INTERNATIONAL TELECOMMUNICATION UNION

CCITT

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

X.138

(09/92)

DATA COMMUNICATION NETWORKS

NETWORK ASPECTS

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

Recommendation X.138

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.138 was prepared 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.

© ITU 1993

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the ITU.

Recommendation X.138

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

(1992)

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 networks providing data transmission services;

(e) that Recommendation X.96 specifies call progress signals in public data networks;

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

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

(h) that Recommendation X.140 defines general Quality of Service parameters for communication via public data networks;

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

(j) that Recommendation X.135 specifies speed of service (delay and throughput) performance values for public data networks when providing international packet switched service;

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

(l) that Recommendation X.137 specifies availability performance values for public data networks when providing international packet switched service,

(m) that Recommendation X.139 specifies echo, drop, generator and test DTEs for measurement of performance values,

unanimously declares

that this Recommendation describes the various measurement architectures which can be used to measure performance values for packet switched public data networks.

1 Introduction

This Recommendation describes architectures for measuring performance values for packet-switched public data networks. Section 3 defines a measurement method using controlled and monitored sources and sinks. Section 4 describes a means of synchronizing measurement equipment. Annex A provides detailed information on the calculation of packet switched network performance statistics, Annex B provides information on factors which can affect their observation and Annex C provides examples of the use of measurements.

Note that other measurement devices and procedures that adhere to the definitions of Recommendations X.134, X.135, X.136 and X.137 are also acceptable for estimating the performance of packet-switched public data networks.

1.1 *General*

The network performance may need to be measured for reasons including the following:

- to assist network planning;
- network performance parameters may be included in contractual arrangements;
- Quality of Service parameters may need to be measured for other Recommendations;
- network operators may require general service descriptions.

1.2 *Parameters to be measured*

The speed of service (delay and throughput) performance values for packet-switched public data networks defined in Recommendation X.135 are as follows:

- Call set-up delay (Recommendation X.135, § 2)
- Data packet transfer delay (Recommendation X.135, § 3)
- Throughput (Recommendation X.135, § 4)
- Clear indication delay (Recommendation X.135, § 5)
- Clear confirmation delay (Recommendation X.135, § 5)

  Note – Recommendation X.135 does not specify values for this parameter.

The Quality of Service parameters defined in Recommendations X.136 and X.137 are as follows:

- Call set-up error probability (Recommendation X.136, § 2)
- Call set-up failure probability (Recommendation X.136, § 2)
- Residual error rate (Recommendation X.136, § 3)
- Reset stimulus probability (Recommendation X.136, § 3)
- Reset probability (Recommendation X.136, § 3)
- Premature disconnect stimulus probability (Recommendation X.136, § 3)
- Premature disconnect probability (Recommendation X.136, § 3)
- Call clear failure probability (Recommendation X.136, § 4)
- Availability (Recommendation X.137)

It should be noted that the expected values for many of the various Quality of Service parameters defined in Recommendations X.136 make their estimation within reasonable bounds likely to take an excessive number of samples (during which time the structure or performance of the network may have changed). Thus it may not be practical to estimate these parameters between a single pair of interfaces. However, it may be possible to estimate the values with more certainty averaged over a whole network.

1.3 *Accuracy requirements of measurements*

1.3.1 *Measurement objectives*

There is always a cost versus sample size trade off which must be resolved when attempting to estimate a given parameter, and as the cost of taking an observation can have a considerable effect on the final estimate achieved, this Recommendation does not recommend a minimum number of observations. Additionally, in certain situations, estimating the parameter with great precision may not be as critical as in others.

Estimates of means, variances, percentile, modes, maxima and minima are all examples of parameter estimation. In reporting any such estimated parameters it is always recommended that a measure of the precision of the estimate be included. The variance of the estimate, or a confidence interval about the estimate, are two common methods of expressing the estimate's precision.

1.3.2 *Reference events*

The time of occurrence of various reference events is the basis for the definitions of the speed of service parameters listed above. In this Recommendation and in Recommendation X.139, times are specified with respect to exit and entry events relative to the test equipment. Thus the time of occurrence of an entry event into such a DTE is the time at which the last bit of the closing flag of the layer two frame carrying the packet enters the DTE from a circuit section and the time of occurrence of an exit event from such a DTE is the time at which the first bit of the address field of the layer two frame carrying the packet enters the circuit section from the DTE.

Comprehensive lists of X.25/X.75 packet layer reference events are given in Tables 1/X.134 and 2/X.134, and the definition of a packet layer reference event is given in Recommendation X.134, § 3.

2 Measurement architectures

This section provides an overview of the various architectures which may be used to measure the Quality of Service parameters for packet-switched public data networks as specified in Recommendations X.135, X.136 and X.137, together with general considerations for measuring the performance parameters.

2.1 *General considerations and measurement methodologies*

A general measurement methodology involves setting up a call to a data sink and generating a known and sufficient quantity of packet traffic. The protocol and user information signals transferred across the user/network (DTE/DCE) interfaces are observed in real time and a chronological event history is compiled. This history can then be analysed to provide a measurement of the performance parameters.

Thus, measurements of packet-switched networks require a source, a sink, and one or more monitors. A source transmits call set-up requests, data packets, or call clearing requests through the portions under test. A sink receives and acknowledges call processing or data from the portions under test. The function of the monitor is to record (or record and timestamp) the relevant reference events. The monitor function(s) should be placed as near as possible to the boundaries of the portions under test. Differences in location between the monitor functions and the appropriate boundaries must be compensated for in the calculation of the performance of the portions.

Sources and sinks can be either controlled or non-controlled. Controlled sources and sinks are under the control of the test programme and must respond quickly to events exiting the portions under test. Examples of controlled sources or sinks are stand-alone test equipment, special software within network equipment (for example within packet switches), and special programs within customer applications. Non-controlled sources or sinks are sources or sinks not under the direct control of a test programme. Non-controlled sources and sinks may not always respond quickly to network events. The most important examples of non-controlled sources and sinks are live customer applications, generating and receiving traffic according to their own needs.

A monitor function can be provided by stand-alone test equipment "T" connected at the appropriate X.25 or X.75 interface. Alternatively, a monitor function can reside in the test device that provides the source or sink function. Network equipment (such as packet switches) and customer equipment (for example, DTEs) can also be programmed to record reference events and serve the monitor function.

Various combinations of monitors and controlled and non-controlled sources and sinks can be used to measure packet network performance. Figure 1/X.138 illustrates some of these possibilities. The architectures are identified by specifying whether the source and sink are controlled (C) or non-controlled (N), and whether the two portion boundaries are monitored (M) or unmonitored (U). The notation (C,M/N,U) indicates a controlled source, a non-controlled sink, monitored at the source boundary, and no monitoring at the sink boundary. When both the source and the sink are controlled and there are time synchronized monitor functions at both boundaries (C,M/C,M), all the parameters defined in Recommendations X.135, X.136 and X.137 can be measured without further assumptions. Other test architectures are more limited because they cannot be used to measure all of the parameters.

Some applications of the (C,M/C,U) architecture using drop, generator and test DTEs are described in Recommendation X.139.

A means of synchronizing monitoring equipment for use in conjunction with a (C,M/C,M) architecture, to measure throughput and delay, is described in § 4 of this Recommendation.

Subsection 2.2 lists the primary performance parameters (as specified in Recommendations X.135, X.136 and X.137) and identifies the test architectures that can be used to measure them. In some cases the test architectures can be used to measure parameters only if certain additional assumptions are made and such assumptions are described.

2.2 *Performance parameters and measurement architectures*

2.2.1 *Call set-up delay*

Call set-up delay can best be measured with monitors at both portion boundaries. If the sink is known to accept set-up requests with constant or insignificant delay and if the probability of call set-up error events is insignificant, call set-up delay can be measured without a sink-side monitor by subtracting the known sink delay from the one-sided measurement.

2.2.2 *Data packet transfer delay and clear indication delay*

Both data packet transfer delay and clear indication delay require a sink-side monitor time synchronized with either a source-side monitor or with the source itself.

2.2.3 *Clear confirmation delay*

Clear confirmation delay only requires a monitored source (or sink). As it is a local parameter, clear confirmation delay is not discussed further in this Recommendation.

2.2.4 *Throughput capacity*

Throughput capacity measurements require controlled sources and sinks rapidly transmitting and acknowledging data packets. In effect, the throughput capacity of the source and sink must be greater than or equal to the throughput capacity of the portions under test.

As noted in Recommendation X.135, § 4.2, steady state throughput is the same when measured at every pair of portion boundaries of a virtual connection. Thus, assuming no user data bits are lost, added, or inverted in transfer, a steady state throughput measurement can be made at any single portion boundary within a virtual connection.

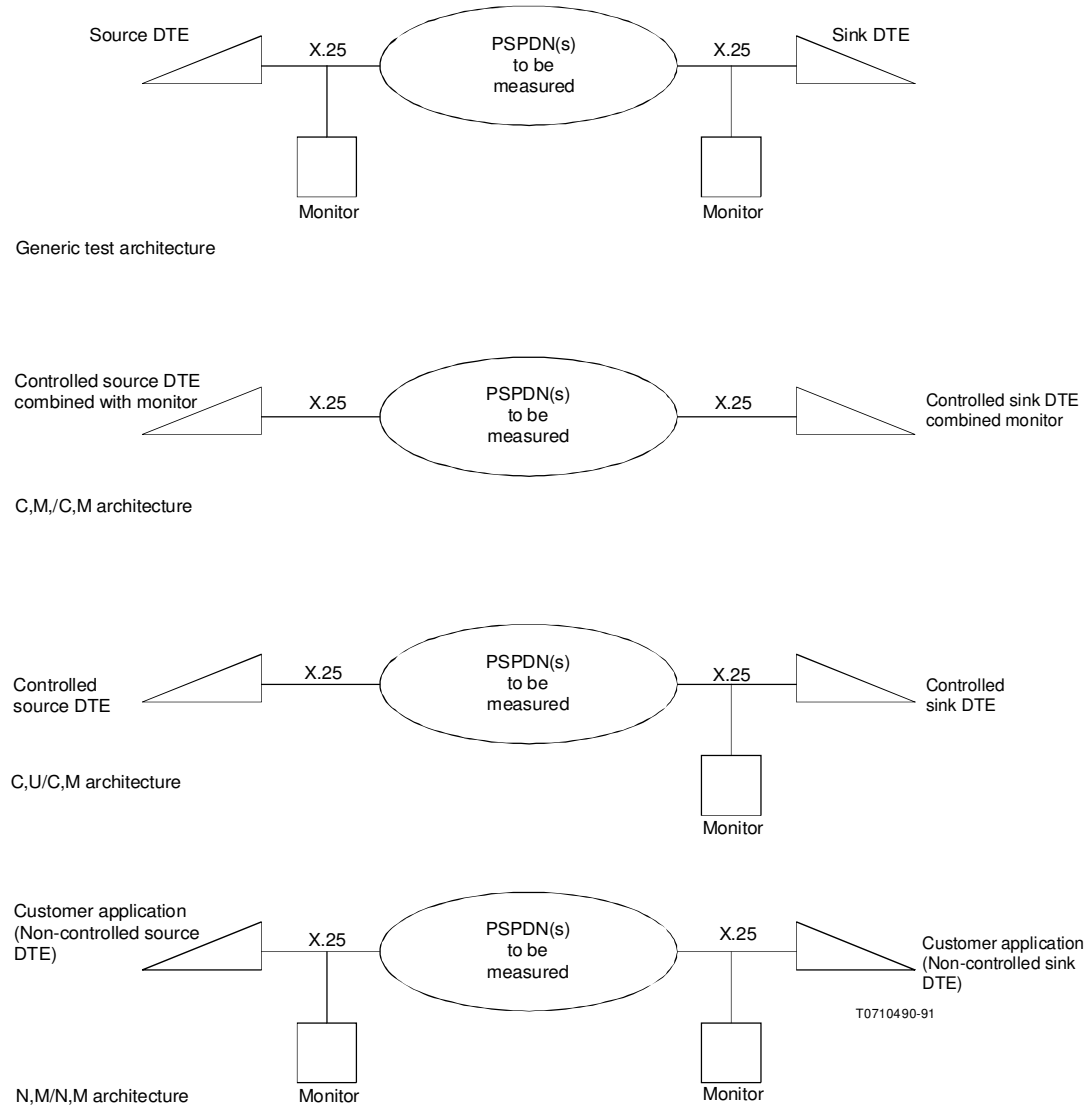


FIGURE 1/X.138
Example test architectures

2.2.5 *Call set-up error probability*

Call set-up error probability can only be measured if there is monitoring at both portion boundaries.

2.2.6 *Call set-up failure probability*

Call set-up failure probability can best be measured with monitors at both boundaries. The sink device must be fast enough that it does not significantly contribute to the probability of timing out. If the sink can be relied upon to accept every call set-up attempt, call set-up failure probability can be measured without a monitor at the sink-side.

2.2.7 *Residual error rate*

Residual error rate requires monitoring at both boundaries or a controlled source transmitting a known sequence of user data. Both these architectures allow the transmitted and received user data to be compared.

2.2.8 *Reset and premature disconnect performance*

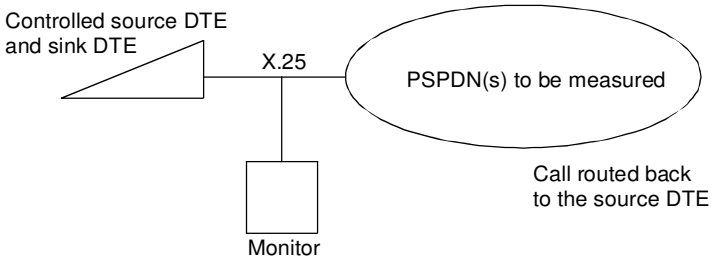
Reset stimulus probability and premature disconnect stimulus probability can be measured by a single monitor at a single boundary. Reset probability and premature disconnect probability require monitors at both boundaries. These allow for distinguishing the resets and clears that exit the portions under test from the resets and clears stimulated at the distant boundary.

