INTERNATIONAL TELECOMMUNICATION UNION

CCITT

X.140

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

(09/92)

DATA COMMUNICATION NETWORKS

NETWORK ASPECTS

GENERAL QUALITY OF SERVICE PARAMETERS FOR COMMUNICATION VIA PUBLIC DATA NETWORKS

Recommendation X.140

FOREWORD

The CCITT (the International Telegraph and Telephone Consultative Committee) is a permanent organ of the International Telecommunication Union (ITU). CCITT is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The Plenary Assembly of CCITT which meets every four years, establishes the topics for study and approves Recommendations prepared by its Study Groups. The approval of Recommendations by the members of CCITT between Plenary Assemblies is covered by the procedure laid down in CCITT Resolution No. 2 (Melbourne, 1988).

Recommendation X.140 was revised by Study Group VII and was approved under the Resolution No. 2 procedure on the 10th of September 1992.

____

CCITT NOTES

1) In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized private operating agency.

2) A list of abbreviations used in this Recommendation can be found in Annex D.

Recommendation X.140

GENERAL QUALITY OF SERVICE PARAMETERS FOR COMMUNICATION VIA PUBLIC DATA NETWORKS

(Malaga-Torremolinos, 1984; amended at Melbourne, 1988; revised 1992)

The CCITT,

considering

(a) that users of data transmission services need general parameters which express their quality of service requirements without reference to any particular service or the means of its provision;

(b) that providers of data transmission services need similar general parameters for representing offered services, and for relating user quality of service requirements to network performance capabilities;

(c) that Recommendations X.130 and X.131 define protocol-specific performance parameters and objec-tives for circuit-switched public data networks;

(d) that Recommendation X.134 specifies portion boundaries and packet-layer reference events for defining packet-switched performance parameters;

(e) that Recommendations X.135, X.136 and X.137 define protocol-specific performance parameters and values for packet-switched public data networks;

(f) that Recommendation X.200 defines the Reference Model of Open Systems Interconnection (OSI) for CCITT applications;

(g) that Recommendation X.213 defines the OSI network service;

(h) that Recommendation X.300 defines general principles and arrangements for interworking among public data networks, and between public data networks and other networks,

unanimously recommends

that the general parameters defined below be used in specifying the end-to-end quality of public data network services as seen from the user's point of view.

1 Scope and application

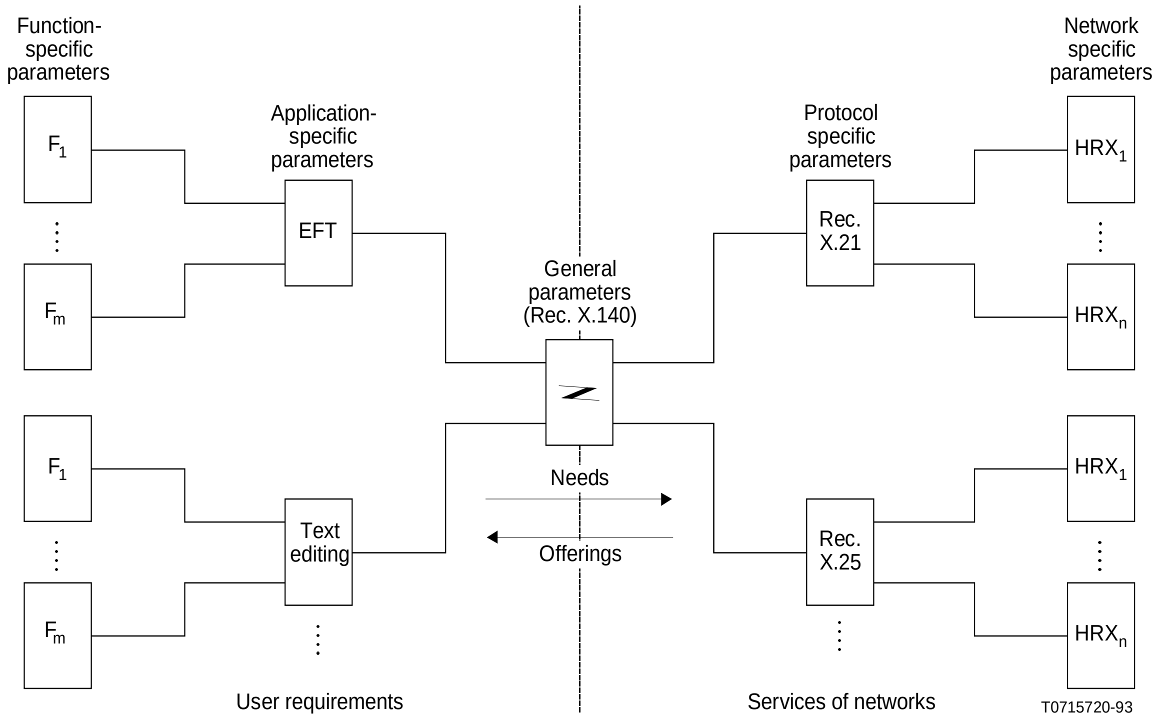
1.1 This Recommendation defines a set of general quality of service (QOS) parameters for public data networks (PDNs). The parameters have two essential characteristics:

1) they focus on performance *effects* which are observable at the network interfaces, rather than their causes within the network; and

2) their definitions are based on protocol-independent events (e.g. access request) rather than protocol-specific interface events (e.g. issuance of an X.21 *call request* signal).

These characteristics make the parameters independent of application, network, and service. With proper specialization, they may be used to specify or measure the quality of any data communication service, irrespective of network internal design or network access protocol. Examples of data communication services to which the parameters may be applied are circuit-switched services, packet-switched services, and leased circuit services. The parameters are applicable to both connection-oriented and connectionless services.

1.2 The general quality of service parameters defined in this Recommendation are specifically designed to be used in relating the performance capabilities of particular network services with user requirements (see Figure 1/X.140). The network specific performance parameters defined in other X-Series Recommendations are focused on specific service interface protocols (e.g. X.21, X.25) and specific network configurations (e.g. X.92, X.110). They are essential for network design and operation and component performance specification, but are not necessarily understandable or relevant to users. Similarly, performance requirements of users are often focused on particular applications (e.g. electronic funds transfer (EFT), text editing) and may not be directly useful to network providers. An example is the data processing parameter "response time". The general parameters provide a "common language" for relating the two. They enable users to specify communication requirements without presupposing any particular service, network, or protocol, and enable providers to describe service performance in terms that are relevant to users, but not specialized to any particular application.



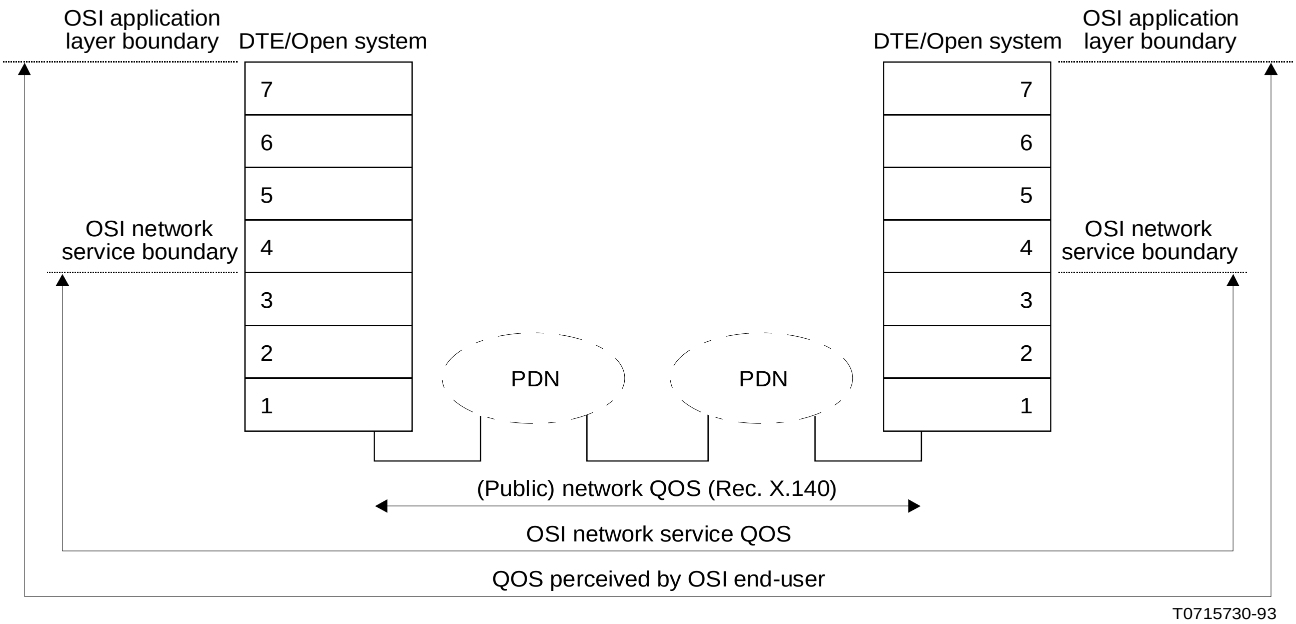
EFT Electronic funds transfer
HRX Hypothetical reference connection

