.800.

[13] See ITU-T Recommendation E.600, *Terms and definitions of traffic engineering*, March 1993, at http://www.itu.int/rec/T-REC-E.600.

[14] When it was state-owned the incumbent operator in the UK published quality of service measurements. It stopped doing so "for commercial reasons" in preparation for privatisation. It resumed doing so in October 1987, under pressure from the regulator.

[15] An industry group could be either "self-regulatory" or "co-regulatory". The distinction between these is often rather blurred. One way of making the distinction sharp is to say that the regulator does not participate in a self-regulatory group (except to require that the group fulfil its mandate and perhaps to provide the secretariat) but that the regulator does participate in a co-regulatory group. The role of such groups has been formalised in

Australia, for example, where industry groups can submit codes of practice to the regulator, who can then "register" them and make them compulsory.

[16] See TRAI, *Consultation Paper on Review of Quality of Service (QOS) Parameters of Basic and Cellular Mobile Telephone services*, February 2005, at http://www.trai.gov.in/trai/upload/ConsultationPapers/13/conpaper23feb05.pdf.

[17] For links to the reports see http://www.trai.gov.in/Reports_list_year.asp.

[18] See TRAI, *Regulation on Quality of Service of Basic and Cellular Mobile Telephone Services, 2005*, July 2005, at http://www.trai.gov.in/trai/upload/Regulations/3/regu1jul05.pdf.

[19] For links to the reports see http://www.subtel.cl/servlet/page?_pageid=58&_dad=portal30&_schema=PORTAL30.

[20] See Subtel, *Fija indicadores de calidad del servicio de acceso a internet y sistema de publicidad de los mismos*, June 2001, at http://www.subtel.cl/pls/portal30/docs/folder/wsubtel_contenidos_sitio/subtel/marcoregulatorio/norm_tec_internet/norma_tecnica_internet/res_669_calidad_de_serv_internet.pdf.

[21] See Subtel, *Fija indicadores de calidad de los enlaces de conexión para cursar el tráfico nacional de Internet y sistema de publicidad de los mismos*, June 2000, at http://www.subtel.cl/pls/portal30/docs/folder/wsubtel_contenidos_sitio/subtel/marcoregulatorio/norm_tec_internet/norma_tecnica_internet/res_698_trafico_internet.pdf.

[22] Figures provided by 4 operators in the UK showed variations by factors of 74 in costs per customer and 34 in costs per minute (though variations by factors of 7 and 5 were perhaps more usual, with the high costs being outliers). Organisations making measurements generally regarded the costs as very low. See:

Oftel, *Reporting Quality of Service Information to Consumers: A consultation on the regulatory options for acquiring and publishing quality of service measures for telecommunication services, November 2003*, at http://www.ofcom.org.uk/consult/condocs/qualitystate/qosoriginal/qos.pdf.

Ofcom, *A Statement on providing quality of service information to consumers: A Consultation on quality parameters including a Notification and Draft Direction*, September 2004, at http://www.ofcom.org.uk/consult/condocs/qualitystate/qos/qos.pdf.

[23] For links to the reports see http://www.askcomreg.ie/about_us/Consumer%5FInformation.53.LE.asp#306.

[24] For the database see http://sistemas.anatel.gov.br/sgiq/Default.asp?SISQSmodulo=6347&SISQSsistema=540.

[25] See Anatel, *Regulamento de Indicadores de Qualidade do Serviço Telefônico Fixo Comutado – RIQ*, October 2005, at http://www.anatel.gov.br/biblioteca/Resolucao/2005/Anexo_res_417_2005.pdf.

[26] For the database see http://sistemas.anatel.gov.br/sgiq/Default.asp?SISQSmodulo=8463&SISQSsistema=164.

[27] See Anatel, *Regulamento de Indicadores de Qualidade do Serviço Móvel Pessoal – SMP*, April 2003, at http://www.anatel.gov.br/biblioteca/Resolucao/2003/anexo_res_335_2003.pdf.

