INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

X.137

(08/97)

SERIES X: DATA NETWORKS AND OPEN SYSTEM COMMUNICATION

Public data networks – Network aspects

Availability performance values for public data networks when providing international packet-switched services

ITU-T Recommendation X.137

(Previously CCITT Recommendation)

ITU-T X-SERIES RECOMMENDATIONS

DATA NETWORKS AND OPEN SYSTEM COMMUNICATION

PUBLIC DATA NETWORKS	X.1–X.199
Services and facilities	X.1–X.19
Interfaces	X.20–X.49
Transmission, signalling and switching	X.50–X.89
Network aspects	**X.90–X.149**
Maintenance	X.150–X.179
Administrative arrangements	X.180–X.199
OPEN SYSTEM INTERCONNECTION	X.200–X.299
Model and notation	X.200–X.209
Service definitions	X.210–X.219
Connection-mode protocol specifications	X.220–X.229
Connectionless-mode protocol specifications	X.230–X.239
PICS proformas	X.240–X.259
Protocol Identification	X.260–X.269
Security Protocols	X.270–X.279
Layer Managed Objects	X.280–X.289
Conformance testing	X.290–X.299
INTERWORKING BETWEEN NETWORKS	X.300–X.399
General	X.300–X.349
Satellite data transmission systems	X.350–X.399
MESSAGE HANDLING SYSTEMS	X.400–X.499
DIRECTORY	X.500–X.599
OSI NETWORKING AND SYSTEM ASPECTS	X.600–X.699
Networking	X.600–X.629
Efficiency	X.630–X.649
Naming, Addressing and Registration	X.650–X.679
Abstract Syntax Notation One (ASN.1)	X.680–X.699
OSI MANAGEMENT	X.700–X.799
Systems Management framework and architecture	X.700–X.709
Management Communication Service and Protocol	X.710–X.719
Structure of Management Information	X.720–X.729
Management functions	X.730–X.799
SECURITY	X.800–X.849
OSI APPLICATIONS	X.850–X.899
Commitment, Concurrency and Recovery	X.850–X.859
Transaction processing	X.860–X.879
Remote operations	X.880–X.899
OPEN DISTRIBUTED PROCESSING	X.900–X.999

For further details, please refer to ITU-T List of Recommendations.

ITU-T RECOMMENDATION X.137

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

Summary

This Recommendation provides availability parameters and associated worst-case objectives within the X.130-Series Recommendations on PSPDN performance.

Source

ITU-T Recommendation X.137 was revised by ITU-T Study Group 7 (1997-2000) and was approved under the WTSC Resolution No. 1 procedure on the 9th of August 1997.

FOREWORD

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. The ITU-T 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 World Telecommunication Standardization Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

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

INTELLECTUAL PROPERTY RIGHTS

The ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. The ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, the ITU had/had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

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

(revised in 1997)