FIGURE 1/X.140

"Common language" function of the general QOS parameters

1.3 The general parameters are principally intended to describe communication performance at interfaces between public data networks and customer DTEs. The detailed characteristics of such network-user interfaces depend on the type of network service and the user application. The quality of X.21-based circuit-switching networks is described in terms of signals occurring at the physical DTE/DCE interfaces (e.g. *call request*, *incoming call*). The quality of X.25-based packet-switching networks is described in terms of corresponding events (or state transitions) occurring at the packet layer of X.25. Specific relationships between the X.140 parameters and the circuit-switching and packet-switching network performance parameters defined in the X.130-Series Recommendations are described in Annexes A and B, respectively.

1.4 Many applications of public data networks will conform to the Reference Model of Open Systems Interconnection for CCITT applications (Recommendation X.200). In that model, QOS parameters are defined at abstract boundaries between layers. Public data networks provide support for the OSI network service (see Recommendation X.213). The QOS parameters defined in the OSI network service reflect those aspects of public data network quality of service that are observable and significant to OSI network service users. The general relationship between PDN quality of service and OSI network service quality is illustrated in Figure 2/X.140. Specific relationships between the X.140 QOS parameters and the network layer QOS parameters are defined in Annex C. Relationships among the general parameters, the X.130-Series parameters, and the OSI network service performance QOS parameters are summarized in Figure 3/X.140.



Note 1 – QOS parameters and values for specific PDN's (CSPDN, PSPDN) are specified by separate Recommendations.

Note 2 – The signalling of QOS information in various interworking situations is not the subject of this Recommendation.

Note 3 – Interfaces are defined in § 1.3.

Note 4 – Application to private networks may be possible at the discretion of individual Administrations.

FIGURE 2/X.140

QOS relationships in an OSI context

1.5 Because the X.140 parameters are based on protocol-independent events, they may also be applied at higher layers in the OSI model. Application of the X.140 parameters at the OSI end user interfaces is illustrated in Figure 2/X.140. Details of parameter specialization, relationships to application-specific parameters, and the mapping of end user QOS values into corresponding lower layer values are for further study.

1.6 A requirement also exists to describe QOS at higher layers (above the network layer) in non-OSI applications. An example is the X.28/X.29 PAD facility. Use of the X.140 parameters to express QOS characteristics in such applications, and possible relationships to PDN QOS parameters, are also for further study.

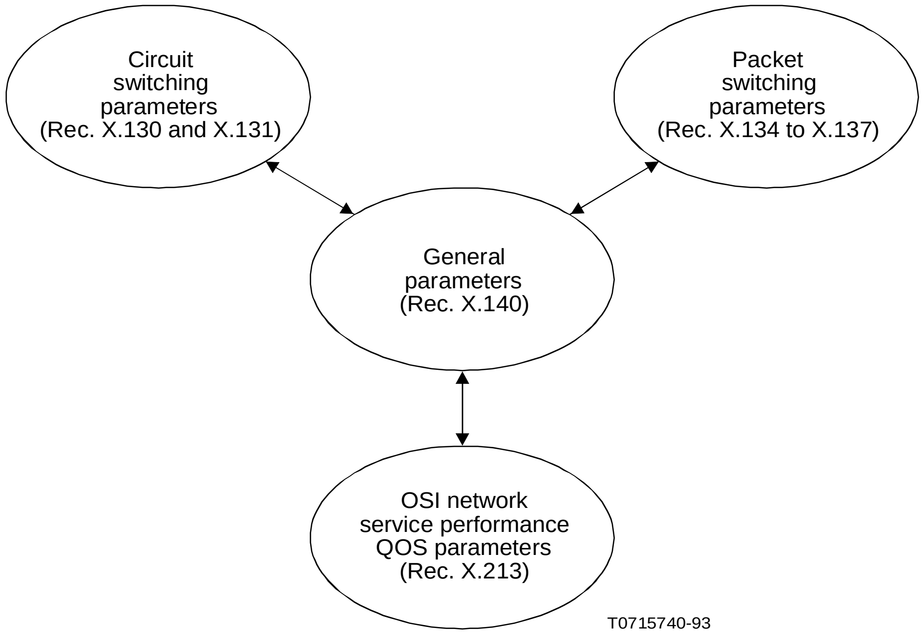


FIGURE 3/X.140

Relationships among the general parameters, the X.130-Series parameters, and the OSI network service performance QOS parameters

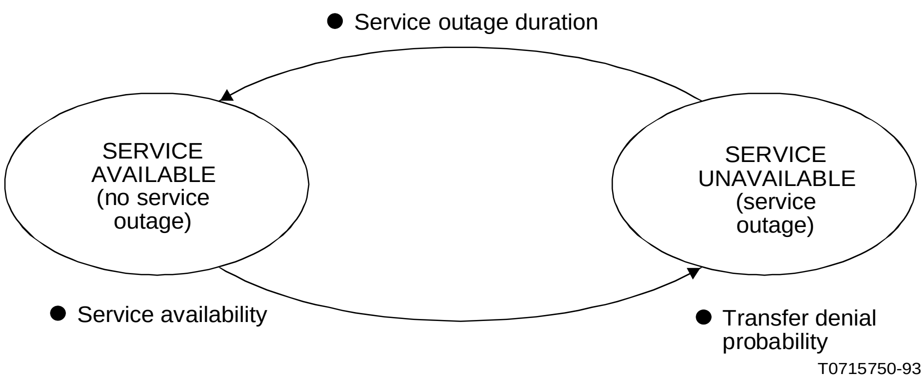
1.7 Some public networks will have the capability to signal QOS requests and conditions, or to permit users to "negotiate" certain QOS characteristics of the network. This Recommendation does not describe any public data network capabilities of this kind, nor does it specify how they might be used. The provision and use of such capabilities will be the subject of other Recommendations (for example, Recommendations describing how public data networks may support the OSI network service). In the case of interworking between networks, such capabilities are described in Recommendation X.300.

1.8 This Recommendation does not specify values for the general QOS parameters. Values may be specified either by the service user, in characterizing a particular data communication requirement, or by the service provider, in characterizing a particular service offering. Values may be measured by either the users or the provider.

1.9 To ensure comparability, stated values for the general parameters should be accompanied by supplementary information which clearly identifies their scope of application and statistical meaning. User delays may be "factored out" of stated delay and transfer rate values using the method defined below. The same method can be used to factor out provider delays in cases where an assessment of user performance is desired.

1.10 Figure 4/X.140 identifies the general QOS parameters defined in this Recommendation. The parameters are of two types: primary parameters and availability parameters. The primary parameters describe performance during periods of normal service operation, in the absence of service outages. The availability parameters describe the frequency and duration of service outages.