[28] Doubts about whether the benefits of publication outweigh the costs may be why in 2003 measurements were published in only 6 of the 15 member countries of the EU, despite apparent obligations to publish under the directives. At the same time 10 member countries imposed targets for fixed telephony universal operators and 3 member countries imposed targets also on other fixed telephony operators. See European Commission, *Ninth report on the implementation of the Electronic Communications Regulatory Package, Annex 2*, COM(2003) 715, December 2003, at http://europa.eu.int/eur-lex/lex/LexUriServ/site/en/com/2003/com2003_0715en01.pdf.

[29] See TRC, *Explanatory Memorandum on the Quality of Service Framework*, February 2006, at http://www.trc.gov.jo/Static_English/doc/Technical/Explanatory_Memorandum_on_the_Quality_of_Service_Framework.pdf.

[30] See TRAI, *TRAI's Direction to all Service Providers on Provision of Interconnection*, June 2005, at http://www.trai.gov.in/trai/upload/PressReleases/79/pr7jun05.pdf.

[31] See TRAI, *TRAI issues show cause notices to mobile operators*, March 2006, at http://www.trai.gov.in/trai/upload/PressReleases/5/pr7mar06.pdf.

[32] In Chile customers are entitled to reductions in monthly rentals if faults persist for at least 12 hours and to rebates if faults persist for at least 3 consecutive days. In India customers are entitled to rebates if faults persist for at least 3 consecutive days (with rebates amounting to 7 days for faults repaired in at most 7 days, 15 days for faults repaired in at most 15 days, and 1 month otherwise). See:

Subtel, *Acerca del contrato de servicio*, August 2006, at http://www.subtel.cl/servlet/page?_pageid=58&_dad=portal30&_schema=PORTAL30.

TRAI, *Regulation on Quality of Service of Basic and Cellular Mobile Telephone Services, 2005*, July 2005, at http://www.trai.gov.in/trai/upload/Regulations/3/regu1jul05.pdf.

[33] There are 13 measurements for retail services and 24 measurements for wholesale services. Of the 24 measurements for wholesale services, 21 are made competitor-by-competitor, so that compensation can be directed to the correct competitors. See:

CRTC, *Telecom Decision CRTC 2005-20 Finalization of quality of service rate rebate plan for competitors*, March 2005, at http://www.crtc.gc.ca/archive/ENG/Decisions/2005/dt2005-20.htm.

CRTC, *Telecom Decision CRTC 2005-17 Retail quality of service rate adjustment plan and related issues*, March 2005, at http://www.crtc.gc.ca/archive/ENG/Decisions/2005/dt2005-17.htm.

[34] The fines amounted to more than BRL4300000 (USD1600000). The problems largely, but not solely, concerned the service supply time, the fault repair time, the complaint resolution time and the call setup success ratio. See:

Anatel, *Anatel multa operadoras de telefonia fixa em R$ 2,6 milhões por descumprirem metas de qualidade*, May 2005, at http://www.anatel.gov.br/index.asp?link=/biblioteca/Releases/2005/release_03_05_2005(5).pdf.

Anatel, *Anatel multa Telemar em R$ 107,5 mil por descumprir metas de qualidade*, April 2005, at http://www.anatel.gov.br/Tools/frame.asp?link=/biblioteca/releases/2005/release_07_04_2005.pdf.

Anatel, *Descumprimento de Metas de Qualidade na telefonia fixa gera mais R$ 1,3 milhão em multas*, March 2005, at http://www.anatel.gov.br/Tools/frame.asp?link=/biblioteca/releases/2005/release_22_03_2005.pdf.

Anatel, *Anatel aplica mais de R$ 300 mil em multas contra a Telemar e a Brasil Telecom*, March 2005, at http://www.anatel.gov.br/Tools/frame.asp?link=/biblioteca/releases/2005/release_17_03_2005(2).pdf.

