INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

E.744

(10/96)

SERIES E: TELEPHONE NETWORK AND ISDN

Quality of service, network management and traffic engineering – Traffic engineering – ISDN traffic engineering

Traffic and congestion control requirements for SS No. 7 and IN-structured networks

ITU-T Recommendation E.744

(Previously CCITT Recommendation)

ITU-T E-SERIES RECOMMENDATIONS

TELEPHONE NETWORK AND ISDN

OPERATION, NUMBERING, ROUTING AND MOBILE SERVICES	
INTERNATIONAL OPERATION	E.100–E.229
OPERATIONAL PROVISIONS RELATING TO CHARGING AND ACCOUNTING IN THE INTERNATIONAL TELEPHONE SERVICE	E.230–E.299
UTILIZATION OF THE INTERNATIONAL TELEPHONE NETWORK FOR NON-TELEPHONY APPLICATIONS	E.300–E.329
ISDN PROVISIONS CONCERNING USERS	E.330–E.399
QUALITY OF SERVICE, NETWORK MANAGEMENT AND TRAFFIC ENGINEERING	
NETWORK MANAGEMENT	E.400–E.489
International service statistics	E.400–E.409
International network management	E.410–E.419
Checking the quality of the international telephone service	E.420–E.489
TRAFFIC ENGINEERING	E.490–E.799
Measurement and recording of traffic	E.490–E.505
Forecasting of traffic	E.506–E.509
Determination of the number of circuits in manual operation	E.510–E.519
Determination of the number of circuits in automatic and semi-automatic operation	E.520–E.539
Grade of service	E.540–E.599
Definitions	E.600–E.699
ISDN traffic engineering	**E.700–E.749**
Mobile network traffic engineering	E.750–E.799
QUALITY OF TELECOMMUNICATION SERVICES: CONCEPTS, MODELS, OBJECTIVES AND DEPENDABILITY PLANNING	E.800–E.899
Terms and definitions related to the quality of telecommunication services	E.800–E.809
Models for telecommunication services	E.810–E.844
Objectives for quality of service and related concepts of telecommunication services	E.845–E.859
Use of quality of service objectives for planning of telecommunication networks	E.860–E.879
Field data collection and evaluation on the performance of equipment, networks and services	E.880–E.899

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

ITU-T RECOMMENDATION E.744

TRAFFIC AND CONGESTION CONTROL REQUIREMENTS FOR SS NO. 7 AND IN-STRUCTURED NETWORKS

Summary

This Recommendation has four main purposes:

1) Provide guidelines for implementing standardized traffic and congestion control procedures in SS No. 7 and IN-structured networks.

2) Provide general guidelines on requirements that should be placed on network elements so that networks can be made robust to unforeseen problems.

3) Provide general principles and guidelines on how network level and network element level control should be structured to maximize effectiveness and avoid network problems.

4) Provide high-level requirements for traffic and congestion controls for consideration in developing specific procedures in other ITU-T Recommendations.

Requirements and implementation considerations are covered for MTP, SCCP and IN traffic and congestion control procedures, and requirements are provided for adequate overload controls in network elements. In addition, considerations for making networks robust against failure propagation are discussed.

Source

ITU-T Recommendation E.744 was prepared by ITU-T Study Group 2 (1993-1996) and was approved under the WTSC Resolution No. 1 procedure on the 8th of October 1996.

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 (Helsinki, March 1-12, 1993).

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.

CONTENTS

Page

Recommendation E.744

TRAFFIC AND CONGESTION CONTROL REQUIREMENTS FOR SS NO. 7 AND IN-STRUCTURED NETWORKS

(Geneva, 1996)

1 Scope

This Recommendation covers requirements and implementation considerations for traffic and congestion control procedures in SS No. 7 and IN-structured networks. It addresses traffic and congestion control procedures defined in other ITU-T Recommendations, and it describes the requirements and considerations needed to ensure the defined procedures work properly when they are implemented in a network.

2 References

The following 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 are 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 currently valid ITU-T Recommendations is regularly published.

