INTERNATIONAL TELECOMMUNICATION UNION

CCITT

X.137

(11/1988)

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

SERIES X: DATA COMMUNICATION NETWORKS: TRANSMISSION, SIGNALLING AND SWITCHING, NETWORK ASPECTS, MAINTENANCE AND ADMINISTRATIVE ARRANGEMENTS

Data communication networks – Network aspects

AVAILABILITY PERFORMANCE VALUES FOR PUBLIC DATA NETWORKS WHEN PROVIDING INTERNATIONAL PACKET-SWITCHED SERVICES

Reedition of CCITT Recommendation X.137 published in the Blue Book, Fascicle VIII.3 (1988)

NOTES

1 CCITT Recommendation X.137 was published in Fascicle VIII.3 of the *Blue Book*. This file is an extract from the *Blue Book*. While the presentation and layout of the text might be slightly different from the *Blue Book* version, the contents of the file are identical to the *Blue Book* version and copyright conditions remain unchanged (see below).

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

© ITU 1988, 2008

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Recommandation X.137

AVAILABILITY PERFORMANCE VALUES FOR PUBLIC DATA NETWORKS WHEN PROVIDING INTERNATIONAL PACKET-SWITCHED SERVICES

(Melbourne, 1988)

The CCITT,

considering

(a) that Recommendation X.1 specifies the international user classes of service in public data networks;

(b) that Recommendation X.2 specifies the international data transmission services and optional user facilities in public data networks;

(c) that Recommendation X.25 specifies the DTE/DCE interface for packet mode terminals connected to public data networks by dedicated circuit;

(d) that Recommendation X.75 specifies the packet-switched signalling system between public data networks providing data transmission services;

(e) that Recommendation X.323 specifies general arrangements for interworking between packet-switched public data networks;

(f) that Recommendation X.96 specifies the international routing principles and routing plan for public data networks;

(g) that Recommendation X.110 specifies the international routing principles and routing plan for public data networks;

(h) that Recommendation X.213 defines the OSI Network Layer service;

(i) that Recommendation X.140 defines general quality of service parameters for communication via public data networks;

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

(k) that Recommendation X.135 specifies speed of service performance values for public data networks when providing international packet-switched service;

(l) that Recommendation X.136 specifies accuracy and dependability (including blocking) performance values for public data networks when providing international packet-switched service.

unanimously declares

(1) that the availability parameters defined in this Recommendation shall be used in the planning and operation of international packet-switched data communication service provided in accordance with Recommendations X.25 and X.75,

(2) that in such service, the performance values specified in this Recommendation shall be taken as worst-case limits under the conditions specified herein.

1 Introduction

1.1 This Recommendation is the fourth in a series of four CCITT Recommendations (X.134-X.137) that define performance parameters and values for international packet-switched data communication services. Figure 1/X.137 illustrates the scope of these four Recommendations and the relationships among them.

1.2 Recommendation X.134 devides a virtual connection into basic sections whose boundaries are associated with X.25 and X.75 interfaces; defines particular collections of basic sections, called virtual connection portions, for which performance values will be specified; and defines a set of packet layer reference events (PEs) which provide a basis for performance parameter definition. The basic sections consist of network sections and circuit sections. They are delimited, in each case, by physical data terminal equipment (DTE) or data switching equipment (DSE) interfaces. Two types of virtual connection portions are identified: national portions and international portions. Each PE is defined to occur when a packet crossing a section boundary changes the state of the packet layer interface.

1.3 For comparability and completeness, packet-switched network performance is considered in the context of the 3 × 3 performance matrix defined in Recommendation X.140. Three protocol-independent data communication functions are identified in the matrix: access, user information transfer, and disengagement. These general functions correspond to call set-up, data (and interrupt) transfer, and call clearing in packet-switched virtual call services conforming to the X.25 and X.75 Recommendations. 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 or correctness, and degree of certainty with which the function is performed.

1.4 Recommendation X.135 defines protocol-specific speed of service parameters and values associated with each of the three data communication functions. Recommendation X.136 defines protocol-specific accuracy and dependability parameters and values associated with each function. The Recommendation X.135 and X.136 parameters are called "primary parameters" to emphasize their direct derivation from packet layer reference events.

