ETSI TECHNICAL REPORT

ETR 003

August 1990

Source: ETSI TC-NA

Reference: RTR/NA-042102

ICS: 33.080

Key words: quality of service, network performance

Network Aspects (NA); General aspects of quality of service and network performance in digital networks, including ISDN

ETSI

European Telecommunications Standards Institute

ETSI Secretariat

Postal address: F-06921 Sophia Antipolis CEDEX - FRANCE
Office address: 650 Route des Lucioles - Sophia Antipolis - Valbonne - FRANCE
X.400: c=fr, a=atlas, p=etsi, s=secretariat - **Internet:** secretariat@etsi.fr

Tel.: +33 92 94 42 00 - Fax: +33 93 65 47 16

Contents

Blank page

1 Foreword

ETSI Technical Reports (ETRs) are informative documents resulting from ETSI studies which are not appropriate for European Telecommunication Standard (ETS) or Interim - European Telecommunication Standard (I-ETS) status. An ETR may be used to publish material which is either of an informative nature, relating to the use or application of ETSs or I-ETSs, or which is immature and not yet suitable for formal adoption as an ETS or I-ETS.

This ETR has been produced by the Network Aspects (NA) Technical Committee of the European Telecommunications Standards Institute (ETSI).

2 References

[1] CCITT Recommendation G.821 (1984): Error performance of an international digital connection forming part of an integrated services digital network

[2] CCITT Recommendation E.800 (1988): Quality of service and dependability vocabulary

[3] CCITT Recommendation I.211 (1984): Bearer services supported by an ISDN

[4] CCITT Recommendation I.350 (1988): General aspects of quality of service and network performance in digital networks, including ISDN

3 General

3.1 Purpose of this Technical Report

The purpose of this Technical Report is to describe the methods for deriving Quality of Service from Network Performance. The Technical Report entitled "The Relationship between Network Component Performance and Overall Network Performance" complements this document by giving the mechanisms for deriving overall Network Performance from the performance characteristics of individual components.

This Technical Report has been developed to:

- Provide descriptions of Quality of Service and Network Performance,
- Distinguish between Quality of Service and User Opinion,
- Illustrate how the Quality of Service and the Network Performance concepts are applied in digital networks, including ISDNs,
- Describe the features of and the relationship between Network performance, Quality of Service and User Opinion,
- Show how this can provide a feedback mechanism from Quality of Service to Network Performance,
- Indicate and classify performance concerns for which parameters may be needed,
- Identify generic performance parameters.

The generic term "performance" refers to:

- Quality of Service,
- Network Performance

3.2 Description of Quality of Service (QoS) and Network Performance (NP)

3.2.1 Description of Network Performance (NP)

Network Performance is a statement of the performance of the connection element or concatenation of connection elements employed to provide a service. It is defined and measured in terms of parameters which are meaningful to the network provider and are used for the purposes of system design, configuration, operation and maintenance. NP is defined independently of terminal performance and user actions. It is also service independent in that it must be able to support all the services the particular network level is required to transport.

NP is defined in CCITT Recommendation E.800 [2] as:

"The ability of a network or a network portion to provide the functions related to communications between users."

Note: The performance of a network and its component parts contributes to serveability performance and service integrity performance as defined in CCITT Recommendation E.800 [2], and is characterized by a set of measurable and calculable parameters."

3.2.2 Description of Quality of Service (QoS)

Quality of Service is defined in CCITT Recommendation E.800 [2] as follows:

"The collective effect of service performances which determine the degree of satisfaction of a user of the service."

Note: The Quality of Service is characterized by the combined aspects of service support performance, service operability performance, serveability performance, service integrity and other factors specific for each service."

For a given service, QoS is a statement of the performance of the service as offered or specified to the customer. It is defined and measured in terms of parameters which are stated in user understandable language appropriate to the particular service concerned, and which are user verifiable. These parameters will depend upon the service definition, and upon the point at which the service is accessed by the user.

Teleservice QoS parameters describe the QoS required by the user at the H reference point taking into account QoS compensation means provided in the terminals.

Bearer Service QoS parameters describe the QoS required from the network at the T, S or R reference points for the different services.