– ITU-T Recommendation Q.703 (1996), *Signalling System No. 7 – Signalling link*.

– ITU-T Recommendation Q.704 (1996), *Signalling System No. 7 – Signalling network functions and messages*.

– ITU-T Recommendation Q.705 (1993), *Signalling System No. 7 – Signalling network structure*.

– ITU-T Recommendation Q.714 (1996) *Signalling System No. 7 – Signalling connection control part procedures*.

– CCITT Recommendation Q.724 (1988), *Specifications of Signalling System No. 7 – Signalling procedures*.

– ITU-T Recommendation Q.764 (1993), *Signalling System No. 7 – ISDN user part signalling procedures*.

– ITU-T Recommendation Q.1214 (1995), *Distributed functional plane for intelligent network CS-1*.

– E.410-Series Recommendations, *International network management.*

– ITU-T Recommendation E.412 (1996), *Network management controls*.

– ITU-T Recommendation E.743 (1995), *Traffic measurements for SS No. 7 dimensioning and planning*.

– CCITT Recommendation E.502 (1992), *Traffic measurement requirements for digital telecommunication exchanges.*

3 Definitions

For the purposes of this Recommendation, the following definitions apply:

3.1 congestion control: The automatic action taken in a network to prevent or reduce the overload of one or more network elements. Congestion control action may be implemented in other network elements beyond those experiencing overload.

3.2 overload control: The blocking, diverting or postponing, by or within a network element, of a processing load greater than its capacity or than can be handled at the relevant Grade of Service (GOS) objective.

3.3 successful message: A message is considered to be successfully delivered if it is delivered, uncorrupted, to the appropriate destination within such a time that it does not cause malfunctioning of the application.

4 Abbreviations

For the purposes of this Recommendation, the following abbreviations are used.

GOS	Grade of Service
GTT	Global Title Translation
IN	Intelligent Network
ISDN	Integrated Services Digital Network
ISUP	ISDN User Part
MTP	Message Transfer Part
SCCP	Signalling Connection Control Part
SIB	Status Indication Busy
SS No. 7	Signalling System No. 7
SSP	Service Switching Point
STP	Signalling Transfer Point
TFA	Transfer-allowed (signal)
TFC	Transfer-controlled (signal)
TFP	Transfer-prohibited (signal)
TFR	Transfer-restricted (signal)
TUP	Telephone User Part
UPT	Universal Personal Telecommunication

5 Introduction

5.1 Purpose

A key objective in operating SS No. 7 and IN-structured networks is the maximization of the robustness of the network, that is, the ability of a network to withstand both traffic overloads and failures of network elements. While many steps can be taken to avoid the occurrence of these network stresses, cases of overload or failure will nonetheless occur. During such stresses, the operation of the network should be maintained, and the level of completed traffic kept as high as possible until the stress can be relieved.

To achieve this goal, the network resources have to be managed in a coordinated way so as to avoid, for example, that situations of focused overload spread out to other parts of the signalling/IN network or that momentary inability to face excessive demand on the signalling network will negatively affect the IN-structured network or vice versa. The actions in the signalling/IN network are distinct from those of the circuit-switched network. Interaction between management action on signalling/IN and the circuit-switched network are for further study.

This Recommendation is required because without properly-implemented traffic and congestion controls, the SS No. 7 and IN-structured networks may not provide sufficient protection to properly and efficiently handle the congestion and overloads that are likely to arise. Some of the characteristics motivating these controls are the following:

- It is difficult to prepare accurate traffic forecasts for the quantities of calls and of signalling messages.
- SS No. 7 networks are subject to short-term overloads that may not be addressed adequately using dimensioning and provisioning procedures only.
- The distributed nature of the SS No. 7 and IN-structured networks, such that overloads in one location may be caused by excessive traffic offered at distant points, making single-node controls inadequate.
- If one service provider stimulates more traffic than can be carried, congestion could seriously affect the service provided to other, unrelated calls being handled by different service providers.
- Service logic processors, having very short holding times, will be much more subject to peak loading, and will provide severely degraded performance even if overloaded by only a small proportion.