2.2.9 *Call clearing failure probability*

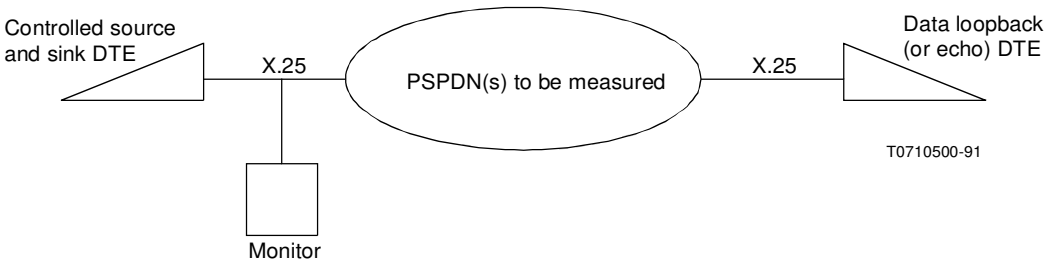
Call clearing failure probability can best be measured with monitors at both boundaries. If the transmission of the clear request by a controlled test device is reasonably well synchronized with the sink-side monitor, the monitor can anticipate clearing and observe call clearing failures.

2.3 *Loopbacks*

Loopbacks provide an alternative measurement architecture allowing a single test device to serve as both a source and a sink. Figure 2/X.138 illustrates the two possibilities for using loopbacks in packet-switched network performance measurements.



Routing loopback architecture



Data loopback architecture

FIGURE 2/X.138
Loopback architectures

Routing loopbacks are established by the packet network(s) by routing virtual circuits through one or more switching functions (or through multiple networks) back to the originating interface. The result is a virtual circuit that originates on one logical channel and terminates on a different logical channel on the same test device. A monitor at the portion boundary can then be used to measure all of the primary performance parameters. If the routing loopback is significantly different from an ordinary virtual connection through the portions (for example, in the number of switches or distance traversed), the performance calculations must compensate for those differences.

Data loopback can be used to measure data packet delay, throughput capacity and residual error rate. A data loopback can be provided by special software within network equipment, by stand-alone test equipment, or by special test programmes within customer applications. A data loopback device terminates the virtual connection, removes the data from the incoming data packets and returns that data through the same virtual connection in new outgoing data packets. The data loopback should rapidly acknowledge data packets and return the user data without significant delay or error. If the portions under test have symmetric delays and residual error rates, data packet delay and residual error rate are half of what is calculated by comparing outgoing and incoming data packets at a source-side monitor. The throughput capacity measurements made at a monitored boundary yield the smaller value of the throughput capacity for the two directions.

An application of these techniques, using echo and test DTEs is described in Recommendation X.139.

2.4 *Summary of measurement architectures*

The possible combinations of controlled and non-controlled sources and sinks and monitored and unmonitored boundaries yield twelve different architectures. Routing loopbacks and data loopbacks create two more possible architectures. In Table 1/X.138 these 14 architectures are listed with an indication of their ability to measure each primary parameter.

TABLE 1/X.138

Summary of measurement architectures

Measurement	Primary parameters											
Architecture	csd	dpd	tc	cid	cep	cfp	rer	rsp	rp	pdsp	pdp	ccfp
CM/CM	Y	Y,1	Y	Y,1	Y	Y	Y	Y	Y	Y	Y	Y
NM/CM	Y	Y,1	N	Y,1	Y	Y	Y	Y	Y	Y	Y	Y
CM/NM	Y,2	Y,1	N	Y,1	Y	Y,2	Y	Y	Y	Y	Y	Y
NM/NM	Y,2	Y,1	N	Y,1	Y	Y,2	Y	Y	Y	Y	Y	Y
CU/CM	N	Y,3	Y	Y,3	N,4	N,4	Y,5	Y,6	N,7	Y,6	N,7	Y,3
CU/NM	N	Y,3	N	Y,3	N,4	N,4	Y,5	Y,6	N,7	Y,6	N,7	Y,3
CM/CU	Y,8	N	Y	N	N	Y,9	N	Y,6	N,7	Y,6	N,7	N
NM/CU	Y,8	N	N	N	N	Y,9	N	Y,6	N,7	Y,6	N,7	N

TABLE 1/X.138 (*cont.*)

Measurement	Primary parameters											
Architecture	csd	dpd	tc	cid	cep	cfp	rer	rsp	rp	pdsp	pdp	ccfp
NM/NU	N,10	N	N	N	N	Y,9	N	Y,6	N,7	Y,6	N,7	N
CM/NU	N,10	N	N	N	N	Y,9	N	Y,6	N,7	Y,6	N,7	N
NU/NM	N	N	N	N	N	N	N	Y,6	N,7	Y,6	N,7	N
NU/CM	N	N	N	N	N	N	N	Y,6	N,7	Y,6	N,7	N
rlb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
dlb	N,13	Y,11	Y	N,13	N,13	N,13	Y,12	Y,6	N,7	Y,6	N,7	N,13

csd Call set-up delay

rsp Reset stimulus probability

dpd Data packet transfer delay

tc Throughput capacity

cid Clear indication delay

cep Call set-up error probability

cfp Call set-up failure probability

rer Residual error rate

rp Reset probability

pdsp Premature disconnect stimulus probability

pdp Premature disconnect probability

ccfp Call clear failure probability

rlb Routing loopback

dlb Data loopback

1. Assumes the two monitors are synchronized.
2. Assumes the non-controlled sink is reasonably fast in responding.
3. Assumes the source's creation of packets is synchronized with the sink-side monitor.
4. Cannot observe event "d" (see Figure 2/X.136).
5. Assumes the user data created by the source is known in advance.
6. At the monitored boundary only
7. Cannot distinguish between events originating within the portions and events caused by stimuli at the distant boundary.
8. Assumes there are no call set-up error events and the sink accepts the call with known or insignificant delay.
9. Assumes no unsuccessful call set-ups due to the sink devices.
10. Cannot exclude delays due to the sink devices.
11. Assumes dpd is equal in both directions.
12. Assumes rer is equal in both directions.
13. The data loopback function only operates in the user data transfer phase.

3 Measurement method using controlled and monitored source and controlled and monitored sink (C,M/C,M)

Three types of test procedures (access, data transfer, and disengagement) are used to obtain values for the primary parameters described in this Recommendation. Values for call set-up delay, call set-up error probability and call set-up failure probability are obtained using access trials. Values for data packet transfer delay, throughput capacity, residual error rate, reset stimulus probability, reset probability, premature disconnect stimulus probability and premature disconnect probability are obtained using data transfer trials. Clear indication delay and call clear failure probability values are obtained using disengagement trials. The availability parameters, service availability and mean time between service outages, are measured using a combination of the three types of performance trials.

Each performance trial consists of two procedures:

– *Data extraction:* Packet layer reference events associated with the trial are created (or observed), time stamped and recorded at the appropriate section boundaries.

– *Data reduction:* The recorded reference event histories are processed, consistent with the performance parameter definitions, to determine the trial's outcome.

3.1 *General assumptions and constraints*

The following text describes specific performance trials and data extraction and reduction procedures. Other equivalent (or superior) procedures can be designed and used. General assumptions and constraints which underlie these test procedures are specified throughout the text.

3.2 *Test devices*

These procedures use the (C,M/C,M) test mode (described in § 2). The procedures assume that the test devices conform fully to the protocols employed at the monitored interfaces, that is, in addition to generating the packet level reference events appropriate for the particular trial, the test devices should respond correctly to the events generated by the section under test. The test devices should correctly record and accurately time stamp every reference event on the relevant logical channels. All such reference events should be recorded and time stamped regardless of whether the event was expected. The complete record of reference events will be used in the determination of the trial's outcome.

3.3 *Trial sequences*

Each of the data extraction procedures defines an elementary sequence of steps in which a single access trial, one or more data transfer trials, or a single disengagement trial is conducted. To satisfy any common measurement objective, multiple trials must be completed. The trials described here may be conducted in any arbitrary sequence linked together using the trial linkages described below. These linkages facilitate repeated trials for a single parameter and economical testing of several parameters.

The following two procedures show how to achieve the initial states of the access, data transfer and disengagement trials. There are many possible final states of the B_i and B_j boundaries at the end of these trials. A generic approach is taken which ensures the attainment of the desired initial states regardless of the current state of the B_i and B_j boundaries. The procedures are written with reference to achieving a state at a particular boundary.

3.3.1 *Procedure for establishing state p1*

The following steps will create state p1 at a boundary:

1) Issue a SABM/SABME if the data link layer is not available and up (r1);

2) At an X.25/X.75 boundary issue a DTE/STE clear request on the chosen logical channel;

3) Wait for confirmation of clearing.

3.3.2 *Establishing p4/d1 at a boundary*

The following steps will create the state p4/d1 at a boundary:

1) Get to state p1 (see above) (1);

2) At the B_i boundary issue a call request with the chosen logical channel, called and calling addresses;

3) At remote boundary B_j, wait for an incoming call or call request packet (2) and issue the corresponding call accepted or call connected packet;

4) Wait for the corresponding call connected packet at the B_i boundary.

Note 1 – A channel allowed for incoming calls must be cleared at the called DTE.

Note 2 – The virtual connection may be established on an logical channel other than the one originally chosen at the called DTE.

3.4 *Trial failures and recovery*

None of the data extraction or data reduction procedures include detailed recovery algorithms for recovering from failed trials (for example, failed call set-up attempts, resets, restarts or residual error). The trial linkages assume nothing about the states of the interfaces following a trial, so recovery routines are not absolutely necessary. However, all of the procedures (data extraction and data reduction) could be made more robust by implementing failure recovery mechanisms.

3.5 *Matching packet level reference events*

The data reduction procedures all require matching of corresponding packet level reference events at the two test boundaries. In general, any reasonable method for matching protocol events (PEs) is sufficient. If the clocks in the test devices are synchronized, timing information can be used in the method. If the section under test preserves data packet sequence numbers that information can be used in the method that matches data packets. Data packet matching may also be done by comparing user data fields. If the data extraction procedures also include failure recovery mechanisms, the data reduction procedures must be more sophisticated in their ability to match packets (recognizing and compensating for lost, extra or errored data packets).

3.6 *Testing during service outages*

By definition, the performance of the primary parameters is not to be evaluated during service outages. In each of the data reduction procedures a decision is made as to whether the service was available during the test. If it was available, the trial results are included in the cumulative statistics used in evaluating the primary parameters. If the service was unavailable, the trial results may be used only in measuring availability performance.

Determining the availability state in the data reduction phase is a difficult problem. A single trial failure is insufficient to declare unavailability, so in the absence of other evidence, the service is assumed to be available. However, if this trial was contained in a five minute interval where one or more primary parameters performed worse than their decision criteria (Table 2/X.137), the service should be declared unavailable. Thus the availability decision

performed in each individual data reduction procedure should take into consideration all trial results within plus or minus five minutes. Decisions about service availability can be delayed until all of the individual trials have been analysed. If this approach is used, the primary performance parameter cumulative counters should be corrected retroactively wherever outages are discovered.

3.7 *Access trial*

Values for call set-up delay, call set-up error probability and call set-up failure probability can be obtained using the following procedures.

3.7.1 *Access trial data extraction*

Figure 3/X.138 shows the access trial data extraction procedure. Boundaries B_i and B_j are at X.25 or X.75 interfaces bounding the set of virtual connection sections under test.