3.2.3 Description of User Opinion (UO)

User opinion reflects the effect on the user of the QoS. It is a subjective reaction, and can be considered in terms such as dissatisfaction, difficulty or irritation, for example. It is not expected that this should be the subject of an ETS; nevertheless it is an important part of the feedback mechanism (see section 7.5) and as such needs to be considered.

4. Purpose of QoS and NP

4.1 General

Bearer Services and Teleservices as described in the CCITT I.200 Series Recommendations are the objects which network and service providers offer to their customers. A major attribute of these services is the set of QoS parameters which a particular service offers. These parameters are user oriented and take into account the elements involved in a particular service as given in Figure 2 of CCITT Recommendation I.211 [3].

Figure 1 shows that service performance can be considered in two areas: QoS and User Opinions. QoS will have parameters defined and appropriate values specified which can be internationally standardized and be objectively measured.

Bearer Services and Teleservices are supported by a range of Connection Types, each of which comprises several connection elements. The performance of the Connection Types is characterized by a set of NP parameters. These parameters are network oriented.

For any particular service, Figure 2 illustrates how the concepts of QoS and NP are applied in the ISDN environment, giving a general Reference Configuration including also Private Telecommunications Networks (PTNs).

4.2 Multiple service provider environments

It should be recognized when establishing standards that services may be provided by multiple providers. In such an environment different levels of QoS may be supported. Therefore in practice users may experience a variety of ranges of QoS. It is therefore important to establish minimum QoS levels for each service and performance levels for connection elements providing inter-network or international connections.

4.3 Purpose of QoS

A typical user is not concerned with how a particular service is provided, or with any of the aspects of the network's internal design, but only with the resulting service quality. From the user's point of view, Quality of Service is expressed by parameters which:

- focus on user-perceivable effects, rather than their causes within the network

- do not depend in their definition on assumptions about the network internal design,

- take into account all aspects of the service from the user's point of view which can be technically measured at the service access point,

- may be assured to a user at the service access point by the service provider(s),

- are described in network independent terms and create a common language understandable by both the user and the service provider.

4.4 Purpose of NP

A network provider is concerned with the efficiency and effectiveness of the network, in providing services to customers. Therefore, from a network provider's point of view, NP is expressed by parameters which provide information for:

- network development and the possible introduction of new services

- network planning, both nationally and internationally,

- operation and maintenance.

5. Principles of QoS and NP parameters

5.1 General principles

5.1.1 Distinction between QoS and NP parameters

The QoS parameters provide a valuable framework for network design, but they are not necessarily usable in specifying performance requirements for particular connections. Similarly, the NP parameters ultimately determine the QoS, but they do not necessarily describe that quality in a way that is meaningful to users. Both types of parameters are needed, and their values must be quantitatively related if a network is to be effective in serving its users. The definition of QoS and NP parameters should make mapping of values clear in cases where there is not a simple one-to-one relationship between them.

Table 1 shows some of the characteristics which distinguish QoS and NP. Section 5 explores in more detail the process of mapping NP parameters to QoS parameters.

TABLE 1: Distinction Between QoS and NP

Quality of Service	Network Performance
User oriented	Provider oriented
Service attribute	Connection element attribute
Focus on user-observable effects	Focus on planning, development (design), operations and maintenance
Between (at) service access points	End-to-end or network connection elements capabilities

5.1.2 Description of Primary and Derived Performance Parameters

Primary Performance Parameter: A parameter measured on the basis of direct observations of events at service access points or connection element boundaries.

Derived Performance Parameter: A parameter derived from observed values of one or more relevant primary performance parameters and decision thresholds for each relevant primary performance parameter. See Table A.2 in the Annex for an example of this process - in this case the derived performance parameters are accessibility, retainability, network availability and network capability outage duration.

5.1.3 Relationship between Primary and Derived Performance Parameters

The following text is quoted from CCITT Recommendation I.350 [4] and may need further consideration.

A number of event types can be directly observed at service access points or connection element boundaries. Examples of such events are:

- the layer 3 protocol state transition associated to the transfer of a "SETUP" message or a "DISCONNECT" message across a connection element boundary,

- the correct receipt of an information bit (or a specified number of information bits) at an interface.