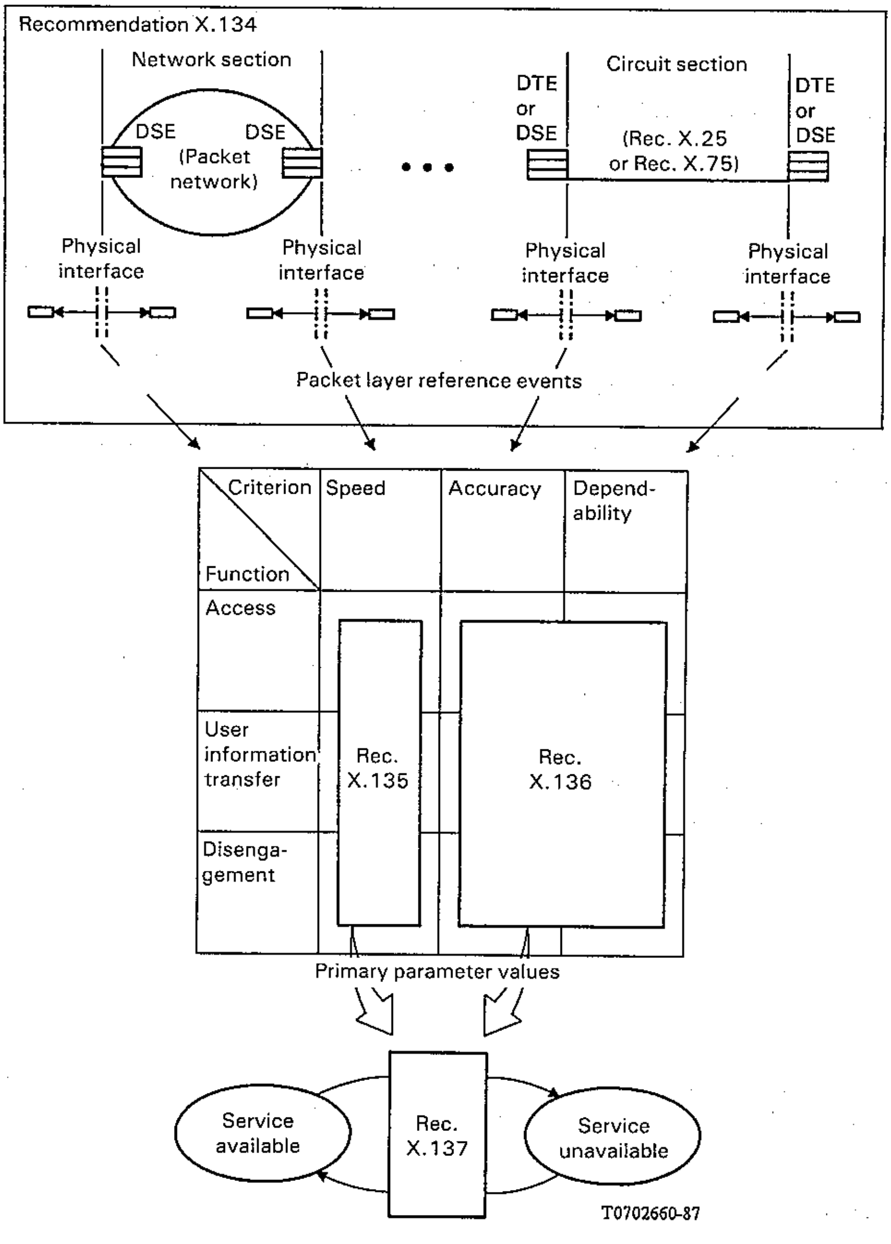


FIGURE 1/X.137

Packet-switched service performance description framework

1.5 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. This Recommendation specifies the availability function and defines the availability parameters and values that characterize the resulting binary random process.

1.6 Two availability parameters are defined in this Recommendation: service availability and mean time between service outages. Each parameter can be applied to any basic section or portion of a virtual connection. This generality makes the parameters useful in performance allocation and concatenation.

1.7 This Recommendation specifies availability values for national and international portions of two types (Table 1/X.137). Performance values for data terminal equipment are not specified, but the parameters defined in this Recommendation may be employed in such specification to assist users in establishing quantitative relationships between network performance and quality of service (see Recommendation X.140).