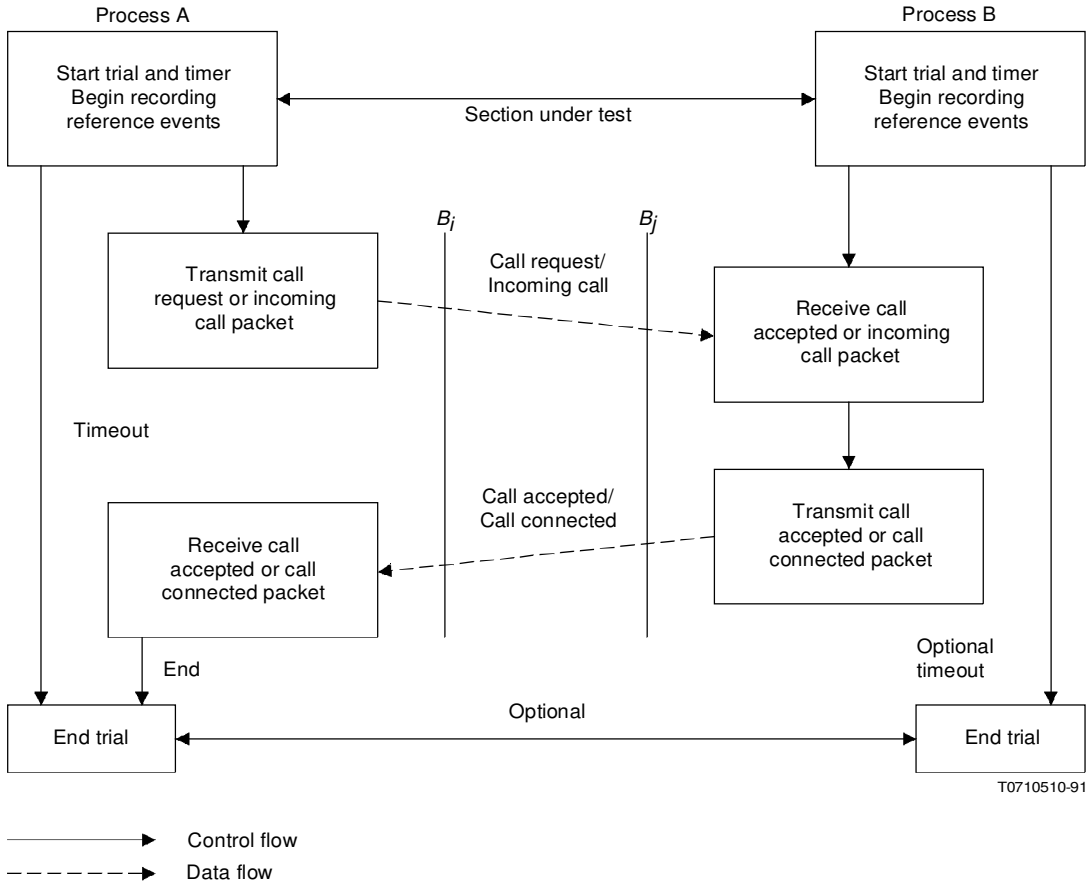


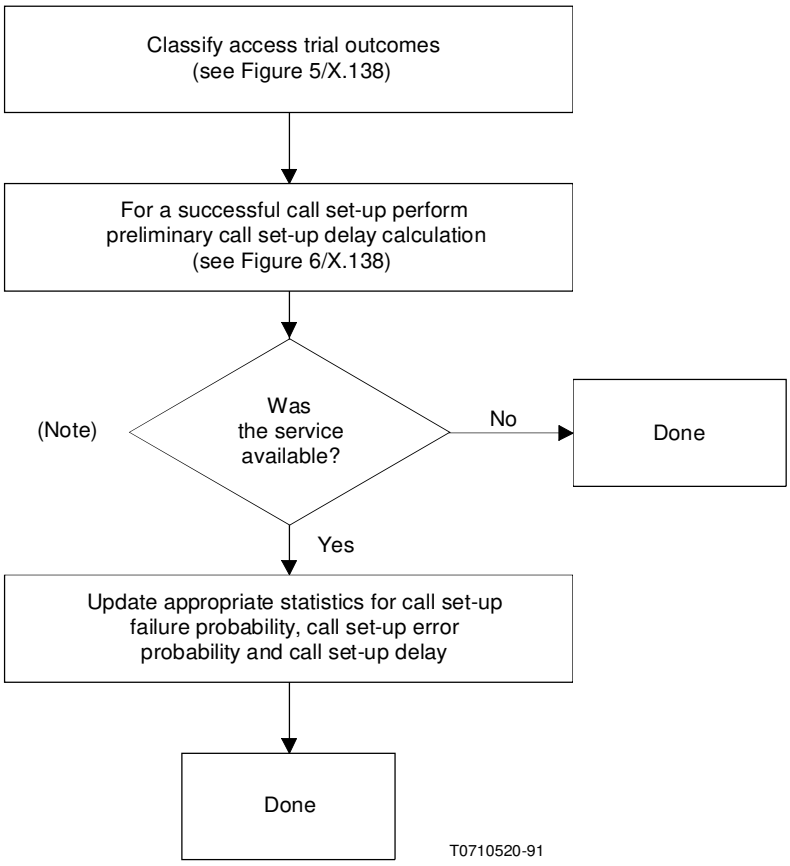
FIGURE 3/X.138
Access trial data extraction procedure

Logical channels of boundaries B_i and B_j must initially be in state p1. Process A will transmit a call request (or incoming call) packet and wait for the corresponding call accepted (or call connected) packet. Process A should wait no less than the 200 second call set-up failure threshold.

Timing of Process A and Process B must be sufficiently well synchronized so that Process B will be ready to receive the corresponding call request (or incoming call) packet. (In the case of a data loopback test arrangement, no synchronization is required). Process B will receive that packet and respond with the appropriate call accepted (or call connected) packet. The time required for this response will be subtracted from call set-up delay calculations; however, this response interval should be as small as possible to avoid substantially increasing the probability of exceeding the 200 second call set-up failure threshold.

3.7.2 *Access trial data reduction*

Figure 4/X.138 shows the access trial data reduction procedure. Each access trial is classified as illustrated in Figure 5/X.138. Figure 6/X.138 illustrates the steps needed to calculate call set-up delay for successful call set-up attempts. If the service was available during this trial, cumulative call set-up statistics may be updated. Cumulative counters can be kept for total call set-up errors, total call set-up failures and total call set-up delay (during successful attempts). Estimates of call set-up error probability, call set-up failure probability and mean call set-up delay can then be created by dividing these cumulative counters by a cumulative count of call set-up attempts.



Note – If the appropriate statistical considerations have been addressed, this decision may be used in estimating availability.

FIGURE 4/X.138
Access trial data reduction procedure

Figure 5/X.138 shows the use of the record of packet level reference events (identified as events "A", "B", "C" and "D") to determine if the call set-up attempt was successful or unsuccessful. Event "D" is said to have occurred only if the call accepted (or call connected) packet was received at B_i within the 200 second call set-up failure threshold. Otherwise it is assumed not to have occurred. If this process classifies any access trials as "Unsuccessful trial – cause outside portion boundaries" this is indicative of defective test devices which must be corrected.

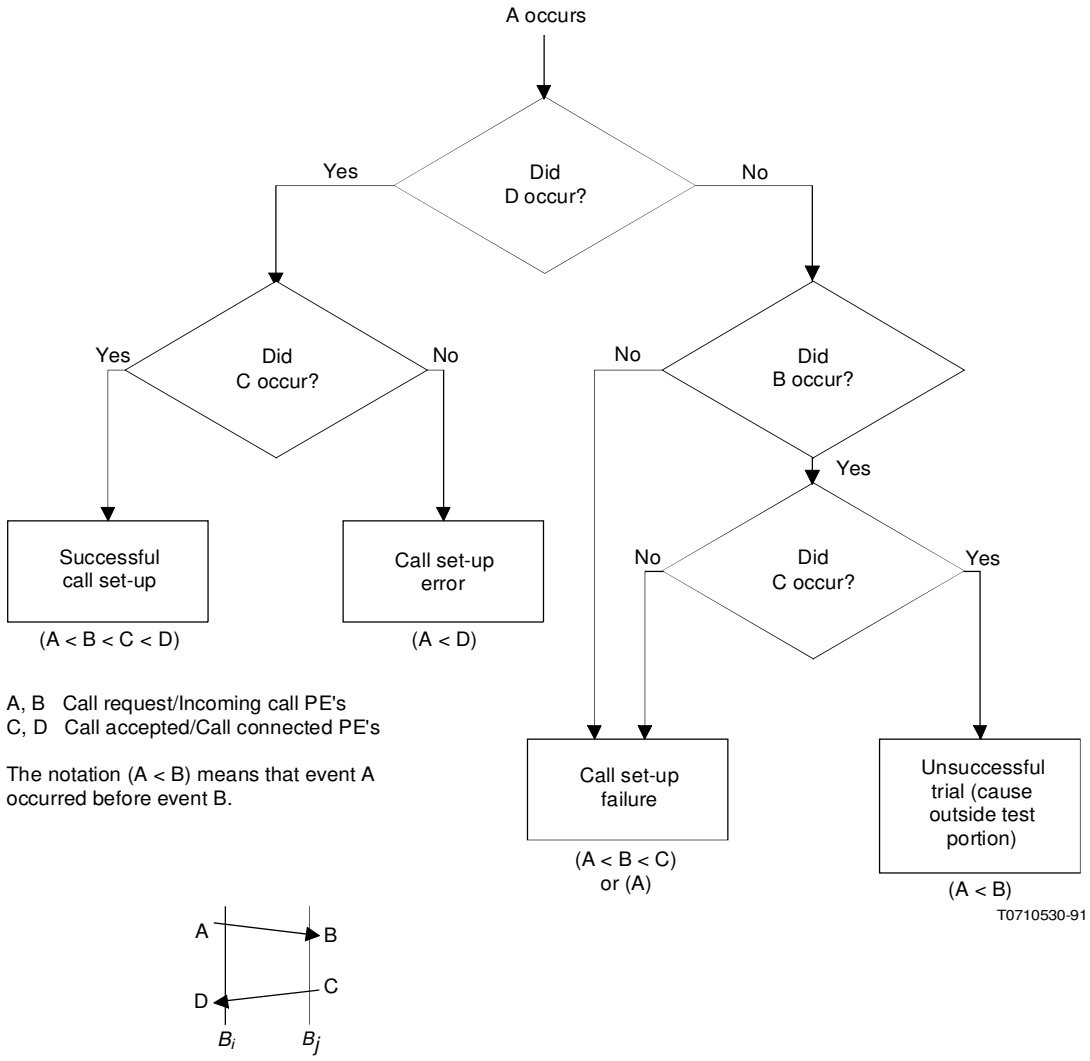
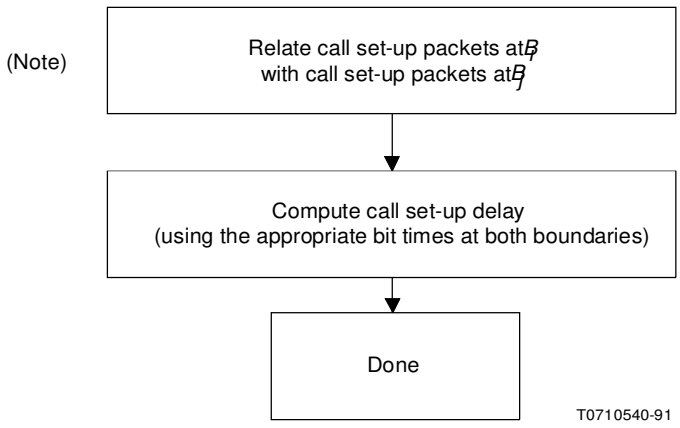


FIGURE 5/X.138
Classifying an access trial outcome

Calculating call set-up delay (see Figure 6/X.138) first requires matching packets recorded at boundary B_i with packets recorded at boundary B_j (see § 3.3). The exact timer values used in calculating call set-up delay depend on the location of the boundaries B_i and B_j, and the direction the packets are moving across those boundaries.



Note – Any reasonable method may be utilized.

FIGURE 6/X.138
Call set-up delay calculation

3.8 *Data transfer trial*

Values for data packet transfer delay, throughput capacity, residual error ratio, reset stimulus probability, reset probability, premature disconnect stimulus probability and premature disconnect probability can be obtained using the following procedures.

3.8.1 *Data transfer trial data extraction*

Figure 7/X.138 shows the data transfer trial data extraction procedure. In this case, a data loopback architecture is used. Boundaries B_i and B_j are at X.25 or X.75 interfaces bounding the set of virtual connection sections under test.

Logical channels at boundaries B_i and B_j must be in state p4/d1. Process A will transmit N-data packets. If throughput capacity is being tested, Process A should not delay the transmission of successive data packets except in response to window closings (in the figure: X = 0).

Process B should be ready to receive and record the reception of the corresponding data packets. (Warning: if packet segmentation or recombination occurs in the sections under test, the expected number of data packets at B_j may be different from the data packets transmitted by Process A). If throughput capacity is being tested, Process B should not delay acknowledging frames and packets received. Clocks in Process A and Process B must be sufficiently well synchronized (unless a data loopback is utilised) so that the difference between the two clocks is an insignificant fraction of the likely data packet transfer delay value. Synchronization accuracy of one or two milliseconds is usually sufficient.

Using a combination of timers and control mechanisms the total time allowed for Process B to receive the last data packet from Process A should be at least the 200 second residual error rate threshold.

If the trial is being used to evaluate reset probability, reset stimulus probability, premature disconnect probability and premature disconnect stimulus probability, both Process A and Process B should respond to resets, reset stimuli, premature disconnects and premature disconnect stimuli as specified in Recommendations X.25 and X.75. Process A should re-establish the virtual connection if it is prematurely disconnected. After a reset or connection re-establishment, Process A should resume transmission of data packets.

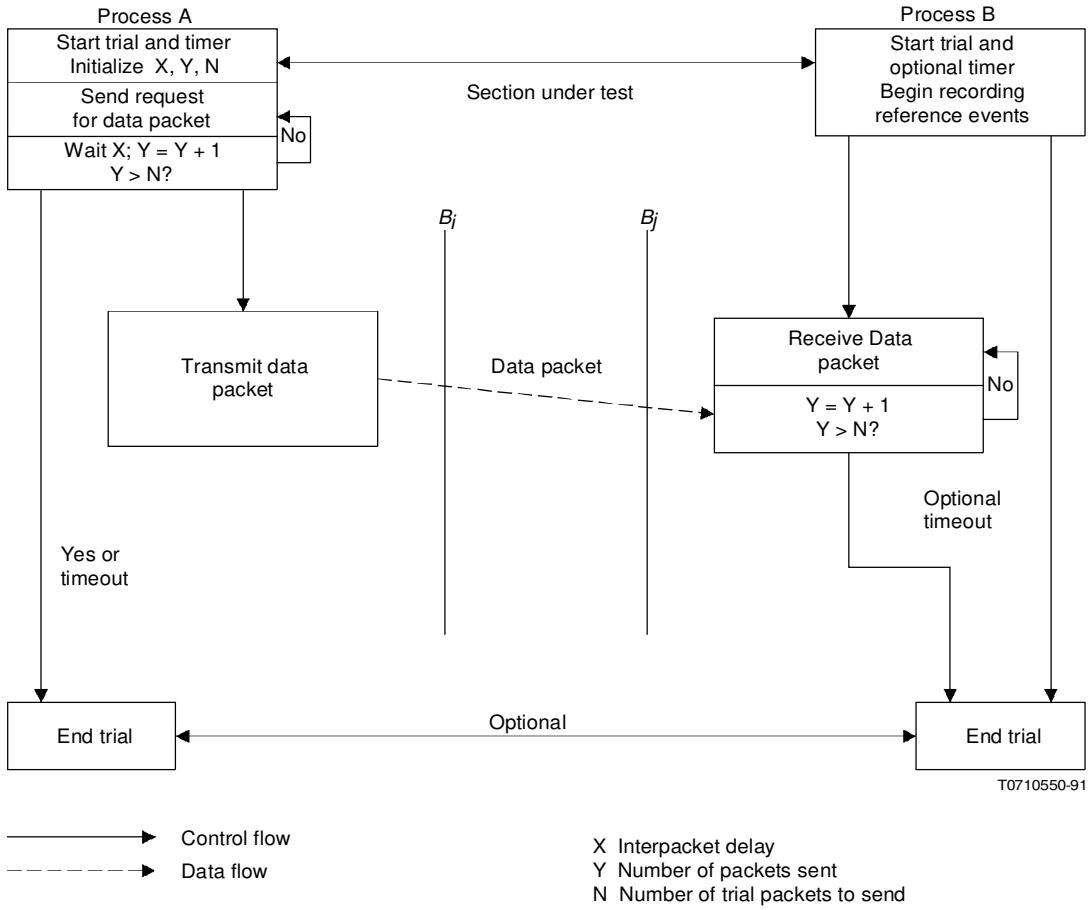


FIGURE 7/X.138

Data transfert trial data extraction procedure

User data bits lost or errored in conjunction with resets or clears are not counted as residual errors. User data bits retransmitted by Process A in order to recover from a reset or clear are not counted as residual errors.