Function / Criterion	Speed	Accuracy	Dependability
Access	• Access delay	• Incorrect access probability	• Access denial probability
User information transfer	• 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	• Disengagement delay	• Disengagement denial probability	



b) Availability parameters

FIGURE 4/X.140

Summary of user-oriented QOS parameters

1.11 Three protocol-independent data communication functions are considered in defining the primary parameters: access, user information transfer, and disengagement. These general functions correspond to connection set-up, data transfer, and connection clearing in connection-oriented services. They are also applicable to connectionless services. Each function is considered with respect to three general performance concerns (or "performance criteria"): speed, accuracy, and dependability. These express, respectively, the delay or rate, degree of correctness, and degree of certainty with which the function is performed.

1.12 An associated two-state model provides a basis for describing overall service availability. A specified availability function compares the values for a subset of the primary parameters with corresponding outage thresholds to classify the service as "available" (no service outage) or "unavailable" (service outage) during scheduled service time. The availability parameters characterize the resulting binary random process.

1.13 The remainder of this Recommendation is comprised of three sections. Section 2 defines the set of user-oriented QOS parameters. Section 3 describes a method of separating delays into user and network components and determining "responsibility" for timeout performance failures. Section 4 specifies supplementary information which should be provided in conjunction with any statement of parameter values.

2 Parameter definitions

This section provides definitions for the fourteen user-oriented QOS parameters.

2.1 access parameters

Performance of the access function is described by three parameters: access delay, incorrect access probability, and access denial probability.

2.1.1 *Access delay*

Access delay is the value of elapsed time between an access request and successful access.

An access request is any interface signal that notifies the network of a user's desire to initiate a data communication session.

Elapsed time values are calculated only on access attempts that result in successful access. The successful access outcome is indicated in one of two ways:

1) by network issuance of a *ready for data* or equivalent signal to the calling user before access timeout, in networks that provide such a signal; or
2) by the fact that at least one bit of user information is input to the system before access timeout, in networks that do not provide a *ready for data* or equivalent signal. In connection-oriented services, there is the additional requirement that the intended called user must have been contacted and committed to the data communication session during the access attempt. This requirement distinguishes successful access outcomes from incorrect access outcomes, as discussed in § 2.1.2 below.

Access delay is divided into user-dependent and network-dependent components. Values for the network-dependent components are specified in network-specific Recommendations (e.g. Recommendation X.135).

2.1.2 *Incorrect access probability*

Incorrect access probability is the ratio of total access attempts that result in incorrect access to total access attempts in a specified sample.

Incorrect access is essentially the case of a "wrong number". It occurs when the network establishes a physical or virtual circuit connection to a user other than the one intended by the call originator, and then does not correct the error before the start of user information transfer. Incorrect access can only occur in connection-oriented services, since the network does not establish a connection between users in connectionless services. Incorrect access is distinguished from successful access (in connection-oriented services) by the fact that the intended called user is not contacted and committed to the data communication session during the access attempt. Values for network-specific parameters corresponding to incorrect access probability are contained in network-specific Recommendations (e.g. X.136).

2.1.3 *Access denial probability*

Access denial probability is the ratio of total access attempts that result in access denial to total access attempts in a specified sample.[1)]

Access denial (also termed network blocking) can occur in two ways:

1) the network issues a blocking signal to the originating user during the access period (preventing the start of user information transfer); or
2) the network delays excessively in responding to user actions during the access period, with the result that user information transfer is not initiated before access timeout. Access denial is distinguished from service outage by the fact that some active response (i.e. interface signal) is issued by the network during the access attempt.

[1)] This ratio and all other probability ratios defined in this Recommendation are actually *estimates* of the true probability values.

An access attempt can also fail as a result of user blocking. Such failures are excluded from network performance measurement. User blocking is defined as any case where an access attempt fails as a result of incorrect performance or non-performance on the part of a user. Examples of user blocking include the following:

a) either the originating or the called user issues a termination (or blocking) signal to the network during the access period (preventing the start of the user information transfer); or

b) the originating or the called user delays excessively in responding to network actions during the access period, with the result that user information transfer is not initiated before access timeout. An example of the latter is the case where the called user does not answer an incoming call.

Access timeout occurs (i.e. an access attempt is considered to have failed for performance assessment purposes) whenever the duration of an individual access attempt exceeds a specified value. A procedure for distinguishing access denial from user blocking is described in § 3. Values for network-specific parameters corresponding to access denial probability are contained in network-specific Recommendations (e.g. X.136).

Note – Delay to access denial is not included as a parameter because its effect on the users is considered to be insignificant.

2.2 **user information transfer parameters**

Performance of the user information transfer function is described by six parameters: user information transfer delay, user information transfer rate, user information error probability, extra user information delivery probability, user information misdelivery probability, and user information loss probability.

2.2.1 *User information transfer delay*

User information transfer delay is the value of elapsed time between the start of transfer and successful transfer of a specified user information unit (e.g. block).

The start of user information unit transfer occurs, for any given user information unit, when two conditions have been met:

1) all bits in the unit are physically present within the network facility; and

2) the network has been authorized to transmit them. Authorization may either be an explicit user action (e.g. typing carriage return at a buffered CRT terminal) or an implicit part of inputting the user information itself (e.g. typing a single character at an unbuffered asynchronous terminal).

The successful transfer outcome is declared (on end of transfer) when an information unit is transferred from the source user to the intended destination user within the specified transfer timeout period, and the delivered unit has exactly the form and content intended by the source. The form or content of an information unit successfully delivered to a destination user may differ from that input by the source if desired conversions are performed within the network.

The end of user information unit transfer records the output of user information units to the destination user in essentially the same way as the start of transfer records their input at the source. It is defined to occur when:

a) all bits in the unit are physically present within the destination user facility; and

b) the destination user has been notified that the information is available for use. The notification may be explicit or implicit.

The user information unit used in defining user information transfer delay is a contiguous group of user information bits delimited at the source user-network interface for transfer to a destination user as a unit. The specific number of bits in such a unit may be defined by the provider in specifying an offered service, or by the user in specifying a service requirement. User information transfer delay is divided into user-dependent and network-dependent components. Values for the network-dependent components are specified in network-specific Recommendations (e.g. X.135).

2.2.2 *User information transfer rate*

User information transfer rate is the total number of successfully-transferred user information units in an individual transfer sample divided by the input/output time for that sample.

The input/output time for a transfer sample is the larger of the input time or the output time for that sample (see Figure 5/X.140). The sample input time begins when the transfer sample (defined above) begins, and ends when either:

1) all digits in the sample have been input to the network, and the network has been authorized to transmit them; or

2) sample input/output timeout occurs.

The sample output time begins when the first user information digit in the sample is delivered by the network to the destination user. It ends when either:

– the last digit of user information in the sample is delivered to the destination user; or

– sample input/output timeout occurs.

As noted earlier, either the input or the output of a transfer sample may be delayed excessively by a user (rejected sample). Such failures are excluded from network performance measurement. As in the case of user information transfer denial probability, rejected samples are distinguished from valid transfer samples using the procedure described in § 3.

Note – A "maximum user information transfer rate" which excludes the effect of user input/output delays can be calculated using the procedure described in § 3. Values for network-specific parameters corresponding to user information transfer rate are contained in network-specific Recommendations (e.g. X.135).

2.2.3 *User information error probability*

User information error probability is the ratio of total incorrect user information units to total successfully transferred user information units *plus* incorrect user information units in a specified sample.

A transferred user information unit is defined to be an incorrect user information unit when the value of one or more digits in the unit is in error, or when some, but not all, digits in the unit are lost digits or extra digits (i.e. digits that were not present in the original signal).

Bit error ratio is a limiting case of user information error probability in which the user information unit length, on which the error performance is based, is a single binary digit.