TABLE 1/X.137

Virtual connection portion types for which performance values are specified [a)]

Portion type	Typical characteristics
National A	Terrestrial connection via an access network section
National B	Connection via an access network section with one satellite circuit; or via an access network section and one or more transit network sections
International A	Connection via a direct terrestrial internetwork section
International B	Connection via two satellite circuits and one transit network section; or via one satellite circuit and two or more transit network sections

[a)] The values specified for type B portions also apply to virtual connection portions not explicitly identified as type A or type B.

1.8 Worst-case values for each of the two availability parameters are specified below for each virtual connection portion type identified in Table 1/X.137. The term “worst case” means that these values should be met during the normal busy hour in the worst-performing virtual connection portion used in providing international packet-switched service. The performance of a virtual connection portion may be better than the worst-case values specified in this Recommendation. Design objectives that take into account more demanding user applications and network performance and connectivity enhancements are for further study.

Numerical methods for combining individual portion performance values to estimate end-to-end performance are also provided in this Recommendation. DTE to DTE values for two particular hypothetical reference connections are derived using these methods in Annex B.

2 Availability function

Eight performance parameters from Recommendations X.135 and X.136 are used in computing the availability of a virtual connection: throughput capacity (X.135), call set-up failure probability (X.136), call set-up error probability (X.136), residual error rate (X.136), reset probability (X.136), reset stimulus probability (X.136), premature disconnect probability (X.136), and premature disconnect stimulus probability (X.136). Five particular linear combinations of these parameters are called the availability decision parameters. Each decision parameter is associated with an outage threshold. These decision parameters and their outage thresholds are listed in Table 2/X.137.

TABLE 2/X.137

Outage criteria for the availability decision parameters

Availability decision parameters	Outage criteria
Call set-up failure probability (cfp) Call set-up error probability (cep)	(cfp + cep) > 0.9
Throughput capacity (tc)	tc < 80 bit/s
Residual error rate (rer)	rer > 10^{-3}
Reset probability (rp) Reset stimulus probability (rsp_1, rsp_2)	(rsp_1 + rp + rsp_2) > 0.015
Premature disconnect probability (pdp) Premature disconnect stimulus probability ($pdsp_1$, $pdsp_2$)	($pdsp_1$ + pdp + $pdsp_2$) > 0.01

Note – These outage criteria are provisional.

Performance is considered independently with respect to each availability decision parameter. If the value of the parameter is equal to or better than the defined outage threshold, performance relative to that parameter is defined to be acceptable. If the value of the parameter is worse than the threshold, performance relative to that parameter is defined to be unacceptable.

The packet layer reference events that are used in defining the decision parameters do not occur if a data link layer at a section boundary is unavailable. During a continuous time interval the data link layer of a circuit section is defined to be available for packet layer service if and only if:

1) the link is in the information transfer phase for at least 99% of the time interval; and

2) all continuous periods when the link is not in the information transfer phase are less than 1 second in length; and

3) all continuous busy (flow-controlled) conditions are less than 10 seconds in length.

Otherwise the data link layer is considered unavailable for providing packet layer service.

The data link layer of a circuit section can be unavailable for the following reasons:

1) a nonfunctional physical circuit; or

2) a data link layer controller either unable or unwilling to establish the information transfer phase; or

3) a data link layer controller either unable or unwilling to clear a busy condition.

A virtual connection section is defined to be available (or to be in the available state) if:

1) the performance is acceptable relative to all decision parameters; and

2) both data link layers at the boundaries of the section are available.

The virtual connection section is defined to be unavailable (or in the unavailable state) if:

1) the performance of one or more of the five decision parameters is unacceptable; or
2) one or both of the data link layers at the boundaries of the section are unavailable due to causes inside the section. (Data link layer unavailability due to causes outside the section are excluded, i.e. failures of data link controllers or physical circuits outside the section in question.)

The intervals during which a virtual connection section is unavailable are identified by superimposing the unacceptable performance periods for all decision parameters as illustrated in Figure 2/X.137.

