INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

E.734

(10/96)

SERIES E: TELEPHONE NETWORK AND ISDN

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

Methods for allocating and dimensioning Intelligent Network (IN) resources

ITU-T Recommendation E.734

(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.734

METHODS FOR ALLOCATING AND DIMENSIONING INTELLIGENT NETWORK (IN) RESOURCES

Summary

This Recommendation provides an overview of typical considerations which are useful in network planning for the initial deployment and subsequent growth of networks based on Intelligent Network architectures, including considerations of interworking with existing PSTN/ISDN infrastructures.

It addresses methods for analysing network infrastructure requirements and detailed dimensioning of Intelligent Network nodes to accommodate expected traffic loads for services offered by a given network.

Source

ITU-T Recommendation E.734 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.734

METHODS FOR ALLOCATING AND DIMENSIONING INTELLIGENT NETWORK (IN) RESOURCES

(Geneva, 1996)

1 Scope

This Recommendation provides an overview of typical considerations which are useful in network planning for the initial deployment and subsequent growth of networks based on Intelligent Network (IN) architectures.

Because the various relationships between Intelligent Network functional entities can be implemented using different kinds of subnetworks, an important consideration is interworking with existing PSTN/ISDN infrastructures. Moreover, this Recommendation deals primarily with implementations using SS No. 7 signalling subnetworks but does not imply that these are the only implementations.

Networks providing IN services may also provide basic PSTN/ISDN services. Network elements providing IN services may also provide non-IN functions. Dimensioning of IN-structured networks thus should consider all demands on network elements. Recognizing that appropriate procedures exist for dimensioning the basic PSTN/ISDN, this Recommendation identifies the extra calculation procedures required to address the needs for IN services.

Note also that this Recommendation is developed in the context of IN Capability Set 1.

In this Recommendation, references to various networks (including Intelligent Networks and signalling subnetworks) and any implied network boundaries are not intended to have any ownership or regulatory significance. Such issues may have operational implications, but such matters are outside the scope of this Recommendation.

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 were valid. All Recommendations and other references are subject to revision: all users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T recommendations is regularly published.

- ITU-T Recommendation E.492 (1996), *Traffic reference period.*
- CCITT Recommendation E.508 (1992), *Forecasting new telecommunication services.*
- CCITT Recommendation E.520 (1988), *Number of circuits to be provided in automatic and/or semiautomatic operation, without overflow facilities*.
- CCITT Recommendation E.522 (1988), *Number of circuits in a high-usage group*.
- ITU-T Recommendation E.724 (1996), *GOS parameters and target GOS objectives for IN services.*
- ITU-T Recommendation E.733 (1996), *Methods for dimensioning resources in Signalling System No. 7 networks*.

- ITU-T Recommendation Q.1200 (1993), *Q-series intelligent network Recommendation structure*.
- ITU-T Recommendation Q.1205 (1993), *Intelligent network physical plane architecture*.
- ITU-T Recommendation Q.1211 (1993), *Introduction to intelligent network Capability Set 1*.
- ITU-T Recommendation Q.1215 (1995), *Physical plane for intelligent network CS-1*.

3 Definitions

Definitions of Intelligent Network terms can be found in the Recommendations listed in clause 2 – in particular, see Recommendation Q.1215 for definitions of the following terms:

- Service Switching Point (SSP);
- Service Control Point (SCP);
- Service Data Point (SDP); and
- Intelligent Peripheral (IP).

4 Abbreviations

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

CS-1	Capability Set 1
GOS	Grade of Service
IN	Intelligent Network
IP	Intelligent Peripheral
ISDN	Integrated Services Digital Network
ISUP	ISDN User Part
MTP	Message Transfer Part
PSTN	Public Switched Telephone Network
SCCP	Signalling Connection Control Part
SCE	Service Creation Environment
SCEF	Service Creation Environment Function
SCF	Service Control Function
SCP	Service Control Point
SDF	Service Data Function
SDP	Service Data Point
SIB	Service Independent Building Block
SMAF	Service Management Access Function
SMF	Service Management Function
SP	Signalling Point
SRF	Specialized Resource Function
SS No. 7	Signalling System No. 7
SSF	Service Switching Function

SSP	Service Switching Point
STP	Signalling Transfer Point
TCAP	Transaction Capabilities Application Part
UPT	Universal Personal Telecommunication

5 Introduction

This clause describes the aims of this Recommendation, which deals with resource allocation and dimensioning methods for Intelligent Networks.

Resource allocation encompasses the methods for locating IN-specific elements and for partitioning Intelligent Network functionality (such as service logic) among these elements. Dimensioning determines the sizing of network components.

In the case of Intelligent Networks, as for the PSTN/ISDN, traffic engineering activities involve the interplay of three factors: performance objectives, traffic forecasts, and resource allocation/dimensioning. Recommendation E.724 deals with GOS parameters and performance objectives; this Recommendation deals with traffic forecasts and resource allocation/dimensioning.

Subclause 5.1 discusses the new traffic engineering considerations needed to address Intelligent Network deployment; subclause 5.2 briefly describes the approach to resource allocation/dimensioning.

Clause 6 elaborates on how to assess service loads and message flows in the network.

Clause 7 discusses the subject of resource allocation in more depth; subclause 7.1 covers the general factors to be considered while subclause 7.2 details particular issues relating to SSPs, SCPs, SDPs, and IPs.

Clause 8 discusses various aspects of dimensioning the circuit-switched subnetwork and IN-specific components as well as the supporting signalling subnetwork.

Clause 9 provides for each network some general guidelines on re-dimensioning. The intent is to facilitate the exchange of information between network operators and equipment manufacturers – and not to provide an exhaustive analysis of network planning tools for particular networks and specific IN implementations.

5.1 New traffic engineering factors in Intelligent Networks

The deployment of Intelligent Network services introduces some new aspects into traffic engineering. This encompasses PSTN/ISDN subnetwork issues and the important subject of re-dimensioning.

IN-based services are normally incremental to existing PSTN/ISDN networks, but traffic engineering and resource allocation/dimensioning for networks offering IN-based services involve several new factors to be considered in planning and managing IN networks.

In an Intelligent Network environment, the following factors have an impact on traffic models and traffic forecasts:

– *Distribution of the Busy Hour Call Attempt rate (BHCA)*

  The dimensioning process for IN should take into account that the BHCA for IN services may have characteristics which differ from those adopted for PSTN. Some IN services, like Private Virtual Networks, have peaks during normal working hours, while services like

Televoting are generally concentrated in the evening. Therefore, it is necessary to analyse the distribution of the BHCA in order to derive traffic loads for the dimensioning process.

– *Traffic characteristics*

  The characteristics of the distribution of the call arrival process for some IN services may deserve special attention. The Poissonian assumptions adopted for the PSTN may not apply for all IN services. Also, IN services which may lead to mass calling situations must be carefully considered.

– *Traffic loads*

  In the PSTN, the dimensioning of network elements is made based on call attempt rates and average holding times. These values are approximately known and used as input values. In an IN environment, however, the services may have differing degrees of complexity. A simple IN service such as number translation, with high call attempt rate, may have the same impact on the network as a complex service with a relatively low call attempt rate.

– *Call processing logic*

  With IN-based services, call processing logic is no longer provided only by switching nodes acting sequentially on a given call. Instead, call processing logic may be distributed over other nodes such as SCPs and IPs. In general, while the call processing logic may be described independently of the network architecture, planning for the implementation of a given service requires the specification of service logic functions and the assignment, or allocation, of these functions to particular network nodes. Thus, IN-based services generate a new requirement within service design and implementation: namely, allocation of call processing logic to appropriate network nodes.