3.8.2 *Data transfer trial data reduction*

Figure 8/X.138 shows the data transfer trial data reduction procedure. Using the complete record of protocol events, resets and premature disconnects are counted as illustrated in Figure 9/X.138. The procedure of Figure 10/X.138 is used to estimate the residual error rate for trials in which no resets or premature disconnects occur. Figure 11/X.138 illustrates the steps needed to calculate data packet transfer delay and throughput capacity for trials in which the residual error rate estimate is zero. The preliminary information derived in these three steps is used together with other trial results to determine if the service was available during this trial. If the service was available, cumulative data transfer statistics may be updated.

Cumulative counters can be kept for total reset events, total reset stimuli, total premature disconnect events, and total premature disconnect stimuli. Estimates of reset probability, reset stimulus probability, premature disconnect probability and premature disconnect stimulus probability can then be created by dividing those counters by a cumulative count of the time during which data transfer was tested. Long-term cumulative counters can also be kept for the number of user data bits transmitted, the number of user data bits received, the number of user data bits lost, the number of errored user data bits received, the number of extra user data bits received, the number of user data bits successfully transmitted and received. Residual error rate can then be estimated using the equations in Figure 8/X.138. For successful data transfer attempts, total data packet transfer delay and total data packets transferred can be accumulated. The ratio of these two numbers is an estimate of mean data packet transfer delay. For successful throughput capacity trials, total bits transferred and total time in those trials (as defined in Recommendation X.135) can be accumulated. Dividing these numbers yields an estimate of the throughput capacity.

Figure 9/X.138 uses the record of packet level reference events to evaluate the resets and premature disconnects that occurred during the trial. The equations presented depend on the fact that a reset (or premature disconnect) event between B_i and B_j will cause two reset (or clear) packets to exit the section(s) under test while a reset (or premature disconnect) stimulus will cause one reset (or clear) packet to enter the sections and one to exit.

Figure 10/X.138 shows an acceptable approximation for calculating residual error rate. The approximation is based on the assumption that in a single trial only one type of residual error can occur: that is, lost bits and errored bits do not occur in the same trial, errored bits and extra bits do not occur in the same trial, and lost bits and extra bits do not occur in the same trial. For purposes of estimating residual error rate, this approximation is assumed to be sufficiently accurate. Other, more sophisticated methods of comparing transmitted bits with received bits may yield a more accurate estimate of residual error rate. User information received at B_j more than 200 seconds after it was transmitted at B_i is defined to be lost information (see Recommendation X.136, § 3.3.3).

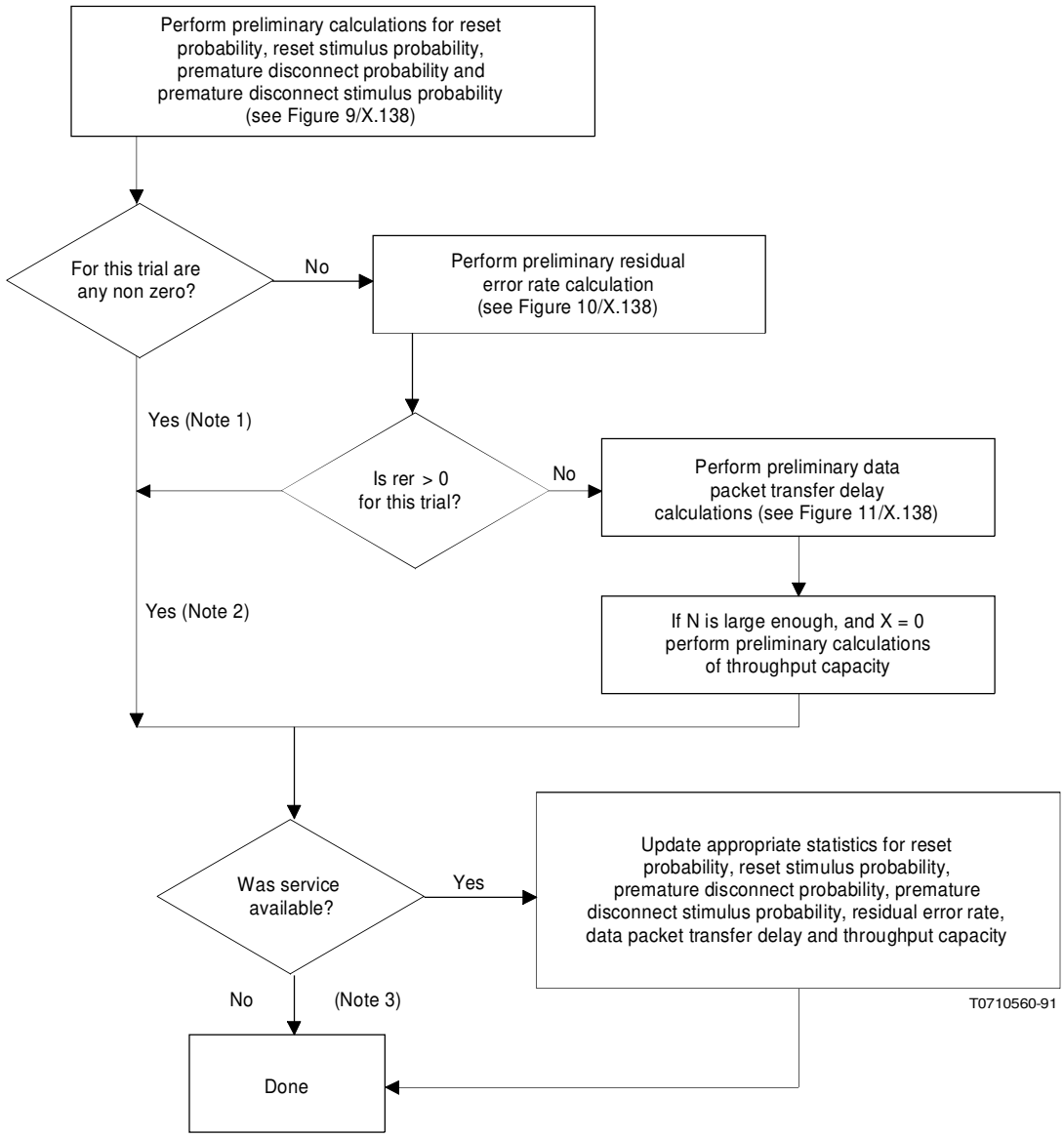
Calculating data packet transfer delay and throughput capacity (see Figure 11/X.138) first requires matching packets recorded at boundary B_i with packets recorded at boundary B_j. The exact timer values used in calculating data packet transfer delay and throughput capacity depend on the location of the boundaries B_i and B_j, and the direction the packets are moving across these boundaries.

3.9 *Disengagement trial*

Clear indication delay and call clear failure probability values can be obtained using the following procedures.

3.9.1 *Disengagement trial data extraction*

Figure 12/X.138 shows the disengagement trial data extraction procedure. Boundaries B_i and B_j are at X.25 or X.75 interfaces bounding the set of virtual connection sections under test.



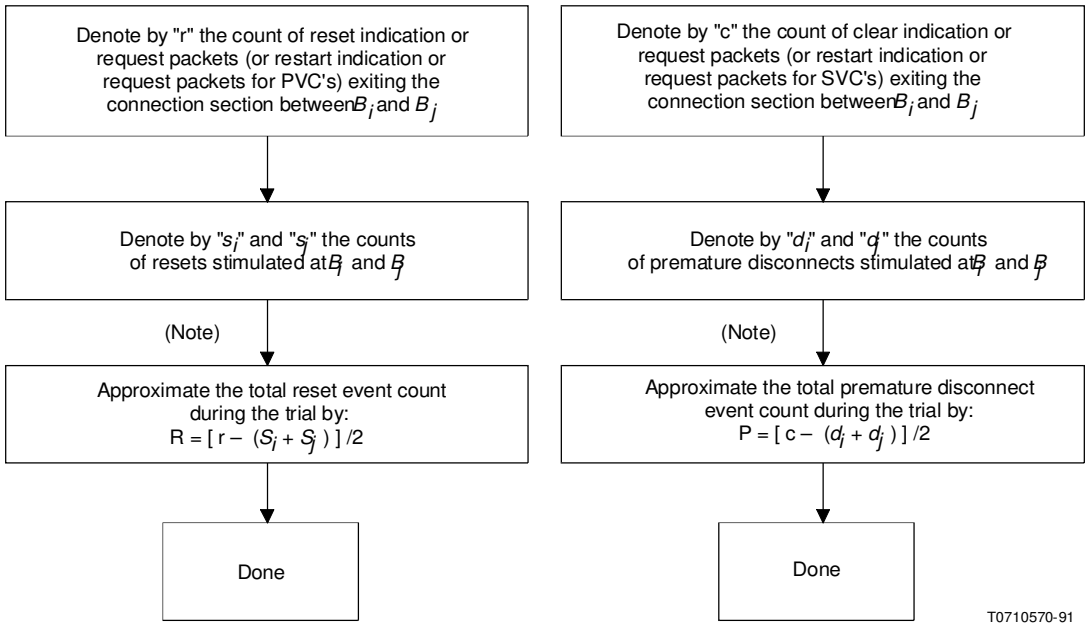
Note 1 – More sophisticated procedures may recover from resets and premature disconnects, and permit residual error rate, data packet transfer delay and throughput capacity calculations.

Note 2 – More sophisticated procedures may recover from residual errors and permit data packet transfer delay throughput capacity calculations.

Note 3 – This decision may be used in estimating service availability if the appropriate statistical considerations have been met.

FIGURE 8/X.138

Data transfert trial data reduction procedure



PVC Permanent virtual circuit
SVC Switched virtual call

Note – These assume that the sections outside the test section do not initiate resets, restarts or clears on their own, but do respond correctly to reset and premature disconnect stimuli.

FIGURE 9/X.138

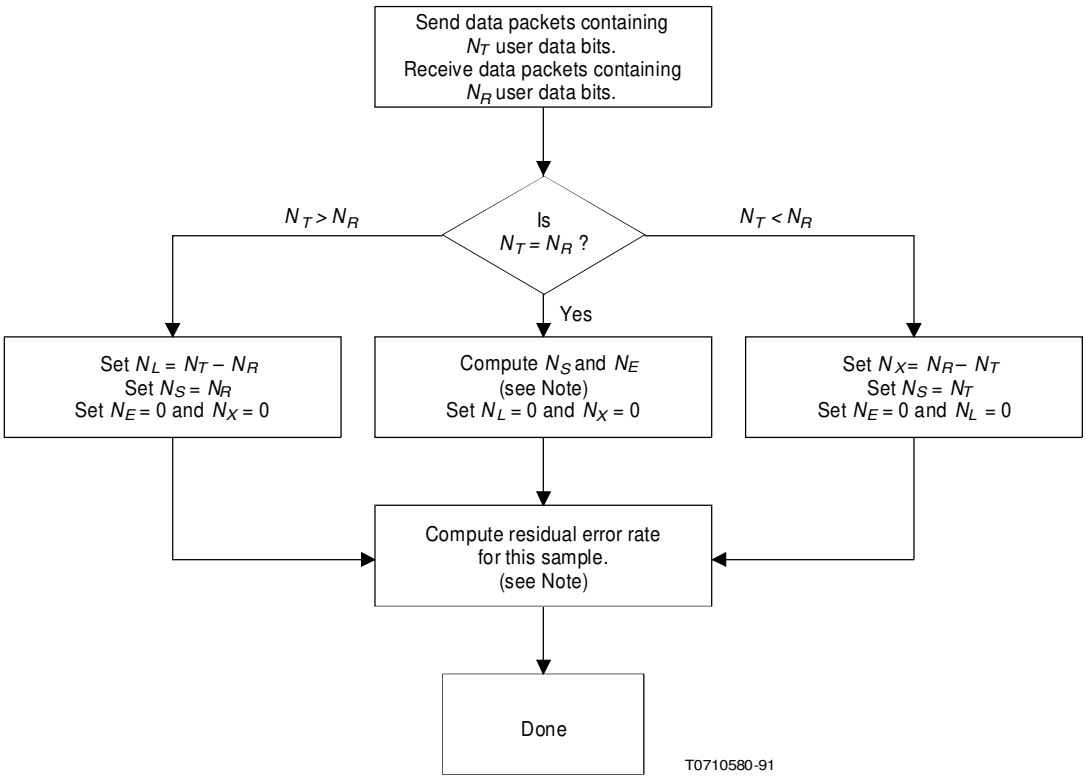
Reset probability, reset stimulus probability, premature disconnect probability and premature disconnect stimulus probability calculations (approximate method)

Logical channels at B_i and B_j must initially be in state p4/d1. Process A will transmit a clear request (or clear indication) packet. Process B should be ready to receive and record the reception of the corresponding clear request (or clear indication) packet. Clocks in Process A and Process B must be sufficiently well synchronized so that their difference is an insignificant fraction of the likely clear indication delay value. Synchronization accuracy of one or two milliseconds is usually sufficient.