[35] For example, there is a theoretical justification for the view that penalties proportional to profits may increase investment in quality but that penalties proportional to revenues may increase or decrease investment in quality according to circumstances, provided that (in both cases) they are less than the maximum values permitted by the law. See Weisman, D. L., "Price regulation and quality," *Information Economics and Policy*, March 2005, Volume 17, Issue 2, Pages 165-174.

[36] For links to the reports see http://www.acma.gov.au/ACMAINTER:STANDARD::pc=PC_1378.

[37] See ACMA, *Telecommunications (Customer Service Guarantee) Standard 2000 (No. 2)*, February 2004, at http://www.acma.gov.au/acmainterwr/consumer_info/csg/csg_consolidatedstandard04.pdf.

[38] See ACMA, *Telecommunications Industry Monitoring and Reporting Arrangements*, August 2004, at http://www.acma.gov.au/acmainterwr/aca_home/issues_for_comment/discussion/archive/telecommunications_industry_monitoring_and_reporting.pdf.

[39] See ACMA, *Telecommunications Performance Report 2004-05*, November 2005, at http://www.acma.gov.au/acmainterwr/_assets/main/lib100449/tpr 2005.pdf.

[40] For links to the codes see http://www.acif.org.au/ACIF_documents/codes.

[41] For links to the codes see http://www.iia.net.au/index.php?option=com_content&task=section&id=3&Itemid=33.

[42] For the bar charts see http://www.topcomm.co.uk/.

[43] See Ofcom, *A Statement on providing quality of service information to consumers: A Consultation on quality parameters including a Notification and Draft Direction*, September 2004, at http://www.ofcom.org.uk/consult/condocs/qualitystate/qos/qos.pdf.

[44] For the coverage maps see http://www.topnetuk.co.uk/.

[45] For the database see http://www.sui.gov.co/reportesSUI/SUI_ReporteTele.htm.

[46] See CRT, *Resolución 1250 de 2005*, June 2005, http://www.crt.gov.co/Documentos/Normatividad/ResolucionesCRT/00001250.pdf.

[47] See CRT, *Resolución 575 de 2002*, December 2002, http://www.crt.gov.co/Documentos/Normatividad/ResolucionesCRT/00000575.pdf.

[48] See CRT, *Resolución 338 de 2000*, December 2000, http://www.crt.gov.co/Documentos/Normatividad/ResolucionesCRT/00000338.pdf.

[49] For links to the reports see http://www.crt.gov.co/crt_2001-2004/paginas/internas/biblioteca/regulatorio_b.htm.

[50] For the database see http://svartifoss2.fcc.gov/eafs/adhoc/table_year_tab.cfm?reportType=4305.

[51] See:

Ai, C., and Sappington, D. E. M., "The Impact of State Incentive Regulation on the U.S. Telecommunications Industry," *Journal of Regulatory Economics*, September 2002, Volume 22, Issue 2, Pages 133-159.

Banerjee, A., "Does Incentive Regulation "Cause" Degradation of Retail Telephone Service Quality?," *Information Economics and Policy*, June 2003, Volume 15, Issue 2, Pages 243-269.

Clements, M. E., "Local Telephone Quality of Service: A Framework and Empirical Evidence," *Telecommunications Policy*, June-July 2003, Volume 28, Numbers 5-6, Pages 413-426.

Roycroft, T. R., and Garcia-Murrilo, M., "Trouble Reports as an Indicator of Service Quality: The Influence of Competition, Technology, and Regulation," *Telecommunications Policy*, November-December 2000, Volume 24, Numbers 10-11, Pages 947-967.

[52] See Sappington, D. E. M., "The Effects of Incentive Regulation on Retail Telephone Service Quality in the United States," *Journal of Network Economics*, December 2003, Volume 3, Issue 4, Pages 355-375, at http://www.rnejournal.com/articles/sappington-RNE_5_dec_03final.pdf.

[53] See Uri, N. D., "Service Quality Effects of Incentive Regulation on Access Service in Telecommunications in the United States," *European Journal of Law and Economics*, November 2003, Volume 16, Issue 3, Pages 369-390.