– *Call processing data*

  In the same way that IN-based service logic may be distributed over various network nodes, the specific data used by the call processing logic may also be distributed, and will not necessarily duplicate the distribution of service logic. Thus, IN-based services generate a second allocation requirement: namely, allocation of the storage locations for data required for call processing.

– *Communication paths*

  The distributed nature of the call processing logic and data generates additional signalling traffic (typically on an SS No. 7 subnetwork) and additional circuit-switched traffic (for example, to establish communication paths to IPs). These new parcels of traffic need to be characterized, and new dimensioning procedures may be required when IN-based services are added to existing PSTN/ISDN.

– *New network elements and functions*

  The new network architecture for IN-based services includes several new network elements (such as IPs and SCPs) and new functions for existing nodes (such as local and transit exchanges). These new nodes and functions require appropriate dimensioning guidelines for use in planning, implementing and providing new IN-based services.

  The Intelligent Peripheral is a new element associated with the delivery of IN-services requiring interaction with the user as part of the call set-up process, and the Service Control Points (SCPs) provide the distributed service logic to support IN-based services. Other components, such as Signalling Transfer Points (STPs) and their interconnecting links may already be present in some pre-IN networks or may be dedicated to message processing only for IN-based services.

- *Additional complexity of traffic estimation*

  Traffic will clearly be affected by the specific services developed and the way the services are invoked by users, but the resource-allocation process will affect the distribution of traffic loads (expressed in messages and calls per unit time originating or terminating at a given node). Service developer traffic forecasts may not be readily available (for example, if the service has been initially developed in a different network). Equally, users may find novel applications for the services in ways not anticipated by the developer or network operator. Thus planning, allocation and dimensioning procedures must be made flexible enough to provide, as quickly as possible, the resources required as user demand grows and changes.

- *Re-dimensioning*

  The rapid introduction of new IN-based services requires an understanding of the impact of such new services on the different portions of the supporting IN infrastructure. Indeed, a new traffic engineering procedure is introduced to answer these kinds of questions; namely, the re-dimensioning procedure. Further detail on re-dimensioning is given in clause 9.

- *Other issues*

  The IN functions SMAF, SCEF and SMF should also be considered, because these functions will generate messages which may provide part of the processing load for other IN functions and which may be carried by some network operators in the SS No. 7 signalling network.

5.2 General approach to dimensioning

Figure 1 shows the various network operation actions taken over different time scales to respond to changes in service demand and traffic loads:

- traffic management (short term);
- re-dimensioning (medium term);
- dimensioning (long term).

Planning for Intelligent Network services usually begins by developing an estimate of the projected demand, which will probably be based on a marketing forecast of the projected sales of those IN services; the forecast, in terms of sales, customers and perhaps usage, must be translated into traffic matrices giving node-to-node circuit and message demand volumes (see 6.1, including Figure 2, for further detail).

When the demand is determined, it is mapped or allocated to the equipment and functional nodes in the network, while at the same time the functional capabilities required in the network are allocated to the nodes. When this is done, the traffic demand has been specified, by type and location, and routing rules established for the services. Individual network elements can then be dimensioned, and an assessment is made regarding the required number of circuits and signalling links.

On an ongoing basis, a process of re-dimensioning (sometimes called servicing) is used to ensure maximum utilization of the existing equipment and determine appropriate reallocation when service demand changes before additional equipment can be installed. (See clause 9 for more detail.)

Traditional network traffic management is an activity closer to real time. Re-dimensioning activities are carried out at time intervals intermediate between dimensioning and network traffic management.

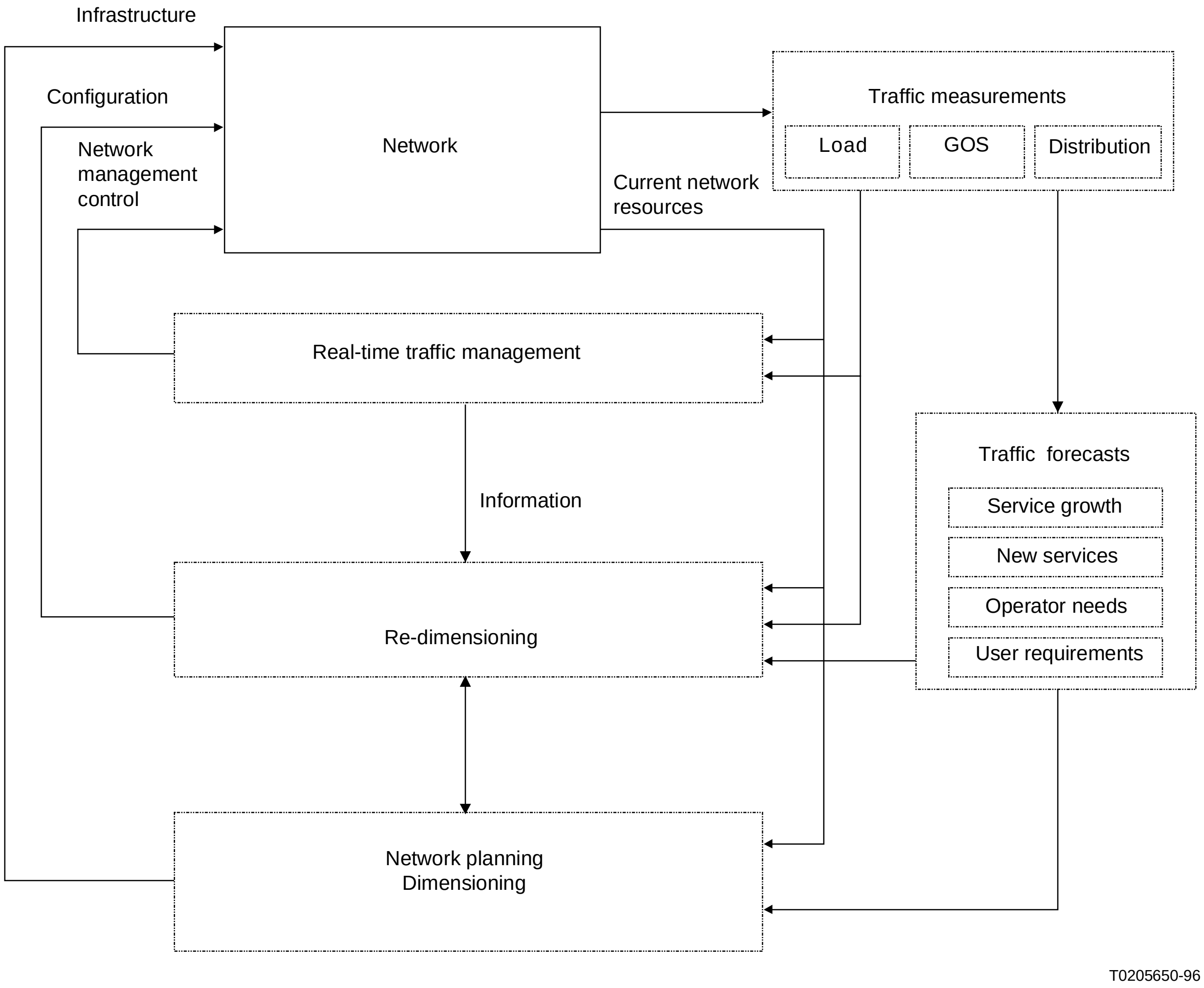

FIGURE 1/E.734

Network dimensioning/re-dimensioning in an Intelligent Network environment

6 Intelligent Network traffic forecasting

6.1 General

This clause identifies the mapping between message flows and service attempts, and what service forecasts are needed, keeping in mind that there are both call and non-call related flows.