This Recommendation has four main purposes:

One of the main purposes of this Recommendation is to provide guidelines and performance characterizations that should be useful in implementing standardized traffic and congestion control procedures in SS No. 7 and IN-structured networks. A number of different traffic and congestion control procedures are provided in ITU-T Recommendations. These procedures generally leave key parameter values to be specified as part of the implementation. In this Recommendation information is provided as to how the different procedures perform as a function of the parameter value settings, and suggestions are made as to how the parameters should be chosen in different types of networks.

A second purpose of this Recommendation is to provide general guidelines on requirements that should be placed on network elements so that networks can be made robust to unforeseen problems. In particular, the needs for overload control are addressed and considerations are made of how such controls must interrelate to network level congestion controls. Principles are provided regarding what network element capabilities should be present in order to avoid fault and congestion propagation problems and other forms of network instability.

The third purpose of this Recommendation is to provide general principles and guidelines regarding how network level and network element level control should be structured to maximize effectiveness and avoid network problems. An important consideration here is in regards to interconnecting different networks. Network level controls and network element level controls must work properly in each network as well as between networks. To keep the different systems and controls harmonized, some basic principles need to be followed. This allows for various vendor products and network implementations to be interconnected with a high-level of confidence the traffic and congestion control procedures will work properly and no fault conditions will result in network instabilities.

The fourth purpose of this Recommendation is to provide high-level requirements for traffic and congestion controls. Specific description and definition of the control mechanisms are covered by the Q and E.410-Series Recommendations.

5.2 Concept of robustness

The robustness of SS No. 7 and IN-structured networks is the ability of those networks to ensure:

- system sanity, i.e. to perform all critical functions necessary for the system and network to perform properly regardless of overload level;
- that traffic overload and failures are contained and will not propagate through the network; and
- maximization of successful traffic throughput under overload and failure conditions.

When a system becomes overloaded, it must reduce traffic in some manner. Overload controls within a system are intended to protect the system without reliance on actions from any other network element. The other way to protect a system is to have congestion controls between network elements, where an overloaded element requests others to reduce their traffic load to it.

In order to achieve network robustness, overload and congestion control techniques need to be considered for all network elements to accommodate both unusually high or unbalanced traffic loads, and failure conditions leading to stressful load conditions. These overload and congestion control techniques encompass both routing and traffic volume controls. Depending on the severity of the problem, the controls should be progressive and selective in their action. In the case of IN-structured networks, in addition to controls in the IN function layer, some control activation may be appropriate in the signalling and bearer portions of the network, such as call admission controls in the circuit-switched network.

To achieve a high-level of network robustness, a layering of controls is needed to address the many potential sources of network stress. Certain congestion and overload requirements may be established in the SS No. 7 transport function using the MTP and SCCP layers of the protocol. Additional network congestion control requirements can be defined for the IN service layer. Finally, overload requirements may be established for the individual elements in the network. The following clauses outline the requirements for these three different areas.

Another aspect of robustness is the ability of a network to prevent propagation of a problem to other parts of the network. The robustness strategy is to provide capabilities in the SS No. 7 and IN-structured networks to ensure that failures and overloads are contained and will not propagate from one element to another.

6 Requirements for MTP and SCCP traffic and congestion control procedures

6.1 MTP signalling link flow control procedures

The signalling link (level 2) flow control procedure is defined in Recommendation Q.703. The flow control procedure is initiated when congestion is detected at the receiving end of the signalling link. The congested receiving end notifies the transmitting end of its congestion with Status Indication Busy (SIB) signal units and withholds acknowledgements of all incoming signal units. This action allows the receiving end to stop acknowledgements and keep the transmitting end from failing the link due to a time-out on positive acknowledgements (i.e. expiration of timer T7). However, if the congestion condition lasts too long (3 to 6 seconds), the transmitting end will fail the link due to expiration of congestion timer T6.