[54] See Zimmerman, P. R., "Regional Bell Operating Company Entry Into Long-Distance and Non-Price Discrimination Against Rival Interexchange Carriers: Empirical Evidence from Panel Data," *Applied Stochastic Models in Business and Industry*, November 2003, Volume 19, Issue 4, Pages 269–290.

[55] For example, in the UK by September 1989 the main incumbent operator could repair 86% of faults in 1 working day, compared with 65% in September 1988. The improvement was attributed to network modernisation spurred by competition. At about the same time the operator, under pressure from the regulator, started to pay compensation for excessive times for service supply and fault repair. If the committed time for service supply or fault repair was missed by 2 working days, customers were paid GBP5 (USD8.4) per working day, with up to GBP1000 (USD1700) for residential customers and GBP5000 (USD8400) for business customers if they suffered financially.

[56] For suitable correlations, showing that in the US more network modernisation has occurred with incentive regulation than with rate of return regulation, see Ai, C., and Sappington, D. E. M., "The Impact of State Incentive Regulation on the U.S. Telecommunications Industry," *Journal of Regulatory Economics*, September 2002, Volume 22, Issue 2, Pages 133-159.

[57] For a summary of the work of ITU and other organisations on quality of service see ITU, *Quality of Service and Network Performance*, March 2004, at http://www.itu.int/pub/T-HDB-QOS.02-2004.

[58] Other statistics besides averages, such as variances, can also help with showing variability but not immediately in identifying the cause of variability.

[59] See Milne, C. B., and Mitchell, J., "Quality of Service indicators for the telephone service to residential consumers, Final Report to the Commission of the European Communities", Polytechnic of Central London, October 1989.

[60] See European Commission, *Directive 98/10/EC of the European Parliament and of the Council of 26 February 1998 on the application of open network provision (ONP) to voice telephony and on universal service for telecommunications in a competitive environment*, February 1998, at http://portal.etsi.org/public-interest/Documents/Directives/Standardization/Directive_1998_10.pdf.

[61] See European Commission, *Directive 2002/22/EC of the European Parliament and of the Council of 7 March 2002 on universal service and users' rights relating to electronic communications networks and services (Universal Service Directive)*, March 2002, at http://portal.etsi.org/public-interest/Documents/Directives/Standardization/Universal_Service_Directive.pdf.

[62] See OECD, "Quality of service indicators," Chapter 5 of *Performance indicators for public telecommunications operators*, November 1990, at http://www.oecd.org/dataoecd/34/59/2366923.pdf.

[63] See OECD, "Quality of service," Chapter 7 of *Communications Outlook 2003*, May 2003, at http://www1.oecd.org/publications/e-book/9303021E.pdf.

[64] An example is sometimes provided by the call setup success ratio. For 35 operators in 20 developing countries, a customer uses between 200 and 1,600 minutes per year of mobile telephony. If the average call duration is 2 minutes, the customer using 200 minutes per year may not distinguish between call setup success ratios of 99% and 99.9%, and the customer using 1,600 minutes per year may not distinguish between call setup success ratios of 99.9% and 99.99%. When the customer is actually selling the use of the telephone informally, the call setup success ratio could well be unimportant.

[65] Similarly 33% of complaints to the regulator in Germany in 1999 were about incomprehensible bills, 41% of complaints to the regulator in the US in 2000 were about inaccurate information, and 28% of complaints to the ombudsman in Australia in 2005 concerned prepaid or postpaid billing, almost independently of the services (fixed telephony, mobile telephony and internet). See:

OECD, "Quality of service," Chapter 8 of *Communications Outlook 2001*, May 2001, at http://www1.oecd.org/publications/e-book/9301021E.pdf.

ACMA, *Telecommunications Performance Report 2004-05*, November 2005, at http://www.acma.gov.au/acmainterwr/_assets/main/lib100449/tpr 2005.pdf.