The proportion of errored seconds is a particular case of user information error probability in which the user information unit length is defined as one second. The number of digits contained in each user information unit in this case is numerically equal to the digit rate per second. This parameter is usually expressed in the form of the percentage of its complement, i.e. as a percentage of error-free seconds (% EFS). A similar parameter, the percentage of error-free deciseconds (% EFdS), can be defined based on a user information unit length of 100 ms.

Values for network-specific parameters corresponding to user information error probability are contained in network-specific Recommendations (e.g. X.136).

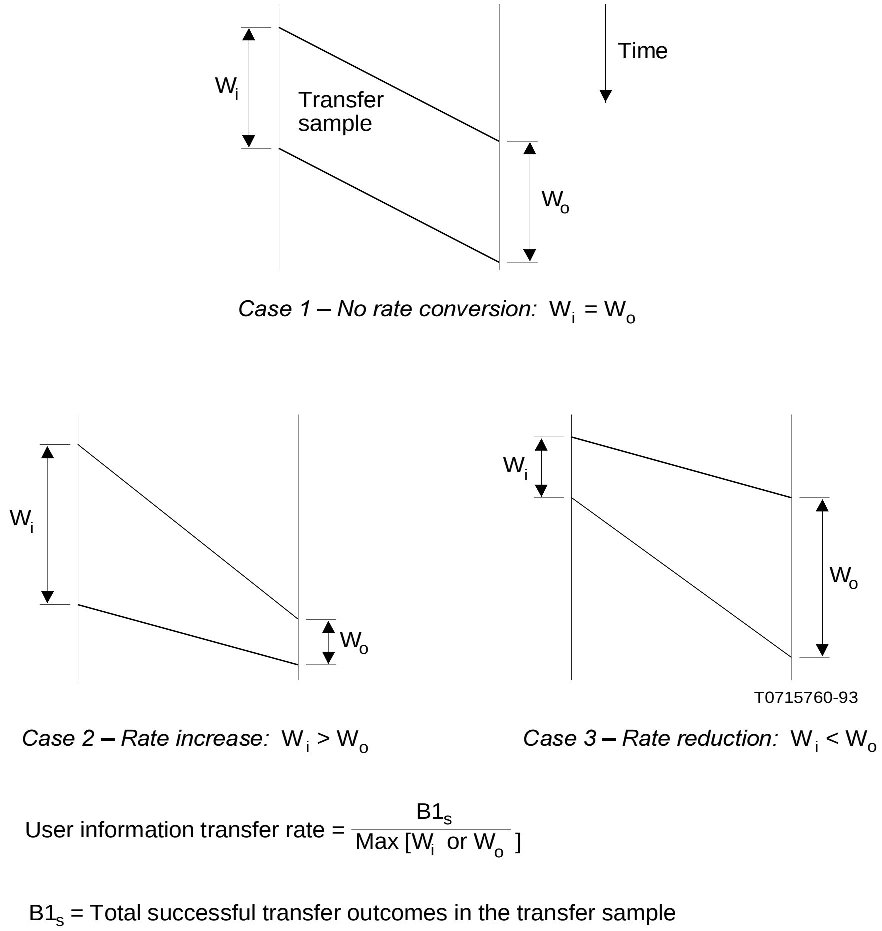


FIGURE 5/X.140

User information transfer rate

2.2.4 *Extra user information delivery probability*

Extra user information delivery probability is the ratio of total (unrequested) extra information units to total information units received by a destination user in a specified sample.

An information unit received by a particular destination user is declared to be an extra information unit when none of the bits in the unit were input to the system by the source user for delivery to that destination. Unless misdelivered user information units are explicitly identified in a measurement process, they will be counted as extra information units. Values for network-specific parameters corresponding to extra user information delivery probability are contained in network-specific Recommendations (e.g. X.136).

2.2.5 *User information misdelivery probability*

User information misdelivery probability is the ratio of total misdelivered user information units to total user information units transferred between a specified source and destination user in a specified sample.

A misdelivered user information unit is a user information unit transferred from a source user to a particular destination user that was actually intended for delivery to a different destination user. It is considered inconsequential whether the information is correct or incorrect in content. Values for network-specific parameters corresponding to user information misdelivery probability may be contained in network-specific Recommendations (under study).

2.2.6 *User information loss probability*

User information loss probability is the ratio of total lost user information units to total transmitted user information units in a specified sample.

A transmitted user information unit is declared to be a lost user information unit when none of the bits in the unit are delivered to the intended destination user within the specified timeout period, and the network is responsible.

User information may also remain undelivered as a result of user information refusal – i.e. non-delivery attributable to excessive delay on the part of a user. An example is a destination user’s exercise of flow control. Such outcomes are excluded from network performance measurement.

Transfer timeout occurs (i.e. a transfer attempt is considered to have failed for performance assessment purposes) whenever the duration of an individual transfer period exceeds a specified value. A procedure for distinguishing user information loss from user information refusal is described in § 3. Values for network-specific parameters corresponding to user information loss probability are contained in network-specific Recommendations (e.g. X.136).

2.3 **disengagement parameters**

Performance of the disengagement function is described by two parameters: disengagement delay and disengagement denial probability.

2.3.1 *Disengagement delay*

Disengagement delay is the value of elapsed time between the start of a disengagement attempt for a particular user and successful disengagement of that user.

The disengagement request notifies the system of a user’s desire to terminate an established data communications session. It is complementary to the access request in most networks.

Elapsed time values are calculated only on disengagement attempts that result in successful disengagement. The successful disengagement outcome is indicated in one of two ways:

1) by network issuance of a *clear confirmation* or equivalent signal to the requesting user before disengagement timeout, in networks that provide such a signal; or

2) by the fact that the user is able to initiate a new access before disengagement timeout, in networks that do not provide a *clear confirmation* or equivalent signal.

Disengagement delays may be defined independently for each participating user when significantly different values are expected. Disengagement delays are divided into user-dependent and network-dependent components. Values for the network-dependent components are specified in network-specific Recommendations (e.g. X.135).

2.3.2 *Disengagement denial probability*

Disengagement denial probability is the ratio of total disengagement attempts that result in disengagement denial to total disengagement attempts in a specified sample.

The disengagement denial outcome is indicated in one of two ways:

1) by the absence of a *clear confirmation* or equivalent signal within the disengagement timeout period (in networks that provide such a signal); or

2) by the inability of the user to initiate a new access within the specified disengagement timeout period (in networks that do not provide a *clear confirmation* or equivalent signal).

In some networks, a disengagement attempt can also fail as a result of user disengagement blocking. User disengagement blocking is defined as any case where a disengagement attempt fails as a result of incorrect performance or non performance on the part of a user. Examples of user disengagement blocking include the following:

1) a user issues a disengagement blocking signal to the network during the disengagement period (preventing the termination of a connection-oriented data communication session); or

2) a user delays excessively in responding to network actions during the disengagement period, with the result that disengagement is not completed before disengagement timeout.

   Such failures are excluded from network performance measurement.

Disengagement timeout occurs (i.e. a disengagement attempt is considered to have failed for performance assessment purposes) whenever the duration of an individual disengagement attempt exceeds a specified value. A procedure for distinguishing disengagement denial from user disengagement blocking is described in § 3. Values for network-specific parameters corresponding to disengagement denial probability are contained in network-specific Recommendations (e.g. X.136).

2.4 **availability parameters**

Three parameters are defined to describe overall service availability: service availability, user information transfer denial probability, and service outage duration.

2.4.1 *Service availability*

Service availability is the ratio of the aggregate time during which satisfactory or tolerable service is, or could be, provided to the total observation period.

In practice, the observation period may consist of several non-contiguous smaller time intervals. The time during which satisfactory or tolerable service is available includes all time that is not within the service outage duration as defined above. The criteria by which the service is judged to be unacceptable are for further study. Such study will embrace consideration of the parameters, including call-related events, that are relevant, and consideration of the observation period(s) and performance thresholds for unacceptability. Values for the network-specific parameter(s) corresponding to service availability are contained in network-specific Recommendations (e.g. X.137).