The purpose of activating this flow control procedure is to prevent acknowledging messages at the receiving end and then subsequently discarding them due to an overload condition in the level 3 processing. Thus, the activation mechanism for MTP level 2 flow control should stop acknowledgement of signal units at a condition that indicates level 3 congestion, but the probability of message loss is still very small. A typical detection mechanism would be a receive buffer threshold for acknowledged messages queued for level 3 processing.

The sending of SIBs should be stopped when the queued level 3 processing workload has been reduced to a level close to zero. The objective is to keep the probability of going back into level 3 processing congestion very low, or if level 3 congestion is going to occur, the objective would be to maximize the time it takes that to happen.

The invocation of the level 2 flow control procedure may also cause the invocation of level 3 flow control (TFC procedure) due to a build-up of messages in the transmit buffer at the remote end of the link. If this happens then there will be a subsequent reduction of traffic offered to the congested link.

The control thresholds for activation and deactivation must be chosen so that the probability of the transmit end timing out T6 (time in congestion) and failing the link should be very small.

6.2 MTP signalling route management procedures

The unavailability, restriction and availability of a signalling route is communicated by means of the transfer-prohibited (TFP), transfer-restricted (TFR) (a national option), and transfer-allowed (TFA) procedures, which are described in 13.2/Q.704, 13.4/Q.704 and 13.3/Q.704 respectively.

Congestion of a signalling link results in a transfer-controlled (TFC) procedure being invoked at signalling transfer points (STPs) and traffic flow control actions being taken at the sources of signalling traffic. For international signalling the TFC procedure is given in 13.6/Q.704 and the traffic flow control procedure is given in 11.2.3/Q.704. TFC and traffic flow control procedure options for national networks with and without congestion priority are given in 13.7/Q.704 and 11.2.4/Q.704, 13.8/Q.704 and 11.2.5/Q.704, respectively.

6.2.1 Requirements for TFP, TFR and TFA procedures

When an STP recognizes it is unable to transfer signalling traffic to a particular destination, it sends a corresponding TFP message to its adjacent signalling points by either a broadcast or response method. It is important that this notification gets to adjacent signalling points as soon as possible, since all messages sent to an STP that cannot transfer them are discarded. Therefore, a requirement should be set for the time it takes an STP to recognize it cannot transfer signalling traffic to a destination and begin sending TFPs by either the broadcast or response method. For the broadcast method a requirement should also be placed on the time it takes to complete the broadcast (considerations should also be made to ensure broadcasting does not cause overload). The specific requirements will depend on the particular network configuration. Guidelines for the considerations to make and how to set these requirements are for further study.

When a signalling point receives a TFP message for a particular destination, it does a forced rerouting (clause 7/Q.704) to route traffic to the affected destination on an alternative signalling route. It is important that the time interval from reception of TFP to activating forced rerouting be relatively fast, since messages sent on the old route will be discarded. Therefore, a requirement should be set for the response to the receipt of a TFP message. Guidelines for the considerations to make and how to set these requirements are for further study.

When an STP cannot use the normal link set (or combined link set) to a destination "X" due to a "long-term" failure, or there is congestion on an alternate link set currently being used to destination "X", the STP sends TFR messages to its adjacent signalling points relating to destination "X". When a signalling point receives a TFR for a destination "X", it does a controlled rerouting (clause 8/Q.704) to an alternative equal priority link set that is not restricted to destination "X", if such a link set is available. In sending and responding to TFRs in connection with a "long-term" failure situation, a rapid response time is not crucial, since typically some waiting time is spent to ascertain that the failure is "long-term" before sending the TFR. In these situations messages are not being lost, and the impact is increased delay due to a longer path to the affected destination. On the other hand, when a TFR message is sent due to congestion on an alternate link set, it is important to react quickly to minimize congestion and lost messages. Therefore, a requirement should be established for the time it takes an STP to send TFR messages to its adjacent signalling points when it is reacting to a congestion condition, and a requirement is also needed for the time it takes signalling points to respond after receiving a TFR message. Guidelines for the considerations to make and how to set these requirements are for further study.