Recommendation E.508, Forecasting for New Services, provides related information.

Planning for IN services generally starts with a forecast of IN service usage. The overall forecast may be broken down by service and, if possible, should contain details about the customer distribution, utilization rates, etc. What is required for dimensioning purposes is the following set of data:

– the expected traffic volume (measured in terms of call rate);

– signalling message rates; and

– service usage (by service).

A process is required to translate marketing or service forecasts into a traffic demand as shown in Figure 2.

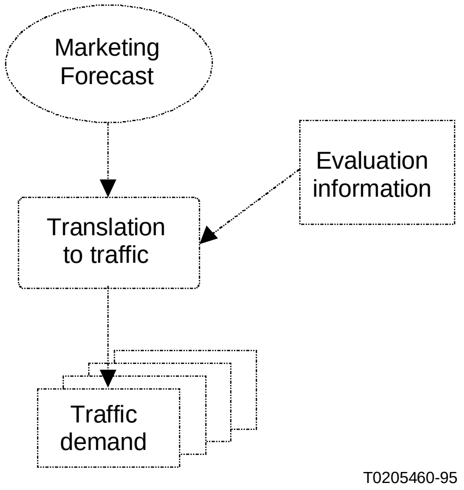

FIGURE 2/E.734

Traffic demand evaluation

To assist the conversion process, a set of evaluation data is desirable. The data may have several forms to help different aspects of generating the traffic demand. For example, service descriptions may be used to develop multiplication factors to convert service invocations in the marketing forecast into TCAP messages to a particular SCP. Other data may give average holding times by service, call type, user type, etc., for use in estimating loads on circuits and processing capacity.

The IN service to be introduced may be described by service logic and represented as a chain Of Service Independent building Blocks (SIBs). By calculating the load for each SIB, the load represented by the service can be obtained by summing the loads for each of its SIBs.

In some cases, instead of a service-by-service forecast, the customer demand may be expressed in the utilization of SIBs. A SIB can result in a distinct set of information flows and circuit connections and so may be used as a scale factor in estimating traffic demand. For example, a "User Interaction" SIB may require the set-up of a connection to an IP to collect additional dialled digits. Thus an estimate of the usage of this SIB may be translated into circuit and switching load (call rate, average holding time, etc.) and the signalling load (ISUP messages, if any, and TCAP messages).

When IN services are offered in a network, operational measurements may be used to provide additional forecast evaluation data. For example, average query rates to SCP and average holding times for IP circuits may be calculated and then used to improve the accuracy of the traffic demand estimates.

It is possible that marketing data may be applicable for time periods which do not represent the reference period data required for proper network dimensioning. Some process, beyond the scope of this Recommendation, is required to translate the demand data to the appropriate reference period load data. Information on selecting appropriate reference period data is contained in Recommendation E.492.

The specific method for converting the marketing demand to traffic demand is for further study.

6.2 Understanding traffic flows

The planning process in an IN environment contains aspects that must be carefully handled. One of these aspects is a direct consequence of the platform provided for the delivery of new services in a fast and flexible manner. The large number of expected new services may be characterized by a complex traffic profile. There is the necessity of a precise traffic characterization already in the service specification phase.

The basic components of IN service definitions are the SIBs. A SIB can result in distinct information flows in the physical plane. The user interaction SIB, for example, impacts the network in different forms depending on factors such as the number of announcements to be played, amount of information to be collected, and the location of the IP.

A new service is defined as a chain of these SIBs. Generally, the service logic depends on decision events and consequently the chain of SIBs representing the service logic approaches a tree structure. According to the service logic and the situations involved, quite a complex graph of SIBs may be expected.

The traffic of an IN service is generally given as a number of call attempts per time unit. The service logic offers different possibilities for the execution of the service, e.g. successful call, fail situations, time-out mechanisms, etc. These possibilities impact the IN network infrastructure in different ways – depending on the resources involved. This diversity of situations must be studied carefully.

For planning purposes, the impact of the introduction of a new IN service can be assessed by representing all possible paths in the service logic linearly – forming a set of scenarios of the service logic. This concept is shown in Figure 3.

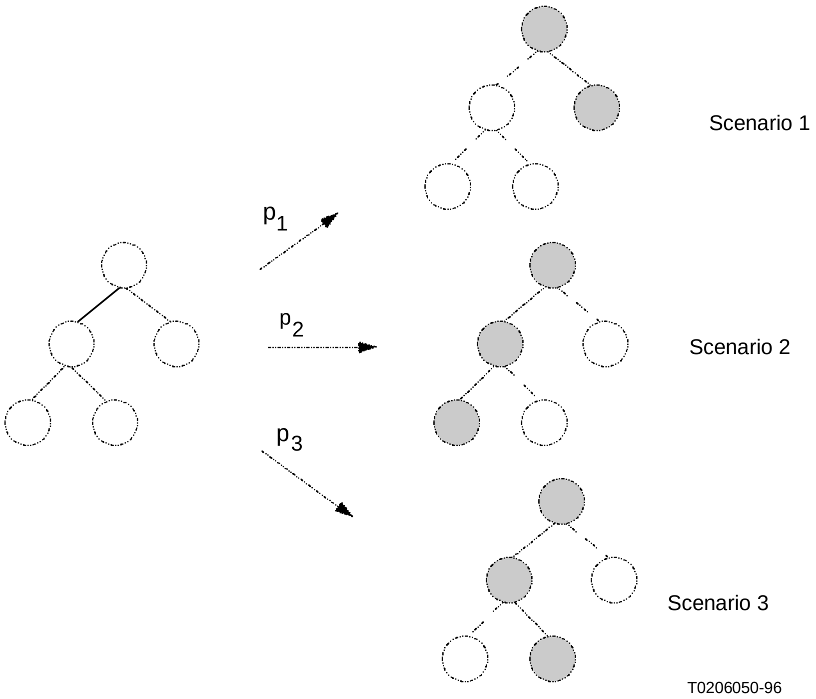

FIGURE 3/E.734

Linearization of the service logic

For each one of these paths a probability is assigned. Therefore, it is possible to derive the call attempt rate for each path. This allows a better view of the impact of each situation on the overall network structure.

It should be noted that the above probabilities are affected by factors such as the experience of the user, country characteristics, and market factors. Therefore, to assess such probabilities prior to launching new IN services, there is the need to interact with marketing people. After the deployment of the services, these probabilities may be better estimated by measurements.

For most planning purposes, all possible scenarios do not need to be completely enumerated, and just the most significant ones – covering the majority of calls – may be considered.

6.3 Development of IN service traffic matrices

Traffic matrices represent traffic in only one plane of the IN architecture and appropriate steps need to be identified to map traffic demand in one plane to the nodes and links in another plane.

As shown in Figure 4, the IN service demand must in general be integrated with the structure of the basic PSTN, so that proper dimensioning can be done. For example, if calls are routed to particular nodes to reach an IP function, they may use the same circuit groups which are used for basic PSTN calls. With quasi-associated signalling for ISUP, links from SSPs to STPs will carry both ISUP and TCAP traffic.

There may also be additional switching or signalling load offered by other services. For example, virtual private network services may use some of the same network elements as PSTN or IN calls. If there are other sources of calls or messages, the traffic demand for these should also be included in the total demand.