2.4.2 *User information transfer denial probability*

User information transfer denial probability is the ratio of total transfer denials to total transfer samples during a specified observation period.

A transfer sample is a discrete observation of network performance in transferring user information between a specified source and destination user. A transfer sample begins on input of a selected user information digit at the source user interface, and continues until the outcomes of a given number of transfer attempts have been determined.

A transfer denial is a transfer sample in which the observed performance is worse than a specified minimum acceptable level. Transfer denials are identified by comparing the measured values for four supported quality of service parameters with specified transfer denial thresholds. The four supported parameters are user information error probability, user information loss probability, extra user information delivery probability, and user information transfer rate. Transfer denial includes cases where the network unilaterally terminates user data transmission (e.g. reset or clearing due to network congestion).

A transfer sample may also indicate performance worse than the minimum acceptable level if:

1) the source or destination user intentionally disengages during the sample transfer period; or

2) a user delays excessively in inputting or accepting the sample data (e.g. through exercise of flow control). Such failures (called rejected samples) are excluded from network performance measurement.

A transfer sample input/output timeout occurs (i.e. a transfer sample is considered to have failed for performance assessment purposes) whenever the duration of an individual sample input or output period exceeds a specified value. A procedure for distinguishing a transfer denial from a rejected sample is described in § 3. Values for network-specific parameters corresponding to user information transfer denial probability are contained in network-specific Recommendations (e.g. X.137).

2.4.3 *Service outage duration*

Service outage duration is the duration of any continuous period of time for which satisfactory or tolerable service is not available. It is recognized that the determination of an outage condition requires a finite observation period.

A service outage includes any period during which the user is unable or would be unable to elicit any response from the network; i.e. the network is "dead". It also includes any period during which the service provided by the network is unacceptable because of, for example, poor error performance or throughput. The criteria by which the service is judged to be unacceptable are for further study. Such study will embrace consideration of the parameters, including call-related events, that are relevant, and consideration of the observation period(s) and performance thresholds for unacceptability. Values for the network-specific parameter(s) related to service outage duration are contained in network-specific Recommendations (e.g. X.137).

3 Distinguishing network and user components of performance

This paragraph describes a method of separating delays into network and user components and determining "responsibility" for timeout performance failures. This is accomplished by dividing selected performance periods into alternating "responsibility intervals" of two types:

1) intervals in which the network is responsible for creating the next event in a sequence of interface events leading to the accomplishment of a specified data communication function (e.g., access);

2) intervals in which a user is responsible for creating the next event in such a sequence.

A simple illustration of this concept is provided in Figure 6/X.140. The four interface events in a typical connection establishment sequence divide the connection establishment period into three responsibility intervals: two network-dependent intervals surrounding one user-dependent interval. User responsibility intervals must normally be "factored out" in specifying network performance objectives, since their durations are not under network control.

Figure 7/X.140 illustrates the responsibility transfer concept in more detail. Two general types of responsibility transfer events are identified and are defined below. Both are defined with respect to particular data communication functions and associated performance periods, which are defined in § 3.3.

3.1 *Network-user responsibility transfer*

A network-user responsibility transfer occurs upon issuance of any interface signal that:

1) initiates user activity needed to accomplish a specified function;

2) solicits a subsequent user response indicating that the required activity has been completed; and

3) suspends network activity on the function pending the expected response. Examples are network issuance of an *incoming call* signal (in Recommendation X.21) or packet (in Recommendation X.25) to a called user.

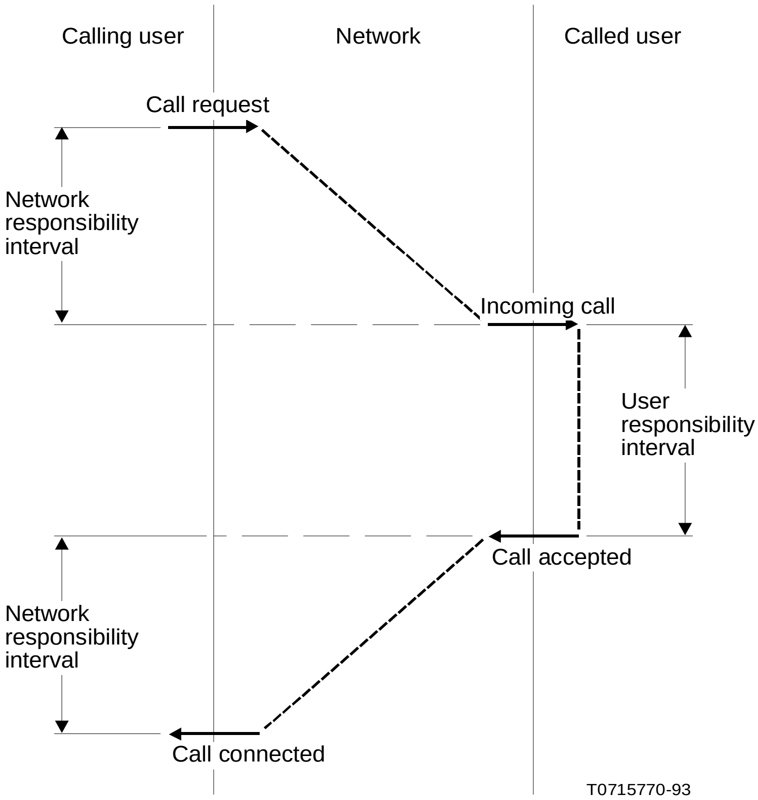


FIGURE 6/X.140
Illustration of network-user responsibility intervals

3.2 *User-network responsibility transfer*

A user-network responsibility transfer occurs on issuance of any interface signal that:

1) initiates network activity needed to accomplish a specified function;

2) solicits a subsequent network response indicating that the required activity has been completed; and

3) suspends user activity on the function pending the expected response. Examples are user issuance of *call request* and *call accepted* signals and packets in Recommendations X.21 and X.25, respectively.

3.3 *Use of the responsibility transfer events*

The responsibility transfer events may be used in defining user and network responsibility intervals within four specific performance periods:

1) the period between the beginning and the end of an access attempt;

2) the period between the beginning and the end of a block transfer attempt;

3) the period between the beginning and the end of a disengagement attempt (for a specified user);

4) the period delimiting the larger of the input time or the output time for an individual transfer sample (as discussed in § 2.2.2).

Defining user and network responsibility intervals within the access, block transfer, and disengagement performance periods enables the specification of separate network and user values for access delay, user information transfer delay, and disengagement delay.