1 Introduction and scope

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

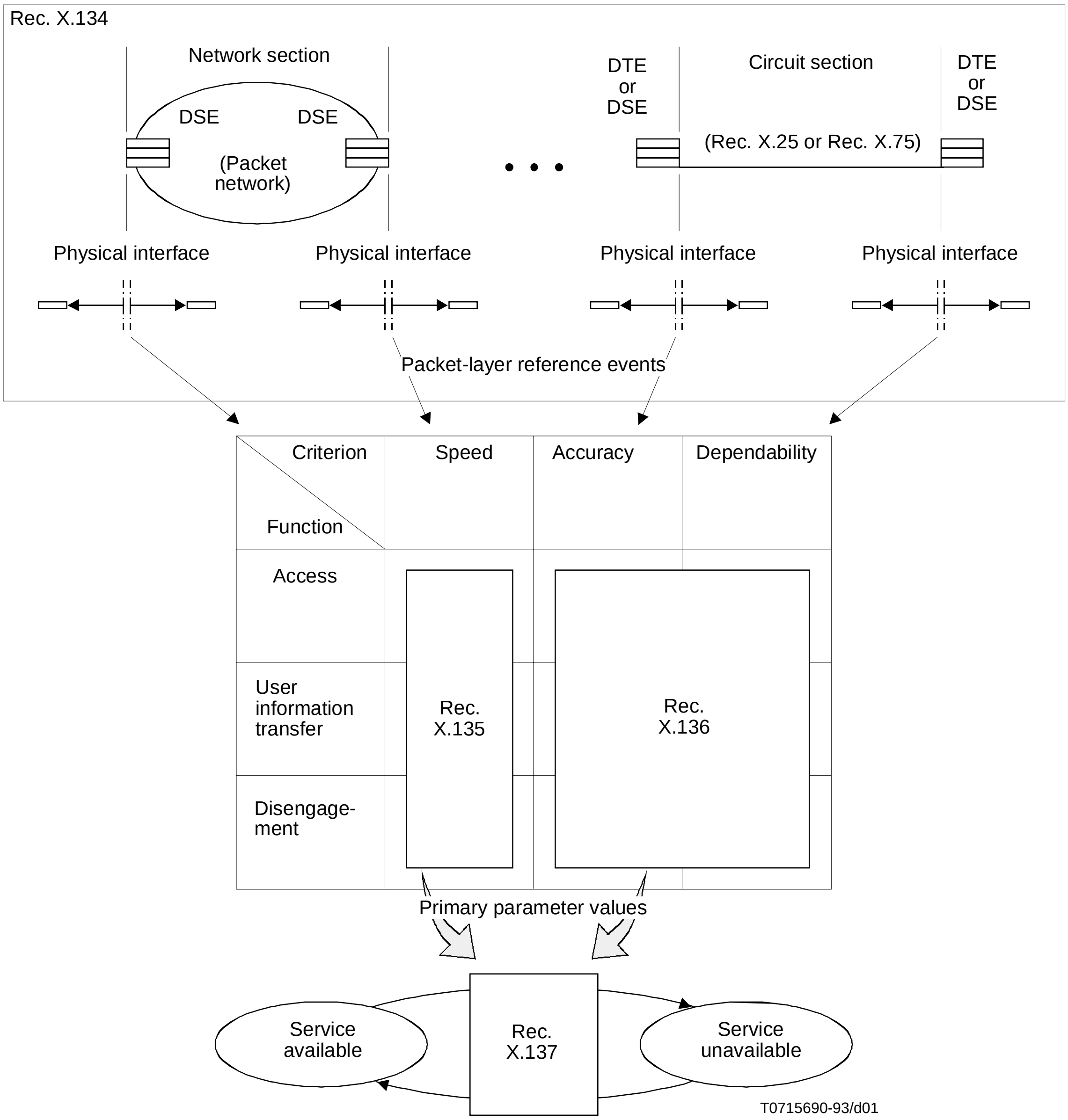


Figure 1/X.137 – Packet-switched service performance description framework

1.2 Recommendation X.134 divides 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 provides 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 Recommendations X.25 and X.75. 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.

1.4 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. X.135 and X.136 parameters are called "primary parameters" to emphasize their direct derivation from packet layer reference events.

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 (see Table 1). 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. The term "worst case" means that these values should be met during any hour of the day 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 References

The following ITU-T Recommendations, and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; all users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published.

– ITU-T Recommendation X.1 (1996), *International user classes of service in, and categories of access to, public data networks and Integrated Services Digital Networks (ISDNs).*

– ITU-T Recommendation X.2 (1996), *International data transmission services and optional user facilities in public data networks and ISDNs.*

– ITU-T Recommendation X.25 (1996), *Interface between Data Terminal Equipment (DTE) And Data Circuit-Terminating Equipment (DCE) for terminals operating in the packet mode and connected to public data networks by dedicated circuit.*

– ITU-T Recommendation X.75 (1996), *Packet-switched signalling system between public networks providing data transmission services.*

– ITU-T Recommendation X.96 (1993), *Call progress signals in public data networks.*

– ITU-T Recommendation X.110 (1996), *International routing principles and routing plan for public data networks.*

– ITU-T Recommendation X.134 (1997), *Portion boundaries and packet-layer reference events: Basis for defining packet-switched performance parameters.*

– ITU-T Recommendation X.135 (1997), *Speed of service (delay and throughput) performance values for public data networks when providing international packet-switched services.*

– ITU-T Recommendation X.136 (1997), *Accuracy and dependability performance values for public data networks when providing international packet-switched services.*

– ITU-T Recommendation X.138 (1997), *Measurement of performance values for public data networks when providing international packet-switched services.*

– ITU-T Recommendation X.139 (1997), *Echo, drop, generator and test DTEs for measurement of performance values in public data networks when providing international packet-switched services.*