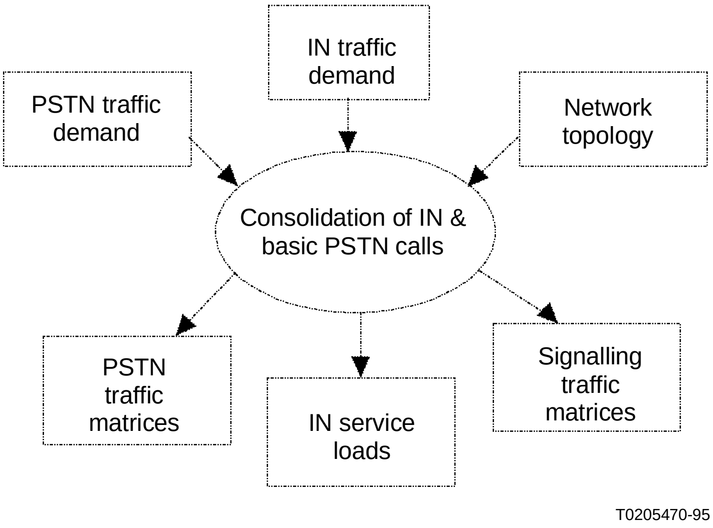

FIGURE 4/E.734

Generation of traffic matrices

The IN traffic demand, in terms of calls, messages and service invocations, is a function also of time and geographic or machine location. This data must be translated by mapping it onto the nodes of the circuit plane and the signalling plane. There are three sets of resultant matrices:

- PSTN traffic matrices (erlangs, call attempts, average holding times, etc.) representing the demand placed on the circuit-switched subnetwork;
- signalling traffic matrices (messages, average message length, peakedness of offered messages, etc.) representing the message load offered to the signalling subnetwork; and
- service loads forecast by IN node indicating the processing load expected at each of the nodes. This load can then be used in the node dimensioning process.

7 Resource allocation

7.1 Introduction

This clause provides an overview of the typical factors which should be taken into account in planning the location of various IN network elements. The relative importance of these factors will vary from network-to-network, and in particular situations additional factors may well need to be considered.

The IN architecture is quite flexible, allowing different degrees of concentration of node capacity and node types in one location. Also, the relatively low cost of transmission allows the sending of calls or messages to distant elements for processing, if this choice is made. Therefore in designing an IN-structured network there is a wide range of options open for consideration. Choices range from full centralization of all processing functions to complete distribution of the functions.

In typical networks, the buildings are usually considered fixed, except in the very long term. So for planning and administration purposes, the real issue is deciding in which location there should be one (or more) of the network elements. The purpose of resource allocation is to decide which functions are provided in each node and also how the customer demand (and service capability) will be allocated to these network elements.

Many different criteria might be used in deciding which process to use in allocating call processing logic functions and call data to the network elements available from the generic IN-service architecture. Some of the following criteria may be useful in making a traffic assessment of the alternatives:

- computing efficiency – where functions are most efficiently performed: for example, so as to minimize signalling load, computing load, signalling links, and unnecessary repetition of logic processing in different nodes;
- computing load – to maximize the capacity utilization of existing equipment in the network;
- minimization of delays:
  - post selection;
  - service data updating:
    - service subscription;
    - service billing;
    - call-associated data – for example, call counts for selective blocking or televoting applications;
    - non-call-associated data – for example, time-of-day routing tables, universal personal telecommunications (UPT) location updates;
- service logic updating;
- other delays (such as UPT registration delay) may need to be identified and minimized;
- cost minimization;

- capital (investment);
- operating;
- load balancing – when multiple elements are replicated for capacity or survivability;
- network protection;
- transmission costs;
- flow control requirements (identified in Recommendation E.744, which also provides guidelines for the implementation of flow controls).

There may be other, non-traffic related constraints that must be considered in deciding on a proper allocation of IN functions. Such factors might include (amongst others):

- security and privacy requirements;
- regulatory or ownership constraints; and
- user-friendliness of the service (if the location of logic or data will affect how users interact with the service).

Analysis of these non-traffic factors is outside the scope of this Recommendation.

7.2 Methods for locating IN specific elements

7.2.1 Introduction

This subclause provides information on methods for locating IN network components. Specific information is provided for SSPs, SCPs, SDPs, and IPs; appropriate extensions of these methods can be used to deal with other IN components such as service nodes and adjuncts.

7.2.2 Methods for locating SSPs

The SSPs are elements deployed as a software extension of digital switches. The location of the SSPs depends on the strategy chosen for the introduction of the IN platform and on the availability of the implementation functionality on the digital switch software.

The following are the three basic methods for locating SSPs:

a) Deployment of SSPs in an overlay network – This alternative allows the deployment of a separate network with switches handling only IN calls. The IN calls are routed to one of these special switches containing the SSF functionality to invoke the IN service logic. An advantage of the overlay method may be lower global cost and deployment of the platform in a relatively short time. Additionally, there is no high dependency on the digitalization of the existing network. Another point to be considered is the isolation of the IN network from the standpoint of propagation of errors and the impact of failures on the non-IN network.

   A potential disadvantage of this method is that the necessary routing of calls to the specialized switch will likely result in the use of long-distance resources and hence higher transmission costs. The specialized switches are also potential bottlenecks as the IN traffic increases.

b) Deployment of SSPs at the transit level – In networks with a hierarchical structure, the SSF functionality can be implemented in a transit switch where both non-IN and IN calls are handled. This method allows a moderate cost for the implementation of the SSPs and possibly a better utilization of the transmission resources as compared to the overlay architecture. Of course, a digital platform at the transit level is required.

   The implementation of SSPs at the transit level may result in a high dependency between the IN services and the telephone network. The occurrence of problems in the IN software or the effect of failures can spread beyond the IN platform and may have consequences for the provision of telephone services. The deployment of new software in a transit exchange is therefore a delicate operation.

c) Deployment of SSPs at the local level – The SSF functionality, in this case, is deployed in the local switch. The implementation cost is the higher among the presented alternatives, but transmission costs are less. This method of deployment has the advantage that errors are more likely to be limited to the coverage area of the local switch.

   This solution requires a modern network infrastructure and high level of digitalization in the network. For a wide coverage a considerable amount of investment may also be necessary.

The above three methods are the theoretical ones, but the most common configurations will probably involve some combination of them. In regions with a low penetration of the IN services or with a low degree of digitalization, the first solution is feasible. For some areas with a medium degree of digitalization and not so high IN traffic, the second solution applies. The third solution may be found in metropolitan areas with high degrees of digitalization and high IN traffic. These are aspects to be considered in the SSP location process, and the optimal solution depends on the particular conditions of the deployment environment.

Once the environment factors described above have been considered, the location process may be started. By way of example, one possible approach to determining the location of an SSP is to treat the task as an optimization process which considers the balance between additional transmission costs to route IN calls and the costs associated with the SSP infrastructure. This approach entails assessing whether transmission cost savings can be achieved by locating SSPs in points where higher demand for IN services is forecast (typically based on marketing information). The main idea of this method is shown in Figure 5.