[66] For an analysis suggesting that this is so, by relating investment types to the quality of particular aspects of services in the US, see Clements, M. E., "Local Telephone Quality of Service: A Framework and Empirical Evidence," *Telecommunications Policy*, June-July 2003, Volume 28, Numbers 5-6, Pages 413-426.

[67] For definitions of 5 measurements designed by an industry group (in a process lasting over a year) and intended to achieve comparability between fixed telephony operators see Topcomm, "Quality of Service Definitions (Applicable to Quarter 1, 2006)," July 2006, at http://www.topcomm.co.uk/topcomm_QOS.pdf.

[68] A co-ordinated effort to define measurements for regulators and operators spanning several countries might provide comparability between operators in different countries. It could also reduce costs, for regulators and operators alike. It could be a response to the increasing tendency of mobile telephony operators to become multinational. Perhaps ECTEL, with the same dominant operator in all its countries, can aspire to such co-ordination. (Measurement definitions at the level of the ETSI ones for the EU leave too many options open to ensure comparability.)

[69] Targets were originally things to aim for. Now in widespread usage (as in this paper) they are things for which there are penalties if they are missed. Arguably targets that represent minimum required quality levels should be called something other than 'targets', but no other term is much more satisfactory.

[70] Also, in some countries of the EU, where calling parties pay for calls, the mobile telephony operators are seen as dominant in the market for receiving calls using mobile access facilities.

[71] Subjective voice quality assessments distinguish differential tests from absolute tests accordingly. See ITU-T Recommendation P.800, *Methods for subjective determination of transmission quality*, August 1996, at http://www.itu.int/rec/T-REC-P.800.

[72] In fact the requirement in the Philippines is to have no more 1 error per 1000 bills per year, so it is particularly stringent if bills are sent every 3 months.

[73] See:

ITU, "ITU-T Study Group 2 (Study Period 2005 - 2008)", August 2006, at http://www.itu.int/ITU-T/studygroups/com02/index.asp. ITU-T Study Group 2 is responsible for "studies relating to: principles of service provision, definition and operational requirements of service emulation; numbering, naming, addressing requirements and resource assignment including criteria and procedures for reservation and assignment; routing and interworking requirements; human factors; operational aspects of networks and associated performance requirements including network traffic management, quality of service (traffic engineering, operational performance and service measurements); operational aspects of interworking between traditional telecommunication networks and evolving networks; and evaluation of feedback from operators, manufacturing companies and users on different aspects of network operation".

ITU, "ITU-T Study Group 12 (Study Period 2005 - 2008)", August 2006, at http://www.itu.int/ITU-T/studygroups/com12/index.asp. ITU-T Study Group 12 is responsible for "Recommendations on the end-to-end transmission performance of terminals and networks, in relation to the perceived quality and acceptance by users of text, data, speech, and multi-media applications". It is the lead study group on performance and quality of service. (Though this work includes the transmission implications of all networks and all telecommunications terminals, and also includes work on performance and resource management, it has a special focus on IP quality of service, interoperability and implications for NGN.)

ITU, "ITU-T Study Group 19 (Study Period 2005 - 2008)", August 2006, at http://www.itu.int/ITU-T/studygroups/com19/index.asp. Study Group 19 is responsible for "studies relating to: network aspects of mobile telecommunications networks, including International Mobile Telecommunications 2000 (IMT-2000) and beyond, wireless Internet, convergence of mobile and fixed networks, mobility management, mobile multimedia functions, internetworking, interoperability and enhancements to existing ITU-T Recommendations on IMT-2000."

[74] See:

ETSI EG 202 057-1 V1.2.1, *Speech Processing, Transmission and Quality Aspects (STQ); User related quality of service parameter definitions and measurements; Part 1: General User related quality of service*, August 2005, at http://pda.etsi.org/pda/queryform.asp.

ETSI EG 202 057-2 V1.1.1, *Speech Processing, Transmission and Quality Aspects (STQ); User related quality of service parameter definitions and measurements; Part 2: Voice telephony, Group 3 fax and modem data services*, September 2002, at http://pda.etsi.org/pda/queryform.asp.