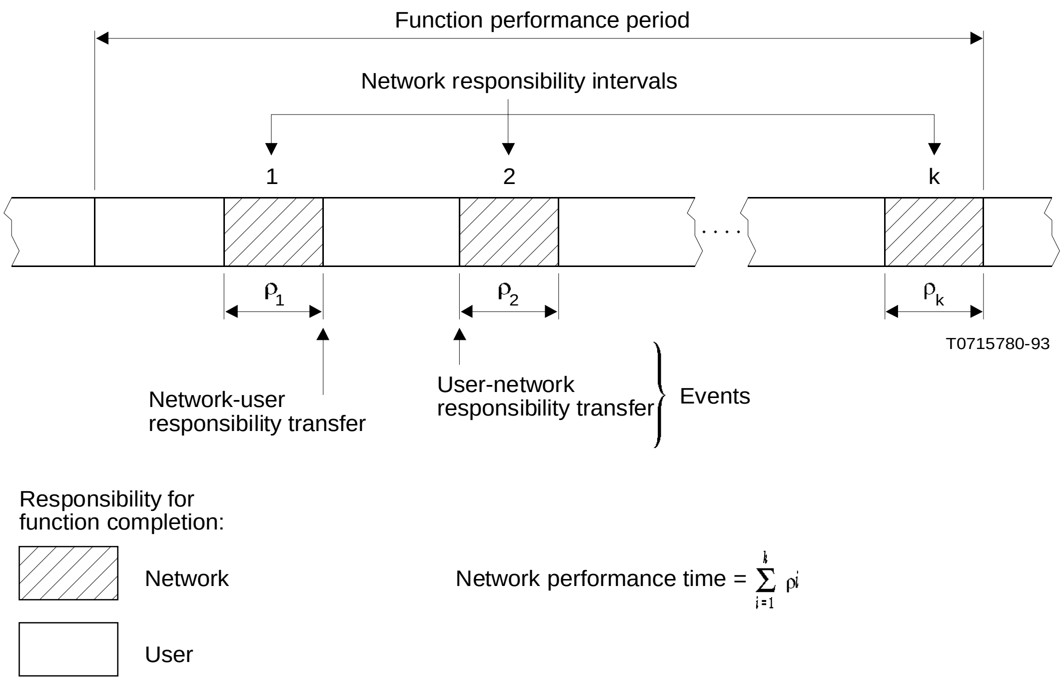


FIGURE 7/X.140

Use of responsibility transfer events in calculating network performance times

Separation of the above performance periods into user and network components also provides a method of establishing "responsibility" for timeout performance failures i.e. whether the user or the network should be charged with the failure when a performance trial is not completed within the established timeout period (and no blocking signals are issued). This decision is made by comparing the user performance time for the trial that failed with a specified maximum user performance time. If the observed user performance time exceeds the specified maximum, the failure is attributed to the user; otherwise, the failure is attributed to the network. This procedure is used in distinguishing access denial from user blocking (§ 2.1.3); user information loss from user information refusal (§ 2.2.6); and disengagement denial from user disengagement blocking (§ 2.3.2). It is also used in distinguishing the transfer denial and rejected sample outcomes in defining user information transfer denial probability (§ 2.4.2) and user information transfer rate (§ 2.2.2).

4 Supplementary information

This section specifies supplementary information which should be provided in conjunction with any statement of values for the general QOS parameters. The specified information is of two types:

1) information which identifies the intended scope of application of the parameter values;

2) information which identifies the particular statistical meaning each value expresses.

Significant differences between user requirement specifications, provider service specifications and measurement reports are noted.

4.1 *Scope of application*

The intended scope of application of stated QOS values should be defined by specifying the following interface and usage characteristics:

1) user-network interfaces to which the values apply;

2) interface event sequences (e.g. call request, incoming call, call accepted, call connected, etc.) by which the specified data communication service is provided in a typical instance;

3) service refusal actions allowed by the user/network interface protocol (e.g. network clearing in response to a user call request);

4) population of users (or communication instances, such as calls) to which the values apply;

5) operating conditions (or range of conditions) under which the values may be expected to hold.

Particular characteristics may be specified generally or more precisely, depending on the type of specification. User requirement specifications define a service need (and any constraints imposed by the user application) without reference to a particular service offering. User-network interfaces and interactions should be defined generally in such specifications, with particular mechanical, electrical, or procedural characteristics identified only where necessary. The population of users and any user-controlled operating conditions (e.g. service time interval, offered traffic) should be defined explicitly. User delays, user information unit lengths, user input/output rates and the selection of user facilities such as abbreviated address calling should also be explicitly defined.

Provider service specifications describe the quality of a particular offered service, often without references to any particular user application. The user-network interfaces and interactions are normally defined explicitly in such specifications (e.g. by reference to an interface specification such as Recommendations X.21 or X.25). The user (or call) population and the operating conditions may be defined more generally, since they refer to potential rather than actual usage.

Measurement reports summarize the actual performance of a network service. Both interface and usage conditions should be specified in detail in such reports to ensure comparability. Details which should be specified include:

1) the particular user-network interfaces instrumented;

2) the exact user-network interaction sequence observed during the measurements, and the performance significance assigned to each interface event;

3) the sampling plan used to select the measurement points, times, and conditions; and

4) the values or ranges of relevant operating conditions. Confidence limits and levels associated with each measured value should also be stated.

4.2 *Statistical meaning*

In order to clearly define the statistical meaning of a stated QOS parameter value, it is necessary to:

1) identify the particular distribution characteristic the parameter represents; and

2) specify values for any variables which may influence the parameter definition.

Any generally accepted statistical measure may be used in stating values for the general QOS parameters. “Mean” and “95 percentile” values are specified for the protocol-specific performance parameters in the X.130-Series Recommendations.

The principal variables which may influence the definitions of the general QOS parameters are listed below:

1) *performance timeouts*

   – access

   – disengagement

   – transfer

   – sample input/output

2) *maximum user performance times*

   – access

   – disengagement

   – transfer

   – sample input/output

3) *transfer denial criteria*

   – transfer sample size

   – user information error probability threshold

   – user information loss probability threshold

   – extra user information delivery probability threshold

   – user information transfer rate threshold

4) *service outage criteria* (for further study)

   – observation period(s)

   – defining events

   – supported parameters

   – unacceptable performance thresholds.

The performance timeouts establish upper bounds on the associated delay distributions. The maximum user performance times provide a basis for identifying and eliminating user-caused failures. The transfer denial and service outage criteria distinguish "unacceptable" performance periods from periods of "satisfactory or tolerable" service.

Specifications should also indicate whether stated values are "target" or "minimum (or maximum) acceptable" values.

ANNEX A

(to Recommendation X.140)

Relationships between the general QOS parameters and the circuit-switched service performance parameters

This annex describes relationships between the general QOS parameters defined in Recommendation X.140 and the X.21-based, circuit-switched service performance parameters for which limits are specified in Recommendations X.130 and X.131. It illustrates one application of the general parameters and provides a framework for relating user QOS needs with the performance capabilities of circuit switching PDNs. Such relationships may be defined either to allocate a user requirement among network elements, in cases where the network performance values are selectable, or to derive resultant QOS values from network performance values, in cases where the latter are fixed.

In the example presented here, it is assumed that quality is to be specified at a pair of DTE/DCE physical interfaces conforming to Recommendation X.21. The call set-up and clearing sequences presented are derived from Annex B of that Recommendation.

Table A-1/X.140 lists the general parameters and the circuit-switched parameters in the rows and columns of a matrix and indicates qualitative relationships between them. Specific network performance parameters are listed for call processing delays (Recommendation X.130) and call blocking (Recommendation X.131).

TABLE A-1/X.140

Qualitative relationships between the general parameters and the circuit-switched parameters

General parameters \ Circuit-switched parameters (X.21 protocol)	Delay (Rec. X.130)		Blocking (Rec. X.131)
	Network post selection delay	Network clear indication delay	Probability non-connection due to congestion (blocking probability)
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		♦	
Disengagement denial probability			
Service availability			
User information transfer denial probability			
Service outage duration			

Within the delay and blocking categories, a mark at a particular row/column intersection indicates that the corresponding parameters are interdependent and should be considered jointly in service performance specification. Each general parameter is influenced by a corresponding circuit-switched parameter, and may influence its value if the latter is selectable. Detailed relationships between the general parameters and the corresponding X.130 delay and X.131 blocking parameters are described below.

Figure A-1/X.140 illustrates the relationship between the access delay and the X.130 parameter network post selection delay. Access delay here describes the total time between the user's issuance of an X.21 *call request* and the network's subsequent issuance of *ready for data*. The X.130 parameter network post selection delay describes two specific network-dependent components of access delay.

Figure A-2/X.140 illustrates the relationship between disengagement delay and the X.130 parameter network clear indication delay. Two independent disengagement delays are identified:

1) Originator disengagement delay – The total time between *DTE clear request* and *DCE ready* at the clearing DTE interface.

2) Non-originator disengagement delay – The total time between *DTE clear request* at the clearing DTE interface and *DCE ready* at the cleared DTE interface.

Network clear indication delay contributes directly to non-originator disengagement delay, but does not include the delays associated with issuance of the *DTE clear confirmation* and *DCE ready* signals at the cleared DTE interface.

Access denial probability corresponds to the probability of blocking in the X.21 application. It includes cases where no *network congestion* signal is issued.