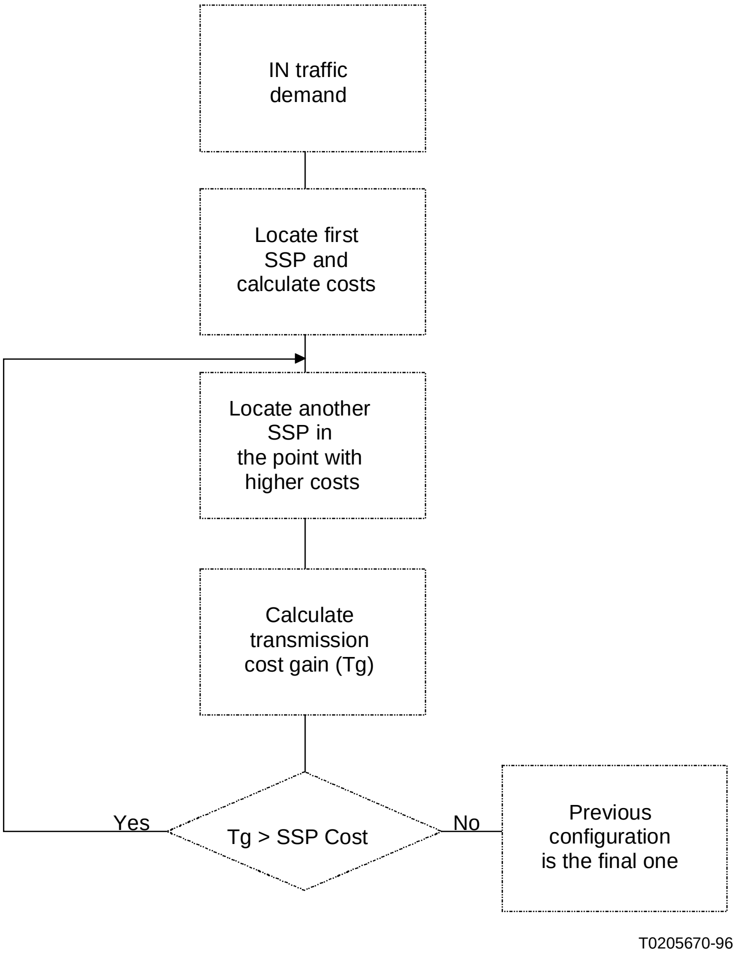

FIGURE 5/E.734

SSP location method

7.2.3 Methods for locating SCPs

The SCPs are centralized elements which contain the service logic subprograms and data used to provide IN service. When higher reliability is required, SCPs are generally implemented in pairs, geographically separated. The location process for the SCPs, as in the case of the SSPs, also depends on the existing infrastructure, which is related to the existence of a signalling network for connecting the SCP with the SSPs.

After the consideration of the structure of the signalling network, the location process may be started. As in the previous subclause, one approach would be to determine whether transmission and processing cost savings in the signalling network justify the deployment of additional SCPs.

7.2.4 Methods for locating SDPs

SDPs are elements which contain the user and network data accessed during the execution of IN service logic. Again, as in the cases of locating SSPs and SCPs, the process for determining where to locate an SDP can be based on considerations of transmission cost savings – in this case savings in the transmission costs for the signalling traffic between the SDP and the SCP.

If the SDP is accessed by a single SCP, the problem may be solved by the use of a high-speed interface, leading to the integration of the SDP into the SCP. In the case of an SDP accessed by more than one SCP, the location process for the SDPs also depends on the existing infrastructure. And in this case the necessary infrastructure is related to the existence of a signalling network for connecting the SDP with the SCPs.

7.2.5 Methods for locating IPs

IPs are elements which execute user interactions and are usually located near to the SSPs. In this case the location process is a function of the configuration adopted for the IPs in the IN platform structure. In the Q-Series there are several options for the connection of an IP and several interfaces between IP/SCP and IP/SSP are possible. Cost factors to be considered are the signalling costs and the transmission costs. The signalling costs are those related to the transmission of the signalling information between the IP and the SCP, while the transmission costs are associated with the connection between the IP and the user.

As in the case of SSPs, the task of locating IPs can be treated as an optimization process which considers the balance between the additional transmission costs to route the user to the IP and the costs associated to the IP infrastructure.

8 Dimensioning of Intelligent Networks

The main goal of network dimensioning is to determine the number and size of the network elements required to carry a forecast load while satisfying given grade-of-service performance requirements. It is part of the network planning process to calculate systems and equipment requirements, evaluate network development costs, determine investment schedules and organize engineering work. As a result of dimensioning activities, the network structure will be modified and its capacity adjusted to support new service offerings and growth in existing services.

8.1 Load determination

When information has been obtained regarding the service demand and traffic loads, an initial view can be developed regarding an allocation of network resources to provide the functional capabilities. This load can then be mapped or allocated to equipment and functional nodes in the network to develop traffic matrices as described in 6.3. The topology of the existing network (existing nodes, circuit groups and their size), traffic models, traffic routing rules, performance requirements, dimensioning models, technical or general constraints and the availability of new equipment are inputs to this overall process.

8.2 Resource allocation and element dimensioning

After the traffic matrices have been determined and the network resources allocated, an assessment can be made regarding the required number of circuits and signalling links, and the various network elements can then be dimensioned.

8.3 Network modelling

After the network elements have been dimensioned, it may be desirable to validate the dimensioning with an analytic or simulation model of the network. In this way it can be confirmed that the network is meeting end-to-end performance requirements. In the case of new services, the impact on various network parameters such as circuit, link and node loading should be assessed. (Specific techniques to perform this modelling are for further study.)

8.4 Signalling subnetwork considerations

Signalling networks achieve the high availability required by providing diverse extra capacity to handle the load of any failed component. The amount of redundant capacity depends on the signalling network architecture. The nodes and links should be dimensioned to meet objectives specified for failure conditions that fully utilize the redundant capacity. Intelligent network dimensioning objectives follow these principles.

From the network provider point of view, the most important criteria for node dimensioning are delay and congestion. Other factors to be considered are reliability and survivability. The network architecture and the node architecture are also to be considered.

GOS parameters and target GOS objectives for IN services are described in Recommendation E.724. These parameters guide the calculations of the number of links and node devices required to handle IN loads.

The dimensioning of signalling networks under IN conditions must take account of the special characteristics of IN services and their impact on the underlying signalling network. These aspects may be summarized as:

- Message length – Some signalling messages carrying IN information use higher layer capabilities (TCAP) of the signalling network protocol. These messages are longer than those of traditional PSTN/ISDN applications. The corresponding increase of load in the signalling links must be carefully considered.
- Processing requirements – Some IN signalling messages are more complex than those of the PSTN/ISDN and require more processing capacity from the involved signalling network elements. A study should be carried out to determine the impact of the IN messages on the utilization of the processors and to check that the utilization does not exceed established capacity limits.

8.5 Dimensioning of network components

This subclause discusses in more detail dimensioning of circuit groups, components of the supporting signalling subnetwork, and the various network elements.

8.5.1 Dimensioning of circuit groups

Once the circuit network traffic matrices have been determined, the nodes and links in the circuit network may be dimensioned. The links are usually voice circuits interconnecting the switching nodes and also connecting switching nodes with the IP nodes providing IN services

Figure 6 outlines the process of dimensioning circuit groups – which, in general, must take into account the signalling link requirements to support the identified service load on the overall network. The traffic matrices developed in 6.3 are used together with appropriate dimensioning algorithms (such as for dynamic routing) according to routing rules, any specified constraints, and appropriate GOS requirements. Circuit group dimensioning techniques are provided in Recommendation E.520. The result of the analysis is the quantity of circuits required in each group for which traffic data is provided in the circuit network traffic matrix.

Dimensioning of communication links in the signalling network is addressed further in 8.5.2 and 8.5.3.