ETSI EG 202 057-3 V1.1.1, *Speech Processing, Transmission and Quality Aspects (STQ); User related quality of service parameter definitions and measurements; Part 3: quality of service parameters specific to Public Land Mobile Networks (PLMN)*, April 2005, at http://pda.etsi.org/pda/queryform.asp.

ETSI EG 202 057-4 V1.1.1, *Speech Processing, Transmission and Quality Aspects (STQ); User related quality of service parameter definitions and measurements; Part 4: Internet access*, August 2005, at http://pda.etsi.org/pda/queryform.asp.

[75] See ETSI TR 102 276 V1.1.1, *User Group; Users' Quality of Service Criteria for Internet Access in Europe*, October 2003, at http://pda.etsi.org/pda/queryform.asp.

[76] See:

ETSI TS 102 250-1 V1.1.1, *Speech Processing, Transmission and Quality Aspects (STQ); quality of service aspects for popular services in GSM and 3G networks; Part 1: Identification of Quality of Service aspects*, October 2003, at http://pda.etsi.org/pda/queryform.asp.

ETSI TS 102 250-2 V1.4.1, *Speech Processing, Transmission and Quality Aspects (STQ); quality of service aspects for popular services in GSM and 3G networks; Part 2: Definition of Quality of Service parameters and their computation*, March 2006, at http://pda.etsi.org/pda/queryform.asp.

ETSI TS 102 250-3 V1.3.2, *Speech Processing, Transmission and Quality Aspects (STQ); quality of service aspects for popular services in GSM and 3G networks; Part 3: Typical procedures for Quality of Service measurement equipmen*t, September 2005, at http://pda.etsi.org/pda/queryform.asp.

ETSI TS 102 250-4 V1.1.1, *Speech Processing, Transmission and Quality Aspects (STQ); quality of service aspects for popular services in GSM and 3G networks; Part 4: Requirements for Quality of Service measurement equipment*, October 2003, at http://pda.etsi.org/pda/queryform.asp.

ETSI TS 102 250-5 V1.3.1, *Speech Processing, Transmission and Quality Aspects (STQ); quality of service aspects for popular services in GSM and 3G networks; Part 5: Definition of typical measurement profiles*, November 2005, at http://pda.etsi.org/pda/queryform.asp.

ETSI TS 102 250-6 V1.2.1, *Speech Processing, Transmission and Quality Aspects (STQ); quality of service aspects for popular services in GSM and 3G networks; Part 6: Post processing and statistical methods*, October 2004, at http://pda.etsi.org/pda/queryform.asp.

[77] See ETSI TR 102 274 V1.1.2, *Human Factors (HF): Guidelines for real-time person-to-person communication services*, January 2004, at http://pda.etsi.org/pda/queryform.asp.

[78] See:

ITU-T Recommendation E.721, *Network grade of service parameters and target values for circuit-switched services in the evolving ISDN*, May 1999, at http://www.itu.int/rec/T-REC-E.721.

ITU-T Recommendation E.771, *Network grade of service parameters and target values for circuit-switched public land mobile services*, October 1996, at http://www.itu.int/rec/T-REC-E.771.

[79] See ITU-T Recommendation G.114, *One-way transmission time*, May 2003, at http://www.itu.int/rec/T-REC-G.114.

[80] See ITU-T Recommendation G.1010, *End-user multimedia quality of service categories*, November 2001, at http://www.itu.int/rec/T-REC-G.1010.

[81] See ITU-T Recommendation Y.1541, *Network performance objectives for IP-based services*, February 2006, at http://www.itu.int/rec/T-REC-Y.1541.

[82] For example, with rounding to 2 significant figures, fault repair times of 1.1 and 1.9 days would be treated as different and written as “1.1 days” and “1.9 days” respectively, but fault repair times of 15.1 days and 15.9 days would be treated as the same and written as “16 days”.