Parameters related to the time interval between specific events and the frequency of events can be measured. These directly measurable parameters or primary performance parameters describe the QoS (at service access points) or the NP (at connection element boundaries) during periods when the service or connection is available.

Derived performance parameters describe performance based on events which are defined as occurring when the performance of a primary parameter (either singly or in combination) crosses a particular threshold. These derived threshold events identify the transitions between the available and the unavailable states. Parameters related to the time interval between these derived threshold events and their frequency can be identified. These derived performance parameters describe the QoS and the NP for all time intervals; i.e. during periods when the service or connection is available or unavailable.

Note: Primary performance parameters are measured for all time intervals, since the transitions between available and unavailable states depend upon the value of these parameters. However, the values of primary performance parameters would not be specified for a service or connection in the unavailable state.

5.1.4 Measurability of QoS and NP parameter values

Due to the distinction between QoS and NP a number of general points should be noted when considering the development of parameters:

- The definition of QoS parameters should be clearly based on events and states observable at service access points and independently of the network processes and events which support the service;

- The definition of NP parameters should be clearly based on events and states observable at connection element boundaries, such that end-to-end NP can be calculated;

- The use of events and states in the definition of parameters should provide for measurement at the boundaries identified above. Such measures should be verifiable in accordance with generally accepted statistical techniques.

5.2 QoS parameter principles

For the definition of parameters for QoS in the ISDN, the concept of Bearer Services and Teleservices needs to be borne in mind. In particular there is a difference between the kinds of parameters which would describe the QoS of a Bearer Service and that of a Teleservice. This is because the point of observation of, or access to, the service is different in each case which are illustrated in Figure 2.

In the case of Teleservices the interface between the user and the service provider may be a man-machine interface. In case of Bearer Services this interface corresponds to the S/T reference points. As a result, some of the parameters for describing the QoS of a Teleservice will be different from those which describe the QoS of a Bearer Service.

In describing the QoS of Teleservices, the performance of the Terminal Equipment (TE) has to be taken into account. For a Teleservice there should be a mapping between the QoS of the Teleservice and the combination of the customer equipment (including the terminal performance) and the overall (end-to-end) NP of the connection elements supporting this service.

For a Bearer Service there should be a mapping between the QoS of the Bearer Service and the overall (end-to-end) NP of the connection elements supporting this service.

5.3 NP parameter principles

When developing NP parameters the following points should be borne in mind:

- NP parameters must be measurable at the boundary of the network connection element(s) to which they are applied. The definitions should not be based on assumptions about either the internal characteristics of a network (or portions) or the internal causes of impairments observed at the boundaries;

- The division of a network portion into sub-components should only be done if they must be specified separately in order to ensure satisfactory end-to-end performance or, where appropriate, to derive fair and reasonable allocations among providers. No network provider should bear a disproportionate cost in establishing and operating a service.

6. Generic performance parameters

Nine generic primary performance parameters are listed below. These parameters may be used as a basis for developing specific QoS and NP parameters:

- Access Speed

- Access Accuracy

- Access Dependability

- Information Transfer Speed

- Information Transfer Accuracy

- Information Transfer Dependability

- Disengagement Speed

- Disengagement Accuracy

- Disengagement Dependability

Note: Dependability is the performance criterion that describes the degree of certainty (or surety) with which the function is performed regardless of speed or accuracy, but within a given observation interval.

The definition of Dependability in the Supplement No.6 of the CCITT E-Series Recommendations is used there for another concept.

Section 5.1.2 defines derived performance parameters in addition to primary parameters. Derived performance parameters are determined utilizing a function of the primary performance parameter values. CCITT Recommendation G.821 [1] defines one such function, which identifies transitions between available and un available states based on a threshold for severely errored seconds. The generic derived performance parameter associated with such a function is availability.

Examples of specific primary and derived performance parameters for Bearer Service QoS and those for circuit-switched and packet-switched NP are provided in Annex A.

