INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

E.529

(05/97)

SERIES E: OVERALL NETWORK OPERATION, TELEPHONE SERVICE, SERVICE OPERATION AND HUMAN FACTORS

Quality of service, network management and traffic engineering – Traffic engineering – Determination of the number of circuits in automatic and semi-automatic operation

---

Network dimensioning using end-to-end GOS objectives

ITU-T Recommendation E.529

(Previously CCITT Recommendation)

ITU-T E-SERIES RECOMMENDATIONS

OVERALL NETWORK OPERATION, TELEPHONE SERVICE, SERVICE OPERATION AND HUMAN FACTORS

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.529

NETWORK DIMENSIONING USING END-TO-END GOS OBJECTIVES

Summary

This Recommendation contains the network dimensioning guidelines for circuit-switched networks using end-to-end GOS objectives. The dimensioning methods for networks with fixed traffic routing, time-dependent traffic routing, state-dependent traffic routing and event-dependent traffic routing are described.

Source

ITU-T Recommendation E.529 was prepared by ITU-T Study Group 2 (1997-2000) and was approved under the WTSC Resolution No. 1 procedure on the 26th of May 1997.

FOREWORD

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

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

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

NOTE

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

INTELLECTUAL PROPERTY RIGHTS

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

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

CONTENTS

NETWORK DIMENSIONING USING END-TO-END GOS OBJECTIVES

(Geneva, 1997)

1 Scope of this Recommendation

This Recommendation is intended to provide network dimensioning guidelines for circuit-switched networks enabling the network operator to meet end-to-end GOS objectives from the user-plane perspective. Traffic engineering methods for network dimensioning will be presented in this Recommendation. Network dimensioning guidelines from the control-plane perspective are treated in other E-Series Recommendations. Transmission network design is beyond the scope of this Recommendation.

2 References

The following ITU-T Recommendations and 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.

– CCITT Recommendation E.170 (1992), *Traffic routing*.

– ITU-T Recommendation E.301 (1993), *Impact of non-voice applications on the telephone network*.

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

– CCITT Recommendation E.500 (1992), *Traffic intensity measurement principles*.

– ITU-T Recommendation E.501 (1997), *Estimation of traffic offered in the network*.

– CCITT Recommendation E.523 (1988), *Standard traffic profiles for international traffic streams*.

– CCITT Recommendation E.524 (1992), *Overflow approximations for non-random inputs*.

– CCITT Recommendation E.525 (1992), *Designing networks to control grade of service*.

– ITU-T Recommendation E.526 (1993), *Dimensioning a circuit group with multi-slot bearer services and no overflow inputs*.

– ITU-T Recommendation E.527 (1995), *Dimensioning at a circuit group with multi-slot bearer services and overflow traffic*.

– CCITT Recommendation E.721 (1991), *Network grade of service parameters and target values for circuit-switched services in the evolving ISDN*.

3 Terms and definitions

This Recommendation defines the following terms.

3.1 **node**: A switching centre or a hypothetical switching centre representing a network.

3.2 **circuit group**: A group of circuits which is engineered as a unit. See Figure 3-1.

3.3 **traffic stream**: A class of calls with the same traffic characteristics.

3.4 **O-D pair**: An originating node to destination node pair for a given traffic stream.

3.5 **end-to-end GOS**: Overall GOS of a traffic stream quantified between its O-D pair defined in the network concerned.

3.6 **route**: A concatenation of circuit groups providing a connection between an O-D pair. See Figure 3-1.

3.7 **route set**: A set of routes connecting the same O-D pair. See Figure 3-1.

3.8 **routing pattern**: A route set and rules to select one route out of the set for a traffic stream.

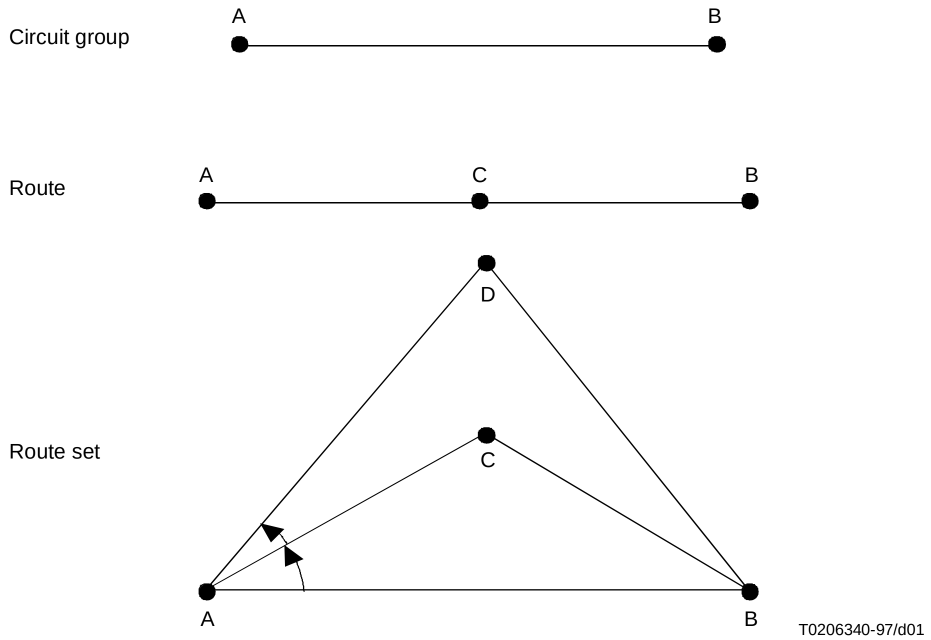


Figure 3-1/E.529 – Terminology

4 Abbreviations

This Recommendation uses the following abbreviations:

AAR	Automatic Alternative Routing
ARR	Automatic Rerouting
ECCS	Economic Hundred Call Seconds
EDR	Event-Dependent Routing
GOS	Grade of Service
LLR	Least Loaded Routing
N-ISDN	Narrow-band Integrated Services Digital Network
PSTN	Public Switched Telephone Network
SDR	State-Dependent Routing
TDR	Time-Dependent Routing

5 Introduction

PSTN and N-ISDN support switched point-to-point bearer connections and provide on-demand bearer services. In such a circuit-switched network, a call is routed, according to traffic routing rules used in the network, from its originating node to destination node in the network.

Routing structures, routing schemes, and route selection for traffic routing are described in Recommendation E.170. There are various traffic routing methods available, from fixed to flexible methods. The traffic routing methods, together with network capacity and other traffic control mechanisms, such as service protection methods, provide an end-to-end GOS for each traffic stream in the network. Recommendation E.525 describes available service protection methods and provides guidelines for applications of the methods.

The purpose of this Recommendation is to present network dimensioning methods that are applicable to networks with routing and traffic control methods. This Recommendation focuses on end-to-end call blocking probability as the GOS parameter that is taken into account in the network dimensioning process, and on traffic modelling techniques and methods available for computing end-to-end call blocking probabilities. Further, this Recommendation focuses on non-hierarchical routing, considering that end-to-end GOS criteria are usually applied in non-hierarchical routing, irrespective of network architectures. The definition of hierarchical and non-hierarchical routing is given in 2.1/E.170.

This issue of this Recommendation gives example dimensioning methods for a network in which a single routing scheme is used. Dimensioning methods for a network in which multiple routing schemes are interworked are for further study.

5.1 Overview of network operation

For network engineering purposes, a network that provides a circuit-switched service is often represented by a network model which comprises switching nodes interconnected by circuit groups. A circuit group is of either the one-way or both-way type. Figure 5-1 gives an example of a network. Figure 5-2 illustrates a model for the overall architecture of network operations. The central box represents a network, which can have various configurations, and the traffic routing tables within the network. Network configurations include metropolitan area networks, national intercity networks, and global international networks, which support both hierarchical and non-hierarchical routing and mixes of the two. Routing and network management ensure that performance objectives are met under all conditions including load shifts and failures, and network servicing and network forecasting ensure that network dimensioning meets performance objectives at minimum cost. Figure 5-2 illustrates network management, network servicing and network forecasting as interacting feedback loops around the network. The input driving the network is a stochastic traffic load, consisting of predictable average demand components added to unknown forecast error and load variation components. The feedback controls regulate the grade of service provided by the network through capacity and routing adjustments.

Network management provides for real-time monitoring of network performance through collection and display of real-time traffic and performance data, and allows traffic controls, such as code blocking, call gapping, and re-route controls, to be inserted when circumstances warrant. Network management also allows routing management which takes account of the capacity provided by network servicing and network forecasting, and on a real-time basis adjusts routing patterns as necessary to correct service problems. The updated routing patterns are sent to the switching nodes either directly or via an automated routing management system.

Network forecasting accounts for both the current network and the forecast loads in planning network changes, and then network servicing makes routing and circuit group size adjustments, if network performance under the actual loads becomes unacceptable because of errors in the forecast. Network dimensioning methods are embedded mainly in both network forecasting and network servicing procedures so that an appropriate level of circuit group capacity is determined and implemented through the network servicing process. The major objective of network dimensioning using end-to-end GOS objectives is to determine network capacity, call routing patterns, and other traffic control parameters, such as circuit reservation to meet given end-to-end GOS requirements, for a given set of traffic load assumptions in an economical way. Under exceptional circumstances, circuit group capacity can be added in network servicing on a short-term basis to alleviate service problems, but is normally planned, scheduled, and managed over a period of one year or more. The network forecasting strategy, which is performed in the forecasting step, minimizes reserve capacity while maintaining an acceptable level of network servicing. The model in Figure 5-2 illustrates network forecasting as a process which predicts the required network capacity to meet the future demand. It considers both the demand forecast, which is subject to error, and the existing network. In dealing with forecast uncertainties, network forecasting provides sufficient network capacity to meet these demands with a minimum of network servicing. Annex A describes example methods for network forecasting and network servicing for dynamic routing networks.

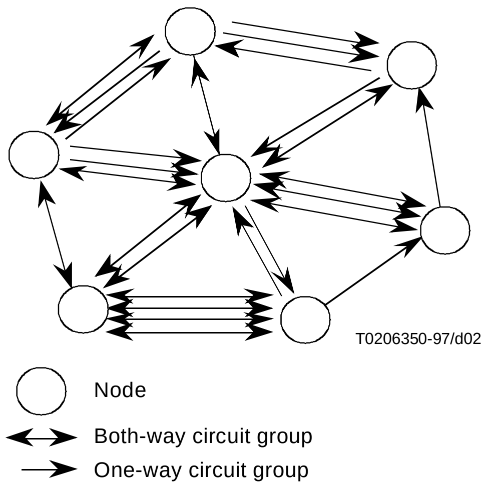


Figure 5-1/E.529 – Network example

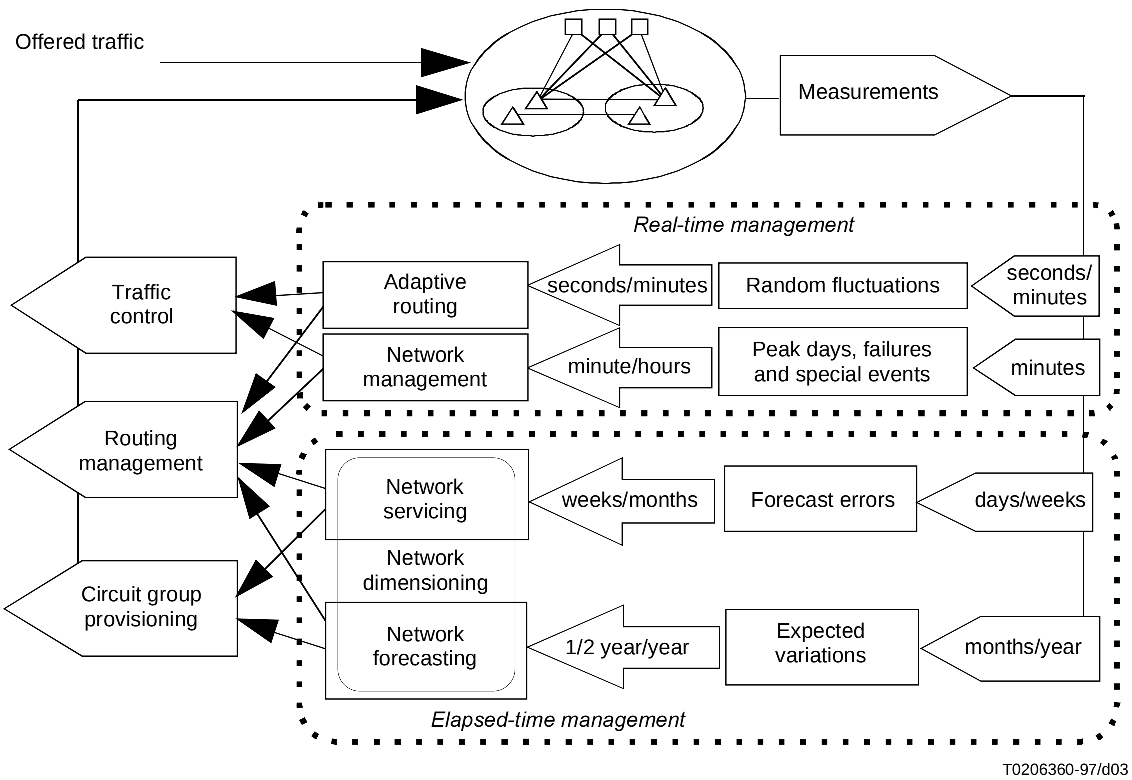


Figure 5-2/E.529 – Overall architecture of network operation

5.2 Outline of this Recommendation

This Recommendation treats network dimensioning of logical networks, and is therefore concerned with the network forecasting process loop. Noting that network topology may be determined by taking account of various factors that include network provider's policies, determination of the topology of the logical network is beyond the scope of this Recommendation. Design of physical networks including transmission network dimensioning, assignment of logical circuit groups to the transmission network, and transmission route assignment to transmission facilities are beyond the scope of this Recommendation. An overview and categorization of routing methods for traffic engineering purposes are given in clause 6. Network dimensioning elements and dimensioning problems are described in clause 7. The end-to-end GOS objectives and their use are described in clause 8. Guidelines for establishing the traffic load assumption are described in clause 9. Principles of network dimensioning methods are presented in clause 10. Methods for computing end-to-end call blocking probabilities of the various types of network are provided in clauses 11 to 14. Annex A describes example methods for network forecasting and network servicing for dynamic routing networks. Annexes B to D provide example methods for network dimensioning.