The X.140 access and disengagement parameters may be used to describe the quality of X.21 leased circuit services by simply specifying their values as zero.

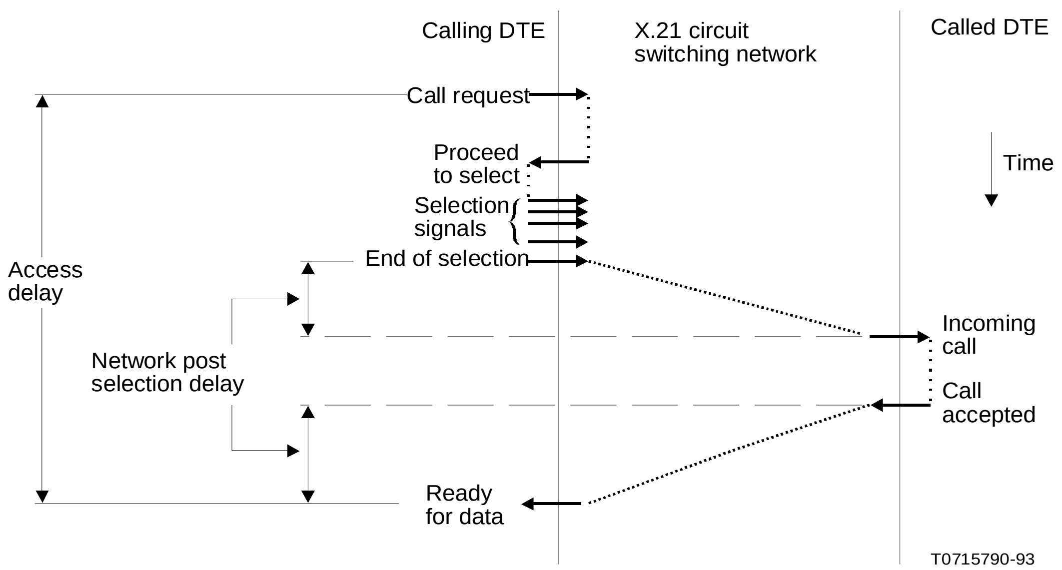
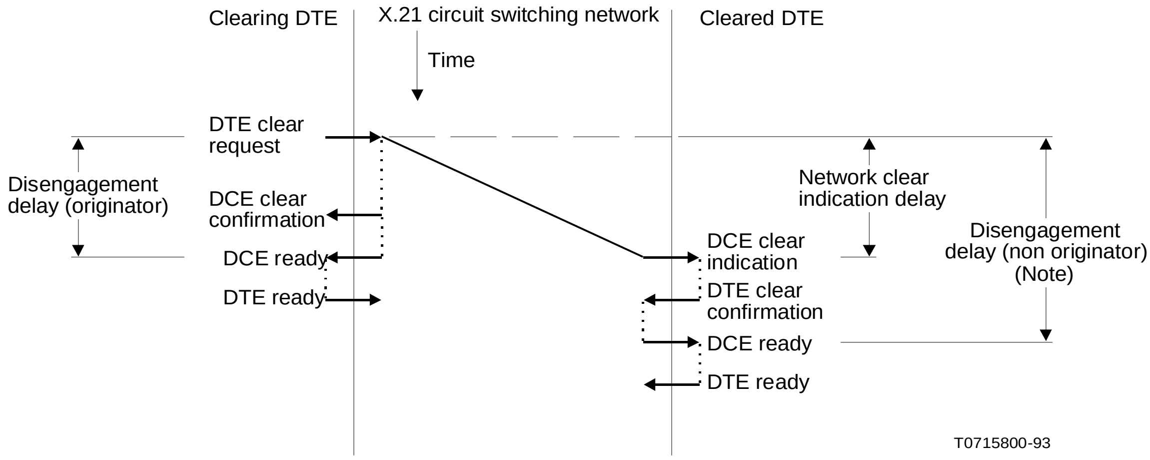


FIGURE A-1/X.140

Relationship between access delay and X.130 network post selection delay



Note – Specification of separate disengagement delays for each participating user is optional.

FIGURE A-2/X.140

Relationship between disengagement delay and X.130 network clear indication delay

ANNEX B

(to Recommendation X.140)

Relationships between the general QOS parameters and the packet-switched service performance parameters

This annex describes relationships between the general QOS parameters defined in Recommendation X.140 and the X.25-based, packet-switched service performance parameters for which limits are specified in X.130-Series Recommendations. It illustrates a second application of the general parameters and provides a framework for relating user QOS needs with the performance capabilities of packet-switched PDNs. Such relationships may be defined either to allocate a user requirement among network elements, in cases where the network performance values are selectable, or to derive resultant QOS values from network performance values, in cases where the latter values are fixed.

In the example presented here, it is assumed that quality is to be specified in terms of packet-layer reference events observed at the physical boundaries separating communicating DTEs from their adjacent access circuit sections. It is assumed that Recommendation X.25 procedures are used on the access circuit sections. The section boundaries and the specific packet reference events are defined in Recommendation X.134. The packet-switched service performance parameters are defined in Recommendations X.135 to X.137. Methods for the measurement of packet-switched service performance parameters are defined in Recommendations X.138 and X.139. The call set-up and clearing sequences presented are derived from state diagrams presented in Annex B of Recommendation X.25.

Table B-1/X.140 lists the general parameters and the packet-switched service parameters in the rows and columns of a matrix and indicates qualitative relationships between them. Each set of parameters is grouped in four categories: access parameters, user information transfer parameters, disengagement parameters, and availability parameters.

A mark at a particular row/column intersection in the matrix indicates that the corresponding parameters are interdependent and should be considered jointly in service performance specification. Each general parameter is influenced by one or more packet-switched service parameters, and may influence their values if the latter are selectable. Detailed relationships between the general parameters and corresponding X.135 to X.137 parameters are described below.

The relationship between access delay and call set-up delay is illustrated in Figure B-1/X.140. As described in Recommendation X.135, call set-up delay can be defined either at a single virtual connection section boundary or between two section boundaries. When defined at the calling DTE boundary B_1, call set-up delay differs from access delay in only one respect: it includes the modulation time (*X*) of the call request packet on the calling DTE access circuit section, while access delay does not. When defined between the calling and called DTE boundaries B_1 and B_n, call set-up delay differs from access delay in one additional respect: it excludes the called DTE response time (i.e. the call set-up delay at boundary B_n).

The general parameters user information transfer delay and disengagement delay correspond closely with the packet-switched service parameters data packet transfer delay and call clearing delay, respectively, when each parameter is defined at the X.25 DTE boundaries. Data packet transfer delay includes the modulation time (Y) of the data packet on the originating DTE access circuit section, while user information transfer delay does not. Similarly, call clearing delay includes the modulation time (*Z*) of the clear request packet on the clearing DTE access circuit section, while disengagement delay does not.

The general parameters incorrect access probability, access denial probability, and disengagement denial probability are essentially identical to the packet-switched service parameters call set-up error probability, call set-up failure probability, and call clear failure probability, respectively, as defined at the X.25 DTE boundaries. The packet-switched service parameter throughput capacity expresses the maximum continuously achievable (steady-state) value of the general parameter user information transfer rate; the former parameter also differs from the latter in that its definition allows it to be measured at a single boundary.