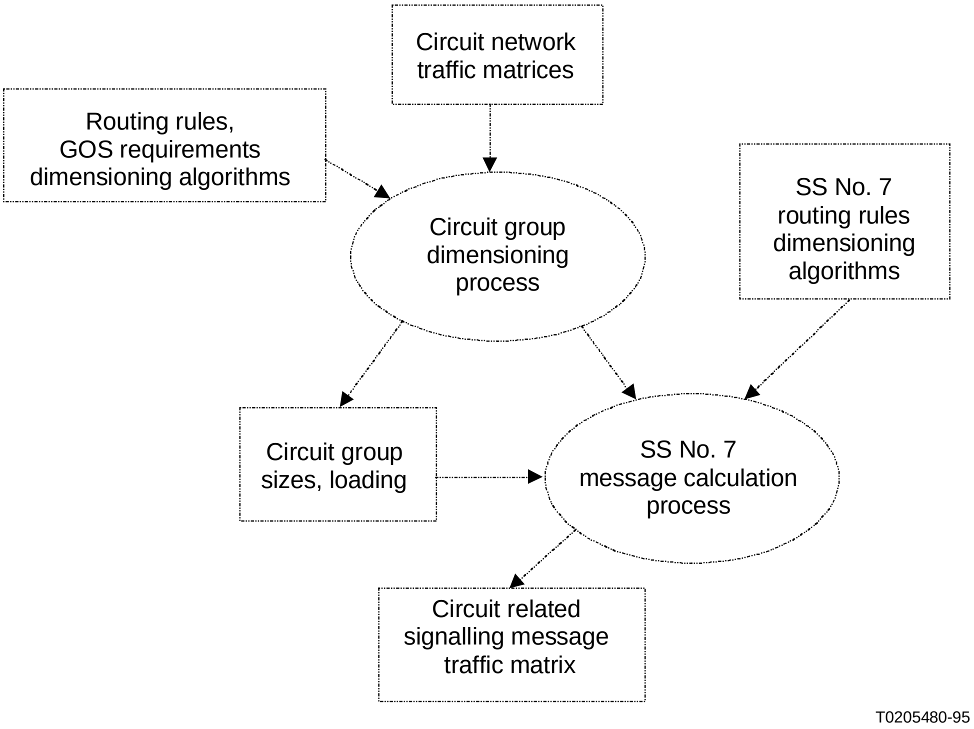

FIGURE 6/E.734

Dimensioning of circuit groups and communication links

8.5.2 SS No. 7 link dimensioning

The results of the circuit group dimensioning process (circuit quantities, carried load) are fed into the SS No. 7 message calculation process. This process uses the SS No. 7 routing rules and the call loads to calculate the expected signalling message loads, average message lengths, etc.

The quantity of SS No. 7 links required to handle the projected message demand is usually calculated using link occupancy as a criterion; additional information on link dimensioning is available in Recommendation E.733.

8.5.3 Dimensioning of non-SS No. 7 communication links

In order to provide IN services in the network being planned there may be links which do not use the SS No. 7 protocol; examples might be the links between the SCP and SDP functions. Where alternate protocols are used for communicating, appropriate procedures should be used to dimension the communication links.

8.5.4 STP dimensioning

Within the framework of IN dimensioning, suitable models and engineering rules have to be described to determine the STP system capability – in terms of traffic handling and number of links the system can manage – to meet target performance requirements. The system dimensioning depends on the offered call attempts, the functional aggregations foreseen at MTP and SCCP levels and the system architecture. The signalling network traffic and the signalling network architecture are significant factors.

STP systems can have a centralized architecture, a distributed architecture, or a hybrid architecture. The main differences reside in the way processors are grouped and managed and on the existence of common memories. The number of processors and the capacity needed to manage a specific signalling traffic load and the number of distinct connections respecting a maximum message transfer delay and a minimum node availability must be determined.

For a defined network configuration, to obtain a signalling load it is necessary to first determine the signalling relations, the routing of the signalling load through the signalling network and the aggregation of signalling loads in the nodes.

An STP must be dimensioned to provide for the termination of all required SS No. 7 links, as well as the processing capacity sufficient for the STP functions.

8.5.5 Dimensioning of circuit switching nodes

When the circuit groups have been dimensioned, the impact of this switching load on the node may be calculated. This is useful to ensure that sufficient capacity exists to handle the expected volume of traffic.

The procedures for dimensioning circuit-switching nodes are typically proprietary to the particular equipment used.

Note that for those circuit switching nodes which use SS No. 7 signalling or which also provide other capabilities (such as SSP), the node dimensioning will be dependent on, and therefore logically must follow, the dimensioning activity for the signalling links.

8.5.5.1 SSP dimensioning

As an SSP contributes to the overall network performance (as regards delay and blocking), to dimension SSPs, GOS targets should be met. Parameters influencing SSP performance have to be identified, rules derived to determine such influences, and methods elaborated to dimension SSPs.

The dimensioning procedure determines the SSP capacity to process the traffic load, satisfying GOS requirements. This dimensioning depends on the network architecture and on the processor architecture. After calculating the traffic load, the required capacity is derived and the evaluation of expected delays is made, to check the compliance with GOS requirements.

SSP architecture is dealt with in two aspects: the distribution of the SSF in the whole network and the assignment of different SSF functions to a number of processors.

The IN architecture has consequences on the SSP dimensioning: the functionalities vary for stand alone SSPs or SSPs integrated to local exchanges. The external connections to other network elements may differ depending on the IN architecture.

Different types of processor architectures are possible for both stand alone and integrated SSPs. For each processor architecture the relevant capacities must be identified: for example, the amount of real time available for processing IN messages, equipment that can be connected to other equipment, capacity in messages or bits per second on the bus or links between the different blocks and processors.

The dimensioning methodology is similar to the one used for STPs. According to the distribution of functionalities, the number of processors, bus, links and processing capacity are evaluated.

The SSP must be dimensioned to meet several requirements:

- sufficient circuit capacity to terminate the number of circuits calculated previously;
- sufficient storage resources to handle the requirements of the expected demand for SSP functionality; and
- sufficient processing capacity for the expected total SSP load (circuit switching and IN processing). There may be difficulty in obtaining detailed information on the processing

times of IN operations in a specific device. In this case, a mean processing time for an individual operation may be available or measurable. Alternatively an overall average processing time may be estimated for all operations.

8.5.5.2 Switching nodes with no SSF functionality

In dimensioning switching nodes with no SSF functionality, the traffic impact resulting from the IN service load must be considered. This information is captured in the traffic matrices described in 6.3.

Relevant information on performance design objectives applicable to such switching nodes can be found in Recommendation Q.543.

8.5.6 SCP dimensioning

The main SCP functionalities should be described and an SCP model developed. The SCP distribution (co-located with STP or stand alone) and functional architecture (centralized or distributed) also have to be considered.

As the physical entity of the network with IN capabilities where the service logic is performed – receiving SSP requests and determining destination numbers, formulating and sending responses to SSPs, processing billing information and statistics, sending, transferring or receiving data from the Service Management Function (SMF), processing SDP queries, etc., – the SCP has an important role in the overall network performance.

The SCP must be dimensioned to meet requirements similar to those for the SSP:

– termination capacity to connect the number of signalling links required;

– storage capacity sufficient for the service logic, call data and database volume; and

– processing capacity sufficient for the expected load.

Generally, an SCP will be processing-limited and its capacity given in terms of the number of queries (or calls) per second that can be processed while meeting the response delay objectives. In assessing SCP performance, a model can be useful – in particular, in considering the functional architecture of the specific SCP (centralized or distributed).

(It is worth noting here that prior to dimensioning a given SCP, the overall SCF functionality requirements of the network must be assessed and a reasonable distribution of service control logic made to determine the number and location of the SCPs to be deployed in the network.)

8.5.7 SDP dimensioning

The main tasks of SDPs concern information retrieval requests or information update requests, i.e. searching tasks and index building for data updates, additions or deletions.