6 Categories of routing methods

In this clause, categories of routing methods and example methods in each category are described. The categorization and description of routing methods in this Recommendation are based upon Recommendation E.170. However, routing methods described in this Recommendation are only for purposes of traffic engineering, and are intended to capture features of the routing methods that have significant impact on network dimensioning. An exhaustive list of all routing methods or specification of particular routing methods is not intended.

Routing patterns describe the route set choices and route selection rules for each O-D pair, for a connection request for a particular service. Routing patterns can be hierarchical, non-hierarchical, fixed, and dynamic, and are used for each of a multiplicity of services on the telecommunications network [CH1]. A call in a network consists of a pair of symmetric point-to-point unidirectional communication channels in forward and backward directions, respectively. One or more traffic streams may be defined for an O-D pair in the network. In a network **with progressive call control**, a call is sent from one node to another node, and control of the call is also passed to the next node. A route from one node to the next node is one or more circuit groups providing a connection between the two nodes. A node will maintain a route set for a traffic stream. A particular routing pattern may be applied to a particular traffic stream or class of traffic streams. A route is a concatenation of circuit groups providing a connection between an O-D pair. As a result of routing, a call will be established on a route or be blocked due to lack of resources in the network. In a network where an originating node can select a route for a traffic stream, such as a network with **originating call control**, a set of routes and a route selection rule will be maintained at the originating node. The route set and route selection rule for a traffic stream constitute a routing pattern for the traffic stream. If **crankback** or Automatic Rerouting (ARR) is used at a transit node, the preceding node may maintain control of calls even if the calls are blocked at all the routes of the transit node. By using a multi-circuit group crankback capability, an originating node can maintain control of calls and test all the possible routes to the destination with fixed sequences of routes, e.g. [AS1].

6.1 Fixed routing methods

In a fixed routing method, a routing pattern is fixed for a traffic stream. A fixed routing network is a network in which a routing pattern is fixed for every traffic stream.

6.1.1 Fixed sequential alternative routing

This routing method is based on fixed sequences of routes. The set of routes and selection sequence are determined on a preplanned basis. Automatic Alternative Routing (AAR) is a particular type of progressive call control or originating call control, where if a node has more than one route to the next node, it may select a route according to a fixed route selection sequence. Typically, if a direct route to the next node exists, the route selection rule may first test the availability of the direct route for a newly arriving call. When sufficient capacity is not available for a call at the last route in the sequence, the call is blocked and lost. Fixed sequential alternate routing can incorporate crankback or ARR at a transit node, so that the preceding node may maintain control of calls even if the calls are blocked at all the routes of the transit node, or a multi-circuit group crankback capability, so that an originating node can maintain control of calls.

6.1.2 Fixed load sharing routing

This routing method is based on a fixed set of proportions each of which indicates a proportion of calls routed to a corresponding route. For instance, a route may be assigned to a newly arriving call randomly, according to the route selection probability reflecting the proportions. The set of routes and the proportions are determined on a preplanned basis. The simplest case is that only one route in the route set is tested for a newly arriving call and, when the idle capacity of the route is insufficient, the call is blocked. In this routing method, either progressive call control or originating call control can be used, e.g. [GAR].

6.2 Time-dependent routing methods

Time-dependent routing methods are a type of dynamic routing method such that routing patterns will be altered at a fixed time during the day (or week). These time-dependent routing patterns are determined on a preplanned basis and will be implemented consistently over a time period by all nodes in the network.

6.2.1 Time-dependent sequential alternative routing

In this routing method, a sequence of routes is predetermined for each of the time periods of a day (or week), and calls are routed based on the fixed route sequence for each time period. The route sequences of time periods may be different from each other. Crankback may be used in the network for originating call control. Calls are blocked when a final choice route for the calls is found busy.

6.2.2 Time-dependent load sharing routing

In this routing method, route selection proportions are predetermined for each of the time periods of a day (or week), and calls are routed based on the fixed proportion for each time period. In this routing method, either progressive call control or originating call control can be used.

6.3 State-dependent routing methods

In state-dependent routing methods the routing patterns will vary automatically according to the state of the network. For a given state-dependent routing method, an algorithm to determine the routing patterns in response to changing network status is preplanned and used over a long time period. Information on network status may be collected at a central processor or distributed to switching nodes in the network. The information exchange may be performed on a periodical or on-demand basis. In state-dependent routing methods either progressive call control or originating call control can be used. Typically in cases where progressive call control is used, calls are blocked at a transit node where a circuit to the destination node is unavailable.

All state-dependent routing methods use the principle of routing calls on the best available route on the basis of network state information, often using a Least Loaded Routing (LLR) method, as now described. In the LLR routing method residual capacity of the routes for respective traffic streams is calculated, and the route with the largest residual capacity in the route set for the traffic stream is selected for a newly arriving call. If any route with sufficient free circuits is not available, the call is blocked. This route selection rule may be applied only for alternative route set for calls overflowing their first choice direct circuit group [AS3], [CAM], [CH2].

In state-dependent least cost routing a well-defined cost of the routes for respective traffic streams is calculated, based on various factors such as congestion state of circuit groups. For a newly arriving call, the route whose cost is minimal in the route set for the traffic stream and below a given threshold is selected. If such a route is not found, the call is blocked

[KRI]. In state-dependent sequential alternate routing the residual capacity of circuit groups is measured periodically, for example based on five-minute traffic measurements, and the sequence of alternative routes for a traffic stream is updated based on the estimated residual capacity [GAU]. Typically, the direct circuit group is the first choice, if any.

6.4 Event-dependent routing methods

In this routing method the routing patterns will be updated locally on the basis of whether calls succeed or fail on a given choice. In the event-dependent routing method a fresh call is offered first to its direct route, when it exists, and is always routed to that route if there is a free circuit. Otherwise, overflow from the direct circuit group is offered to a currently selected alternative route. If a call fails to be established at the alternative route, the call is blocked and, further, the alternative route is reselected by choosing a route at random from a set of available alternative routes for the traffic stream. The alternative route can be updated randomly, as in the **sticky random routing** method used in dynamic alternative routing [KEY], or cyclically, in which case the next route in the routing table is selected. The alternative route is not reselected if the call is successfully routed on either the direct route or other alternative routes.

There are various alternative rules to reselect a route. An alternative route set for a traffic stream may be changed in a time-dependent manner considering the time-variation of traffic load. This approach is implemented in state and time-dependent routing [INO], [KAW], which combines EDR and TDR methods.

7 Network dimensioning elements

At the initial stage of the network dimensioning process, an appropriate network model, as shown in Figure 5-1, will be constructed. The network model to be studied will be determined based on various factors such as an existing network, the requirement on precision of network dimensioning and policies of network operators. For instance, to find appropriate network configurations and capture overall capacity requirements, a simple network model may be established by representing a part of network by a single node and representing the entire capacity between two adjacent nodes, or parts of the network, by a single circuit group. In a model of the international network, multi-gateway configurations in a country may be reflected in this way, and the tie circuit groups between the gateways may be included as a network element, while the sizes of those tie circuit groups may be treated as a design variable depending on the network dimensioning practice.

As an essential input data for network dimensioning, the traffic demand model will correspond to the derived network model, and the traffic load assumptions will be provided. This must be specified for every defined traffic stream of every O-D pair in the network. The traffic load assumptions will be derived based upon traffic measurements, forecasting, etc. When a network model represents a part of a network, transit traffic streams traversing the part may be taken into account in the network dimensioning practice, and thus are modelled as an appropriate traffic stream in the network model.

In a derived network model, a routing method must be specified for every traffic stream. In this Recommendation, however, it is assumed that a routing method for every traffic stream is predetermined and not changed in the network dimensioning process.

Further, the crankback capability of each node and the type of service protection method applied at each circuit group are assumed to be pre-specified for each of the network elements during the construction of the network model.

In this context, design variables to be determined in the network dimensioning process using end-to-end GOS objectives are as follows:

1) size of every circuit group in the network;

2) routing pattern for every traffic stream;

3) traffic control parameters, such as circuit reservation levels [AKI], [STA] on each circuit group. (Recommendation E.525 provides guidelines for service protection methods and Recommendation E.412 provides information regarding traffic control methods.)

Regarding routing patterns, a set of routes and a rule to select one route out of the route set must be determined at every node for every traffic stream, including the originating traffic streams and transit traffic streams for the node. Figure 7-1 provides a minimum set of inputs and outputs of the network dimensioning process.

The objective of network dimensioning in this Recommendation is to determine the design variables for a given traffic load assumption to meet a given end-to-end GOS requirement in the most economical way. A simple example is to minimize the total number of circuits in the network whilst satisfying a given GOS objective. More detailed approaches take into account costs and revenues in the dimensioning process [LE1].

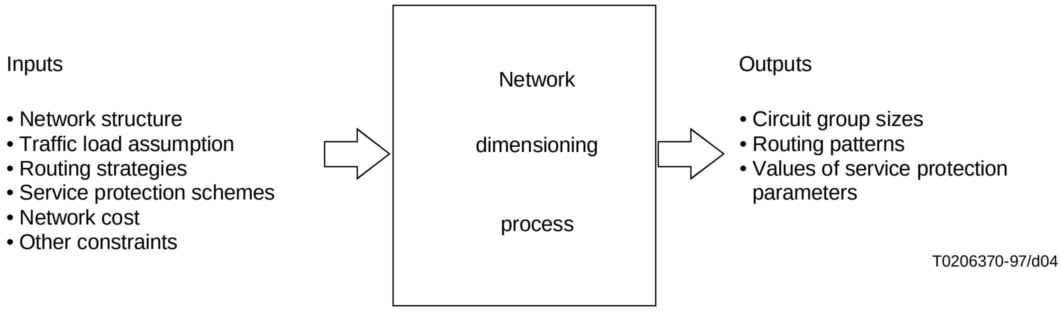


Figure 7-1/E.529 – Inputs and outputs of network dimensioning process

8 End-to-end GOS objectives

The definition of end-to-end blocking probability is given in Recommendation E.721. In this Recommendation, call blocking due to the lack of user plane resources is focused upon and the blocking due to lack of control plane resources is outside the scope of this Recommendation. Recommendation E.721 provides the standard target values of end-to-end call blocking probabilities for local, toll and international connections.

In network dimensioning, objectives for end-to-end call blocking probabilities for various traffic load conditions as well as for normal conditions may be specified. In this case, the objectives may be used simultaneously as constraints to find the appropriate network capacity which meets all objectives in the most economical manner.

The definition of "end-to-end" in international, national transit and metropolitan networks is given in Table 8-1.

Table 8-1/E.529 – Definition of "end-to-end"

Network	Definition
International	Gateway-to-gateway
National transit	Transit node-to-transit node
Metropolitan	Subscriber node-to-subscriber node

9 Modelling of offered traffic

To establish a traffic load assumption for network dimensioning, the traffic demand model must be specified. The following are typical elements of the traffic demand model.

9.1 Traffic variables

The most important traffic variable to represent a traffic stream is a traffic intensity expressed in erlangs. The peakedness factor of a traffic stream may be included as a traffic variable when appropriate. In N-ISDN, the bandwidth of a call in a traffic stream will be an important traffic variable.

9.2 Models for variations in traffic load

In network dimensioning, estimated traffic load variations such as local and global traffic surges may be taken into account in order to dimension a network capable of handling the traffic variations with appropriate end-to-end GOS levels. The concept of normal load and high load is provided in Recommendation E.500, and the concept is taken into consideration in Recommendation E.721 to specify end-to-end GOS target values for circuit-switched services in the evolving N-ISDN.

Table 9-1 compares traffic routing strategies with respect to:

a) design to accommodate traffic load variations;

b) traffic representation for network dimensioning;

c) characterization of traffic route selection; and

d) characterization of traffic routing variation [CH3].