The total time allowed for Process B to receive the call-clearing packet from Process A should be at least than the 180 second call clear failure threshold.

3.9.2 *Disengagement trial data reduction*

Figure 13/X.138 shows the disengagement trial data reduction procedure. Each disengagement trial is classified as shown in Figure 14/X.138. Figure 15/X.138 shows the steps needed to calculate clear indication delay for successful call clearings. The preliminary information derived in these two steps is used together with other trial results to determine if the service was available. If the service was available during this trial, cumulative call-clearing statistics may be updated. Cumulative counters can be kept for total call-clearing failures and total clear indication delay (during successful attempts). Estimates of call clear failure probability and clear indication delay can then be created by dividing these cumulative counters by a cumulative count of call clearing attempts.



$$\textit{Note} - \text{Residual error rate} = \frac{(N_L + N_X + N_E)}{(N_L + N_X + N_E + N_S)}$$

$$N_E = \sum_{i=1}^{N_T} b_i^T \oplus b_i^R \qquad \text{(where } \oplus \text{ is addition modulo 2)}$$

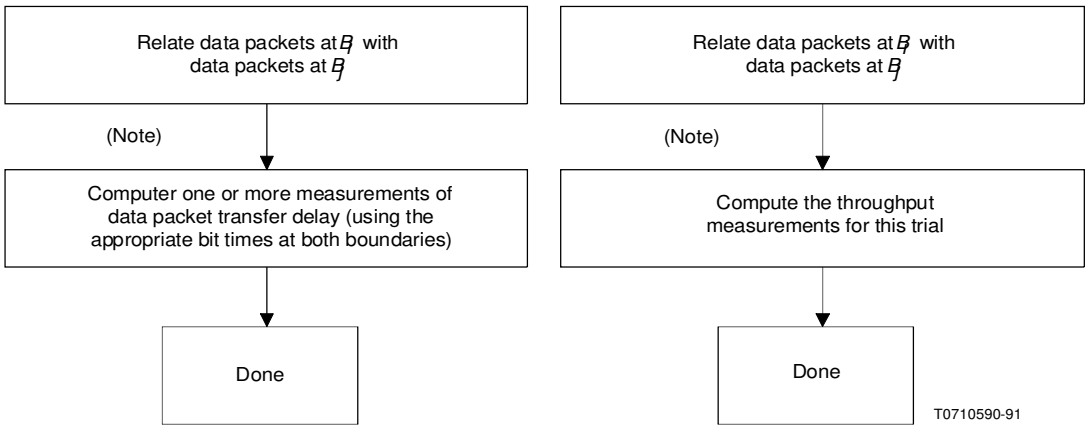
$$N_S = N_T - N_E$$

FIGURE 10/X.138

Residual error rate calculation (approximate method)

Figure 14/X.138 uses the record of packet level reference events (identified as events "A" and "B" in this figure) to determine if the call clearing was successful or unsuccessful. Event "B" is said to have occurred only if the clear indication (or clear request) packet was received at B_j within the 180 second call clear failure threshold. Otherwise it is assumed not to have occurred and the clear attempt is classified as a failure.

Calculating clear indication delay (Figure 15/X.138) first requires matching packets recorded at boundary B_i with packets recorded at boundary B_j. The exact timer values used in calculating clear indication delay depend on the location of the boundaries B_i and B_j and the direction the packets are moving across those boundaries.



Note – Any reasonable method may be utilized.

FIGURE 11/X.138

Data packet transfer delay and throughput capacity calculation

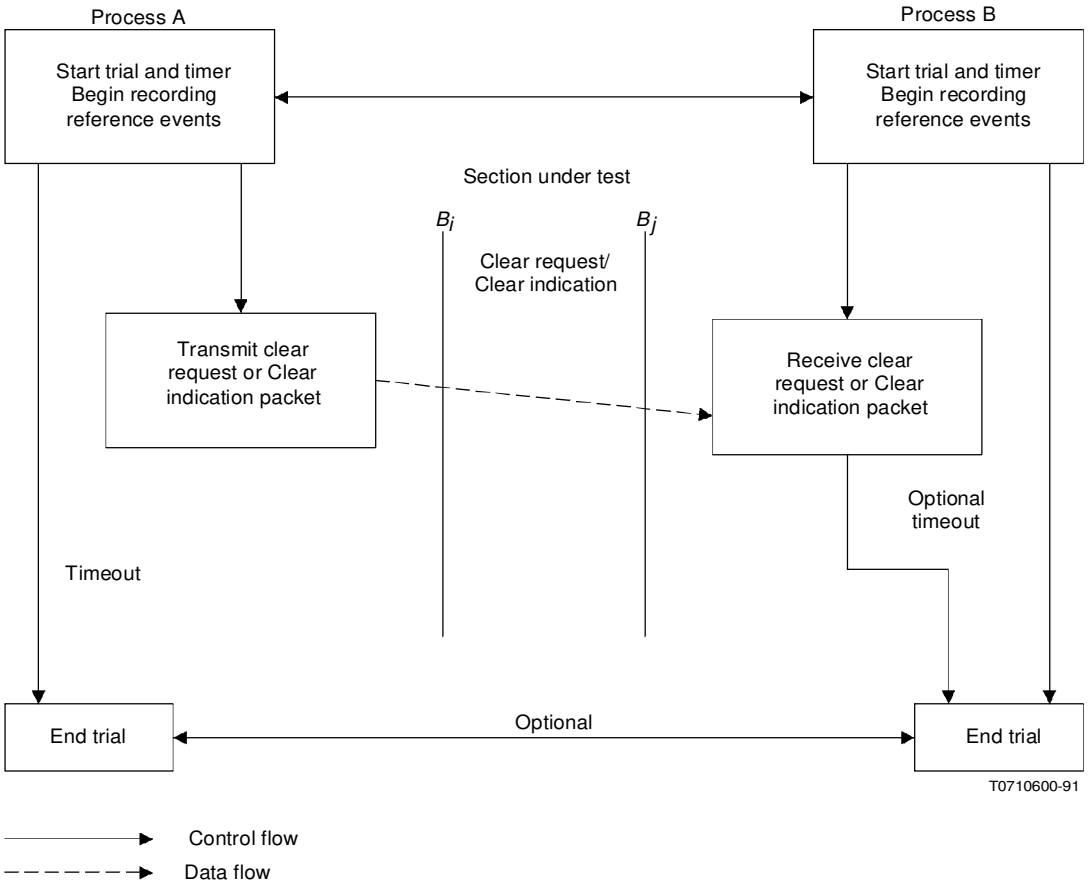
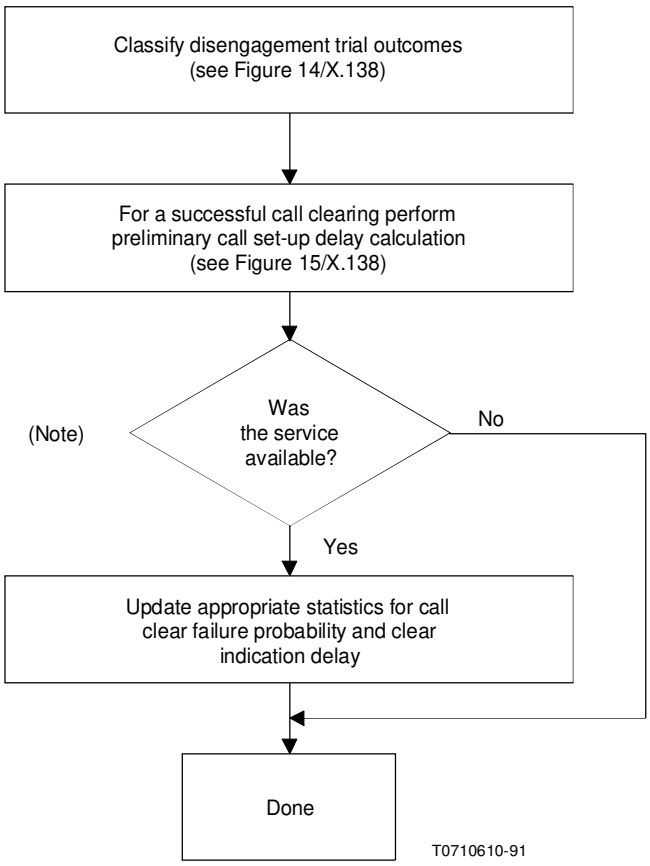


FIGURE 12/X.138

Disengagement trial data extraction procedure



Note – If the appropriate statistical considerations have been addressed, this decision may be used in estimating availability.

FIGURE 13/X.138

Disengagement trial data reduction procedure

3.10 *Estimating availability parameters*

The availability parameters, service availability and mean time between service outages can be measured for a given virtual connection section by an appropriate sequence of availability trials.

The values for N_1, N_2, N_3, N_4 and N_5 are as specified in Table 2/X.137. As these values may change, they are parameterized in this section. The current values are:

a) $N_1 = 4$;

b) $N_2 = 80$;

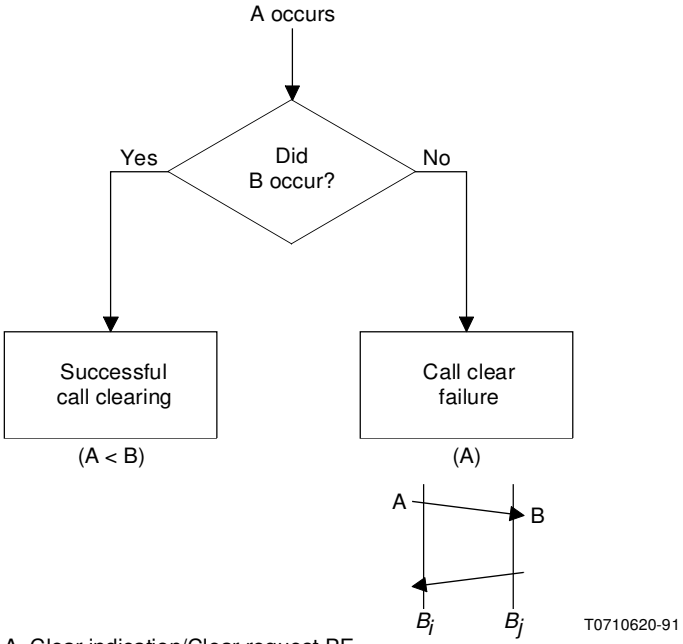
c) $N_3 = 5$;

d) $N_4 = 10^{-3}$;

e) $N_5 = 1$;

3.10.1 *Minimal availability trial data extraction*

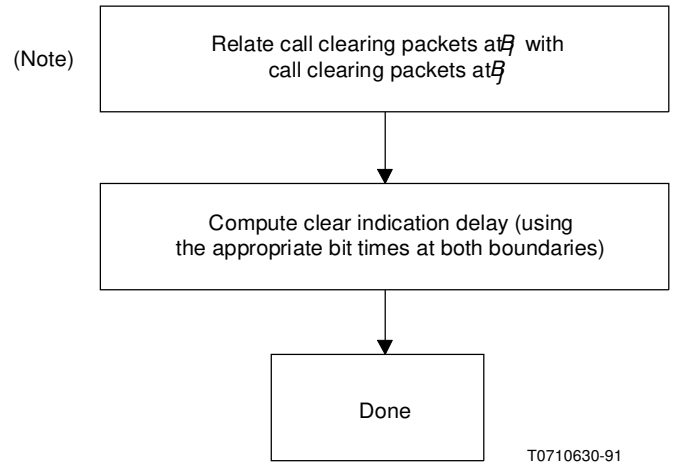
The following test and its decision criteria are defined to be the minimum criteria necessary to sample the availability state of a section.



A Clear indication/Clear request PE
B Clear indication/Clear request PE
The notation (A < B) means that event A occurred before event B.

FIGURE 14/X.138

Classifying a disengagement trial outcome



Note – Any reasonable method may be utilized.

FIGURE 15/X.138

Clear indication delay calculation

Figure 16/X.138 shows a procedure for conducting a minimal availability trial across a section. The trial is divided into two phases: access (Phase I) and user information transfer (Phase II). The trial requires controlled sources and sinks and monitors at each section boundary.