– CCITT Recommendation X.140 (1992), *General quality of service parameters for communication via public data networks.*

– ITU-T Recommendation X.213 (1995), *Information technology – Opens Systems Interconnection – Network service definition.*

– CCITT Recommendation X.323 (1988), *General arrangements for interworking Between Packet-Switched Public Data Networks (PSPDNs).*

3 Abbreviations

This Recommendation uses the following abbreviations:

cep	Call set-up Error Probability
cfp	Call set-up Failure Probability
DCE	Data Circuit-terminating Equipment
DSE	Data Switching Equipment
DTE	Data Terminal Equipment
MTBSO	Mean Time between Service Outages
MTTSR	Mean Time to Service Restoral
OSI	Open Systems Interconnection
pdp	Premature Disconnect Probability
pdsp	Premature Disconnect Stimulus Probability
PE	Packet-layer reference Event
rer	Residual Error Ratio
rp	Reset Probability
rsp	Reset Stimulus Probability
SPRT	Sequential Probability Ratio Test
tc	Throughput Capacity

4 Availability function

Eight performance parameters from Recommendations X.135 and X.136 are used in computing the availability of a virtual connection: throughput capacity (see Recommendation X.135), call set-up failure probability (see Recommendation X.136), call set-up error probability (see Recommendation X.136), residual error ratio (see Recommendation X.136), reset probability (see Recommendation X.136), reset stimulus probability (see Recommendation X.136), premature disconnect probability (see Recommendation X.136), and premature disconnect stimulus probability (see Recommendation 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.

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 Ratio (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 non-functional 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.

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.

5 Availability parameters

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

5.1 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.

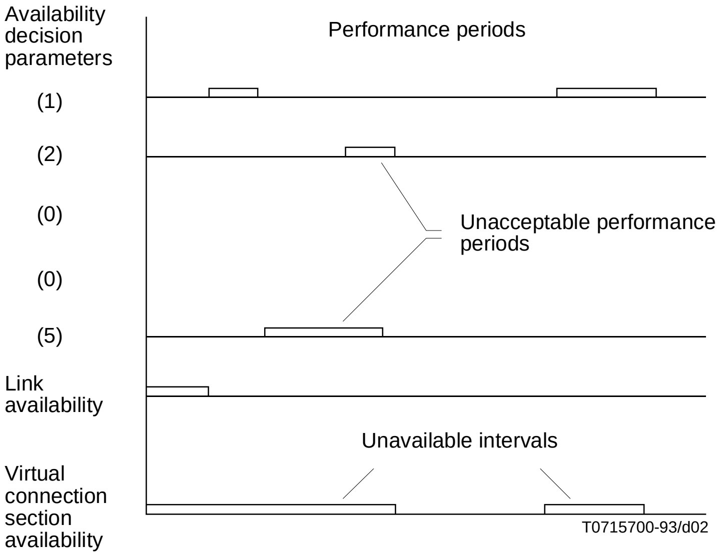


Figure 2/X.137 – Determination of availability states

5.2 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 scheduled 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.

5.3 Values

The contribution 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. 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 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 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.

5.4 Related parameters

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

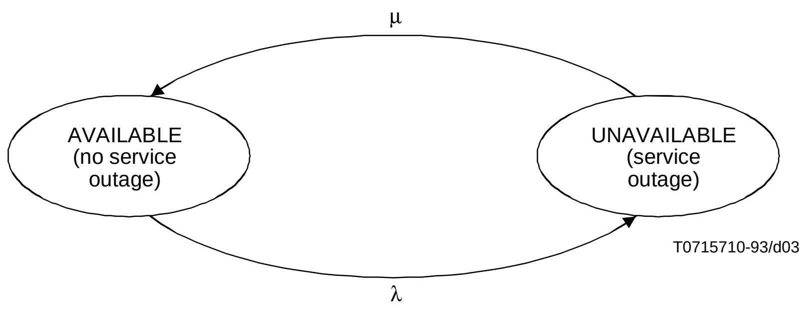
– **mean time to service restoral (MTTSR)** is the average duration of unavailable service time intervals.

– **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.



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

Sampling estimation of availability parameters

A.1 Test of 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 unavailability 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.

In those situations where greater flexibility in managing the Type I and Type II errors is required, an alternative non-minimal test (the SPRT methodology described in A.4) should be used.