Table 9-1/E.529 – Comparison of traffic routing strategies

Routing method	Traffic variations in the design model	Traffic representation	Traffic routing	Traffic routing variation
Fixed traffic routing	None	Single traffic matrix	Single matrix of ordered routes	None
Time-dependent traffic routing	Predictable (seasonal) (hourly)	Several traffic matrices	Several matrices of ordered routes	Predetermined
State-dependent traffic routing	Instantaneous (expected) (random)	Several traffic matrices	Single matrix of available routes	Determined in real-time
Event-dependent traffic routing	Instantaneous (expected) (random)	Several traffic matrices	Single matrix of available routes	Determined in real-time

These aspects of traffic routing strategies are now described further. Traffic load variations can be categorized as:

a) within-the-hour variations;

b) hour-to-hour variations;

c) day-to-day variations; and

d) seasonal variations.

For each node pair for a given hour, the load is modelled as a stationary random process characterized by a fixed mean and variance (and therefore a fixed value of peakedness = variance/mean). From hour-to-hour, the mean loads are modelled as changing deterministically, for example according to their 20-day average values. From day-to-day, for a fixed hour, the mean load can be modelled as a random variable having a gamma distribution with a mean equal to the

20-day average load. From season-to-season, the load variation can be modelled as a time varying deterministic process in the network dimensioning procedure. The random component of the realized seasonal load is the forecast error (equal to the forecast load minus the realized load). Forecast error is accounted for in the network servicing process. Table 9-2 summarizes the types of models used to represent the different traffic variations under consideration:

Table 9-2/E.529 – Traffic models for load variations

Traffic variations	Traffic model
Within-the-hour	Random
Hour-to-hour	Deterministic
Day-to-day	Random
Seasonal	Deterministic

Peakedness methods within the network dimensioning procedure account for the mean and variance of the within-the-hour variations of the offered and overflow loads. As one component of the network dimensioning procedure, Wilkinson's equivalent random method [WIL] can be used to size circuit groups for these two parameters of load. Other methods for modelling within-the-hour traffic variations for sizing of circuit groups is discussed in Recommendation E.524.

Routing pattern design embedded in the network dimensioning method accounts for the hour-to-hour variations of the load. When it is significant to take into account variation of traffic during a day, a 24-hour traffic profile will be prepared for a traffic stream. Recommendation E.523 presents a standard 24-hour voice traffic profile in the international telephone network. Note that non-voice traffic often has a 24-hour traffic profile different from voice traffic, especially in the network where time zone differences result in the peaks of voice and non-voice traffic occurring at different times. Recommendation E.301 presents a standard 24-hour non-voice traffic profile. The hourly traffic profiles represent non-coincidence of busy hours of traffic streams in the network. To obtain circuit reduction taking advantage of non-coincidence of busy hours in the network, the hourly or appropriately specified traffic profiles should be studied.

Accommodating day-to-day variations in the network dimensioning procedure can use an equivalent load technique that models each node pair in the network as an equivalent circuit group engineered to the GOS objective. On the basis of Neal-Wilkinson engineering [HIL], [WIL], the number of circuits N that are required in the equivalent circuit group to meet the required GOS for the forecast load R, with its specified peakedness Z and specified level of day-to-day variation φ, is given as follows:

$$N = \mathrm{TRKRQS}(R,\varphi,Z,\mathrm{GOS}) \qquad (9\text{-}1)$$

where GOS is the required end-to-end call blocking probability, and TRKRQS is a function mapping R, φ, Z, and GOS into the circuit group requirement N.

Holding fixed the specified peakedness Z and the calculated circuits N, we calculate what larger load r requires N circuits to meet the GOS objective by the following equation, if the forecast load had no day-to-day variation:

$$N = \mathrm{TRKRQS}(r,0,Z,\mathrm{GOS}) \qquad (9\text{-}2)$$

where $\varphi = 0$ signifies no day-to-day variation. The equivalent load r then produces the same equivalent number of circuits N when engineered for the same peakedness level but in the absence of day-to-day variation.

The network dimensioning and routing management process accommodate the random components of seasonal variations (i.e. forecast errors) in the procedures. When some realized end-to-end blockings are found to be larger than their objective values, additional circuits and/or routing changes are provided to restore the network blocking to the objective level. Circuits are often not disconnected in the network servicing process, even when load forecast errors are such that this would be possible without service degradation. As a result, the process leaves the network with a certain amount of reserve or idle capacity, even when the forecast error is unbiased [FRA].

In principle, it is possible to take into account multiple traffic matrices, each of which reflects a possible combination of traffic variations of the O-D pairs, to dimension a network or to check if a resultant network has sufficient performance against traffic variations.

10 Principles of network dimensioning methods

We summarize the network dimensioning methods discussed in this clause in Table 10-1. The routing methods listed are provided as examples of routing methods and are not specifically recommended.

Table 10-1/E.529 – Network dimensioning methods

Routing method	Examples	Network dimensioning method
Fixed traffic routing	Fixed hierarchical Fixed non-hierarchical	High usage circuit group-Truitt ECCS [TRU] Final circuit group-Neal/Wilkinson [HIL], [WIL] Unified design model [AS1], [AS2], etc.
Time-dependent traffic routing	Dynamic non-hierarchical routing, etc.	Multi-hour engineering [EIS] Multiperiod stochastic model [KAS] Unified design model [AS1], [AS2], [KRI], etc.
State-dependent traffic routing	Dynamically controlled routing real-time network routing dynamic routing – five minutes, etc.	DCR design model [REG], [BNR] Unified design model [AS4] Traffic flow model [AS5], etc.
Event-dependent traffic routing	Dynamic alternative routing state- and time-dependent routing, etc.	DAR design model [KEY] STR design model [KAW], etc.

10.1 Basic iterative procedure

To find an optimal set of network circuit group sizes, routing patterns and traffic control parameters, an iterative procedure is often used. A basic iterative procedure, as illustrated in Figure 10-1, is as follows:

Step 1) Set an initial set of circuit group sizes, routing patterns and traffic control parameters of the network.

Step 2) Compute end-to-end blocking probabilities for all traffic streams to check whether given GOS objectives are all satisfied.

Step 3) Modify the current circuit group sizes, routing patterns, and traffic control parameters in an appropriate way until an optimal solution that satisfies the GOS objectives is reached.

In this network dimensioning procedure, Step 2) plays an important role from the viewpoint of traffic engineering. To compute end-to-end blocking probabilities, an analysis of the state space of the network may be performed and the end-to-end call blocking probabilities may be derived from the state probability distribution. Note that, in a simple case, the state probability distribution of network may have a product-form solution. In general, however, the product-form solution of the network is not always available, and therefore, approximate methods for computing end-to-end blocking probabilities are usually used. A common principle of the methods available and widely used in practice is the technique of the fixed point model (or the reduced load approximation). In what follows, the principle of the fixed point model is described, and example applications of the principle are presented, for example call routing methods in clause 6 above. All the methods are intended for network dimensioning applications.

It is often assumed in network dimensioning that there is no call blocking in the nodes, and the call set-up time is negligible.

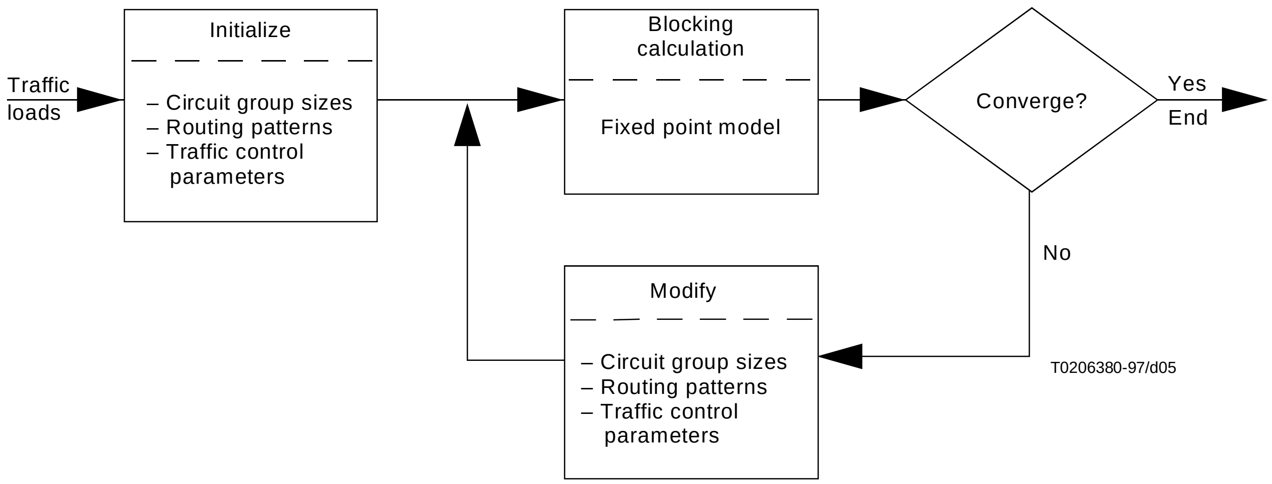


Figure 10-1/E.529 – Three-step basic iterative procedure

10.2 Principle of the fixed point model

The fixed point model, e.g. [KAT], [KE1], [WHI] and [WON] is a type of system decomposition technique which yields an approximation of network performance values and is usually used for network dimensioning problems in traffic engineering. The principle of the technique is to decompose the network into individual circuit groups, analyse the circuit groups separately and then derive the target performance measures of the network from the performance measures of the components, that is, the circuit groups. The technique is based upon the two assumptions.

The first assumption is the circuit group independence assumption. It is assumed that the circuit groups in the network are statistically independent, that is, statistical events such as call acceptance and rejection on a circuit group occur independently of the other circuit groups in the network. Based on this assumption, the network is decomposed into hypothetically independent circuit groups.

The second assumption is that calls arrive at a decomposed circuit group with a hypothetical arrival rate. The hypothetical rate is derived from the traffic characteristics of the calls, traffic characteristics of other circuit groups (e.g. call blocking probabilities of circuit groups, state distributions of circuit groups) and routing patterns. The traffic characteristics of each circuit group are analysed by using the derived hypothetical arrival rate of calls associated with the circuit group. In some cases, the hypothetical traffic offered to the circuit groups may be modelled by a state-dependent arrival rate. Peakedness factors of overflows from a circuit group may be computed and used as a parameter characterizing the hypothetical traffic offered to other circuit groups.

Since traffic characteristics of the circuit groups are related to each other by way of the hypothetical traffic streams, an iterative procedure is used to compute the traffic characteristics of the circuit groups. There may be various alternatives of numerical computation techniques to perform the iteration procedure. They may have different characteristics with respect to convergence. This point, however, is outside the scope of this Recommendation. It should be noted that convergence of the iterative procedure is not guaranteed, although it will occur in many practical cases.

The computation methods presented as examples in the following subclauses are based on the technique of the fixed point model.

10.3 Use of implied cost and shadow price for network dimensioning

10.3.1 The notions of implied cost and shadow price

Traffic engineering makes use of circuit groups indicators such as occupancy, blocking or rejected traffic. The optimization methods introduce indicators such as circuit group marginal overflow, the overflow variation corresponding to a capacity variation. The importance of other indicators has been shown: the implied cost and shadow price [KE2].

These costs measure the effect of a variation in the offered traffic and capacity of a circuit group on the grade of service as regards the whole network. The calculation of implied cost and shadow price assumes that a value is allocated to each traffic unit in erlangs offered to each traffic stream, and this value can either directly represent revenues that are collected by serving the traffic offered in the network, or virtual revenue which integrates the grade of service constraints. The implied cost of a circuit group is then defined as the variation of the total revenue carried by the network due to an infinitesimal variation of the offered traffic to the circuit group. The shadow price is defined as the variation of the total revenue carried by the network due to an infinitesimal variation of the size of a circuit group and an infinitesimal variation of the value of a traffic control parameter such as circuit reservation. This derivative information is useful in a network optimization process for determining the circuit group sizes with a given routing method, in terms of maximizing the defined total network revenue.

Results [KE2] have shown that the implied costs and shadow prices are the solution of a system of linear equations and that their calculation on all the circuit groups in a network is of a similar complexity to the one needed for calculating blocking. This means they can be used for management and operational dimensioning.

Implied costs and shadow prices are indicators that are particularly suited to network servicing and network dimensioning. The shadow prices indicate the circuit groups for which it is most beneficial to increase capacity, and those where capacity can be reduced without disturbing network performance, taking into account all the effects of these variations. To determine how much the capacity should vary, extensions to Kelly's results have been developed to take into account modularity augmentation, which yield the direct comparisons between investment costs and traffic gain on a circuit group basis [TIB]. This method can also be used either for circuit group extension studies in network servicing or within the optimization algorithms. The network structure is immaterial and intervenes only in the calculation of implied costs and shadow prices, and not in their use for dimensioning. The method can process a traffic demand expressed in multi-matrix form.

It has been shown that the implied costs and shadow prices, without having been introduced beforehand, are relevant to the solution of the problem of maximizing the difference between revenue and investment costs, under GOS constraints. The advantage of this approach is to link formally the revenue associated with the flows and GOS constraints [GI1].

The implied costs and shadow prices can also be used to optimize the load sharing ratio in the context of adaptive routing. They then intervene according to two different time scales: in the algorithm for calculating in real-time load sharing parameters for routing control, and in the network dimensioning algorithm required for network servicing and planning [CHI].