In assessing an SDP's query handling capacity, it is usually necessary to analyse its internal architecture. Although the internal physical architecture may differ from one manufacturer to another, an SDP, in general, contains a set of memory devices used by processors to work in parallel. The memory is divided into data memory (for storage of databases), working memory (for processors to handle requests), cache memory (to speed up access to slower memory devices) and queueing memory (to store incoming requests and outgoing responses). The memory and processor units are connected by data links and data buses, which can be high speed connections or internal buses in a large parallel machine. Buses and directs links are dimensioned proportionally to the number of units (memory, processors) they interconnect. One key aspect influencing performance is what caching techniques are being utilized – the performance of caching algorithms is typically dependent on the load distribution of query types.

The objective of SDP dimensioning is to allocate adequate system resources to handle the query load while satisfying a constraint on total delay between the arrival of a request and response departure. An important performance consideration is the caching mechanism, which can affect both delay performance and system capacity. Some optimization can be achieved by a suitable distribution of data through the storage devices.

Input for the dimensioning process is the service usage data (the information accessed during the service logic execution) and the service execution information (the sequence in which the information is accessed).

The dimensioning procedure for the SDP signalling functionalities is analogous to the dimensioning of other signalling connections. The dimensioning of the SDP functionalities refers mainly to processing power and to memory size and speed.

The function of the SDP is to retrieve data from a database. It is reasonable to dimension the processing capacity for the expected load, in terms of the queries to be processed per unit time. There may, however, be problems of data contention (locking, deadlock) and resource contention (input/output channel) to which special attention must be given.

In dimensioning SDPs, it must be kept in mind that the SDP must accommodate not only the query load but also the data-update load.

8.5.8 IP dimensioning

SCPs can request an SSP to connect a user to a resource located in an IP, connected to the SSP from which the request is detected or to another SSP. SRFs may be integrated to IPs in a centralized configuration or, optionally, be mapped to SSPs in a distributed configuration. The different alternatives have to be considered for IP dimensioning.

The IN operations on the IP are user interactions which may impact the network in different ways depending on factors such as the number of announcements to be played, the amount of information to be collected, the type of information received from the users, etc. The IP must be able to handle the expected mix of these operations and the circuit traffic load. In addition, it must support the processing load required to establish and release connections to the users and carry out the processing and logic functions required on each call. The IP must have sufficient processing capacity to handle the expected processing load.

Methods for dimensioning devices and other equipment needed for the SRF within an IN should consider both the SRFs placed in special intelligent peripherals and integrated within the SSPs. When SRF is integrated within SSPs this also has an impact on the SSP dimensioning. The IP dimensioning procedure is carried out to derive the capacity of an IP to deal with a given load/service demand or to provide input data for the general IN optimization.

IP dimensioning input data include IP traffic load of circuit-switched connections and signalling messages which derive from the service demand and service characteristics, internal architecture of the IP, processor capacity, GOS requirements, and cost factors.

Output data from the dimensioning procedure specify the configuration (number of units for each relevant component such as terminating switched groups, terminating signalling links, control components, etc.), the estimates of blocking probabilities and delays as well as the costs associated to the found configuration.

Delay calculations for signalling toward SSPs and for signalling toward SCPs have to be provided. The signalling to establish connections between IPs and SSPs is similar to the signalling for other parts of the circuit-switched network. The cases when the SSP works in a relay mode (transit messages between IP and SCP) are to be considered. Dimensioning criteria are related to set-up and release times. Reliability requirements can be stated. Normal and failure conditions are considered when dimensioning signalling links as well as the different signalling protocol mechanisms which depend on the internal architecture of the IP.

When the SRF is integrated to an SSP, dimensioning is similar to the case of SRF in IP with the exception that there is no need for special circuit groups or signalling links between them. Instead, the integrated SRF within an SSP adds load to signalling and voice traffic links and to the central processor of the SSP. Additional input data are specified such as extra traffic load and load of signalling messages to/from SRF, extra processor load, GOS requirements, and costs factors for SRF devices. Output data include the number of devices of different types, delay and blocking probability for the found configuration.

8.5.9 Other IN specific element dimensioning

There may be other elements to be dimensioned in a particular IN-based network. For such elements – such as non-SS No. 7 signalling links – special dimensioning methods might have to be applied.

9 Re-dimensioning procedure

The process described in the preceding clauses may be used to plan and dimension network elements. However, in practice, traffic demand may fluctuate sufficiently that service to some customers or to some calls may be subject to delay or blockage, while extra capacity exists elsewhere in the network. To address this, the re-dimensioning function provides for reviewing, on a short time scale, such as a week, of the allocation assumptions used in the previous calculations. In this way a revised result, showing different circuit or link quantities or modified allocations rules, may be used to achieve better network throughput as traffic demand varies.

9.1 Re-dimensioning for introduction of a new IN service

One of the aims of Intelligent Network planning is to be able to introduce a new service or respond to a new demand for an existing service in a short delay (typically in a period of time between 3 to 6 months) compared with the periods usually considered short-term for planning or programming activities (1 to 3 years) for general network planning. The problem is then to analyse whether the new traffic can be handled by the existing network and to check whether this traffic could degrade the quality of already existing services.

Re-dimensioning presents opportunities for allowing the introduction of new services without requiring significant investment. Re-dimensioning activities identify potential overloads that the introduction of new subscribers or a new service may cause. The possible problems are identified at the general network level (PSTN and signalling network) as well as at the level of individual IN components, but the solutions should be as local as possible – i.e., performed only on the elements affected by the currently detected bottlenecks, which may be exchanges or circuit groups. Furthermore, re-dimensioning solutions typically do not require significant capital investment and should not adversely impact existing traffic.

Figure 7 shows a typical re-dimensioning procedure.