TFA messages for a destination "X" are broadcast by an STP to its adjacent signalling points when it recognizes it is again able to transfer signalling traffic on a normal link set (combined link set) to destination "X". This can occur when a recovery from either a TFP or TFR condition occurs. The critical situation is when recovering from a TFP condition, since in this case traffic may be being blocked because no other routes are available. When recovering from a TFP condition, either a TFA or a TFR would be broadcast depending on whether a normal or only an alternate link set were available. In either case, it is important to complete the broadcast quickly, and there should be a requirement on the STP for completing this broadcast. Also, there should be a requirement on the time it takes a signalling point to respond after receiving a TFA message for a destination that is identified as inaccessible. Guidelines for the considerations to make and how to set these requirements are for further study.

6.2.2 Requirements for congestion controls

For any of the three types of congestion control (international, national option with congestion priority, and national option without congestion priority), which are specified in 11.2.3/Q.704 to 11.2.5/Q.704, signalling route congestion at STPs results in TFC messages being sent back to the source in response to receiving messages that route over a congested link in the congested route set (TFCs are not necessarily sent for every message arriving to a congested link). When a signalling point receives a TFC it informs its user parts of the congestion status of the related destination, and for the national option with congestion priority the MTP stops sending messages to the affected designation for messages having lower priority than the current route set congestion status. At any signalling point having a congested route set, if a local user part sends a message that goes to a congested link, actions analogous to receiving a TFC are taken. When user parts are notified of a congestion status, they are to take "appropriate action" to reduce signalling traffic. For the national option with congestion priorities, the MTP stops messages from being routed, and the action of the user parts does not affect the signalling link load. For the other two options the MTP does not exercise any control, and there is a total reliance on the user parts to reduce traffic. For the TUP and ISUP user parts the procedure is to decrease and increase traffic in steps as specified in 13.2.2/Q.724 and 2.11.1/Q.764, respectively. The number of steps and their size are not specified.

A congestion control procedure requires a reasonably quick response so that traffic can be reduced and congestion conditions controlled. Therefore, the following are seen to be important requirements to have an effective congestion control:

1) An STP requirement on the time to send a TFC in response to a message arriving at a congested link.

2) A requirement on signalling points for the time to respond to a TFC message, or the analogous action triggered by a message from a local user, and complete the steps specified in the appropriate procedures.

3) Requirements on the user parts for reducing and increasing traffic in response to congestion status information. This is absolutely needed for international signalling and the national option without congestion priorities. For TUP and ISUP this means requirements on the number of steps and size of the steps taken in changing traffic as specified in Recommendations Q.724 and Q.764.

The setting of these requirements is network dependent. Guidelines for the considerations to make and how to set these requirements are for further study.

6.3 SCCP management procedures

The SCCP management procedures are specified in 5/Q.714. Of particular importance to traffic flow and avoiding lost messages are the responses to signalling point and subsystem status changes. When a destination becomes inaccessible or a subsystem becomes prohibited, the SCCP changes its translations to a backup node and subsystem when they are provided. It is important that these changes in translations be rapid, because continued translation to inaccessible point codes or subsystems results in discarded messages. Another type of activity specified in 5.3/Q.714, are broadcasts to inform other users and signalling points of subsystem failure and recovery. It is important that these broadcasts be completed in a reasonable time. Therefore, as part of implementation, requirements need to be set on SCCP response times to notifications of inaccessible point codes and subsystem prohibited conditions, and requirements on the time to broadcast information to other point codes and local users are needed. Guidelines for the considerations to make and how to set these requirements are for further study.

7 Intelligent network traffic and congestion controls

7.1 High-level requirements

The objective of congestion and traffic control in IN-structured networks is to maximize the number of successful calls (for non-call-related messages the concept corresponding to successful calls needs to be defined) consistent with fair treatment between calls, while meeting the following high-level requirements:

Assure balanced treatment

Controls should provide a balanced treatment of different calls and transactions. Balanced treatment will sometimes mean equal probability of making a successful call. Alternatively, if congestion is caused by an overload of calls to one service provider, while call volumes to other providers remain within the planned limits, balanced treatment may mean selective blocking of only the calls causing the congestion. Another possibility is that certain calls (or messages) may be given priority for processing ahead of other calls or messages. In this case some mechanism is required to provide appropriate priority processing depending on what traffic parcels are causing the congestion. Thus congestion controls should provide balanced treatment:

– between services;

– between service providers;

– between users located at different points in the network;

– between user calls and non-call-related messages [such as Universal Personal Telecommunication (UPT) location updates]; and

– between calls of different priority.