10.3.2 Iterative procedure using implied costs and shadow prices

Bounds on the acceptable grade of service can be imposed by choosing an appropriate setting of unit revenue for unit carried traffic for streams in an integrated manner in the revenue maximization approach. Specifically, the following gives a basic procedure for the purpose of imposing the grade of service bounds [KEY]:

1) Dimension the network with the GOS constraints absent.

2) If the GOS conditions are violated by particular streams, set the unit revenue of these streams to 1 and the rest to 0.

3) Augment the network with the resulting new shadow prices until the GOS constraints are all met.

According to this method, the basic iterative procedure for network dimensioning described in 10.1 is modified as follows: in Step 2) computation of the end-to-end call blocking probabilities and also the implied costs and shadow prices are performed all together, and Step 3) involves iterating through the procedure outlined in 1) to 3) above.

11 Network dimensioning methods for fixed routing networks

For purposes of dimensioning a fixed routing network, a busy hour traffic load assumption will be given, and GOS objectives will be specified corresponding to the busy hour traffic load. Recommendation E.501 provides the principle for determining busy hour traffic load. If there exist significant non-coincidences of busy hours among traffic streams in the network, 24-hour traffic profiles of traffic streams will be derived based on traffic measurements and used in network dimensioning, while the routing patterns are fixed. When 24-hour traffic profiles are used, end-to-end call blocking probabilities will be computed for every hour, except for the hours in which the traffic load is insignificant for network dimensioning.

11.1 Blocking calculation methods

11.1.1 Blocking calculation method for fixed sequential alternative routing

Consider a network in which a fixed sequential alternative routing method stated in 6.2 is used at every node, and in which the nodes may have crankback capability. For simplicity of presentation, we assume that the direct route is used as the first choice and only two-circuit group routes are used as an alternative route. Extension of the method presented below to the more general case of multiple-circuit group routes is straightforward. Further, every call is assumed to be established using a single circuit on each circuit group along the route of the call. The following notation is used:

L	=	the number of circuit groups.
S	=	the number of traffic streams.
Nj	=	the number of circuits on circuit group j.
route sp	=	the p-th route for traffic stream s.
Ps	=	the number of routes available for traffic stream s.
circuit group spk	=	the k-th circuit group of route p for traffic stream s; k = 1 or 2 since only 1-circuit group or 2-circuit group routes are considered (the extension to n-circuit group route is straightforward).
a(s)	=	the offered traffic intensity for traffic stream s.
B(sj)	=	blocking probability of traffic stream s at circuit group j; each route circuit group spk corresponds to a unique network circuit group j (j = 1, 2, ..., L), but each network circuit group j may correspond to a number of route circuit groups spk. This relationship is denoted by a mapping X(spk) = j. Thus, B(sj) = B(sX(spk)) = B(spk).
RB(sp)	=	blocking probability of traffic stream s at its p-th route.
EEB(s)	=	end-to-end blocking probability of traffic stream s.

Based on the circuit group independence assumption, the blocking probability at a route with multiple circuit groups in tandem is expressed by the circuit group blocking probabilities. The following equation holds:

$$RB(sp) = 1 - \prod_{k=1,2} (1 - B(spk)), \text{ for all s, p} \qquad (11\text{-}1)$$

Case 1) For a traffic stream for which no crankback is used at any node, the end-to-end blocking probability is expressed as:

$$EEB(s) = 1 - \sum_{p=1}^{Ps} (1 - RB(sp)) \prod B(sh1), \text{ for all s} \qquad (11\text{-}2)$$

Case 2) For a traffic stream for which crankback is used at every transit node, the end-to-end blocking probability is expressed as:

$$EEB(s) = \prod_{p=1}^{Ps} RB(sp), \text{ for all s} \qquad (11\text{-}3)$$

Case 3) For a traffic stream for which crankback is used partially, EEB(s) is expressed in a mixed manner between the above two extremes.

A general formulation for end-to-end blocking probabilities can be found, for example, in [CHA].

In all cases, the circuit group blocking probabilities of individual traffic streams are calculated by using the following function:

$$B(sj) = f_{sj} (N_j, w(1j), w(2j), ..., w(sj)), \text{ for all s, j} \tag{11-4}$$

where w(sj) denotes the hypothetical traffic of traffic stream s offered to circuit group j. $f_{sj}()$ denotes a function that determines the call blocking probability of the traffic stream s at the circuit group j with size of N_j taking into account a service protection method, if applied. In a special case, the function $f_{sj}()$ may represent the Erlang B formula.

The hypothetical traffic w(sj) is often determined by a function of the offered traffic a(s), i = 1, 2, ..., R and circuit group blocking probabilities B(sj), j = 1, 2, ..., L. Therefore, the equation 11-4 gives a recursive relation of the circuit group blocking probabilities. For example, in Case 2) above, the hypothetical traffic w(spk) is expressed as:

$$w(spk) = a(s)(1 - RB(sp)) \prod_{h=1}^{p-1} RB(sh) / (1 - B(spk)), \text{ for all s, j(= X(spk))} \tag{11-5}$$

Equations 11-4 and 11-5 compose the system of fixed point equations for (B, w).

Note that overflow from a route to the succeeding route may be a peaked traffic stream, and traffic carried by a circuit group and then offered to the succeeding circuit group in the route may be a smoothed traffic stream. To characterize the hypothetical traffic that represents such a peaked or smoothed traffic stream, peakedness factor of the traffic stream may be taken into account. Recommendation E.524 presents approximate methods for calculating call blocking probabilities of a circuit group with mixed random and non-random traffic inputs. A method for computing the peakedness factors and higher moments is available for a large class of network models: e.g. [OD1]. An iterative procedure may be applied to compute the call blocking probabilities and peakedness factors of overflow streams at circuit groups: e.g. [PIO], OD2]. Note that the complexity in the computation of peakedness factors is sometimes prohibitive for dimensioning a large network.

11.1.2 Blocking calculation method for fixed load sharing routing

Consider a network in which a fixed proportional routing method is used at every node. Assume that a call is offered first to its direct circuit group and, if no circuit is available at the direct circuit group, the call is routed to an alternative route randomly, with route selection probabilities at the originating node of the call. To compute end-to-end call blocking probabilities, the same principle of the above method for fixed alternative routing can be applied.

The following additional notations are used:

p = 1 = the direct circuit of a traffic stream.

p > 1 = the alternative route p of a traffic stream.

p(sp) = the proportion that the p-th route is selected for a call in stream s where p = 2, 3, ..., P_s.

a(sp) = the offered traffic intensity of traffic stream s to p-th route where p = 2, 3, ..., P_s.

The offered traffic intensity a(sp) is given by a(sp) = R(sp)a(s)LB(sp = 1) where LB(sp = 1) is equivalent to the circuit group blocking probability for traffic stream s on the direct circuit group. The circuit group blocking probabilities B(sj) and thus the route blocking probabilities RB(sp), p = 2, 3, ..., P_s can be computed by the same equations as in 11-4 and 11-1, respectively. The end-to-end call blocking probability of traffic stream s, EEB(s), is given by:

$$EEB(s) = RB(sp = 1) \sum_{p=1}^{P_s} R(sp) \; RB(sp) \tag{11-6}$$

11.2 Determination of design variables

Basically, iterative procedures may be used as in classical non-linear optimization, e.g. [TRU].

12 Network dimensioning methods for time-dependent routing networks

In the dimensioning of a network with time-dependent routing, time-varying traffic for every traffic stream will be provided. Let T and t (= 1, 2, ..., T) denote the number of time periods and the t-th time period of a day (or week), respectively. In order to simplify the network dimensioning procedure, it is often assumed that the network in each time period is statistically independent of the other time periods, and that the network is in steady-state during each time period.

12.1 Blocking calculation methods

For dimensioning a network in which a time-dependent sequential alternative routing method is used at every node, the end-to-end call blocking probabilities at every time period may be computed by applying the method stated in 11.1.1 for each of the time periods. In the case where a time-dependent proportional routing method is used at every node, the same method as stated in 11.1.1 above may be applied to respective time periods.

12.2 Determination of design variables

Annex B describes an example method for the determination of design variables for time-dependent routing networks.

13 Network dimensioning methods for state-dependent routing networks

13.1 Blocking calculation methods

13.1.1 Blocking calculation method for least loaded routing networks

Consider a network in which the direct route is used as the first choice, only two-circuit group routes are used as an alternative route to carry overflow from the direct route and the LLR method is used at every node for its originating calls. Note that, in this routing arrangement, a transit node routes incoming transit calls to the direct route to their destination.

A fixed point model which uses hypothetical state-dependent arrival rates is available and found, for example, in [CHU]. In this method, the hypothetical arrival rate of a traffic stream to a circuit group is allowed to vary with the number of occupied circuits on the circuit group. The rate is calculated by computing the probability that a call in the stream is routed to the circuit group, based on the computed state distribution of the relevant circuit groups in the network. The end-to-end call blocking probabilities are computed from the approximate stationary distribution of the network obtained after convergence of the iterative procedure.

A method of fixed point model with hypothetical proportional routing parameters is also available for a network with LLR [GI2], [CH4]. In this method, the hypothetical stationary proportional routing parameter is introduced, that is, the proportion R(sp) that a call in stream s overflowing its direct circuit group is routed to an alternative route p. The proportion R(sp) is calculated by regarding it as being proportional to the mean residual capacity on the different routes of the traffic stream and computing the mean residual capacity with an iterative procedure. The end-to-end call blocking probabilities are computed as in Formula 11-6 from the approximate route blocking probabilities obtained after convergence of the iterative procedure.

13.1.2 Blocking calculation method for state-dependent least cost routing networks

Consider a network in which only one- or two-circuit group routes are used to carry a call and the state-dependent least cost routing method is used at every node for its originating calls. The cost is assumed to be computed based on the measured number of busy circuits in each circuit group. For this network, a method using hypothetical proportional routing parameters is available [KOU]. The route selection proportions R(sp) are computed as the probability that the route has the least cost. The end-to-end call blocking probabilities are computed as in Formula 11-6 from the approximate route blocking probabilities obtained after convergence of the iterative procedure.

13.2 Determination of design variables

Basically, iterative procedures may be used as in classical non-linear optimization. Annex C provides an example of design methods for a state-dependent routing network [LE2].

14 Network dimensioning methods for event-dependent routing networks

14.1 Blocking calculation methods

Consider a network in which the event-dependent sticky routing method with random reselection is used at every node and only two-circuit group routes are used as an alternative route. The blocking calculation method based on the fixed point model is available for this network [GIB]. In the method, the long-term proportion R(sp) of overflow of traffic stream s offered to route p is considered, and is computed as being proportional to the mean length of the sequence of alternatively routed calls offered to the route p. The mean length of the sequence is computed by using the route blocking RB(sp) expressed in Formula 11-1. The proportions and blocking probabilities are calculated with an iterative procedure. The end-to-end call blocking probabilities are computed as in Formula 11-6 from the approximate route blocking probabilities obtained after the convergence of the iterative procedure.

14.2 Determination of design variables

Basically, iterative procedures may be used as in classical non-linear optimization. Annex D provides an example of design methods for an event-dependent network.

15 History

This is the first issue.

16 Bibliography

[AKI] AKINPELU (J.M.): The overload performance of engineered networks with non-hierarchical and hierarchical routing, *Bell System Technical Journal*, Vol. 63, 1984.

[AS1] ASH (G.R.), CARDWELL (R.H.), MURRAY (R.P.): Design and optimization of networks with dynamic routing, *Bell System Technical Journal*, Vol. 60, No. 8, 1981.

[AS2] ASH (G.R.): Traffic network routing, control, and design for the ISDN era, *Fifth ITC Seminar Traffic Engineering for ISDN Design and Planning*, Lake Como, Italy, 1987.

[AS3] ASH (G.R.), CHEN (J.-S.), FREY (A.E.), HUANG (B.D.): Real-time network routing in a dynamic class-of-service network, *Proceedings of International Teletraffic Congress*, ITC-13, Copenhagen, 1991.

[AS4] ASH (G.R.), CHANG (F.): Transport network design of integrated networks with real-time dynamic routing, *Journal of Network Systems and Management*, Vol. 1, No. 4, 1993.

[AS5] ASH (G.R.), HUANG (B.D.): An analytical model for adaptive routing networks, *IEEE Trans. on Communications*, Vol. 41, No. 11, 1993.

[BNR] BNR, DCR Design Model, *Proceedings of Teletraffic Congress*, ITC-11, Kyoto, Japan, 1985.

[CAM] CAMERON (W.H.), REGNIER (J.), GALLOY (P.), SAVOIE (A.M.): Dynamic routing for intercity telephone networks, *Proceedings of International Teletraffic Congress*, ITC-10, 3.2.3, Montreal, Canada, 1983.

[CHA] CHAN (W.S.): Recursive algorithms for computing end-to-end blocking in a network with arbitrary routing plan, *IEEE Trans. on Commun.*, Vol. COM-28, No. 2, 1980.

[CH1] CHEMOUIL (P.), FILIPIAK (J.), GAUTHIER (P.): Analysis and control of traffic routing in circuit-switched telephone networks, *Comp. Networks and ISDN*, Vol. 11, No. 3, 1986.