7. The Relationship between NP, QoS and User Opinion

7.1 General

The purpose of this section is to give a conceptual framework for linking Network Performance parameters to Quality of Service parameters and User Opinion. It further employs the concept of independent User Opinion parameters which are intended to enable the development of a coherent strategy for feeding back information in such a way that both planned and achieved Network Performance can be judged and, if necessary, altered.

The concept of independent User Opinion parameters arises from the need to express user opinion in terms which are compatible for all QoS parameters. This permits:

- their combination in order to give an index of overall quality of service,

- the ability to establish thresholds of unacceptability for each QoS parameter, from which knowledge of the network performance levels that give rise to such unacceptability (and hence, unavailability) can be gained.

More detail on User Opinion parameters is given in section 7.4.2.

The framework is summarized in Figure 3 which is intended to show the flow of information from network performance parameters to quality of service parameters, where and how they are measured, and how they are modified by the user as they change from complex technical parameters, understandable only to network providers, to concepts readily understood by all users. It also shows how it is possible to use this information to feed user opinion back in to the network planning process to alter planned performance and/or practical operational standards. It is organized to read downwards, starting within the network, extending outward towards the user, and the text below follows that direction.

7.2 Inside the Network

The first row of the Figure 3 depicts Network Performance parameters. The standards adopted for the levels of performance of these parameters need to be carefully derived as they must be suitable for all services to be carried by the network level under consideration.

7.3 Man-Machine-Interface

7.3.1 QoS

The "event level" NP parameters which allow a particular service to function are in terms which are generally irrelevant and/or meaningless to the user. To rectify this, they are converted into Quality of Service parameters, which are potentially a different set of parameters for each individual service (not only different parameters but different required levels of their performance as well) and suitable for use in a service specification.

7.3.2 QoS planned and achieved

At this stage of the discussion all the parameters have been conceived as performance figures which give a picture of the network operating exactly as planned. However, the operation and reliability of the network will not be perfect, and it is therefore necessary to distinguish between "QoS planned" and "QoS achieved", the latter being measured in the same way and using the same parameter set for each individual service as the former, but with the three important differences of:

i) being measured using the real network;

ii) being able to be confirmed by appropriate user surveys;

iii) reflecting the QoS levels the user actually experiences.

7.4 User

7.4.1 User Opinion (UO)

The performance levels of each of the QoS achieved parameters are what the user actually experiences; however the user's opinion and how he or she reacts are dependent on the degree of disruption caused. Clearly this concept "user opinion" is entirely subjective and can be only measured by user surveys. Assessment of user opinion of QoS parameters can be expressed in terms of the percentage of users who had difficulty with the particular aspects of the connection (%D), the percentage of users who were dissatisfied, or independent user opinion surveys (see 7.4.2 below), for example.

A further aspect of the user opinion concept is that it enables the comparative effects of the planned specification and the achieved performance to be assessed. This can be carried out by %D surveys under laboratory conditions, for example, for both QoS planned and achieved, and confirming the results by comparing with those from the real network.

The user's opinion is the final influence in determining the degree to which the service is performing acceptably. If the service is degraded beyond a certain threshold, the service could be said to be performing unacceptably, and consequently to be unavailable. For example, a policy could be adopted of deciding that if 50% of users experience difficulty at a certain parameter performance level, then that constitutes unacceptable performance and therefore unavailability. By extension, downtime per annum or (for non-voice) failure rate tolerance (with various different definitions of failure, depending on the service) could be surveyed and the value of the 50% difficulty threshold established.

However, it remains difficult to assess the overall performance of a service in terms of %D or user satisfaction, as they are not easily combined, and hence the need for independent parameters has arisen.

7.4.2 Independent User Opinion parameters

The concept of independent user opinion parameters was introduced to provide a universal and generalized method for assessing user opinions. It recognizes that different users of different services may have different reactions to the same degraded performance, but assesses these reactions using uniform parameters, which in principle can be combined to give an overall quality index of a given service.

The parameters chosen are the time lost due to imperfect performance (t), the money lost due to imperfect performance (m), and the irritation caused by that imperfect performance (i). These are considered to represent the three main areas of subjective response which could cause such dissatisfaction that the user could, at or beyond a certain threshold, consider the service to be unavailable.