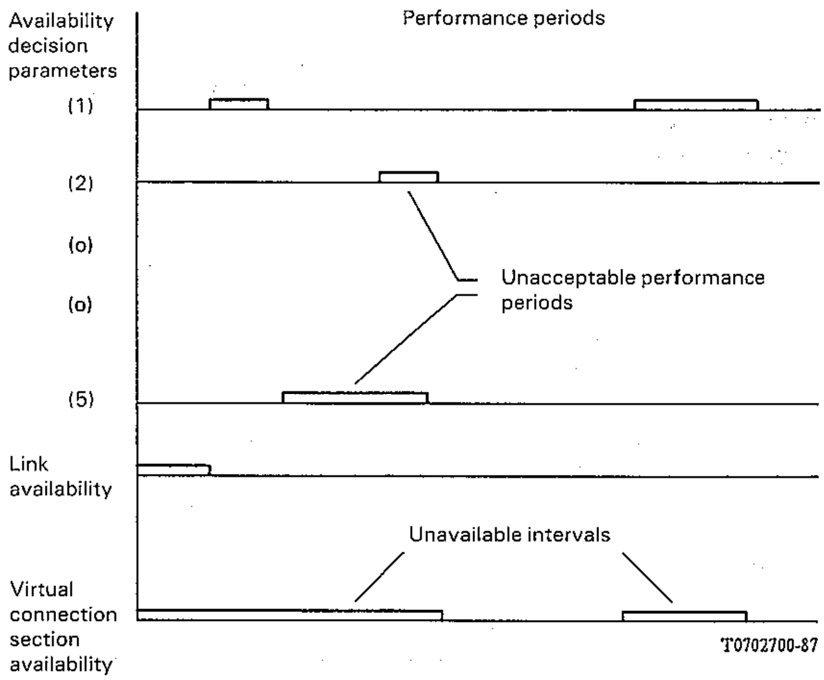


FIGURE 2/X.137

Determination of availability states

In order to exclude transient impairments from being considered as periods of unavailability, a single test of the availability state must exceed 5 minutes. In order to reduce the probability of state transitions during a test of the current availability state, that test should be less than 20 minutes. A minimal availability test meeting these restrictions is defined in Annex B.

3 Availability parameters

This section specifies worst-case values for two availability parameters: service availability and mean time between service outages.

3.1 **Service availability** *definition*

Service availability applies to both virtual call and permanent virtual circuit services. The service availability for a virtual connection portion is the long-term percentage of scheduled service time in which that section is available.

Scheduled service time for a virtual connection section is the time during which the network provider has agreed to make that section available for service. The normal objective would be 24 hours per day, 7 days per week.[1)] A procedure for estimating the availability of a section is described in Annex A.

---

1) Other scheduled service times may be specified in some networks.

3.2 *Definition of* **mean time between service outages**

Mean time between service outages (MTBSO) applies to both virtual call and permanent virtual circuit services. The mean time between service outages for a virtual connection section is the average duration of any continuous interval during which the virtual connection section is available. Consecutive intervals of schedules service time are concatenated. Annex A describes a procedure for estimating the mean time between the service outages of a section.

Mean time between service outages as defined in this Recommendation is closely related to mean time between failures.

3.3 *Values*

The conribution from each network portion to overall service availability and mean time between service outages under the conditions described in this Recommendation shall not be worse than the values specified in Table 3/X.137. The mean time between service outage values for national portions exclude up to 5% of virtual connections to account for geographical and climatic extremes.

TABLE 3/X.137

Worst-case service availability and mean time between service outage values for virtual connection portions

Statistic	Virtual connection portion type			
	National		International	
	A	B	A	B
Service availability (percent)	99.5	99	99.5	99
Mean time between service outages (hours)	1200	800	1600	800
Mean time to service restoral MTTSR [a] (hours)	(6)	(8)	(8)	(8)

[a] The parenthetical values given in Table 3/X.137 represent the mean time to service restoral that would result if the service availability and the mean time between service outage values are achieved as stated in the table. Any improvements in MTBSO should be used to improve service availability and should not be used to degrade MTTSR.

Note – All specified values are provisional.