[CH2] CHEMOUIL (P.), FILIPIAK (J.), GAUTHIER (P.): Performance issues in the design of dynamically controlled circuit-switched networks, *Special issue of IEEE Communications Magazine*, Vol. 28, No. 10, 1990.

[CH3] CHEMOUIL (P.), GAUTHIER (P.): Adaptive traffic routing in telephone networks, *l'Echo des Recherches*, English version, 1990.

[CH4] CHEMOUIL (P.), LEBOURGES (M.), GAUTHIER (P.): Performance analysis of adaptive traffic routing in a metropolitan network: a case study, *Globecom'89*, Dallas, United States, 1989.

[CHI] CHIFFLET (J.), IKEUCHI (N.), LANSARD (P.D.): MICAN: an algorithm to dimension switching networks with adaptive routing using implied costs methodology, *Proceedings of International Teletraffic Congress*, ITC-13, Copenhagen, Denmark, 1991.

[CHU] CHUNG (S.-P.), KASHPER (A.), ROSS (K.W.): Computing approximate blocking probabilities for large loss networks with state-dependent routing, *IEEE Trans. on Networking*, Vol. 1, 1993.

[EIS] EISENBERG (M.): Engineering traffic networks for more than one busy hour, *Proceedings of International Teletraffic Congress*, ITC- 9, Melbourne, Australia, 1979.

[FRA] FRANKS (R.L.), HEFFES (H.), HOLTZMAN (J.M.), HORING (S.), MESSERLI (E.J.): A model relating measurements and forecast errors to the provisioning of direct final trunk groups, *Bell System Technical Journal*, Vol. 58, No. 2, 1979.

[GAU] GAUTHIER (P.), CHEMOUIL (P.): A system for testing adaptive traffic routing in France, *Globecom'87*, 23.4, Tokyo, Japan, 1987.

[GAR] GARCIA (J.-M.), HENNET (J.-C.), TITLI (A.): Optimization of routing in interurban telephone networks, *Large Scale Systems Journal*, Vol. 2, No. 3, 1981.

[GIB] GIBBENS (R.J.), KELLY (F.P.): Dynamic routing in fully connected networks, *IMA J. of Math. Cont. and Info.*, Vol. 7, pp. 77-111, 1990.

[GI1] GIRARD (A.), LIAU (B.): Dimensioning of adaptive routing networks, *IEEE/ACM Trans. on Networking*, Vol. 1, No. 4, pp. 460-468, 1993.

[GI2] GIRARD (A.), BELL (M.-A.): Blocking evaluation for networks with residual capacity adaptive routing, *IEEE Trans. on Commu.*, Vol. 37, No. 12, 1989.

[HIL] HILL (D.W.), NEAL (S.R.): The traffic capacity of a probability engineered trunk group, *Bell System Technical Journal*, Vol. 55, No. 7, 1976.

[HUB] HUBERMAN (R.), HURTUBISE (S.), LE NIR (A.), DRWIEGA (T.): Multi-hour dimensioning for a dynamically routed network, *Proceedings of International Teletraffic Congress*, ITC-11, Tokyo, Japan, 1985.

[INO] INOUE (A.), MASE (K.), YAMAMOTO (H.), SUYAMA (M.): A state- and time-dependent dynamic routing method for telephone networks, *Proceedings of International Teletraffic Congress,* ITC-13, pp.195-200, Copenhagen, Denmark, 1991.

[KAW] KAWASHIMA, (K.), INOUE, (A.): State- and time-dependent routing in the NTT network, *IEEE Communications Magazine*, 1995.

[KAS] KASHPER (A.N.), SCHMITT (J.A.): Multiperiod stochastic design model for international networks, *Proceedings of International Teletraffic Congress*, ITC-13, Copenhagen, Denmark, 1991.

[KAT] KATZ (S.S.): Statistical performance analysis of a switched communication network, *Proceedings of International Teletraffic Congress*, ITC-5, New York, United States, 1967.

[KE1] KELLY (F.P.): Fixed point models of loss networks, *J. Austral. Math. Soc. Ser. B 31*, pp. 204-218, 1989.

[KE2] KELLY (F.P.): Routing in circuit-switched networks: optimization, shadow prices and decentralization, *Adv. Appl. Prob.*, Vol. 20, 1987.

[KEY] KEY (P.B.), WHITEHEAD (M.J.): Cost-effective use of networks employing dynamic alternative routing, *Proceedings of International Teletraffic Congress*, ITC-12, Turin, Italy, 1988.

[KRI] KRISHNAN, (K.R.), OTT, (T.J.): Forward-looking routing: a new state-dependent routing method, *Proceedings of International Teletraffic Congress*, ITC-12, Turin, Italy, 1988.

[KOU] KOUSSOULAS (N.): Performance analysis of circuit-switched networks with state-dependent routing, *IEEE Trans. on Commu.*, Vol. 41, No. 11, 1993.

[KNE] KNEPLEY (J.E.): Minimum cost design for circuit-switched networks, *Technical Note Numbers 36-73, Defense Communications Engineering Centre, System Engineering Facility*, Reston, VA, 1973.

[LE1] LEBOURGES (M.), BONNET (D.), PETIT (D.): Acheminement par régulation économique dans les réseaux multiservices, *l'Echo des Recherches*, No. 161, 1995.

[LE2] LEBOURGES (M.): Les modèles de trafic et leur rôle dans la gestion du réseau téléphonique commuté, PhD Thesis, Paris VI University, France, 1992.

[OD1] ODA (T.): Moment analysis for traffic associated with Markovian queueing systems, *IEEE Trans. on Commun.*, Vol. 39, No. 5, 1991.

[OD2] ODA (T.), FUKUOKA, (H.), WATANABE, (Y.): Robust routing and design of international corporate networks, *Proceedings of International Teletraffic Congress*, ITC-13, Copenhagen, Denmark, 1991.

[PIO] PIORO (M.), WALLSTROM (B.): Multi-hour optimization of non-hierarchical circuit switched communication networks with sequential routing, *Proceedings of International Teletraffic Congress*, ITC-11, Kyoto, Japan, 1985.

[REG] REGNIER (J.), BLONDEAU (P.), CAMERON (W.H.): Grade of service of a dynamic call-routing system, *Proceedings of International Teletraffic Congress*, ITC-10, Montreal, Canada, 1983.

[STA] STACEY (R.R.), SONGHURST (D.J.): Dynamic alternative routing in the British Telecom trunk network, *Proceedings of the International Switching Symposium,* Phoenix, Arizona, United States, 1987.

[TIB] TIBAS (K.), LEBOURGES (M.): Implied costs behaviour and optimal modular dimensioning, *Proceedings of International Teletraffic Congress*, ITC-13, Copenhagen, Denmark, 1991.

[TRU] TRUITT (C.J.): Traffic engineering techniques for determining trunk requirements in alternate routed networks, *Bell System Technical Journal*, Vol. 31, No. 2, 1954.

[WIL] WILKINSON (R.I.): Theories of toll traffic engineering in the United States, *Bell System Technical Journal*, Vol. 35, No. 6, 1956.

[WHI] WHITT (W.): Blocking when service is required from several facilities simultaneously, *AT&T Tech. J.*, Vol. 64, No. 8, 1985.

[WON] WONG (E.W.M.), YUM (T.-S.): Maximum free circuit routing in circuit-switched networks, *Proceedings of IEEE INFOCOM ’90*, 1990.

Annex A

Network forecasting and network servicing methods for dynamic routing networks

A.1 Introduction

This Annex describes example methods for network forecasting and network servicing for dynamic routing networks. The flow diagram of Figure A.1 illustrates designing a network on the basis of forecast loads. Network forecasting accounts for both the current network and the forecast loads in planning network changes, and then network servicing makes routing and circuit group size adjustments, if network performance under the actual loads becomes unacceptable because of errors in the forecast. The network forecasting strategy minimizes reserve capacity while maintaining an acceptable level of network servicing. The model in Figure A.1 assumes that network forecasting is a process which predicts the required network capacity to meet the future demand. It considers both the demand forecast, which is subject to error, and the existing network. In dealing with forecast errors, network forecasting provides sufficient network capacity to meet these demands with a minimum of network servicing.

A procedure called incremental network forecasting designs a minimum-cost augmentation of the existing network, with allowance for disconnects, to meet the forecast loads. Modifications to the routing design and size network steps in the network design model accomplish this change, which uses the existing circuit group capacities as lower bounds on the designed circuit group capacities. In effect, the model takes into account the reluctance to remove circuits. This procedure limits the circuit group augments to those required to meet the forecast loads, achieves a lower reserve capacity, and allows circuit group disconnects of the existing circuit group capacities. This is done by deriving, for each group, an upper threshold for its size, and using this threshold to make initial adjustments to the group size prior to incremental network forecasting.

The size thresholds for a group are based upon the forecast offered loads of the corresponding node-to-node pair, R_k^h, and are determined by choosing a maximum value r_{max} for the ratio r = (circuit group size)/(direct offered load). The upper disconnect threshold is based on the forecast peak load of the direct offered load over the next two years, with some allowance for forecast error, and it corresponds to the ratio r_{max}. The limit r_{max} is chosen by examining the range of values of the ratio in a typical network design. Table A.1 shows the limits for the ratio (circuit group size)/(direct offered load), as a function of the direct offered load.

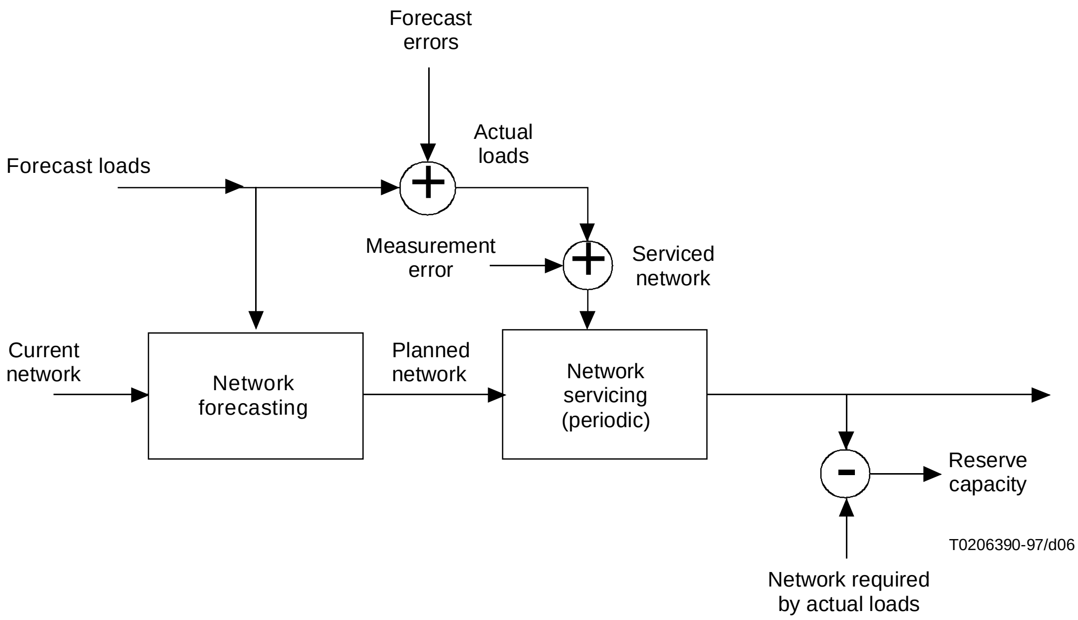


Figure A.1/E.529 – Model of the network forecasting and servicing process

The upper threshold T_{max} for a circuit group is determined in terms of the forecast loads for the corresponding direct offered load:

Table A.1/E.529 – Limit on r_{max} (R) = circuit group size/direct offered load

Load *R* (erlangs)	r_{max} (*R*)
0-5	4.5
5-10	4.5
10-25	3.0
25-50	3.0
50-100	2.5
> 100	1.5

Let:

R_i = peak forecast load for the direct offered load in year i, i = 1, 2.

β_i = forecast uncertainty factor for year i, i = 1, 2, introduced to allow for probable error in the load forecast. Typical values in actual use are $\beta_1 = 1.15$ and $\beta_2 = 1.3$, corresponding to 0.15 coefficient of variation in the forecast.

Then:

$$T_{max} = \max\left[r_{max}(\beta_1 R_1)\beta_1 R_1, r_{max}(\beta_2 R_2)\beta_2 R_2\right] \quad \text{(A-1)}$$

where r_{max} is the appropriate limit established for the ratio (T/R), such as those in Table A.1.

With these upper thresholds, we then define an initial size for each circuit group, which depends on its current size as follows: if the current size of a circuit group is below the upper threshold, its initial size equals its current size. If the current size of the circuit group is above its upper threshold, its initial size equals the upper threshold.

We use the initial network defined in this manner as the starting network for incremental dimensioning (i.e. minimum-cost augmentation) to arrive at the forecast network for each future year. Comparing the result with the current network, the augments and disconnects that must be made to implement the forecast network are determined. Under normal growth conditions, the current circuits are most often used in the initial network, not the upper circuit group limit, and the primary effect is to route traffic on the actual circuits in place, and thus, minimize re-arrangements.