A.1.1 A minimal test of availability

– Phase I: Attempt 4 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).

   A statistical analysis in this Phase I test is presented in A.1.2. As an alternative, the SPRT methodology presented in A.4 may be used in place of the above Phase I test. The SPRT methodology provides more flexibility in controlling Type I and Type II errors.

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 ratio 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.1.2 The statistical basis for the Phase I test with N = 4

By definition, the service is unavailable if the probability of call set-up error plus the probability of call set-up failure is greater than 0.9:

$$cfp + cep > 0.9 \qquad \text{(A-1)}$$

Therefore we take the following as the null hypotheses, H_0, and the alternative hypothesis, H_a:

$$H_0: \quad cep + cfp < z$$

(A-2)

$$H_a: \quad cep + cfp > 0.9$$

Using the X.137 minimal availability test, the probability of Type I and Type II errors are given below:

- Pr (Type I error) $< z^4$ ≈ 0.24 (for $z = 0.7$);
- Pr (Type II error) $< 1 - (0.9)^4$ ≈ 0.35.

Table A.1 presents the probabilities of various events given the actual level of call set-up failure and error probability.

Table A.1 shows the extent to which this test protects against calling an available state unavailable. Also with more than 65% probability, the test will correctly identify the unavailable state.

The SPRT methodology of A.4 below should be used as an alternative, non-minimal test for those situations where greater flexibility in managing the Type I and Type II errors is required.

Table A.1/X.137 – Error performance of the minimal test

Actual cep + cfp	Probability of correctly identifying the available state	Probability of correctly identifying the unavailable state	Probability of identifying the available state as unavailable Pr (Type I error)	Probability of identifying the unavailable estate as available Pr (Type II error)
0.1	0.9999	NA	0.0001	NA
0.2	0.998	NA	0.002	NA
0.3	0.992	NA	0.008	NA
0.4	0.974	NA	0.026	NA
0.5	0.937	NA	0.063	NA
0.6	0.87	NA	0.13	NA
0.7	0.76	NA	0.24	NA
0.8	0.59	NA	0.41	NA
> 0.9	NA	> 0.65	NA	< 0.35
0.95	NA	0.81	NA	0.19
0.99	NA	0.96	NA	0.04
0.999	NA	0.996	NA	0.004
NA Not applicable				

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 duration 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".[2]

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.[3] The only consequence of

[2] Each counter is initially set to zero.

[3] 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.

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.

A.4 SPRT methodology

A.4.1 SPRT test procedure

For Phase I, perform a SPRT[4] of the following pair of hypotheses, utilizing an appropriate value of z ($z < 0.9$). This test will use successive call set-up attempts across the section under test, A. If the SPRT decides that H_0 is true, then proceed to Phase II of the Minimal Test, otherwise if the SPRT decides that H_a is true, then terminate the test and conclude that the service is unavailable due to the sum of call failure probability and call error probability exceeding the service outage threshold of 0.9.

– H_0: cep + cfp < z (service outage criterion is not met);

– H_a: cep + cfp > 0.9 (service outage criterion is met).

A.4.2 SPRT methodology

The hypotheses used are based on the provisional criteria specified in this Recommendation according to which the sum of the call failure probability and call error probability exceeding 0.9 (i.e. cfp + cep > 0.9) determines a service outage. Implicit in this criteria is the assertion that one can in fact distinguish between cfp + cep > 0.9 and cfp + cep < 0.9. However, the best that one can really do is to distinguish between cfp + cep > 0.9 and cfp + cep > z ($0 < z < 0.9$).

The Sequential Probability Ratio Test (SPRT) methodology controls both the Type I and Type II errors simultaneously, and, for deciding between two alternatives, is the most powerful statistical tool available[5]. For simplicity this annex uses the same probability of a wrong decision for both Type I and Type II errors[6]. The assumption of a binomial distribution for the success or failure of an individual call set-up attempt is made in this subclause.

---

4 The SPRT methodology is given in A.4.2.

5 See George G. Roussas, *A First Course in Mathematical Statistics*, by Addison-Wesley.

6 In the formulas below, *error* = Pr (Type I error) = Pr (Type II error). Values of error from 0.01 to 0.10 are commonly used.

The material below discusses the hypotheses to be tested, the decision rule, the upper and lower decision points, the expected number of call set-up attempts, and the least number of successful or failed attempts to end the SPRT.