3.4 *Related parameters*

Four other parameters are commonly used in describing availability performance. These are generally defined as follows:

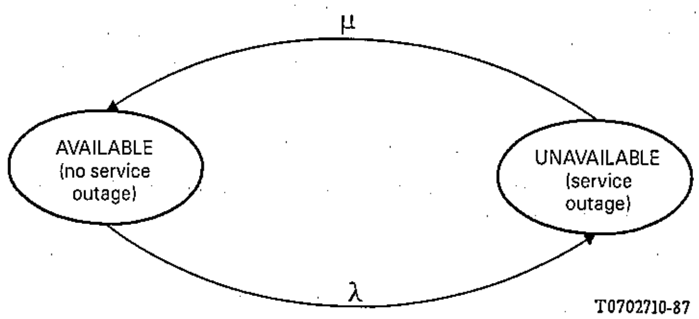
– **mean time to service restoral (MTTSR)** is the average duration of unavailable service time intervals.

– **failure rate (λ)** is the average number of transitions from the available state to the unavailable state per unit available time.

– **restoral rate (μ)** is the average number of transitions from the unavailable state to the available state per unit unavailable time.

– **unavailability (U)** is the long-term ratio of unavailable service time to scheduled service time, expressed as a percentage.

Under the exponential distribution assumption of failure and restoration, the mathematical values for any of these parameters may be estimated from the values for service availability (A) and mean time between service outages as summarized in Figure 3/X.137.



a) *State diagram*

$$MTBSO = \frac{1}{\lambda}$$

$$MTTSR = \frac{1}{\mu}$$

$$A = 100\left(\frac{MTBSO}{MTBSO + MTTSR}\right) = 100\left(\frac{\mu}{\lambda + \mu}\right)$$

$$U = 100 - A = 100\left(\frac{MTTSR}{MTBSO + MTTSR}\right) = 100\left(\frac{\lambda}{\lambda + \mu}\right)$$

b) *Parameter relationships*

FIGURE 3/X.137

Basic availability model and parameters

ANNEX A

(to Recommendation X.137)

Sampling estimation of availability parameters

A.1 *A minimal test for availability*

The definition of availability requires that observed performance for all five decision parameters be compared with outage thresholds. A single success of the following test is defined to be sufficient for declaring the virtual connection section available. A single failure of a section to meet any of the six individual decision criteria is defined to be sufficient for declaring the virtual section unavailable. This test and its decision criteria are defined to be the minimum criteria necessary to sample the availability of the section.

The minimal availability test can be initiated in either direction across the section by equipment and components outside of the section. The test is divided into two phases: access and user information transfer. The access phase is used in conjunction with switched virtual calls only.

Phase I: Attempt N[2)] consecutive call set-ups across A.

Phase II: (If the test did not fail in Phase I) To ensure that the availability test does not fail as a result of insufficient data input, attempt to maintain a virtual connection across A for 5 minutes. Attempt to maintain an average throughput significantly greater than 80 bit/s (e.g. at least 150 bit/s) during that interval.

There are six criteria for deciding if the test has failed or succeeded:

1) The test fails in Phase I if all four call set-up attempts result in either call set-up error or call set-up failure (switched virtual calls only).
2) The test fails in Phase II if the total reset events plus reset stimuli is five or greater.
3) The test fails in Phase II if the throughput is less than 80 bit/s;
4) The test fails in Phase II if the residual error rate is greater than 10^{-3}.
5) The test fails in Phase II if the call and subsequent reestablishments of that call are cleared two or more times due to premature disconnects and/or premature disconnect stimuli (switched virtual calls only).
6) The test fails in Phase I or Phase II if a data link layer at a section boundary is unavailable during a 5-minute interval due to causes inside of A.

If the test passes all six decision criteria, the test is successful and the virtual connection section A is considered to be available during the test. If any of the decision criteria are failed, the virtual connection section A is considered to have been unavailable for the duration of the test.

Because many performance parameters must be supported simultaneously in order for A to be considered available, during normal operation (without a testing procedure like the one described above) it is not possible to prove the section is available (e.g. it may not be possible to observe both access and user information transfer simultaneously). Therefore during normal operation, if the section is correctly performing the currently requested function, the section is assumed to be available.

Service availability and mean time between service outage values can be estimated on the basis of this minimal test (availability performance samples). Such estimation is more practical than measurement based on continuous service observation.