"t, m and i" can be used to assess the threshold of unavailability when used in combination with %D curves, if, for example, 50% difficulty is considered as the irritation threshold of unavailability. Alternatively calculations can be made for particular degradations to establish time or money lost - it would then be a policy decision as to how much time or money could be lost without constituting unavailability.

7.5 Feedback

The level of t, m and i can be established both for QoS planned and QoS achieved levels: the former should obviously give satisfactory results, whereas if the latter gives results which, for some degradations to particular services, exceed the thresholds , then it is clear that network implementation or operation procedures need modification.

If however, QoS planned also results in t/m/i thresholds being exceeded then the original network performance standards require modification. A great deal of careful work would clearly be necessary to ensure that the root cause of the defective network performance is correctly identified: nevertheless the concepts as outlined above enable a network/service provider to be clear as to the nature of the problems which, from the customer's point of view, require solutions. These feedback concepts are shown in flow chart form at the bottom of Figure 3.

7.6 Examples showing the relationship between NP, QoS parameters and User Opinion

Table 2 gives examples illustrating the derivation of QoS parameters from their NP equivalents.

TABLE 2: Relationship between NP, QoS Parameters and User Opinion

Network:	Basic (64kbit/s switched)		
NP parameter:	Connection Setup Delay Measured in: seconds		
Service:	Telephony	Facsimile	64kbit/s unrestricted
QoS Parameter:	Delay to (network) response	Delay to (network) and (terminal response	Delay to (terminal) response
User Opinion Parameters:	% dissatisfaction with response time t, m, i	% dissatisfaction with response time t, m, i	% dissatisfaction with response time t, m, i
	Example Values		
NP Parameter	5 sec		
QoS parameter	5 sec	5 sec 7 sec	5 sec
User Opinion Parameter	4% dissatisfaction	5% dissatisfaction	5% dissatisfaction
time	1	1	1
money	1	1	1
irritation	1	1	1
NP Parameter	10 sec		
QoS parameter	10 sec	10 sec 12 sec	12 sec
User Opinion Parameter	8% dissatisfaction	10% dissatisfaction	10% dissatisfaction
time	2	2	2
money	2	3	3
irritation	3	4	5

cont. Table 2:

Network: NP parameter:	Basic (64kbit/s switched) Connection Setup Delay Measured in: seconds		
Service:	Telephony	Facsimile	64kbit/s unrestricted
QoS Parameter:	Delay to (network) response	Delay to (network) and (terminal response	Delay to (terminal) response
User Opinion Parameters:	% dissatisfaction with response time t, m, i	% dissatisfaction with response time t, m, i	% dissatisfaction with response time t, m, i
	Example Values		
NP Parameter	20 sec		
QoS parameter	20 sec	20 sec 22 sec	20 sec
User Opinion Parameter	30% dissatisfaction	45% dissatisfaction	50% dissatisfaction*
time	5	5	5
money	6	8	10*
irritation	8	10*	12*
NP Parameter	50 sec		
QoS parameter	50 sec	50 sec 52 sec	52 sec
User Opinion Parameter	95% dissatisfaction*	98% dissatisfaction*	99% dissatisfaction*
time	10*	10*	10*
money	12*	18*	20*
irritation	15*	20*	40*

Notes to Table 2:

Note 1: Each of these tables only represents one path from an NP parameter to QoS parameters - other NP parameters can produce the same results in QoS terms. For example, short interruptions could cause the same problems to facsimile and data transmission as does impulsive noise.

Note 2: The time/money/irritation (t/m/i) parameters are in arbitrary units and are assigned arbitrary values for the purpose of example only. A value of zero represents perfection; the higher value the worse the service perception. The different ways in which different services react to the same degree of network performance degradation can thus clearly be identified, and, in particular, the crossing of thresholds of acceptability can be observed. In this example, whenever a parameter reaches 10 it is deemed unacceptable.

Another aspect of this stage is that when all QoS parameters are treated similarly the total t/m/i count can be assessed simply by adding the values together. There is then a global quality index for the particular service.