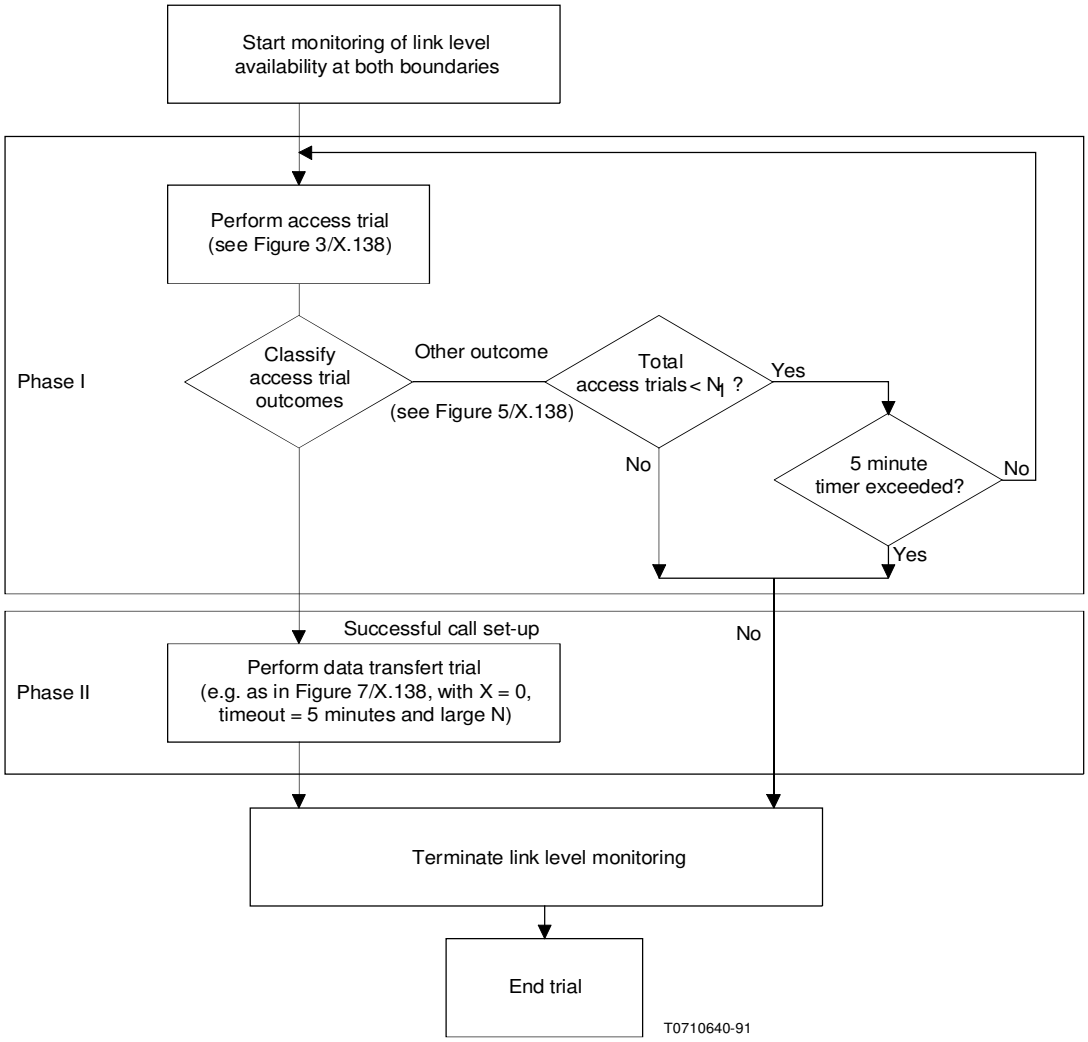


FIGURE 16/X.138

Minimal availability trial data extraction procedure

In Phase I, no more than N_1 successive call set-up attempts (access trials) are performed. Phase I is terminated when any of the following occur: a call set-up attempt succeeds; N_1 successive call set-up attempts are unsuccessful; or the duration of Phase I exceeds 5 minutes. Each call set-up attempt may be performed according to the procedure described in § 3.7.1 and shown in Figure 3/X.138. Phase I is performed only in the case of a switched virtual call.

Phase II of a minimal availability trial attempts to maintain a virtual connection across the tested section for 5 minutes and maintain an average throughput that exceeds N_2 bit/s during that interval. Packet transmission may be performed according to the procedure described in § 3.8.1 and shown in Figure 7/X.138. The values of N and X should be such that the procedure attempts to achieve throughput considerably greater than N_2 bit/s. In the case of a switched virtual call, Phase II is performed only if a call set-up attempt in Phase I is successful. In the case of a permanent virtual circuit, only Phase II is performed.

3.10.2 *Minimal availability trial data reduction*

Figure 17/X.138 shows a procedure for reducing performance data recorded in a minimal availability trial. The procedure implements the availability decision criteria defined in Recommendation X.137.

Criteria defined in X.137 are used to determine availability of the data link layers. Three cases are distinguished: the data link layer at both section boundaries is available; the data link layer at either section boundary is unavailable due to causes inside the section; and the data link layer at one or both boundaries is unavailable due to causes outside the section and the data link layer at neither boundary is unavailable due to causes inside the section. In the first case, availability of the section is determined by results of Phases I and II of the trial. In the second case, the section is declared unavailable. In the third case, the trial is excluded.

Processing of Phase I results is shown in the upper part of Figure 17/X.138. The outcome of each attempt is determined according to the procedure described in § 3.7.2. If all attempts result in either call set-up error or call set-up failure, the virtual circuit section is considered to be unavailable for the duration of the trial. If any call set-up attempt is unsuccessful due to causes outside the portion boundaries (for example, because of test equipment malfunctions), the trial is excluded and is not used to determine availability parameters.

If any access trial is successful and no access trial in unsuccessful due to causes outside the section, availability of the section is determined by the results in Phase II.

Processing of Phase II results is illustrated in the lower part of Figure 17/X.138. The reduction procedure implements the following decision criteria based on Table 2/X.137:

- If the observed number of reset events plus the number of reset stimuli during Phase II is greater than N_3, the section is unavailable;

- If the observed number of call clearing events due to premature disconnects or premature disconnect stimuli is greater than N_5 (in the case of a switched virtual call), the section is unavailable;

- If the measured residual error rate during Phase II is greater than N_4, the section is unavailable;

- If the observed throughput during Phase II is less than N_2 bit/s, the section is unavailable.

If the section is not unavailable according to any of the preceding four criteria, the section is considered to be available during the trial.

3.10.3 *Estimating service availability*

A sufficient estimate of service availability for a virtual circuit section can be computed by performing a sequence of minimal availability trials as described in this section.

A sequence of not less than 300 availability trials is conducted across the section over a long measurement period (for example, six months). Because of the expected durations of service outages, successive trials should be separated by at least seven hours (this serves to keep availability trials uncorrelated). The trials should be uniformly distributed across the scheduled service time. A trial whose outcome is excluded may be replaced by a trial conducted immediately following the excluded trial. The estimate of service availability is 100 times the number of trials in which the section is declared available divided by the number of trials whose outcomes are not excluded.

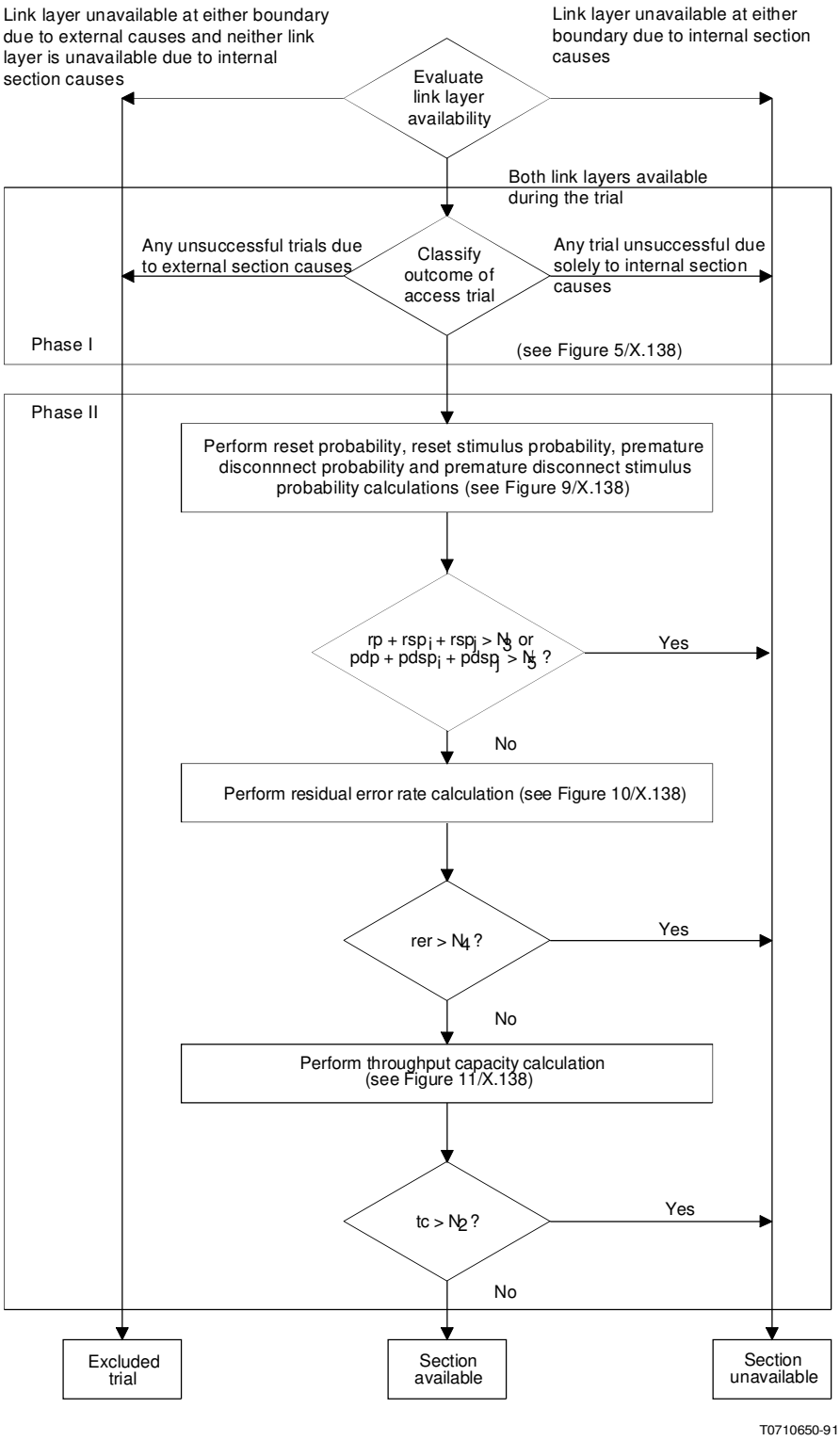


FIGURE 17/X.138

Minimal availability trial data reduction procedure

3.10.4 *Estimating mean time between service outages*

A sufficient estimate of the mean time between service outages for a virtual connection section can be computed by performing a sequence of availability trials. (The assumption of a memory-less (exponential) time to service outage underlies these methods).

Select *k* disjoint time intervals each not less than 30 minutes nor more than three hours. The total amount of time in the *k* intervals should exceed three times the a priori estimate of mean time between service outages. Consecutive availability trials are conducted across the section for the duration of each interval. An additional "post interval" availability trial is conducted immediately following the last trial in an interval. An estimate of the mean time between service outages is obtained from the measured time (A) in the available state and the observed number (F) of transitions from the available state to the unavailable state.

Figure 18/X.138 shows a procedure for reducing the collected performance data and estimating the mean time between service outages. The procedure implements the following specifications:

– If the outcome of any trial in an interval is excluded, discard all trials in the interval.

– If the section is unavailable during the first trial in an interval, assume that the transition to the unavailable state occurred before the interval began and discard all trials in the interval.

– If the section is available during the first trial in an interval and is unavailable during the subsequent trial in any interval, increment the observed number (F) of transitions to the unavailable state by one. Augment the cumulative duration (A) of available states by the duration of all trials in the interval that precede the first unavailability outcome. Discard all trials in the interval that follow the first unavailability outcome.

– If the section is available during all trials in an interval, augment the cumulative duration (A) of available states by the duration of these trials. If the section is unavailable during the post-interval trial, increment the number (F) of transitions to the unavailable state.

After the results of all *k* intervals have been processed, the mean time between service outages is estimated by the ratio A/F if $F > 0$ and by A if $F = 0$.

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

Discarding the remainder of the interval following an unavailability outcome 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, will not bias the result. (If outages tend to be clustered, discarding trials following a transition to the unavailable state will tend to overestimate the mean time between service outages. If outages tend to be negatively clustered, discarding trials following a transition to the unavailable state will tend to underestimate the mean time between service outages). 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 interval should be no longer than three hours. A minimum recommended interval length is 30 minutes, using five-minute availability trials.

An outage that begins during the first availability trial of an interval may or may not result in an unavailable outcome. According to the estimation procedure, if an unavailable outcome occurs, the interval is discarded, the state transition is missed, and the mean time between service outages is overestimated. The post-interval trial identifies any state transition that occurs during the last trial of the interval. It also counts certain transitions that occurred outside the interval. These transitions are counted with the same probability as the probability that transitions during the first trial in an interval are missed. Thus, the two sources of bias tend to cancel out.