A.2 *Procedures for estimating service availability*

A sufficient estimate of the service availability percentage can be computed as follows. Based on an *a priori* estimate of the service availability, choose a sample size "s", not less than 300. Choose "s" testing times during scheduled service time and distribute them across a long measurement period (e.g. 6 months). Because of the expected durations of service outages, choose no two testing times closer together than 7 hours (this serves to keep the observations uncorrelated). The testing times should be uniformly distributed across the scheduled service time. At each predetermined testing time perform the availability test described above. If the test fails, the section is declared unavailable for that sample. Otherwise, the section is declared available. The estimate of the service availability percentage is the number of times the section was declared available multiplied by 100 and divided by the total number of samples.

A.3 *Procedures for estimating mean time between service outages*

A sufficient estimate of the mean time between service outage parameter can be computed by conducting consecutive availability performance samples and by counting the observed changes from the available state to the unavailable state.

Prior to performing any tests, choose *k* disjoint intervals of time each not less than 30 minutes nor more than 3 hours. The total amount of time in the *k* intervals should exceed 3 times the *a priori* estimate of mean time between service outages. For the duration of each pre-defined interval conduct consecutive availability performance samples. The amount of time observed in the available state will be added to a cumulative counter called "A". The number of observed transitions from the available state to the unavailable state will be accumulated in a counter called "F".[3)]

[2)] The number N of consecutive call set-ups depends on the values of the outage decision parameter, to provisionally 4, and so for further study.

[3)] Each counter is initially set to zero.

For each pre-defined interval:

– If all of the consecutive availability samples succeed, then add the total length of the interval to A. Do not change the cumulative value of F.

– If the first availability sample succeeds and any subsequent sample in the interval fails, increase F by one. Add to A the total length of all availability samples prior to the first failure. Following the first failed availability sample the remaining time in the interval may be discarded without testing its availability.

– If the first availability sample fails, assume that the state transition occurred before the interval began. Add nothing to the count of observed availability time, A. Add nothing to the cumulative count of observed state changes, F. The remaining time in the interval may be discarded without testing its availability.

After the results of every pre-defined interval have been accumulated, the ratio, A/F, is an estimate of the mean time between service outages. A statistically more precise estimate can be obtained by increasing the number of observed intervals, *k*.

The estimate of mean time between service outages assumes that, if an outage begins during an availability performance sample, either this sample or the following sample will decide that the section is unavailable. This is a reasonable assumption since service outages, in contrast to transient failures, will last more than 5 minutes.

Discarding the remainder of the interval following a failed availability sample is both practical and statistically justifiable. The virtual connection section must return to the available state before any more available time can be accumulated and before any more transitions to the unavailable state can be observed. First, the expected time to restore service may be large with respect to the remaining time in the interval. It can be inappropriate and counterproductive to continue testing a failed or congested network section. Second, if transitions to the unavailable state are statistically independent, then discarding the remainder of the interval, which may include time in the available state and a proportional number of transitions back into the unavailable state, will not bias the result.[4)] The only consequence of discontinuing the test is the loss of testing time. To minimize that loss, the test intervals should be short with respect to the sum of the expected time to restore service and the expected time between service outages. Thus each test should be no longer than 3 hours.

There are two sources of bias in the estimation procedure described above. First, if an outage begins during the last availability sample of the interval, that transition may or may not cause the sample to fail. If it does not fail, the state transition is missed and the mean time between service outages is overestimated. Second, a state transition to the unavailable state during the first availability sample of the interval may or may not cause that sample to fail. According to the estimation procedure, if the sample does fail, the interval will be discarded, the state transition is missed, and the mean time between service outages is overestimated. These edge effects can be minimized by increasing the length of each interval, consequently increasing the number of availability samples, and thus decreasing the effect of the first and last sample outcomes as a proportion of the total sampled outcomes. A minimum recommended interval length is 30 minutes and size 5 minute availability samples.

Alternatively, both biases can be corrected by replacing the first instruction above with:

– If all of the consecutive availability samples succeed, then add the total length of the interval to A. Take one additional availability sample immediately following the interval. If that sample fails, increase F by one. If that sample succeeds, do not change F. The length of the additional sample has no effect on A.

This modification identifies any state transitions that occurred during the last sample of the interval and eliminates the first source of bias. It also counts certain transitions that occurred outside of the interval. These transitions are counted with the same probability as the probability that the second source of bias inappropriately discards transitions. Thus this modified procedure corrects both sources of bias. Using this modification, the mean time between service outages can be more accurately estimated.