[83] A scale having few points can exaggerate differences between quality levels unfairly: a regulator that endorsed operators as “quality approved” for particular aspects of services would be using a scale having two points, and if 2.0 days was the upper bound on approved times “1.9 days” would be “good” and “2.1 days” would be “bad” despite their closeness.

[84] Weighted sums of measurements are at least better than average rankings of operators, obtained by averaging over the rankings of the operators given by the measurements. Such averages treat minor differences between operators for less important measurements as being as significant as major differences between operators for more important measurements. They have nonetheless occasionally been devised by other monitors and quoted by broadband internet operators.

[85] See ITU-T Recommendation P.800, *Methods for subjective determination of transmission quality*, August 1996, at http://www.itu.int/rec/T-REC-P.800.

[86] See ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, February 2001, at http://www.itu.int/rec/T-REC-P.862.

[87] See ITU-T Recommendation G.107, *The E-model, a computational model for use in transmission planning*, March 2005, at http://www.itu.int/rec/T-REC-G.107.

[88] See ITU-T Recommendation P.562, *Analysis and interpretation of INMD voice-service measurements*, May 2004, at http://www.itu.int/rec/T-REC-P.562.

[89] See ITU-T Recommendation J.144, *Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference*, March 2004, at http://www.itu.int/rec/T-REC-J.144.

[90] There are various useful discussions of the factors influencing drive around tests. See:

ETSI EG 202 057-3 V1.1.1, *Speech Processing, Transmission and Quality Aspects (STQ); User related quality of service parameter definitions and measurements; Part 3: quality of service parameters specific to Public Land Mobile Networks (PLMN)*, April 2005, at http://pda.etsi.org/pda/queryform.asp.

ETSI TS 102 250-5 V1.3.1, *Speech Processing, Transmission and Quality Aspects (STQ); quality of service aspects for popular services in GSM and 3G networks; Part 5: Definition of typical measurement profiles*, November 2005, at http://pda.etsi.org/pda/queryform.asp.

[91] For example, 1537 tests, under conventional assumptions about the distributions, are enough to give 95% confidence that a measured percentage of 99.0% corresponds with an "actual" value between 98.5% and 99.5%, a measured percentage of 96.0% corresponds with an "actual" value between 97.0% and 95.0%, and a measured time is within 10% of the "actual" value (if the observed standard deviation is at most twice the observed mean).

[92] The term 'best effort' refers to IP networks that make their best efforts to deliver IP packets but do not guarantee to do so. For the services such as file transfer for which IP was originally devised, these best efforts are satisfactory, because another protocol (TCP) is placed on top of IP to make delivery reliable. However, TCP adds delay to communication, so real time services usually depend on a different protocol (UDP) instead of TCP. As UDP does not make delivery reliable, if too many IP packets are lost real time services become unusable. In this paper the term 'suitable quality' refers to IP networks that allow assurances of quality for real time and other services, by ensuring in particular that packet delay and loss are tolerable. Though 'best effort' is a conventional term, 'suitable quality' is specific to this paper.

[93] However, getting to a state where retail price controls might be abolished needs time and effort: in the UK it took 22 years and entailed separating the management of the retail services and the wholesale services. See Ofcom, *Retail Price Controls Explanatory statement*, July 2006, at http://www.ofcom.org.uk/consult/condocs/retail/statement/rpcstatement.pdf.

[94] See ITU-T Recommendation G.1000, *Communications quality of service: A framework and definitions*, November 2001, at http://www.itu.int/rec/T-REC-G.1000.

[95] See ITU-T Recommendation I.350, *General aspects of quality of service and network performance in digital networks, including ISDNs*, March 1993, at http://www.itu.int/rec/T-REC-I.350.

[96] See ETSI TS 102 250-2 V1.4.1, *Speech Processing, Transmission and Quality Aspects (STQ); quality of service aspects for popular services in GSM and 3G networks; Part 2: Definition of Quality of Service parameters and their computation*, March 2006, at http://pda.etsi.org/pda/queryform.asp.