A.4.3 Hypotheses

The SPRT uses the following pair of hypotheses, where H_0 corresponds to the outage threshold not being exceeded, and H_a corresponds to the outage threshold being exceeded.

– H_0: $cep + cfp < z$ (service outage criterion is not met);

– H_a: $cep + cfp > 0.9$ (service outage criterion is met).

A.4.4 Decision rule and upper and lower decision points

The SPRT reaches a decision based on observed performance being greater or less than a particular value. These values depend on the number of observations taken (n) and are denoted by UD(n) and LD(n) respectively. Formulas for LD(n) and UD(n) are given below after the decision rule.

A.4.4.1 Decision rule

If, upon making n attempts, the number of failed attempts is greater than UD(n), then the criterion for service outage is met.

$$\left\{ \sum_{j-i}^{\alpha} x_j \geq \mathrm{UD}(n) \right\} \qquad \text{(A-3)}$$

If, upon making n attempts, the number of failed attempts is less than LD(n), then the criterion for service outage is not met.

$$\left\{ \sum_{j-i}^{\alpha} x_j \leq \mathrm{LD}(n) \right\} \qquad \text{(A-4)}$$

Keep attempting calls until a decision is reached.

Formulas for UD(n) and LD(n):

$$\mathrm{UD}(n) = \frac{\left(\log\left(\frac{(1-error)}{(error)}\right) - n * \log\left(\frac{(1-0.9)}{(1-z)}\right)\right)}{\log\left(\frac{0.9*(1-z)}{z*(1-0.9)}\right)} \qquad \text{(A-5)}$$

$$\mathrm{LD}(n) = \frac{\left(\log\left(\frac{(error)}{(1-error)}\right) - n * \log\left(\frac{(1-0.9)}{(1-z)}\right)\right)}{\log\left(\frac{0.9*(1-z)}{z*(1-0.9)}\right)} \qquad \text{(A-6)}$$

A.4.5 Expected number of call set-up attempts

The expected number of call set-up attempts until the SPRT reaches a decision, is useful in determining the length and cost of the test. Under H_0 and H_a respectively, the expected number of call set-up attempts are $E_0(N)$ and $E_a(N)$. Asymptotic approximations for them are as follows, and are based on the use of a binomial probability for the sum of call set-up error and call set-up failure. Calculations resulting in entries in Table A.3 greater than 100 were made up using these approximations. The rest of Table A.3 was constructed using iterative matrix techniques yielding more precise values.

$$E_0(N) \approx \frac{\left((1-2*error)*\log\left(\frac{(error)}{(1-error)}\right)\right)}{z*\log\left(\frac{0.9*(1-z)}{z*(1-0.9)}\right) + \log\left(\frac{(1-0.9)}{(1-z)}\right)} \qquad \text{(A-7)}$$

$$E_a(N) \approx \frac{\left((1-2*error)*\log\left(\frac{(1-error)}{(error)}\right)\right)}{0.9*\log\left(\frac{0.9*(1-z)}{z*(1-0.9)}\right) + \log\left(\frac{(1-0.9)}{(1-z)}\right)} \qquad \text{(A-8)}$$

A.4.6 Least number of failures or successes to end a SPRT

The quantities L and U represent the least number of call set-up attempts required by the SPRT to decide if H_0 or H_a respectively are true. If all L call set-up attempts are successful, then the outage criterion is not met, while if all U of the call set-up attempts fail, then the outage criterion is met. The SPRT test will often continue after the values of U or L are reached, but they are the least values at which a decision in favour of H_a or H_0 respectively could be taken. Tabulated values of U, L, $E_0(N)$ are provided in Tables A.2 and A.3.

Table A.2/X.137 – U/L – Minimum number of call set-up attempts

Percent error			
z	**10%**	**5%**	**1%**
0.85	39/6	52/8	81/12
0.80	19/4	25/5	40/7
0.75	13/3	17/4	26/6
0.70	9/2	12/3	19/5
0.65	7/2	10/3	15/4
0.60	6/2	8/3	12/4
0.55	5/2	6/2	10/4
0.50	4/2	6/2	8/3
0.45	4/2	5/2	7/3
0.40	3/2	4/2	6/3
0.35	3/2	4/2	5/3
0.30	2/2	3/2	5/3
0.25	2/2	3/2	4/3
0.20	2/2	2/2	4/3
0.15	2/2	2/2	3/3
0.10	1/1	2/2	3/3

NOTE 1 – U is the least number of successive call set-up errors or failures needed to terminate the SPRT (deciding in favour of H_a).