Load share

Some elements within the network may be replicated for load capacity, survivability or database segmentation reasons. It is desirable to have load balanced over the possible serving network elements so that delay may be minimized. Under some failure conditions, network elements providing the same function may have different processing capacities. Thus it is desirable for congestion controls to automatically balance loading:

- across more than two network elements;
- across network elements of unequal capacity;
- across network elements with distributed logic or data.

Harmonize across levels and layers

Congestion and traffic control activity is not limited to only one protocol level or network management layer. Controls may be included in the specification of network protocols (for example level 2 and level 3 congestion in SS No. 7), in automatic control procedures in network elements or by specific features in network management systems. A harmonized approach is required. In general, each network element (and sub-element) needs overload control mechanisms to maintain element sanity under overload and to provide control should other network controls fail or not work properly. Overload controls should work harmoniously with congestion controls by allowing those controls to take effect before more drastic action is taken by the overload control. Overload triggers should be set so that adequate time is available for the controls to take effect before the system performance degrades significantly. Guidelines on how to set thresholds are for further study.

Differing control actions will be possible at different network elements (e.g. STP, SSP). These network elements should take control action appropriate to the data available to them. Network throughput can generally be maximized by identifying the cause of congestion or overload at the highest reasonable level and taking control action accordingly.

Thus the IN congestion and overload controls should be a harmonized set of functions providing:

- congestion control appropriate to the scope of data available and to the span of the control; or
- network management system-overload control within each element to maintain sanity and call processing should the above higher level controls not work.

Interwork effectively

The control process (the way a congestion control is activated and modified) should be capable of taking the appropriate control action to handle traffic surges and peaks which cause congestion:

- across network providers;
- across network boundaries (including Global Title Translation – GTT);
- across networks implementing different SS No. 7 and IN procedures.

Provide a stable, responsive and selective control

A proper IN flow control is essentially a feedback control system. It should operate in a stable, controlled manner, avoiding wide swings in the control parameters and actions. The controls should respond quickly to changing user traffic levels. As much as possible controls should automatically react to congestion or overloads. To provide control flexibility to avoid wide swings in control parameters, some form of volume sensitive (rather than on/off) control seems appropriate. To allow control of only those parcels of traffic causing congestion, the controls need selectivity to identify and control calls and messages based on several parameters. Depending on what parcel of traffic is to be controlled, the controls should be flexible to implement controls appropriate to the parcel's traffic characteristics. Thus a congestion control should provide:

1) fast, automatic activation;

2) smoothness of operation;

3) flow control capabilities:

   a) volume-sensitive (as opposed to on/off) controls;

   b) selective, based on parameters such as:

      - dialled number;
      - originating switch;
      - class of service of caller;
      - IN service requested;

4) control appropriate to the characteristics of the traffic in overload.

7.2 Standardized controls

Recommendation E.412 and 5.4.2/Q.1214, define a call rate control capability that should be used in IN-structured networks. Requirements must be established on the implementation of the call rate control capability so it properly protects the network elements. Guidelines on setting these requirements are for further study.

8 Requirements for adequate overload control

It is well-known that when traffic on a real-time processing system rises beyond its capacity, the overall performance of the system degrades. In an ideal case, the throughput rises linearly as a function of the input load until the capacity of the system is reached and then levels off at the system capacity. In the case of actual systems the throughput will rise linearly for light and moderate loads, but near the system capacity the throughput will not increase as fast as the offered load and beyond that point the throughput will usually decline. In most cases, if there are not adequate overload controls, this decline in throughput can be precipitous.