* The t/m/i parameter has crossed the threshold of unacceptability. The service could be deemed unavailable.

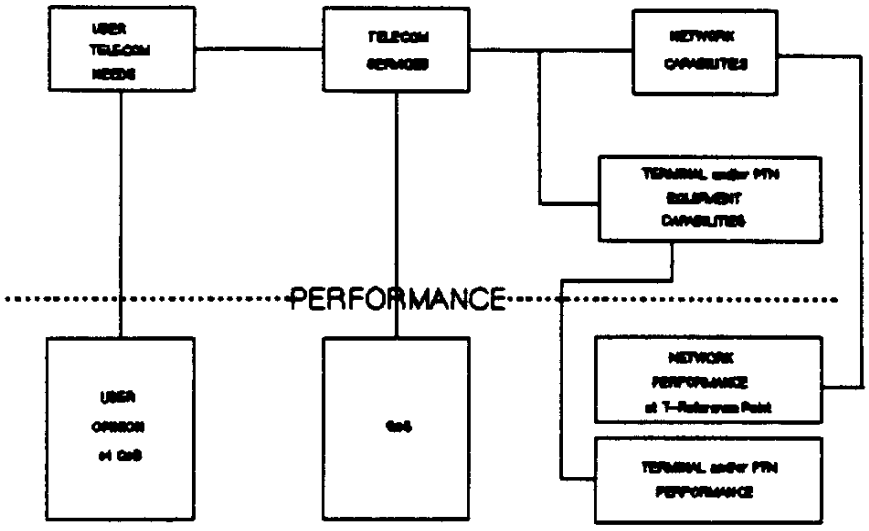

FIGURE 1: Major Areas in Telecommunications

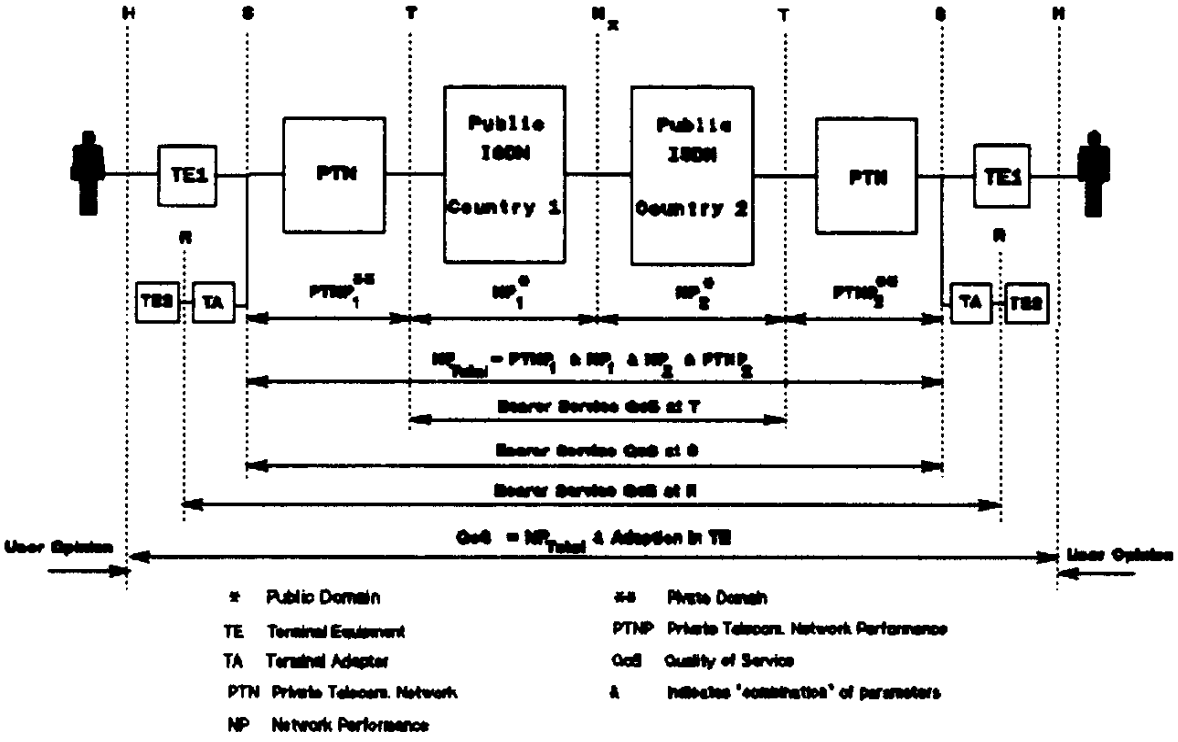