The procedures for planned and network servicing involve modifications to the steps in the design model which optimize route flows and size the circuit groups to allow the initial circuit group capacities to be used as lower bounds on the circuit group capacities. Let Δa_i = capacity augmentation in carried load, on circuit group i, $i = 1, 2, ..., L$. The object is to allocate the traffic flow in each hour to the admissible routes so as to minimize the cost of the required circuit group capacity augmentations. On each circuit group, for a given number of initial circuits and the optimal circuit group blocking determined from economic considerations, there is a maximum load that can be carried on that circuit group without augmentation; this is the unaugmented initial capacity of that circuit group.

Now we outline network forecasting methods for time-dependent routing networks, state-dependent routing networks, and event-dependent routing networks.

A.2 Forecasting methods for time-dependent routing networks

The forecasting methodology for time-dependent routing networks builds on the unified design model described in Annex B. Following the unified design model approach, the flow optimization problem for network forecasting is stated as a linear program in which the decision variables are the flow assignments and the augmentations Δa_i above the existing circuit group capacities a_i (instead of total circuit group capacities as in the network forecasting problem), and the cost to be minimized is the marginal cost of augmentation:

$$\sum_{i=1}^{L} M_i \Delta a_i \quad \text{(A-2)}$$

This formulation ensures that efficient use is made of existing circuit group capacities, by means of routing changes if needed, before circuit group augmentations are proposed.

In the size network step, we are given the traffic routing and loads and find the needed augmentation to those groups that exceed their optimal circuit group blockings. This is accomplished by the following iterative procedure:

1) Begin with assumed circuit group blockings $b'^{h,i}$ (e.g. the circuit group blockings in the unaugmented network) subject to $b'^{h,i} \leq b^{max,i}$, $i = 1, ..., L$, where $b^{max,i}$ is the optimal circuit group blocking determined in the update optimal circuit group blockings step, described in Annex B.

2) Calculate the corresponding carried circuit group loads $y^{h,i}$, under the known routing and assumed circuit group blockings.

3) If for all h, $y^{h,i} \leq a_i$, the capacity of the unaugmented circuit group at its maximum blocking, then the circuit group needs no augmentation; if $y^{h,i} > a_i$, the required augmentation Δa_i is determined by dimensioning the circuit group for load $y^{h,i}$ at blocking $b^{\max,i}$.

4) From the circuit group loads $y^{h,i}$ and circuit group sizes computed in 2) and 3), we recalculate all the circuit group blocking $b^{h,i}$; if $|b^{h,i} - b'^{h,i}|$ is not sufficiently close to zero for all i in all hours h, redefine $b'^{h,i} = b^{h,i}$, $i = 1, ..., L$, $h = 1, ..., H$, and return to 2).

A.3 Forecasting methods for state-dependent and event-dependent routing networks

With state-dependent routing and event-dependent dynamic routing, the following forecasting method can be implemented:

1) Given the circuit group sizes and traffic loads, evaluate the end-to-end blocking probability given the dynamic routing method. Techniques to evaluate blocking described in this Recommendation can be used for this step. If all end-to-end blocking probabilities meet the blocking probability objective, stop.

2) Given the end-to-end blocking and traffic loads, estimate the capacity augment required for each O-D pair not meeting its blocking probability objective. Normally the augmentation is allocated to the direct circuit group of the O-D pair. Return to Step 1).

The incremental circuit-group costs and optimal circuit group blockings are, in general, determined for the forecast loads each year during network forecasting, although in some years their values might change little from the previous year.

A.4 Servicing methods

If the actual network loads exceed forecast values and cause unacceptable blocking, then quick corrective action, called network servicing, is needed. This function is performed in the servicing step. In a fixed hierarchical network, network servicing is usually limited to circuit group augmentations. However, the basic routing patterns are time variable, and hence, routing modifications can be used in network servicing to reduce network augmentation. To the extent that routing changes can be substituted for the installation of circuits, re-arrangements are also reduced.

Network servicing consists of three steps:

1) Detecting the need for network servicing, i.e. determining whether or not all offered loads are receiving adequate service.

2) If network servicing is needed, then determining the best combination of routing changes and circuit group augmentations that restore the desired blocking grade of service at minimum cost of augmentation.

3) Implementing the routing changes and circuit group augments.

End-to-end blocking measurements are needed to determine the level of service being provided and, thus, to detect the existence of service problems. Because of measurement errors and day-to-day traffic variations, such blocking measurements have an inherent statistical variability which must be allowed for by establishing acceptable bands for the measured blockings. The need for a network servicing model is for a simple procedure to determine the corrective action required. There is no attempt to redimension the whole network, or to disconnect circuits, if the network is found to be over provided for the actual network loads. The routing optimization steps described above are used to determine the optimal traffic routing for the realized loads, and the dimension network procedure to determine the circuit group augmentations required to limit circuit group blockings to their maximum permitted values. Thus, a procedure similar to the incremental network forecasting model is used to determine the required changes in network servicing.

Annex B

Example methods for time-dependent routing networks

This Annex describes example methods for determination of design variables for time-dependent routing networks.

Figure B.1 illustrates a method to compute circuit group sizes, routing patterns, and traffic control parameters for time-dependent routing networks, which is now illustrated. The initialize step estimates the optimal circuit group blocking for each circuit group in the network, based on Truitt's ECCS approach [TRU]. Within the routing design and size network modules, the current estimate of the optimal circuit group blocking is held fixed. The routing design step determines the optimal time-dependent routes for each traffic hour by executing the three steps shown within the routing design block in Figure B.1. The optimal routing is then provided to the size network and check blocking steps, which determine the number of circuits required on each circuit group to meet the same optimal circuit group blocking assumed in the router step. Once the circuit groups have been sized by the size network and check blocking steps, the cost of the network is evaluated and compared to that of the last iteration. If the network cost is still decreasing, the update optimal link blockings module computes new estimates of the optimal circuit group blockings. The new optimal circuit group blocking objectives are fed to the routing design module, which again selects optimal routing, and so on. The routing design, size network, and update optimal link blockings steps are now described further.

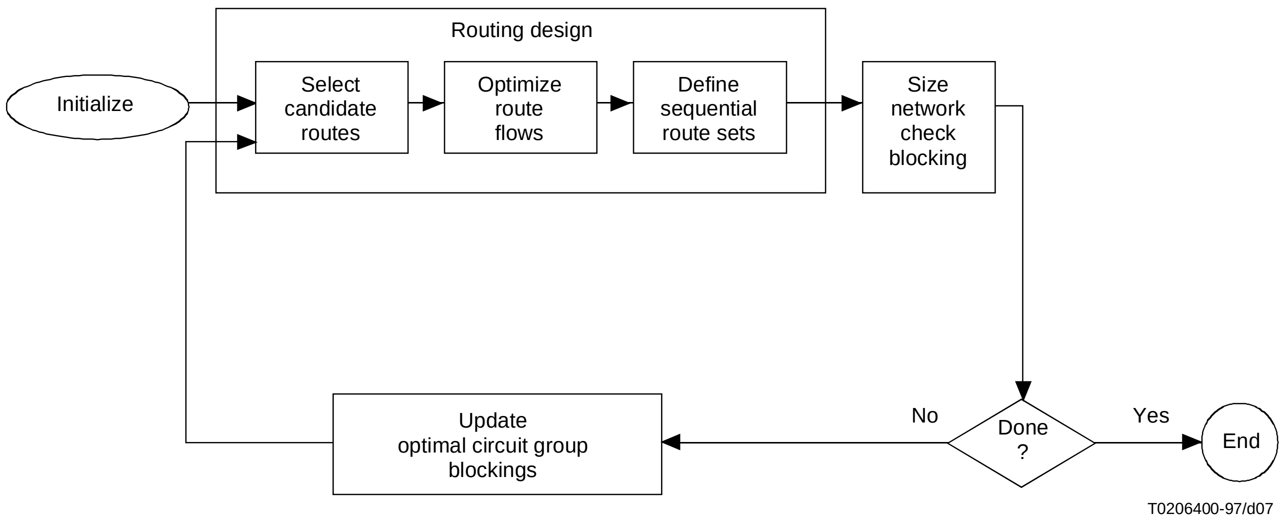
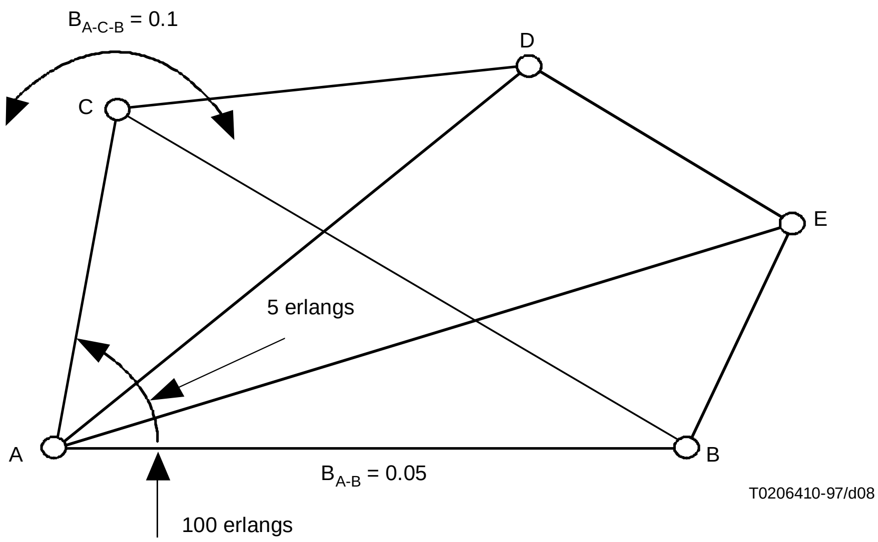


Figure B.1/E.529 – Unified design model

The routing design module consists of three steps: select candidate routes, optimize route flows, and define sequential route sets. The select candidate routes block finds the shortest (least cost) routes between nodes in the network. This step considers circuit-group cost as well as traffic non-coincidence in selecting the best candidate routes. It also considers a variety of transmission constraints such as limiting the distance of two-circuit group routes.

Assume that the select candidate routes module needs to find two candidate routes between nodes A and B and that routes A-B and A-C-B are the least costly candidates among the three possible choices illustrated in Figure B.2. The list of shortest routes, which includes routes A-B and A-C-B for nodes A and B, is then passed to the next step of the router.



Shortest routes: A-B, A-C-B
Route set blocking: $B_{A-B} \times B_{A-C-B} = 0.005$
Blocking objective: = 0.01

Figure B.2/E.529 – Example of optimization procedure

The optimize route flows step assigns flow (carried traffic) to the candidate routes to minimize network cost. The current estimate of the optimal circuit group blocking is held fixed by the routing design in the process of determining the optimal route set. Holding the blocking fixed in the optimize route flows step enables the non-linear optimal flow assignment problem to become a Linear Program (LP).

The LP to optimize route flows solves the following problem:

minimize:

$$\sum_{i=1}^{L} M_i a_i \qquad \text{(B-1)}$$

subject to:

$$\sum_{k=1}^{K} \sum_{j=1}^{J_k^h} P_{jk}^{ih} r_{jk}^{h} \le a_i, \quad i = 1, 2, \dots, L, \quad h = 1, 2, \dots, H,$$

$$\sum_{j=1}^{J_k^h} r_{jk}^{h} \le G_k^h, \quad h = 1, 2, \dots, H, \quad k = 1, 2, \dots, K,$$

$$r_{jk}^{h} \le UPBD_{jk}^{h}, \quad h = 1, 2, \dots, H, \quad k = 1, 2, \dots, K, \quad j = 1, 2, \dots, J,$$

$$r_{jk}^{h} \ge 0, \quad a_i \ge 0,$$

where we define:

L = number of circuit groups.

K = number of node pairs.

H = number of design hours.

M_i = incremental circuit-group cost in terms of dollar cost per carried erlang on.

$R^{h,k}$ = equivalent offered load to node pair k in hour h.

$A^{h,i}$ = offered load to circuit group i in hour h.

a_i = maximum carried load on circuit group i over all hours.

$b^{h,i}$ = blocking probability on circuit group in hour h.

$P_i^{h,jk}$ = 1 if route j for node pair k uses circuit group i in hour h.

= 0, otherwise.

$r^{h,jk}$ = carried load on route j for node pair k in hour h.

$J^{h,k}$ = number of routes for node pair k in hour h.

$G^{h,k}$ = total carried load for node pair k in hour h.

$UPBD^{h,jk}$ = upper bound on flow on route j of node pair k in hour h.

The total carried load for node pair k in hour h is related to the total offered load for node pair k in hour h, as follows. The minimum blocking that can be achieved on node pair k is:

$$E_k^h = \prod_{j=1}^{J_k^h} B_{jk}^h \qquad \text{(B-2)}$$

where $B^{h,jk}$ = blocking probability on route j for node pair k in hour h. Let GOS = end-to-end call blocking probability objective, and the blocking on node pair k in hour h is:

$$f_k^h = \max\left[E_k^h, GOS\right] \qquad \text{(B-3)}$$

Then:

$$G_k^h = R_k^h\left[1 - f_k^h\right] \qquad \text{(B-4)}$$

Thus, the total carried flow is determined by the blocking probability objective, unless $E^{h,k}$ is greater than this desired blocking probability objective. In this latter case, the blocking probability constraint cannot be met, and all routes are required to be at their maximum flow to minimize the blocking. A blocking correction method is then used in the check blocking step stage to correct those routes whose blockings are unacceptable.