In SS No. 7 the MTP protocol uses signalling network management to control traffic after failure and recovery events in the SS No. 7 network. This network management activity requires real-time processing in the network elements and is a mechanism through which a failure in one part of the network can cause overload and malfunctioning in other parts of the network. Thus propagational failures can result leading to widespread outages. One means of preventing this type of failure propagation is to ensure each signalling network element has adequate overload control.

The overall goal of overload control is to detect overload and take actions to minimize performance degradation. Specific requirements for overload control can be stated as follows:

1) Maintain throughput near system capacity at specified overload levels.

2) Ensure system sanity and perform all critical functions necessary for the system and network to work properly regardless of overload level.

3) The system should protect itself from becoming deadlocked.

4) The system should be able to protect itself if network congestion or flow controls do not work.

5) The system should be able to recognize and shed excessive network management work without violating Requirement 2.

The above set of requirements can be considered as implementation system requirements, and they are testable requirements. For Requirement 1 the system can be tested by using traffic generators to load the system from low load to extreme overload and measure the throughput. The second requirement is testable by again loading the system with traffic and also sending network management messages, creating faults in the system, etc., and testing to see if the system performs properly. The third requirement would be tested by setting up conditions where buffers become congested. For example setting up signalling link congestion and disabling the congestion control procedures would be a test for both the third and fourth requirements. The fifth requirement could also be tested by driving the system with appropriate network management message tools.

9 Additional considerations for network operations and administration

9.1 Requirements for screening messages

One of the potential dangers in interconnecting signalling networks is that one network might send other messages the receiving network considers invalid. Software should be designed to cope with this situation, but it is easy for software developers to overlook some conditions and not recognize that an invalid message of some type can cause their software to fail. One way to protect against this potential failure mechanism is to have a screening capability where the traffic enters the system (e.g. and the receive side of a signalling link) and check messages to ensure all parts of the message meet the requirements for that network. The idea is to check all information that is to be processed within the system before sending it to the software that does the processing. This type of screening capability would be defined as part of implementation and should not be confused with screening to prevent unauthorized use of STPs as discussed in clause 8/Q.705.

This type of capability is testable to the extent one can specify the constraints a network has on what the allowed messages are. It is always possible to overlook some condition, so it cannot be guaranteed that every possible fault will be covered, but by carefully defining the allowed message sets one can provide significant protection. The point is that it is not possible to test for every possible invalid message, so an explicit specification of constraints is needed to have a testable capability.

9.2 Requirements for data administration

Signalling and IN-structured networks have stringent availability requirements, and to meet these requirements, redundant components are needed. Thus, in many network topologies there are mated systems that back each other up. As a consequence they share the same data. Networks are constantly changing, so data administration is done to put in new routing or modify existing routing, to add signalling links, to rearrange components, etc. To make these changes, data has to be provided or modified in both a system and its mate. To achieve the needed reliability, the failures of a system and its mate need to be independent. However, data administration activity is one way that correlated failures can be introduced. For example, suppose there were a software bug present that would get triggered by some data administration activity. Then if both a system and its mate had their data changed at the same time, the bug could be triggered at the same time and a mated outage could occur. Therefore, a useful strategy is to stagger common data administration activities between a system and its mate.

To stagger data administration updates, the signalling network must be designed so that a system and its mate can have different data. Simultaneous data administration updates should be avoided but this may not always be possible. Also, within network elements there are mated components that back each other up in order to meet stringent system availability requirements. These components should not have their common data changed simultaneously. This becomes a requirement of system design.

Since most data administration is done through support systems, the need to stagger data administration updates places requirements on those support systems so they can properly implement the desired staggering strategy.

9.3 Principles for component removal

For further study.

9.4 Relationship with network management and test procedures

For further study.

10 Testing requirements

For further study.

11 History

This is the first issue of Recommendation E.744.

ITU-T RECOMMENDATIONS SERIES

Series A	Organization of the work of the ITU-T
Series B	Means of expression
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	**Telephone network and ISDN**
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media
Series H	Transmission of non-telephone signals
Series I	Integrated services digital network
Series J	Transmission of sound-programme and television signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	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
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminal equipment and protocols 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