Figure 2: General Reference Configuration for NP and QoS for a specific Service

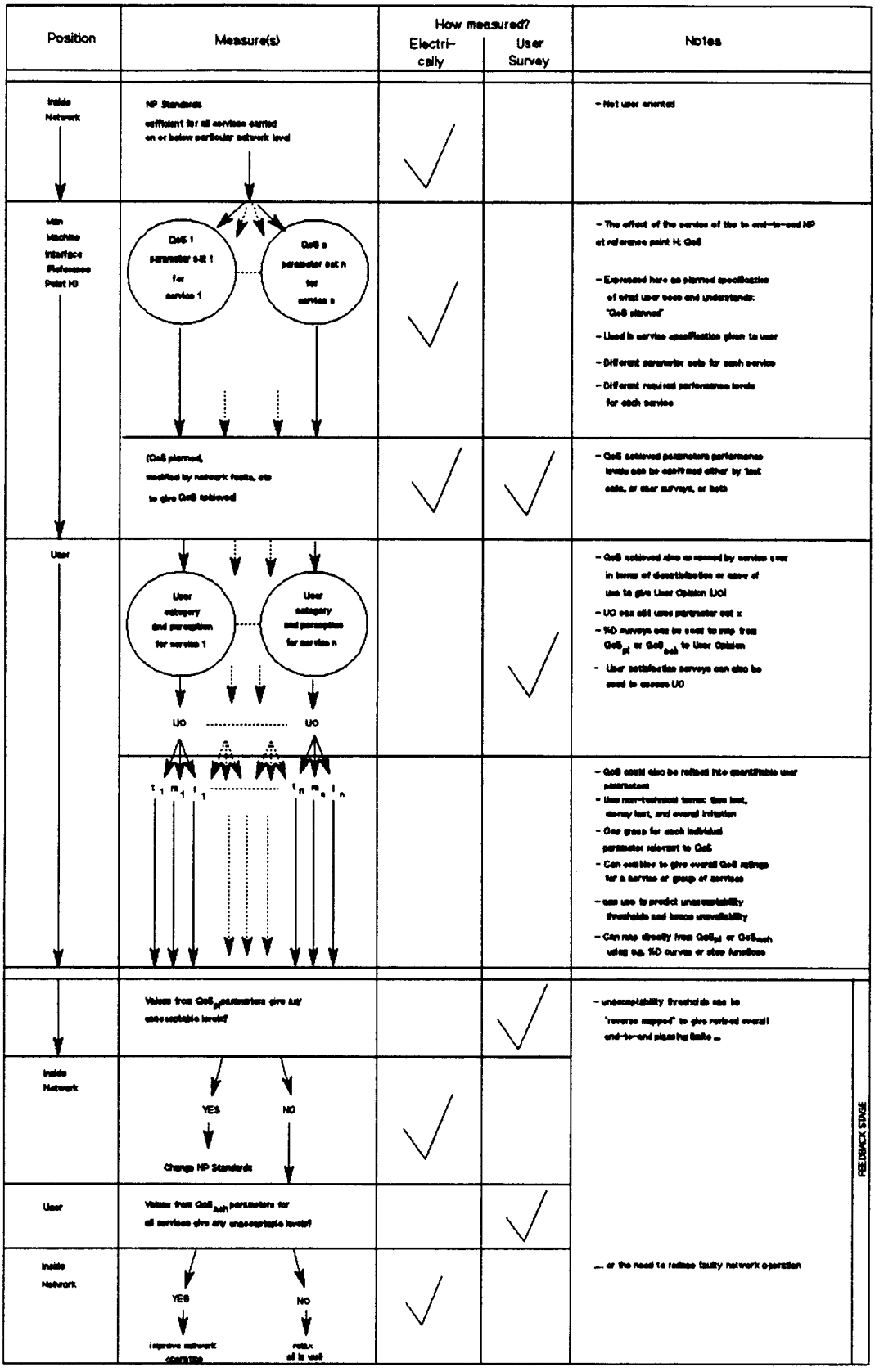

Figure 3: Conceptual Framework for mapping NP to QoS to UO, and feeding back to NP