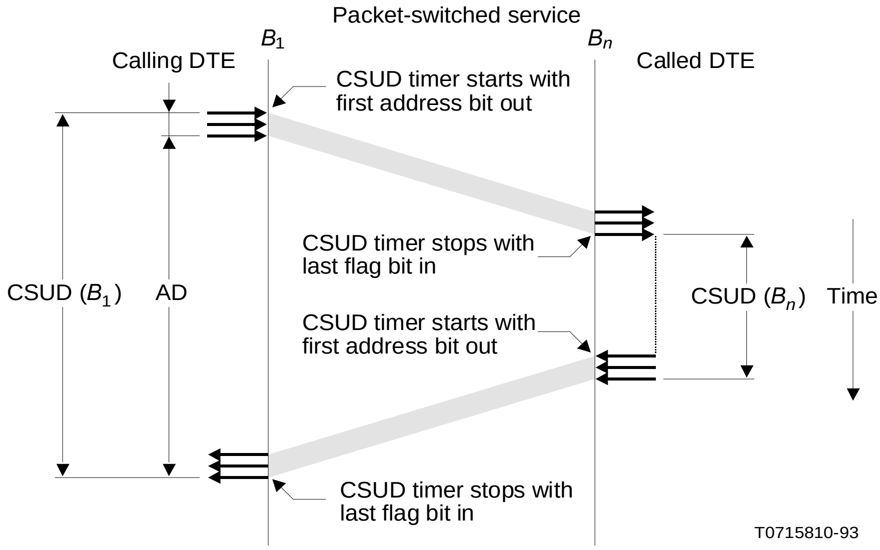
The packet-switched service parameter residual error rate combines the three general parameters user information error probability, extra user information delivery probability, and user information loss probability in a single composite accuracy measure. The mathematical relationship between residual error rate and the three general parameters is specified in Recommendation X.136. There is no packet-switched service parameter that corresponds directly with the general parameter user information misdelivery probability; however, misdelivered data is counted as extra data under the Recommendation X.136 definitions, and is thus reflected indirectly in the definition of residual error rate. The reset and premature disconnect parameters defined in Recommendation X.135 are protocol dependent and thus have no direct counterparts among the general parameters specified in this Recommendation. Their values will normally influence the X.140 parameter user information loss probability.

Recommendation X.137 and this Recommendation define a measure of service availability. The former measure specializes the latter by identifying the particular decision parameters and thresholds that are to be used in defining outages in a packet-switched service. This Recommendation defines a closely-related general parameter, user information transfer denial probability, that provides a sampled measure of unavailability. It is based on a specific definition of outage that differs from that presented in Recommendation X.137 in one respect: the former definition includes only user information transfer parameters among the supported (decision) parameters used in identifying outages, while the latter definition includes call set-up parameters as well. The X.140 parameter service outage duration and the X.137 parameter mean time between service outages provide complementary information on the frequency of transitions between the available and unavailable states.

TABLE B-1/X.140

Qualitative relationships between the general parameters and the packet-switched service parameters

Packet-switched service parameters (X.25 protocol) / General parameters	Call set-up delay	Call set-up error probability	Call set-up failure probability	Data packet transfer delay	Throughput capacity	Residual error rate	Reset stimulus probability	Reset probability	Premature disconnect stimulus probability	Premature disconnect probability	Clear indication delay	Call clear failure probability	Service availability	Mean time between service outages
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											♦			
Disengagement denial probability												♦		
Service availability													♦	
User information transfer denial probability													♦	
Service outage duration														♦



$$\text{CSUD} = \text{CSUD}\,(B_1) - \text{CSUD}\,(B_n)$$
$$\text{AD} = \text{CSUD} + \text{CSUD}\,(B_n) - X$$

AD Access delay
CSUD Call set-up delay

FIGURE B-1/X.140

Relationship between access delay and call set-up delay

ANNEX C

(to Recommendation X.140)

Relationships between the general QOS parameters and the OSI network layer service performance parameters

This annex describes relationships between the general QOS parameters defined in this Recommendation and the OSI network service (NS) performance[2] parameters defined in Recommendation X.213. It illustrates application of the general parameters to a specific (abstract) OSI service interface – the interface between the transport and network layers.

Table C-1/X.140 lists the general parameters and the NS performance parameters in the rows and columns of a matrix and indicates qualitative relationships between them. Each set of parameters is grouped in four categories: access parameters, user information transfer parameters, disengagement parameters, and availability parameters.[3]

Recommendation X.213 defines exact counterparts to five X.140 parameters: access delay, user information transfer delay, user information transfer rate, disengagement delay, and disengagement denial probability.

[2] Recommendation X.213 distinguishes QOS parameters which describe performance from those which describe other service characteristics [network connection (NC) protection, priority, and maximum acceptable cost]. Only the former parameters are addressed here.

[3] Transfer failure probability is included among the user information (data) transfer parameters in Recommendation X.213; that Recommendation does not identify availability as a separate parameter category.

TABLE C-1/X.140

Qualitative relationships between the general parameters and the OSI network layer service performance QOS parameters

OSI network service parameters (X.213) / General parameters	NC establishment delay	NC establishment failure probability	Transit delay	Throughput	Residual error rate	NC resilience	NC release delay	NC release failure probability	Transfer failure probability
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							●	●	
Disengagement denial probability									
Service availability									
User information transfer denial probability									♦
Service outage duration									

● General parameter is identical to the corresponding X.213 parameter when specialized to the OSI network service interface.

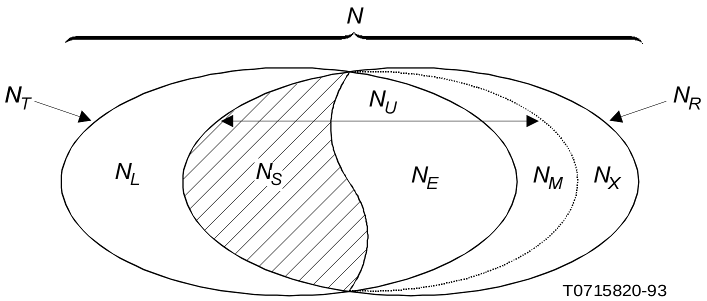
♦ Corresponding parameters are interdependent but not identical.

The X.140 parameters subdivide the X.213 parameters into more detailed components in two cases. The X.213 parameter NC establishment failure probability subsumes two X.140 parameters: access denial probability and incorrect access probability. Values for the X.140 parameters could be added to calculate their X.213 counterpart. The X.213 parameter residual error rate subsumes four X.140 parameters: user information loss probability, user information misdelivery probability, extra user information delivery probability, and user information error probability. The relationships among these probabilities are illustrated in Figure C-1/X.140. Each parameter is normalized so that its possible values range between 0 and 1.

The X.213 parameter NC Resilience is protocol dependent and this has no direct counterpart among the general parameters specified in this Recommendation. Its value will normally influence the X.140 parameter user information loss probability.

The X.140 parameter user information transfer denial probability corresponds closely with the X.213 parameter transfer failure probability; the two differ only in the detailed definition of the supported (decision) parameters used in defining transfer denial (or failure).

Two X.140 parameters have no X.213 counterparts: service outage duration and service availability.



Rec. X.213

$$\text{RER} = \frac{N_L + N_E + N_X}{N}$$

(N_M not distinguished from N_X)

N_T Number transmitted

N_R Number received

N_L Number lost

N_S Number successfully transferred

N_E Number received with errors

N_M Number misdelivered

N_X Number of extras

Rec. X.140

$P(L) = N_L / N_T$

$P(X) = N_X / N_R$

$P(E) = N_E / (N_S + N_E)$

$P(M) = N_M / N_U$

$P(L)$ User information loss probability

$P(X)$ Extra user information delivery probability

$P(E)$ User information error probability

$P(M)$ User information misdelivery probability

RER Residual error rate

FIGURE C-1/X.140

Relationships among the Recommendations X.140 and X.213 transfer failure probabilities

ANNEX D

(to Recommendation X.140)

Alphabetical list of abbreviations used in this Recommendation

AD	Access delay
CRT	Cathode ray tube
CSPDN	Circuit switched public data network
CSUD	Call set-up delay
DCE	Data circuit-terminating equipment
DTE	Data terminal equipment
EFdS	Error-free decisecond
EFS	Error-free second
EFT	Electronic funds transfer
HRX	Hypothetical reference connection
NC	Network connection
NS	Network service
PAD	Packet assembly/disassembly
PDN	Public data network
PSPDN	Packet switched public data network
QOS	Quality of service
RER	Residual error rate