The amount of flow that can be carried on a particular route depends on the blocking on that route and the flow assigned to all other routes comprising the particular route. For instance, in Figure B.2, it would be impossible to carry more than 95 erlangs on route A-B since B_{AB} is 0.05, and the maximum offered load is 100 erlangs. Also the upper bound of flow on route A-C-B would be 4.5 erlangs since B_{ACB} is 0.1, and, if route A-B overflows to route A-C-B, at most 5 erlangs can be offered to route A-C-B. With these upper bounds, the LP may then assign 95 erlangs of carried load to route A-B and 4 erlangs to route A-C-B. This assignment assumes an end-to-end call blocking objective of 0.01, which means that the carried load G = 99 erlangs, and this requires that 99 out of 100 erlangs be assigned to the allowed routes. The prescribed flow could then be realized if we construct a route consisting of route A-B overflowing to route A-C-B, as discussed below. The ability to realize the LP flow assignment with the routes A-B and A-C-B is ensured if we constrain the route flow optimization step to assign at most 95 erlangs to route A-B and at most 4.5 erlangs to route A-C-B. The iterative procedure for computing upper bounds, described in [AS1], estimates at each iteration the sequence of routes that best matches the current LP flow assignment. From that estimate, the model computes new upper bounds that force flow feasibility.

The traffic routing patterns are selected in the third step of the routing design block. For the sequential routing procedure, the determine sequential routes module determines the route sequences that best match the optimal route flows determined in the optimize route flows module. Returning to the example in Figure B.2, assume that the final LP flow assignment is 95 erlangs assigned to route A-B and 4 erlangs assigned to route A-C-B. The determine sequential route sets module now constructs the routes A-B and A-C-B in which all 100 erlangs are offered first to route A-B, and the 5 erlang overflow is then offered to route A-C-B. The overall route blocking achieved is $0.05 \times 0.1 = 0.005$, which is better than the blocking probability objective of 1 per cent. The output from the router block is the optimal route sequence to be used in each hour for each node pair.

After the routing design step has determined the routes, the network must be dimensioned to achieve a circuit group blocking no higher than the assumed optimal circuit group blocking used as input to the router. If the route blocking exceeds the required blocking probability objective, it is corrected in the check blocking step described below.

An iterative method is used in the size network module to determine a consistent set of hourly blockings and offered loads. The iteration uses the present estimates of the circuit group offered loads and optimal circuit group blockings to size circuit group in its peak hour and calculate blocking estimates in side hours. After all groups have been sized, we calculate new estimates of carried load for each circuit group using the blocking estimates and the routing pattern given by the router. Blocking estimates are then recalculated and the process repeated. The iteration is continued until the sum of the absolute blocking changes is less than a prescribed convergence threshold. Network dimensioning is accomplished through the use of a two-parameter traffic model. Fractional circuits are used to stablize the iterative loop and to speed convergence; modular circuit group sizes are determined in a post processor to the model. If a route in the dimensioned network exceeds the required blocking probability objective, the blocking on the first route is decreased in the check blocking step until the blocking probability is met.

After completion of the size network step, the cost of the resulting network is calculated. If the network cost is still decreasing, the update optimal link blockings module computes a new estimate of the optimal circuit group blocking using the following optimization method. Figure B.3 illustrates the cost trade off between carrying traffic on the direct circuit group between A and B, and the alternative network that overflow calls use. The problem is to find the optimal value of circuit group blocking (or, equivalently, the number of circuits) to handle the offered load at a minimum cost. Truitt's model [TRU] is used to derive the condition for the optimum based on the direct route to alternative route cost ratio and the marginal capacity of the alternative route. Truitt's method is commonly used today in both intercity and metropolitan hierarchical network dimensioning. The approach calculates the number of circuits, N^*, (and, hence, the optimal circuit group blocking b) that minimize the total cost of carrying circuit group offered load A on the direct route and the alternative routes. To find this optimal circuit group blocking, the network cost for carrying load A is first written as:

$$Cost = CN + aM = CN + AbM_S \qquad \text{(B-5)}$$

where C is the cost of a direct circuit, N is the number of circuits, a is the overflow load from circuit group AB ($a = A \times b$), and M_S is an equivalent incremental circuit-group cost for the alternate route network. A minimum cost solution is found numerically to the above equation B-5 with respect to N, from which we can then determine the optimal circuit group blocking b.

The new blockings are fed to the routing design module which again selects shortest routes, and proceeds iteratively. See Figure B.3.

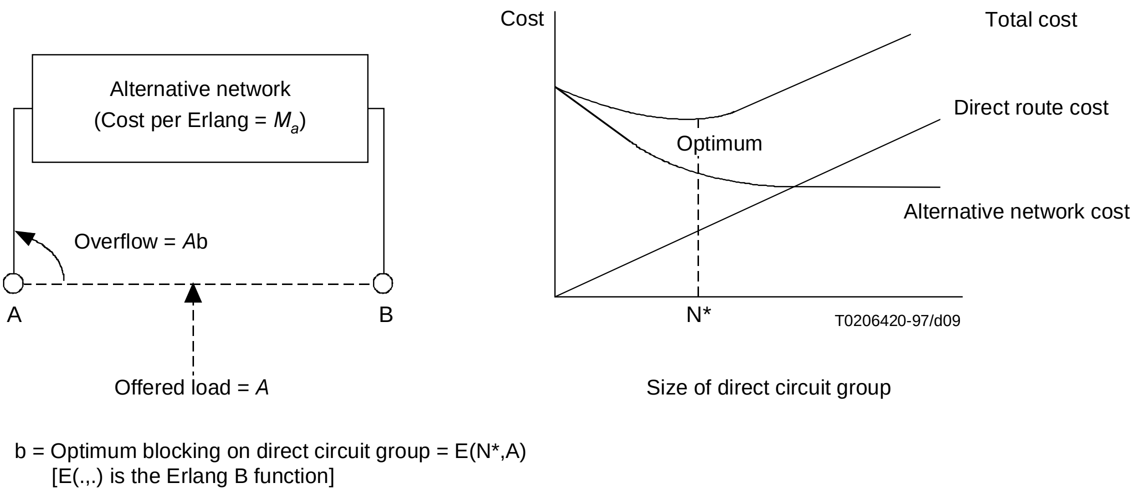


Figure B.3/E.529 – Calculation of circuit group blockings

The unified design model can be modified to be applied to state-dependent routing networks [AS4]. In this application, the unified design model is used to optimize the routing of transport flows, as measured in units of transport demands such 64 kbit/s or 1.5 Mbit/s, and the associated link capacities. For this application, mathematical equations are used to convert traffic demands to transport capacity demands, and the transport flow is then routed and optimized. This model typically assumes an underlying traffic routing structure. For example, the traffic routing model can be fixed hierarchical routing, or non-hierarchical dynamic routing.

In the transport flow optimization model application to dynamic hierarchical network design, a linear programming transport flow optimization model is used for network design. In this application there is an underlying hierarchical traffic routing. End-to-end traffic demands are converted to high usage circuit group and final circuit group node-to-node transport demands by using the hierarchical network design methods, which could include for example the ECCS approach of Truitt [TRU], Wilkinson's equivalent random method [WIL], and Neal-Wilkinson [HIL] models for final circuit group sizing. A linear programming transport flow optimization is solved for network sizing, which routes node-to-node virtual transport load demands by hour on the shortest, least cost routes, and sizes the circuit group to meet the design level of flow.

In the unified design model application to state-dependent routing network design, a linear programming transport flow optimization model is used for network design. End-to-end traffic demands are converted to node-to-node transport demands by using the ECCS approach of Truitt [TRU]. Here Truitt's approach is used to optimally divide the load between the direct circuit group and the overflow network, but in this application of the model we obtain an equivalent virtual transport demand, by hour, as opposed to an optimum circuit group blocking objective. A linear programming transport flow optimization is solved for network sizing, which routes node-to-node virtual transport load demands by hour on the shortest, least cost routes, and sizes the circuit group to meet the design level of flow. Once the circuit group has been sized, the performance of the network is evaluated and if the performance objectives are not met, further modification is made to capacity requirements. Techniques used in the unified design model application to time-dependent routing networks for optimizing traffic flows by hour, as described above, are reused in this application for optimizing virtual transport flows.

Annex C

Example method for state-dependent routing networks

This Annex describes an example method of dimensioning for state-dependent routing networks.

C.1 Network model considered

The method is applied to networks with the following structure:

– The network is composed of traffic originating-destination nodes (X) and of transit nodes (T). These transit nodes can also be origin or destination of traffic streams.

– Circuit groups can be one-way or both way.

Traffic is routed as follows:

– Traffic streams are routed in first choice on the circuit group if it exists and overflow using state-dependent routing on a set of routes with one transit node.

– If the direct circuit group does not exist, streams are first routed using state-dependent routing on a set of routes with one transit node.

– The routing is of LLR type, which means that the choice of the route is done using idle capacity on the alternative routes.

A usual assumption considers that there is no blocking in the nodes.

C.2 Blocking probabilities evaluation

A blocking probabilities evaluation procedure, based on a fixed point model, is used.

For clarity, equations are given for a network with one-way circuit groups, but they can be extended easily in the case of both-way circuit groups or with the mixing of these two types.

The following notations are used:

ν_{ij}: offered traffic for stream *ij*.

C_{ij}: capacity of circuit group *ij*.

B_{ij}: call blocking probability of circuit group *ij*.

ρ_{ij}: offered traffic for circuit group *ij*.

α_{ikj}: probability for a call of traffic stream *ij* to go through transit node *k*.

cr_{ij}: residual capacity of the circuit group *ij*.

γ_{ikj}: residual capacity of route *ikj*.

L_{ij}: end-to-end call blocking probability of traffic stream *ij*.

{ T }: set of nodes used as transit.

Knowing the routing rules, the offered traffic on a circuit group depends on the nature of end-point nodes.

$$i \text{ and } j \notin \{\text{T}\} \quad \rho_{ij} = \nu_{ij} \tag{C-1}$$

$$i \in \{\text{T}\} \text{ and } j \notin \{\text{T}\} \quad \rho_{ij} = \nu_{ij} + \sum_{k \neq j} \nu_{ij} B_{kj} (1 - B_{ki}) \alpha_{kij} \tag{C-2}$$

$$i \in \{\text{T}\} \text{ and } j \in \{\text{T}\} \quad \rho_{ij} = \nu_{ij} + \sum_{k \neq j} \nu_{ij} B_{kj} (1 - B_{ki}) \alpha_{kij} + \sum_{k \neq i} \nu_{ik} B_{ik} (1 - B_{jk}) \alpha_{ijk} \tag{C-3}$$

The system of equations is then solved to obtain ρ_{ij} (offered traffic to the circuit group), B_{ij} (blocking of the circuit group) and α_{ikj} (probability that a call is routed through k).

As an approximation, the traffic offered to a circuit group is considered as a Poisson traffic and Erlang B formula is used to determine the blocking probabilities:

$$B_{ij} = E(\rho_{ij}, C_{ij}) \tag{C-4}$$

State-dependent routing is modelled as load sharing with coefficients computed dynamically. The load sharing coefficients α_{ikj} are determined according to the availability of each route. In the case of state-dependent routing, availability is determined by the residual capacity of the network. Therefore load sharing coefficients may be considered as proportional to the residual capacity of each route on the different routes of the stream.

$$\alpha_{ikj} = \gamma_{ikj} / \left(\sum_{l \neq i,\, l \neq j,\, l \in \{T\}} \gamma_{ilj} \right) \tag{C-5}$$

The mean residual capacity per route is estimated as a function of the mean residual capacity per circuit group. The mean residual capacity per circuit group is:

$$cr_{ij} = C_{ij} - \rho_{ij}(1 - B_{ij}) \tag{C-6}$$

Several formulae can be used to compute the mean residual capacity of the route [HUB]. If we consider that the residual capacities of the circuit groups are constant and equal to the mean, then:

$$\gamma_{ikj} = \min(cr_{ik}, cr_{kj}) \tag{C-7}$$

The end-to-end call blocking probability is evaluated using the link blocking independence assumption as:

$$L_{ij} = B_{ij} \sum_{k \in \{T\}} \alpha_{ikj} \left[1 - (1 - B_{ik})(1 - B_{kj}) \right] \tag{C-8}$$

The algorithm for computing the end-to-end call blocking probabilities L_{ij} consists of the following three steps [LE2].

Step 1

Initialization of blocking probabilities of circuit groups X-X are carried out. If the direct circuit group does not exist, the call blocking probability is set to 1.

1) Blocking probabilities of circuit group XT, TT or TX are initialized to a given value (e.g. 2%).

2) Load sharing coefficients for each traffic stream (i, j) are initialized, with the residual capacity set to the size of the circuit group and taking into account the number of routes.

Step 2

1) Offered traffic for each circuit group is computed using Formulae C-1, C-2 and C-3.

2) Blocking probability for each circuit group is computed using Formula C-4.

3) Residual capacity for each circuit group is computed using Formula C-6.

4) Load sharing coefficients are computed using Formula C-5.