Annex A: Relationship between generic and possible specific QOS and NP parameters

This Annex illustrates the qualitative relationship between the generic parameters defined in Section 4. and a candidate set of specific QoS and NP parameters. Tables A-1, A-2 and A-3 illustrate the relationship between the generic parameters and specific bearer service QoS, circuit switched NP and packet-switched NP parameters, respectively.

TABLE A-1: Qualitative Relationship between Generic Performance Parameter and Candidate Bearer Service QoS Parameters

	Generic Parameters / Bearer Service QoS Parameters	Primary Performance Parameters												Derived Performance Parameter		
		Access Delay	Incorrect Access Probability	Access Denial Probability	User Information Transfer Delay	User Information Transfer Rate	User Information Error Probability	Extra User Information Delivery Probability	User Information Misdelivery Probability	User Information Loss Probability	Disengagement Delay	Incorrect Disengagement Probability	Disengagement Denial Probability	Serviceability Performance	User Information Transfer Detail Probability	Service Outage Duration
Primary	Access Speed	X														
	Access Accuracy		X													
	Access Dependability			X												
	Information Transfer Speed				X	X										
	Information Transfer Accuracy						X	X	X							
	Information Transfer Dependability									X						
	Disengagement Speed										X					
	Disengagement Accuracy											X				
	Disengagement Dependability												X			
Derived	Availability													X	X	X

TABLE A-2: Qualitative Relationship between Generic Performance Parameters and Candidate Circuit-Switched NP Parameters

Generic Parameters \ Circuit Switched NP Parameters		Primary Performance Parameters												Derived Performance Parameters			
		Connection Set-up Delay	Alerting Delay	Connection Set-up Error Probability	Connection Refusal Probability	Propagation Delay	Degraded Minutes	Severely Errored Seconds	Errored Seconds	Disconnect Delay	Release Delay	Premature Disconnect Probability	Connection Clearing Refusal Probability	Network Capability Outage Duration	Network Reliability	Accessibility	Retainability
Primary	Access Speed	X	X													X	
	Access Accuracy			X												X	
	Access Dependability				X											X	
	Information Transfer Speed					X											
	Information Transfer Accuracy						X	X	X							X	X
	Information Transfer Dependability															X	
	Disengagement Speed									X	X						X
	Disengagement Accuracy											X					X
	Disengagement Dependability												X				X
Derived	Availability													X	X	X	X

TABLE A-3: Qualitative Relationship between Generic Performance Parameters and Candidate Packet-Switched NP Parameters

Generic Parameters \ Packet Switched NP Parameters		Primary Performance Parameters											Derived Performance Parameters	
		Virtual Circuit Setup Delay(s)	Virtual Circuit Setup Denial Probability	Data Packet Transfer Delay	Throughput Capacity	Residual Error Ratio	Reset Probability	Reset Duration Probability	Virtual Circuit Clearing Delay	Virtual Circuit Clearing Denial Probability	Virtual Circuit Premature Disconnect Probability	Virtual Circuit Premature Disconnect Duration Probability	Network Capability Outage Duration	Network Availability
Primary	Access Speed	X												
	Access Accuracy		X											
	Access Dependability			X										
	Information Transfer Speed				X	X								
	Information Transfer Accuracy						X	X	X					
	Information Transfer Dependability						X	X	X					
	Disengagement Speed									X				
	Disengagement Accuracy													
	Disengagement Dependability									X	X	X		
Derived	Availability												X	X

History

Document history	
August 1990	First Edition
April 1996	Converted into Adobe Acrobat Portable Document Format (PDF)