NOTE 2 – L is the least number of successive call set-up successes needed to terminate the SPRT (deciding in favour of H_0).

NOTE 3 – *z* is the cutoff value in the null hypothesis (H_0).

NOTE 4 – The column headings represent the specified error rates. Due to approximations used in the SPRT, the error rates are bounded above by error/(1 – *error*). The differences are small over the range of error rates being considered.

Table A.3/X.137 – $E_a(N)$ / $E_0(N)$ – Expected number of call set-up attempts

Percent error			
z	**10%**	**5%**	**1%**
0.85	161.3/143.7	243.2/216.6	413.3/368.1
0.80	51.5/45.3	74.5/65.1	122.7/101.4
0.75	27.4/22.3	39.3/32.5	63.9/52.2
0.70	17.1/14.4	24.5/20.1	40.1/32.3
0.65	12.1/10.2	17.3/13.9	27.9/22.2
0.60	9.2/7.4	13.3/10.8	21.0/16.3
0.55	7.4/6.1	10.0/7.7	16.5/13.0
0.50	5.8/4.9	8.6/6.5	13.0/10.1
0.45	5.4/4.3	7.0/5.4	10.9/8.4
0.40	4.0/3.7	5.6/4.8	8.8/7.2
0.35	3.9/3.4	5.5/4.3	7.0/5.7
0.30	2.6/2.8	4.1/3.7	6.5/5.2
0.25	2.6/2.6	3.7/3.3	5.4/4.6
0.20	2.4/2.5	2.7/2.8	5.0/4.1
0.15	2.4/2.3	2.5/2.7	3.7/3.7
0.10	1.0/1.0	2.4/2.4	3.7/3.7

NOTE 1 – $E_a(N)$ is the expected number of trials needed to terminate the SPRT when the "network" is unavailable.

NOTE 2 – $E_0(N)$ is the expected number of trials needed to terminate the SPRT when the "network" is available.

NOTE 3 – *z* is the cutoff value in the null hypothesis (H_0).

NOTE 4 – The column headings represent the specified error rates. Due to approximations used in the SPRT, the error rates are bounded above by error/(1 – *error*). The differences are small over the range of error rates being considered.

Annex B

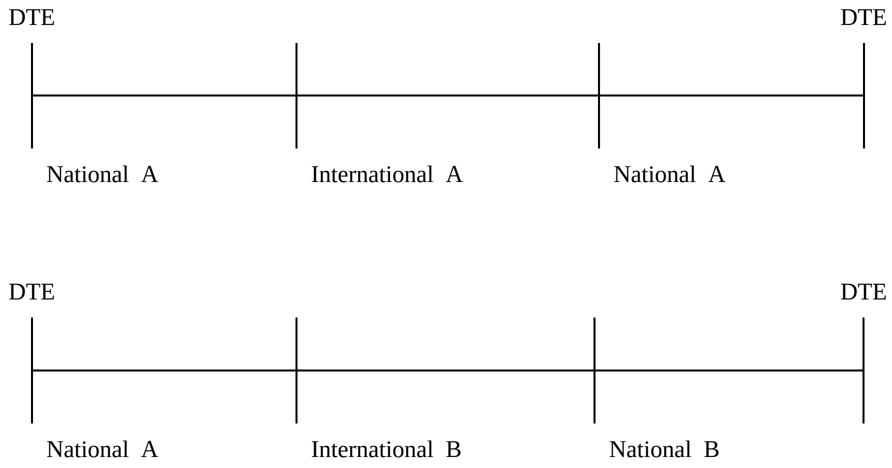
Representative end-to-end availability performance

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 this Recommendation. 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 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 and B.2. 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 this Recommendation.

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 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:

$$T = [T_1^{-1} + T_2^{-1} + \cdots + T_i^{-1} + \cdots + T_N^{-1}]^{-1} \qquad \text{(B-1)}$$

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). The end-to-end performance objective is then:

$$[1200^{-1} + 1600^{-1} + 1200^{-1}]^{-1} = 436 \text{ hours} \qquad \text{(B-2)}$$

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 the 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 communication**
Series Z	Programming languages