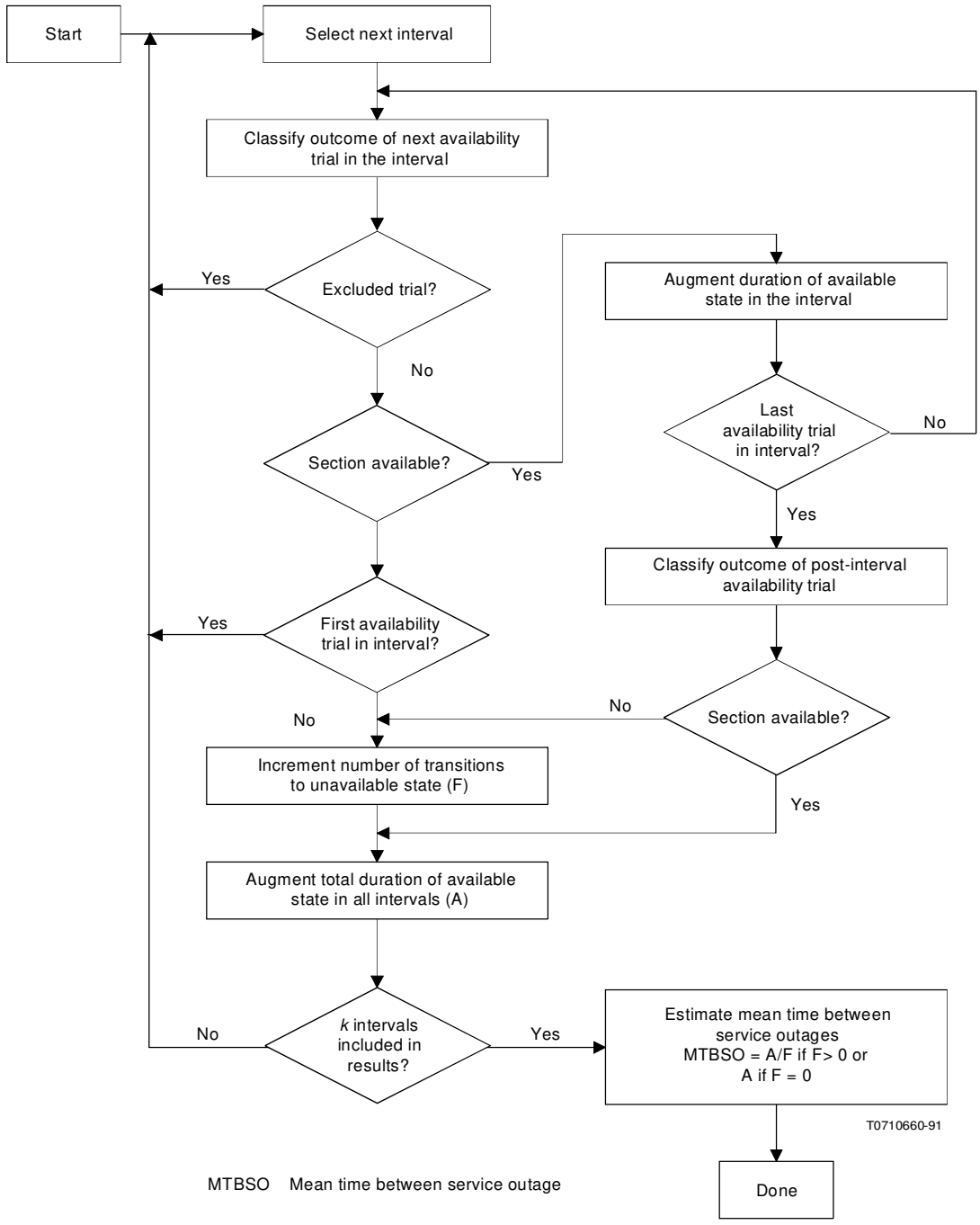


FIGURE 18/X.138

Data reduction procedure for estimating mean time between service outage

4 Method for synchronizing equipment

4.1 *Equipment synchronization*

An optional means by which time synchronization might be obtained between separate items of equipment is as follows:

4.1.1 *General requirements*

A public switched telephone network (PSTN) connection is used for communication between DTEs when synchronization is to be obtained, as shown in Figure 19/X.138.

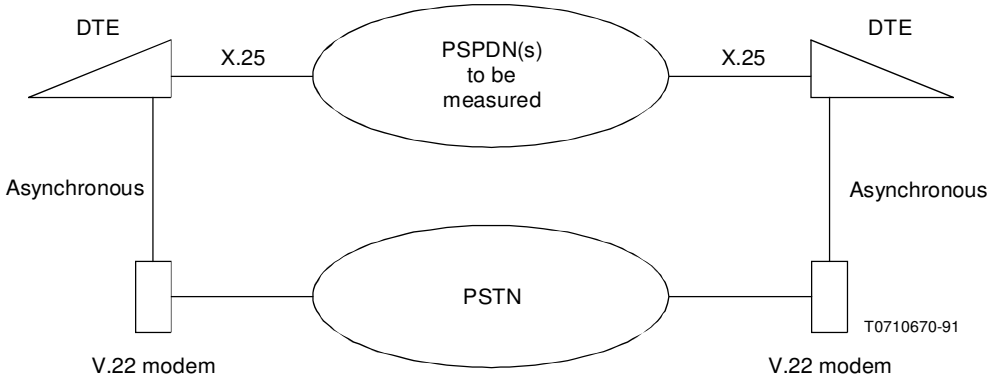


FIGURE 19/X.138

Synchronization configuration

It would be useful if the synchronization path could also be used to convey the results of the measurements made. This is for further study.

DTEs should be synchronized immediately prior to each measurement, and the measurement period should not extend for a period which may result in drift of a magnitude which would give too great an inaccuracy.

4.1.2 *PSTN communication line*

The DTE should be equipped with an asynchronous port in addition to the port used to connect to the packet switched network(s) to be measured. That port is connected to the PSTN using a V.22 modem with automatic answering using V.25.

The connection through the PSTN must be circuit switched (and is assumed to have constant round trip delay). This connection is only required during the time synchronization is being established.

4.1.3 *DTE local clock*

The DTE must have a local clock of sufficient stability to retain the required accuracy of synchronization over the measurement period. An accuracy of at least 1 part in 10^5 (including allowance for initial frequency setting, drift and synchronization imprecision) should be used.

Note – Further consideration of the viability of this accuracy is necessary.

Clocks of this accuracy may result in a drift of $2 \times 10^5 \times$ Time, that is, in a time of ten minutes the drift may add up to about 12 msec.

4.1.4 *Remote command capability*

The DTE should be able to accept an IA5 command from the remote DTE through a separate asynchronous port from the packet measurement port. It should also be able to transmit IA5 commands to the remote DTE through this asynchronous port.

4.1.5 *Clock reset*

Whenever the DTE receives the command TIMER_SET, the internal clock of the DTE is reset to time 0. The response of this action must be within 0 to 5 msec after receiving all the command.

4.1.6 *Timestamp*

The granularity of the timestamp should be less than or equal to 1 msec.

4.1.7 *Synchronization procedure*

DTE A sends a TIMER_SET message terminated by a carriage return to DTE B through the asynchronous path. Let t_0 be the time when the carriage return has been issued.

When DTE B has received the carriage return it sets its clock to zero and issues a TIMER_ACK response terminated by a carriage return to DTE A. These two actions should be performed within a 5 msec delay following the TIMER_SET message reception.

When DTE A has received the carriage return at time t_1, it sets its clock to $\frac{(t_1 + t_m - t_0)}{2}$ where t_m is the time for the asynchronous transmission of the TIMER_ACK message, that is, 83.3 msec at 1200 bit/sec.

This is illustrated in Figure 20/X.138.

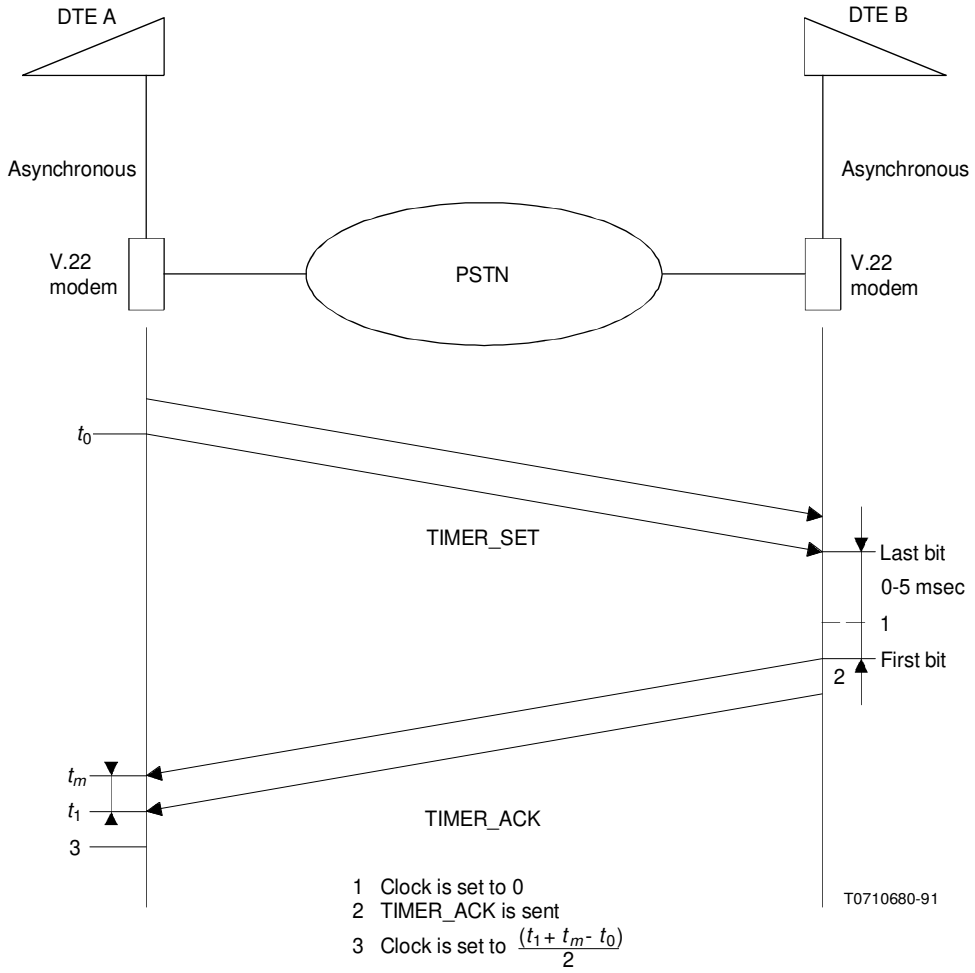


FIGURE 20/X.138

Synchronization procedure

ANNEX A

(to Recommendation X.138)

Calculation of packet switched network performance statistics

This annex provides detailed information on the calculation of packet performance statistics and the measurement conditions (factors) which can affect their observation. Detailed formulas are given in § A.1. Subsection A.2 gives some basic statistical definitions. Subsection A.3 gives a correction factor to the throughput measurement formula.

A.1 *Statistics*

Table A-1/X.138a), b) and c), provides a reference for the calculation of relevant statistics for the performance parameters of Recommendations X.135, X.136 and X.137. For each parameter an equation is given for calculating a single observation x. Most of these equations use the variable names defined in those Recommendations.

TABLE A-1/X.138

Calculation of statistics for performance parameters

a) Speed of service – Delay

Parameter	One observation	Sample Mean	Estimate of variance
Call set-up delay	$d_1 - d_2 = x$, (Note 1)	$\frac{1}{N}\sum_{i=1}^{N} x_i = \bar{x}$	$\frac{1}{N-1}\sum_{i=1}^{N} (x_i - \bar{x})^2$
Data packet transfer delay	$t_1 - t_2 = x$, (Note 2)	$\frac{1}{N}\sum_{i=1}^{N} x_i = \bar{x}$	$\frac{1}{N-1}\sum_{i=1}^{N} (x_i - \bar{x})^2$
Clear indication delay	$t_1 - t_2 = x$, (Note 3)	$\frac{1}{N}\sum_{i=1}^{N} x_i = \bar{x}$	$\frac{1}{N-1}\sum_{i=1}^{N} (x_i - \bar{x})^2$

b) Speed of service – throughput capacity

Parameter	One observation	Weights	Sample mean *tc* estimate (Note 5)
Throughput capacity	$T(\gamma) = \frac{B}{t} = x$; (Note 4)	$\frac{t_i}{\sum_{i=1}^{M} t_i} = w_i$	$\sum_{i=1}^{M} w_i x_i = \bar{x}$

Note 1 – d_1, d_2 are defined in Recommendation X.135, § 2.2.

Note 2 – t_1, t_2 are defined in Recommendation X.135, § 3.1.

Note 3 – t_1, t_2 are defined in Recommendation X.135, § 5.1.

Note 4 – $T(\gamma)$ is the steady state throughput measured with optimal factor levels, γ.

Note 5 – The sample variance of the throughput capacity is

$$\frac{1}{N-1}\sum_{i=1}^{N} (x_i - \bar{x})^2$$

although this is only valid if the time durations, t_i (i = 1 to N) are all equal. The interpretation of the throughput capacity sample variance is clear in this case, but not for the case of a weighted sample. This calculation would apply when the alternative measurement technique (see Recommendation X.135, § 4.2) is used.

c) Accuracy, dependability and availability

Parameter	One observation	Weights	Sample mean
Call set-up error probability	0,1 = x, (Note 1)	–	$\frac{1}{N}\sum_{i=1}^{N} x_i = \bar{x}$
Call set-up failure probability	0,1 = x, (Note 2)	–	$\frac{1}{N}\sum_{i=1}^{N} x_i = \bar{x}$
Residual error rate	$\frac{N_E + N_L + N_X}{N_T} = X$, (Note 3)	$\frac{N_{Ti}}{\sum_{i=1}^{M} N_{Ti}} = w_i$	$\sum_{i=1}^{M} w_i x_i = \bar{x}$
Reset stimulus probability (for a single boundary)	$\frac{S}{N_{vc-s}} = x$, (Note 4)	$\frac{N_{vc-si}}{\sum_{i=1}^{M} N_{vc-si}} = w_i$	$\sum_{i=1}^{M} w_i x_i = \bar{x}$
Reset probability	$\frac{R}{N_{vc-s}} = x$, (Note 5)	$\frac{N_{vc-si}}{\sum_{i=1}^{M} N_{vc-si}} = w_i$	$\sum_{i=1}^{M} w_i x_i = \bar{x}$
Premature disconnect stimulus probability (for a single boundary)	$\frac{d}{N_{vc-s}} = x$, (Note 6)	$\frac{N_{vc-si}}{\sum_{i=1}^{M} N_{vc-si}} = w_i$	$\sum_{i=1}^{M} w_i x_i = \bar{x}$
Premature disconnect probability	$\frac{p}{N_{vc-s}} = x$, (Note 7)	$\frac{N_{vc-si}}{\sum_{i=1}^{M} N_{vc-si}} = w_i$	$\sum_{i=1}^{M} w_i x_i = \bar{x}$
Call clear failure probability	0,1 = x, (Note 8)	–	$\frac{1}{N}\sum_{i=1}^{N} x_i = \bar{x}$
Service availability	0,100 = x, (Note 9)	–	$\frac{1}{N}\sum_{i=1}^{N} x_i = \bar{x}$
Mean time between service outage	$\frac{\mathrm{A}}{\mathrm{F}} = x$, (Note 10)	$\frac{F_i}{\sum_{i=1}^{M} F_i} = w_i$	$\sum_{i=1}^{M} w_i x_i = \bar{x}$

Note 1 – x = 1 if a call set-up error occurs in the call set-up (Recommendation X.136, § 2.1.1).