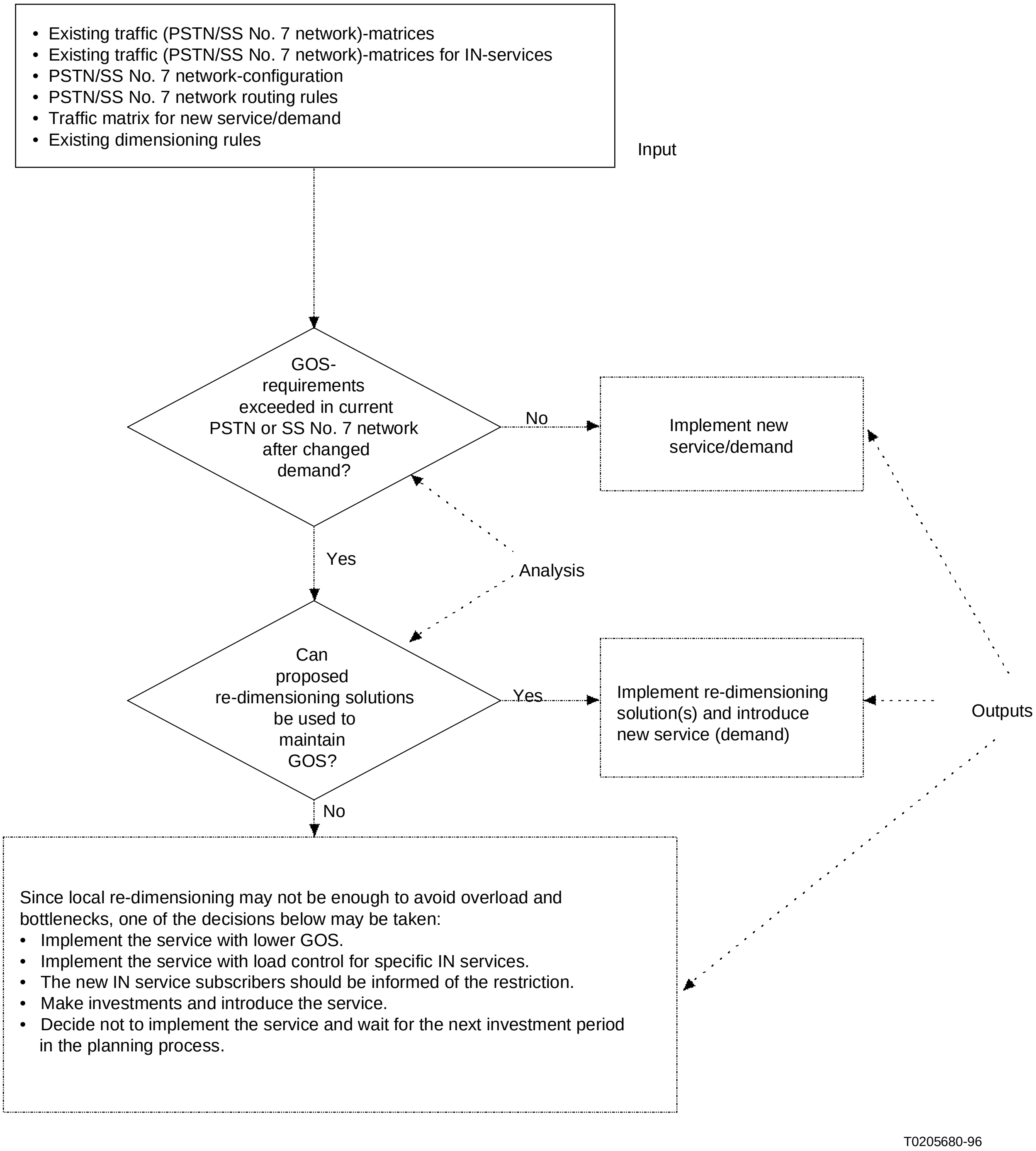

FIGURE 7/E.734

Re-dimensioning procedure

The main parts of the network to be considered when assessing the impact of additional loads from a new IN service demand are:

- exchanges and circuit groups in the underlying PSTN;
- nodes and links in the underlying signalling network;
- IN entities such as SSPs, SCPs, IPs and SDPs.

9.2 Impact on the PSTN

Once the load of the new IN service has been evaluated, an analysis (often called admissibility) has to be carried out to check that the new traffic is compatible with the capacity of the existing or planned network. If there is not sufficient capacity, this activity identifies the network elements where blocking or overload may occur unless action is taken.

The admissibility steps are illustrated in Figure 8.

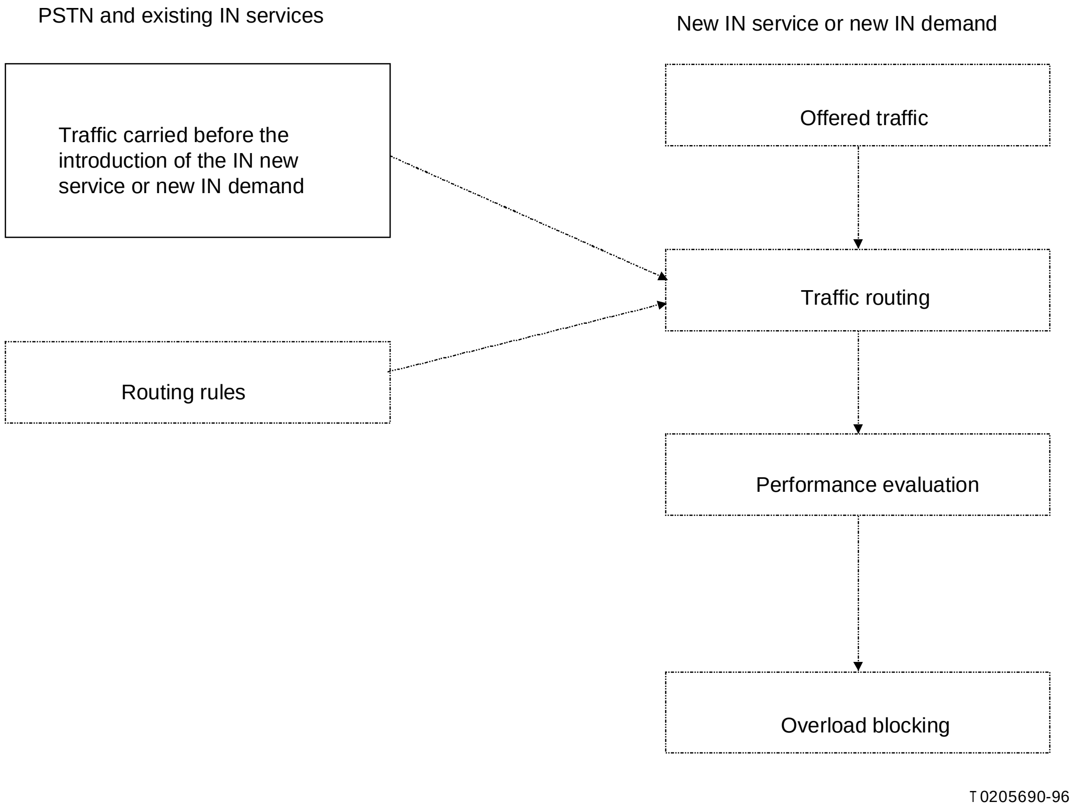

FIGURE 8/E.734

Admissibility analysis

Assuming potential blocking or overloads have been identified, the next stage consists of the implementation of local solutions, some of which are listed here:

- Traffic model re-evaluation: if overload situations are found, a traffic re-evaluation could be necessary, particularly on circuit groups.
- Unsaturation by using circuit group capacities: assuming the admissibility process has detected bottlenecks only on circuit groups, a solution could be the use of the marginal capacity of the circuit groups. One possibility is to check whether there exists capacity on free channels in transmission systems. Another possibility is to use the free capacity of another circuit group with the same extremities but carrying another service.
- Unsaturation by rerouting: if the admissibility process has pointed out problems on transit exchanges, it is possible to modify the load sharing coefficients to obtain a better distribution of the traffic.

9.3 Impact on the signalling network

Increased signalling traffic may lead to bottlenecks on signalling nodes and signalling links. Performance re-evaluation and detection of possible overloads and bottlenecks has to be done, considering the signalling traffic generated by the new IN service. If necessary, a modification of routing rules can be envisaged.

The re-evaluation concerns signalling link loads, SP and STP global load and internal loads as well as the loads of IN physical entities. A queuing analysis/simulation of delays will be of assistance in detecting the main problems.

It is important to determine what kind of traffic and services the links are carrying and derive the length and the number of signalling messages. Measurements on the current links should be carried out.

Signalling equipment rearrangements, involving signalling link set rearrangements or signalling terminal distribution rearrangements, may contribute to resolving bottlenecks. However, care has to be taken to respect engineering rules or physical constraints.

- Load sharing can be another solution, particularly for STP overload situations.
- Network node controlled traffic management is another solution, reducing traffic to avoid overloads of IN resources or the underlying network – a solution applicable mainly for SCP overload.
- Service logic controlled traffic management could be another solution, since some mechanisms influence the traffic which depends on service subscriber specified parameters. This control could result in a limitation of calls to some destinations, a limitation of resources, or service filtering for some customers.

9.4 Impact on the network architecture

Another way to alleviate traffic overload could be modifying the IN architecture. For instance, the implementation of SSF at local exchange level may lead to a better traffic distribution and to shorter connection establishment delays than in the situation where SSF is implemented at the transit exchange level. This is a preventive action, concerning the architecture, usually to be considered in the programming process. If the delay to introduce SSF in all local exchanges may appear too long due to the normal engineering requirements, the situation can perhaps be resolved by implementing SSF only in a limited number of local exchanges.

10 History

This is the first issue of Recommendaiton E.734.

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