[4)] If outages tend to be clustered, discontinuing a test following a transition to the unavailable state will tend to overestimate the mean time between service outages. If outages tend to be negatively clustered, discontinuing a test following a transition to the unavailable state will tend to underestimate the mean time between service outages.

ANNEX B

(to Recommendation X.137)

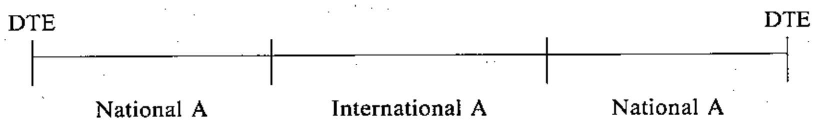
Representative end-to-end availability performance

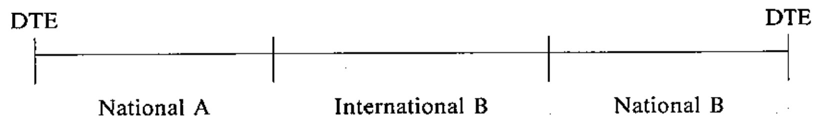
This annex provides two examples to illustrate how end-to-end (DTE to DTE) availability performance can be estimated from the individual virtual connection portion performance values specified in Recommendation X.137. Two example concatenations of Type A and Type B virtual connection portions are defined. The end-to-end service availability and mean time between service outages are calculated for each example. Although alternative network models and statistical assumptions are possible, the methods presented in this annex provide one practical way of estimating the end-to-end performance from the performance of the individual network portions.

B.1 *Definition of the example end-to-end connections*

For ease of reference the two example end-to-end (i.e. DTE to DTE) connections presented in this annex will be referred to as "Type 1" and "Type 2" configurations. These hypothetical, but representative, configurations use the portion boundaries and packet layer reference events described in Recommendation X.134. Table 1/X.137 defines the virtual connection portion types.

The Type 1 configuration is defined to be:





B.2 *End-to-end availability performance for the Type 1 and Type 2 configuration examples*

End-to-end availability performance values have been calculated for the example Type 1 and Type 2 connection configurations and are reported below in Tables B-1/X.137 and B-2/X.137. These calculations have been made by applying the methods described below to the individual network portions that, for convenience in defining these examples, are characterized by the worst-case accuracy and dependability performance values specified in Recommendation X.137.

Assuming that the service availability performance values associated with the individual network portions are statistically independent, then the end-to-end performance values can be calculated by multiplying the percent of time each of the network portions is available.

Example: To compute the end-to-end service availability for the Type 1 configuration, refer to Table 3/X.137 to obtain the individual portion availabilities (National A: percent = 99.5; International A: percent = 99.5). The end-to-end availability in percent is then:

$$99.5 \cdot 99.5 \cdot 99.5 = 98.5$$

The end-to-end performance for the mean time between service outages can be estimated by assuming that the times between service outages in each individual network portion are independent and exponentially distributed. It follows from these assumptions that the end-to-end mean time between service outages performance objective, *T*, can be calculated using the following formula:

$$T = [T_1^{-1} + T_2^{-1} + \dots + T_i^{-1} + \dots + T_N^{-1}]^{-1}$$

where *T* will be in hours if the mean time between service outages for each of the *N* network portions, T_i (I = 1, 2, . . ., *N*), is expressed in hours.

Example: For the Type 1 configuration the National A portion mean time between service outages is 1200 hours and the International A portion is 1600 hours (refer to Table 3/X.137). The end-to-end performance objective is then:

$$[1200^{-1} + 1600^{-1} + 1200^{-1}]^{-1} = 436 \text{ heures}$$

TABLE B-1/X.137

End-to-end availability and mean time between service outage performance for the type 1 configuration example

Type 1 configuration	
Statistic	End-to-end value
Service availability (percent)	98.5
Mean time between service outages (hours)	436

TABLE B-2/X.137

End-to-end availability and mean time between service outage performance for type 2 configuration example

Type 2 configuration	
Statistic	End-to-end value
Service availability (percent)	97.5
Mean time between service outages (hours)	300

ITU-T RECOMMENDATIONS SERIES

Series A	Organization of the work of the ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	**Data networks and open system communications**
Series Y	Global information infrastructure and Internet protocol aspects
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2008