Note 2 – x = 1 if a call set-up failure occurs in the call set-up (Recommendation X.136, § 2.2.1).

Note 3 – outcome totals as specified in Figure 3/X.136.

Note 4 – s is the number of reset stimuli observed at the boundary (Recommendation X.136, § 3.2.1).

N_{vc-s} is the number of virtual circuit seconds observed.

Note 5 – R is the number of reset events observed (Recommendation X.136, § 3.2.2).

N_{vc-s} is the number of virtual circuit seconds observed.

Note 6 – d is the number of premature disconnect stimuli observed at the boundary (Recommendation X.136, § 3.3.1).

N_{vc-s} is the number of virtual circuit seconds observed.

Note 7 – *s* is the number of premature disconnect events observed (Recommendation X.136, § 3.3.2).

N_{vc-s} is the number of virtual circuit seconds observed.

Note 8 – x = 1 if a call failure occurs in the call clearing (Recommendation X.136, § 4.1).

Note 9 – x = 100 if the observation determines the service is available (Recommendation X.137, § 3.1).

Note 10 – A is the cumulative duration of available states.

F is the number of transitions to unavailable state observed or 1 (Recommendation X.137, Annex A.3).

For each parameter an equation is given for converting multiple observations, x_i, into a sample mean x. In those cases where single observations do not depend on either the length of the observations or the number of bits transmitted, the sample mean is the arithmetic mean. In the other cases (throughput capacity, residual error rate, reset stimulus probability, premature disconnect stimulus probability, premature disconnect probability, mean time between service outage) the sample means are calculated with each observation weighted appropriately either by the length of the observation or the number of bits transmitted.

For three of the parameters (call set-up delay, data packet delay and clear indication delay) a formula is given for estimating the variance in the distribution of the parameter. Four other parameters (call set-up error probability, call set-up failure probability, call clear failure probability, service availability) are assumed to be binomially distributed and as such the variance and the sample variance contain no additional information about service performance. For the remaining six parameters whose individual observations depend on the length of observation or on the number of bits transmitted during the observation (throughput capacity, residual error rate, reset stimulus probability, premature disconnect stimulus probability, premature disconnect probability, mean time between service outage), no formula is given for computing a sample variance. In order to evaluate the variability in these parameters a single fixed observation size must be chosen for each parameter.

The first step in developing a set of throughput capacity observations is to choose an optimal (or nearly optimal) set of user controllable factor levels γ. Packet layer window size, data packet length, throughput class, D-bit usage, and inter-packet gaps must all be chosen to maximize the possible throughput. Subsection A.2 gives guidance on how these factors may be adjusted to improve throughput. Each observation of throughput capacity is based on the optimal set of factor level γ. The mean of these observations, weighted by the duration of each observation, yields a single estimate of throughput capacity.

To properly interpret measured performance values the relevant measurement conditions must be known. Table A-2/X.138 identifies general factors that may influence the values for each of the performance parameters defined. Measurements provided in accordance with this Recommendation should contain or reference a specification of the relevant factor levels existing during the measurement. The effects of the specified factors on throughput are described in the following text. Guidelines for multiple test reporting and a throughput measurement correction factor are also given.

TABLE A-2/X.138

Factors that may influence the performance parameter values

Factors	csd	dpd	tc	cid	cep	cfp	rer	rsp	rp	pdsp	pdp	ccfp	sa	mtbso
Signalling rate	x	x	x	x										
Data link layer packet size	x	x	x	x										
Packet layer window size		x	x											
Packet length		x	x				x							
Other virtual connections	x	x	x	x	x		x		x		x		x	x
Time of day	x	x	x	x	x		x		x		x		x	x
Throughput class		x	x											
D-bit usage			x											
Inter-packet gaps		x	x											

csd Call set-up delay

dpd Data packet transfer delay

tc Throughput capacity

rsp Reset stimulus probability

rp Reset probability

pdsp Premature disconnect stimulus probability

cid Clear indication delay

cep Call set-up error probability

cfp Call set-up failure probability

rer Residual error rate

pdp Premature disconnect probability

ccfp Call clear failure probability

sa Service availability

mtbso Mean time between service outage

A.2 *Statistical formulas*

Various statistical formulas are given below to eliminate the confusion that often results from comparison of formulas for sample (estimated) parameters and distributional (population) parameters.

A.2.1 *Sample mean*

As the arithmetical average of a set of observations, the observed mean is used as an estimate of the true mean of the underlying distribution. Under rather general conditions the mean

$$\bar{x} = \frac{\sum_{k=1}^{n} x_k}{n}$$

is the Maximum Likelihood Estimator of the true distributional mean.

A.2.2 *Sample variance*

The estimated variance, $s(X)$ or s, of a sequence of observations, is defined to be

$$s = \frac{\sum_{k=1}^{n} (x_k - \bar{x})^2}{n-1}$$

and is used to obtain estimates of the precision of the estimated mean $\bar{x}$.

The use of the factor $\left(\frac{1}{n} - 1\right)$ assures that the estimator will be unbiased in the sense that $E[s(X)] = V$, where V is the true variance of the underlying distribution.

A.2.3 *xth percentile*

The xth percentile ($0 < x < 100$) of a cumulative continuous distribution F is any number Y satisfying the equation $F(Y) = \frac{x}{100}$. If F is also strictly increasing, Y is unique. For discrete distributions, most percentiles are not unique. In this case, any Y for which $F(Y)$ is minimal, subject to the condition that $F(Y) \geq \frac{x}{100}$, is an xth percentile.

For example, the 95th percentile of a set of measurements would be any number Y for which:

1) at least 95% of the measurements fall below Y;

2) the number of such measurements below Y is a minimal.

A.2.4 *Minimum*

The minimum of a set of measurements is the least value attained by any measurement in the set.

A2.5 *Maximum*

The maximum of a set of measurements is the greatest value attained by any measurement in the set.

A.3 *Throughput measurement correction factor*

In the case where throughput is measured by observation of exit events, a correction factor can be applied to the throughput measurement formula which will account for any difference in the size of A_0 and A_k or A_0 and A_m (as defined in Recommendation X.135, § 4.1).

The correction factor in the case where exit events are observed at B_i is as follows:

Condition	*Correction*
$f(A_0) = f(A_k)$	No correction
$f(A_0) > f(A_k)$	Subtract the insertion time at boundary B_i of the excess user data bits $[f(A_0) - f(A_k)]$ in A_0 from $t_2 - t_1$
$f(A_0) < f(A_k)$	Add the insertion time at boundary B_i of the deficient user data bits $[f(A_0) - f(A_k)]$ in A_0 to $t_2 - t_1$

The analogous table for the case where exit events are observed at B_j is as follows:

Condition	*Correction*
$f(A_0) = f(A_k)$	No correction
$f(A_0) > f(A_k)$	Subtract the insertion time at boundary B_j of the excess user data bits $[f(A_0) - f(A_k)]$ in A_0 from $t_2 - t_1$
$f(A_0) < f(A_k)$	Add the insertion time at boundary B_i of the deficient user data bits $[f(A_0) - f(A_k)]$ in A_0 to $t_2 - t_1$

The error introduced by not applying the correction factor will normally be very small.

ANNEX B

(to Recommendation X.138)

Factors which may affect measured performance

B.1 *Signalling rates*

The signalling rate on a circuit section (usually measured in bits per second) provides an upper bound on the throughput capacity for that section. In general, faster signalling rates yield higher throughput. Call set-up delay, data packet transfer delay and clear indication delay are also affected by the signalling rate. Again, a higher signalling rate on a circuit section generally results in lower delay values.

B.2 *Data Link Layer window size*

As the Data Link Layer underlies all of the logical channels above it, it affects call set-up delay and clear indication delay as well as data packet transfer delay and throughput. Usually, larger Data Link Layer window sizes decreases delay and increases throughput.

B.3 *Packet Layer window size*

Larger Network Layer window sizes may increase throughput and decrease delay.

B.4 *Packet length*

The length of the data packets can affect data packet transfer delay, throughput, and possibly the residual error rate. Data packets with longer lengths have greater delay associated with them, but greater throughput due to increased efficiency. Longer data packets have a theoretically larger probability of having errors in the user data field that could escape detection by the sixteen bit cyclic redundancy check of the Data Link Layer and thus a corresponding increase in the probability of residual errors.

B.5 *Other virtual connections*

The existence of active virtual connections other than the one under test on the same data link may increase the load on that link. Hence a large number of performance parameters (call set-up delay, data packet transfer delay, throughput, call set-up failure probability, reset probability, premature disconnect probability, clear indication delay, service availability, and mean time between service outages) may be affected due to the underlying Data Link Layer contention.

B.6 *Time of day*

Because time of day generally influences network loading, this factor affects performance in a manner similar to the existence of other virtual connections.

B.7 *Throughput class*

This factor may affect the data packet transfer delay and throughput on some networks. It should be set to the maximum allowable when measuring throughput capacity.

B.8 *D-bit usage*

As the D-bit being set to 1 in a data packet requires the receiving DTE to generate a Receiver ready (RR) packet specifically acknowledging the receipt of that data packet, its use can increase the load on the virtual connect and thus decrease throughput.

B.9 *Inter-packet gaps*

The rules governing the time intervals between successive data packets in data packet transfer delay tests (other than those required by the network for flow control purposes) should be specified. Increasing inter-packet gaps tends to decrease delay and throughput.

B.10 *Reporting multiple tests made between different locations*

Full characterization of network performance can require multiple test measurements that are made between different node locations such as cities. In general, the locations where tests are conducted should be reported. If the measured data are significantly affected by the geographic test arrangements then geography should be considered as a factor and incorporated into the appropriate statistical procedures used to estimate the performance parameters.

ANNEX C

(to Recommendation X.138)

Examples of measurement uses

The accuracy required for the measurements will depend upon the use being made of the results. Three common objectives of performance measurement are given below.

C.1 *Examples*

C.1.1 Acceptance testing

Acceptance of a service may depend on demonstrating a given level of performance. This level of performance can be demonstrated by obtaining measurements of the specified performance parameters and performing an hypothesis test to determine whether the performance levels are acceptable.

Typically these are one sided hypothesis tests. A test for data packet transfer delay is given below. The null hypothesis H_0 is that the delay is acceptable, as it will be in most cases, while the alternative hypothesis H_a is that the delay is too great and is therefore unacceptable.

H_0 : Data packet transfer delay mean is $\leq x$ ms.

H_a : Data packet transfer delay mean is $> x$ ms.

H_0 and H_a will be reversed in cases where a higher parameter value is better (for example, service availability).

C.1.2 *Maintenance*

Given that service has been accepted at a particular level of performance, service providers may want to establish maintenance limits. These are thresholds of performance at which the provider takes action to restore performance that has degraded to less than acceptable levels. In the following test of hypotheses the x refers to the value given in the acceptance test above, and rejecting the null hypothesis (no maintenance needed) would be grounds for performing maintenance.

H_0 : Data packet transfer delay mean is $\leq x$ ms.

H_a : Data packet transfer delay mean is $> x$ ms.

C.1.3 *Conformance of data to a particular distribution*

In certain instances it is important to determine whether or not a set of measurements conforms reasonably well to a particular distribution. This type of test is important in determining whether an assumption about the distribution of a certain type of measurement is correct. In the test given below, the null hypothesis is that the distribution is uniform on the closed interval from 0 to 10.

H_0 : The distribution of data packet transfer delay is uniform (0,10).

H_a : Data packet transfer delay is not uniform (0,10).

C.2 *Pairwise and multiple comparisons*

Pairwise and multiple comparisons are useful in assessing the effect of a factor or combination of factors on observed performance. A series of pairwise comparisons of means is not equivalent to a simultaneous comparison of all means. Thus the conclusions that mean *A* is not significantly different from mean *B*, and mean *C* is not significantly different from mean *B*, does not necessarily lead to the conclusion that means *A* and *C* are not significantly different from one another, and certainly not at the same level of significance.

ANNEX D

(to Recommendation X.138)

Alphabetical list of abbreviations used in this Recommendation

ccfp Call clear failure probability

cep Call set-up error probability

cfp Call set-up failure probability

cid Clear indication delay

csd Call set-up delay

DCE	Data circuit-terminating equipment
dlb	Data loopback
dpd	Data packet transfer delay
DTE	Data terminal equipment
MTBSO	Mean time between service outage
OSI	Open systems interconnection
pdsp	Premature disconnect stimulus probability
PE	Protocol event
PSPDN	Packet switched public data network
PSTN	Public switched telephone network
PVC	Permanent virtual circuit
rer	Residual error rate
rlb	Rooting loopback
rp	Reset probability
RR	Receiver ready
rsp	Reset stimulus probability
sa	Service availability
SABM	Set asynchronous balanced mode
SABME	Set asynchronous balanced mode extended
STE	Signalling terminal equipment
SVC	Switched virtual call
tc	Throughput capacity