Step 3

The convergence criterion is the maximum of the differences in circuit group blocking probability from the previous to the current iteration. If the convergence criterion is not reached, Step 2 is iterated.

When the convergence criterion is reached, the end-to-end blocking probability is evaluated using Formula C-8.

C.3 Dimensioning

The dimensioning criterion is a set of given end-to-end blocking probabilities. The general dimensioning procedure is described in Figure 10-1. The algorithm described in C.2 allows to compute the carried traffic for a given size of circuit group.

The procedure is as follows:

Step 1

The network is first assumed to be fully meshed and connectivity between two nodes is determined according to link creation thresholds set-up based on a given planning rules and the offered load of each traffic relation.

At first each X-X circuit group is considered according to the offered traffic and the link creation thresholds.

– If the value of the parameter used as criterion of an X-X node pair is greater than the predetermined creation threshold, the circuit group is not deleted.

– If the value of the parameter used as criterion of an X-X pair is less than the creation threshold, the corresponding circuit group is deleted and the corresponding X-X offered traffic is assumed to be equally shared among all allowable X-T-X routes.

Then the existence of X-T, T-X and T-T (in the latter case, when both T nodes are originating and destination nodes) trunking is analysed.

– If the value of the parameter used as criterion calculated according to the offered traffic (composed of the fresh offered traffic and the apportioned offered traffic of X-X relations having no direct circuit group) of the considered T-T relation is greater than the threshold, the circuit group is not deleted.

– If not, the corresponding circuit group is deleted.

Step 2

Once the network connectivity is determined, first the X-X circuit groups are sized using the Erlang B formula and the appropriate sizing parameter (occupancy, blocking probability or rejected traffic). Then the capacity is adjusted to the upper value to account for the modularity constraint. Erlang B formula is applied again to determine the blocking probability of the X-X circuit groups with the adjusted size.

To size the X-T, T-X and T-T circuit groups, blocking probabilities of circuit groups are initialized (the value of 0.02 has been experimentally determined but could be modified according to the user's need) and the offered traffic of each circuit group is the traffic of the corresponding stream (including the apportioned offered traffic of X-X relations having no direct trunking).

Notations

$Traf\,(\mathrm{XY})$: Traffic offered to X-Y circuit group.

B_{XY}: Blocking probability of X-Y circuit group.

EEB_{XY}: End-to-end blocking probability of X-Y traffic stream.

$RejTraf\,(\mathrm{XY})$: Traffic which is allowed to be rejected from X-Y circuit group.

α_{XYZ}: Load sharing coefficient indicating the portion of X-Y traffic routed via node Z.

Step 2.1

The traffic which is allowed to be rejected from the circuit group is determined in order to meet the end-to-end blocking probability constraint on the traffic stream, regarding the traffic carried on the overflow routes.

$$RejTraf\,(\mathrm{AB}) = Traf\,(\mathrm{AB})\left\{EEB_{\mathrm{AB}} + \sum_{T=1,\ldots,N} \alpha_{\mathrm{ATB}}(1 - B_{\mathrm{AT}})(1 - B_{\mathrm{TB}})\right\} \qquad \text{(C-9)}$$

Step 2.2

The offered and rejected traffics are incremented taking into account the overflow of other traffic streams (A-C or C-B) using A-B as part of their overflow route (see Figure C.1).

a) *Portion of C-B traffic offered to A-B*

$$Traf(AB_{CB}) = (Traf(CB) * B_{CB}) * \{1 - B_{CA}\} * \alpha_{CAB} \quad \text{(C-10)}$$

b) *Portion of A-C traffic offered to A-B*

$$Traf(AB_{AC}) = (Traf(AC) * B_{AC}) * \{1 - B_{BC}\} * \alpha_{ABC} \quad \text{(C-11)}$$

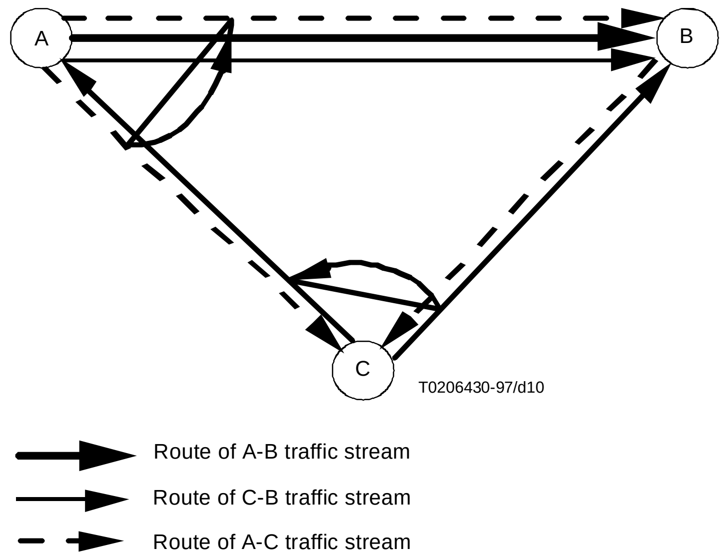


Figure C.1/E.529 – Traffic streams using A-B circuit group

It is assumed that the rejected traffic of A-B resulting from C-B and A-C is incremented only by half the end-to-end rejected traffic of C-B and A-C respectively, the other half being lost on the other leg of the two-leg route.

$$ReTraf(AB_{AC\text{-}CB}) = EEB_{AB} * \{Traf(AC) * \alpha_{ABC} + Traf(CB) * \alpha_{CA}\}/2 \quad \text{(C-12)}$$

Step 2.3

From the offered and rejected traffics, the blocking probability of each circuit group is determined and therefore using Erlang B formula the number of circuits necessary to reach this blocking probability is calculated. Then the circuit group modularity is adjusted to its upper value and the blocking probability and the residual capacity are updated.

Step 2.4

Load sharing coefficients are calculated as function of the residual capacity according to Formula C-5.

The whole procedure is iteratively carried out until convergence is reached.

C.4 Extension to event-dependent routing networks

The method can be applied to network dimensioning for event-dependent dynamic routing networks. The changes concern only the computation of load sharing parameters in the blocking probabilities evaluation procedure. In this case the load sharing coefficients α_{ikj} are determined according to the availability of each route. In the case of event-dependent routing, availability is determined by the inverse of the blocking probability of the network. Therefore load sharing coefficients may be considered as proportional to the inverse of the blocking probability of each route on the different routes of the stream.

$$\alpha_{ikj} = \gamma_{ikj} / \left(\sum_{l \neq i,\, l \neq j,\, l \in \{T\}} \gamma_{jlk} \right) \qquad \text{(C-13)}$$

where:

$$\gamma_{ikj} = 1 / \left[1 - (1 - B_{ik})(1 - B_{kj}) \right] \qquad \text{(C-14)}$$

Annex D

Example method for event-dependent routing networks

This Annex provides an example method for dimensioning event-dependent routing networks. It is based on the implied cost and shadow price methodology [KE2], [KEY].

D.1 Network model considered

Consider a fully interconnected network, where all traffic streams have a direct first route and two-link alternative route. The alternative route is chosen in a "sticky random" way. Circuit reservation is applied at a circuit group so that two priority classes are defined at the circuit group. In this example, direct calls are of priority class and overflow calls are of non-priority class.

The following notations are used:

k: index of a circuit group or the traffic stream that uses the circuit group as the first direct route, $k = 1,2,$ c, K.

ν_k: mean arrival rate of call stream k.

ν: set of the arrival rates ν_k: $\nu = \{\nu_1, \nu_2,$ c, $\nu_K\}$.

w_k: unit revenue of a stream k call.

$p_k\{a,b\}$: the probability that the alternative route $\{a,b\}$ is used given that the direct circuit group k is blocked.

C_{1k}: total capacity of circuit group k.

C_{2k}: capacity of circuit group k up to which non-priority calls can access (a circuit reservation level is given by $r_k = C_{1k} - C_{2k}$).

$\boldsymbol{C}_1$: set of circuit group capacities: $\boldsymbol{C}_1 = \{C_{11}, C_{12},$ c, $C_{1K}\}$.

$\boldsymbol{C}_2$: set of reservation parameters: $\boldsymbol{C}_2 = \{C_{21}, C_{22},$ c, $C_{2K}\}$.

σ_{1k}: circuit group k's mean call arrival rate above the circuit reservation threshold.

σ_{2k}: circuit group k's mean call arrival rate below the circuit reservation threshold.

L_k: end-to-end call blocking probability of stream k.

B_{ik}: the probability that a call of priority level i is blocked at circuit group k, where $i = 1$ for direct calls and $i = 2$ for overflow calls.

A_k: cost of installing a circuit on circuit group k.

W: rate of return from the network.

D.2 Implied costs method

D.2.1 Blocking probabilities for a network with an event-dependent routing

Event-dependent routing can be modelled [KEY] using the parameters:

$$p_k \{a,b\} \propto \frac{1}{1 - (1 - B_{2a})(1 - B_{2b})} \tag{D-1}$$

if $\{a,b\}$ is a valid alternative for stream k.

Let:

$$B_{ik} = E_i (\sigma_{1k}, \sigma_{2k}, C_{1k}, C_{2k}) \tag{D-2}$$

give the call blocking probabilities at circuit group k when the offered streams are Poissonian. Thus we construct a set of fixed point equations for the B's, which can be solved using repeated substitution with appropriate damping. Note that a solution is guaranteed by the Brouwer fixed point theorem, and that multiple solutions can exist at high blocking levels if the circuit reservation parameters are too low. The end-to-end call blocking probabilities L_k are calculated by using B_{ik} and $p_k\{a,b\}$.

The rate of return from the network is given by:

$$W(\nu, \boldsymbol{C}_1, \boldsymbol{C}_2) = \sum_{k=1,\ldots K} w_k \, \nu_k \, (1 - L_k) \tag{D-3}$$

D.2.2 Implied cost calculation

The implied cost methodology [KE2], [KEY] provides a method of obtaining the derivatives of W with respect to ν, $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$. Let c_{ik}, $i = 1, 2$ be the implied cost of circuit group k in blocking state i, which measures the knock-on effect elsewhere in the network if a call is accepted. These costs are defined as:

$$c_{ik} = -(B_{i+1,k} - B_{ik})^{-1} \frac{dW}{d\sigma_{ik}} \tag{D-4}$$

and can be calculated from the following fixed point equations [KE2], [KEY]:

$$c_{ik} = (B_{i+1,k} + B_{ik})^{-1} \left\{ \left[s_k + c_{1k} \right] \nu_k \frac{\partial B_{1k}}{\partial \sigma_{ik}} + \sum_{j,a} \left[\nu_j \, p_j \{a,k\} B_{1j} (1 - B_{2a})(s_j \{a,k\} + c_{2k}) \right] \frac{\partial B_{2k}}{\partial \sigma_{ik}} \right\} \tag{D-5}$$

with $B_{3k} = 1$, and where s_k and $s_k\{a,b\}$ are given by:

$$s_k = w_k - c_{1k} - \sum_{a,b} p_k \{a,b\}(1 - B_{2a})(1 - B_{2b}) s_k \{a,b\}, \quad s_k \{a,b\} = w_k - c_{2a} - c_{2b} \tag{D-6}$$

s_k and $s_k\{a,b\}$ are the surplus values associated with circuit group k and the two-circuit group alternative route $\{a,b\}$. For example, the latter has the interpretation that if a call is carried on the two-circuit group overflow route $\{a,b\}$, it generates revenue w_k directly but at a cost of $c_{2a} + c_{2b}$.

The quantities of interest for capacity allocation are "shadow prices" defined by $d_{ik} = dW/dC_{ik}$, which are a linear transformation of the c's,

$$d_{ik} = -\left\{[s_k + c_{1k}]\nu_k \frac{\partial B_{1k}}{\partial C_{ik}} + \sum_{j,a}\left[\nu_j p_j\{a,k\}B_{1j}(1 - B_{2a})(s_j\{a,k\} + c_{2k})\right]\frac{\partial B_{2k}}{\partial C_{ik}}\right\} \quad \text{(D-7)}$$

We have some freedom to define these derivatives, and shall take the derivatives at integer pairs to be:

$$\frac{\partial f}{\partial C_{1k}} = f(C_{1k}, C_{2k}) - f(C_{1k} - 1, C_{2k}), \; \frac{\partial f}{\partial C_{2k}} = -\left\{f(C_{1k}, C_{2k}) - f(C_{1k}, C_{2k} - 1)\right\} \quad \text{(D-8)}$$

in which case d_{1k} gives the difference in return if the number of circuits is altered by one with the circuit reservation held fixed, whilst d_{2k} measures the difference in return if the circuit reservation parameter is altered by one.

D.3 Multi-hour network dimensioning using shadow prices

D.3.1 The multi-hour dimensioning problem

Suppose the objective of the network optimization is to achieve a specified level of return at minimum cost, then the optimization problem is to:

minimize:

$$\sum_{k=1,\ldots,K} A_k C_{1k} \quad \text{(D-9)}$$

subject to:

$$W(\nu_t, \boldsymbol{C}_1, \boldsymbol{C}_2) \geq (1 - b_t) \sum_{k=1,\ldots K} w_{kt}\, \nu_{kt}, \quad t = 1, 2, \text{c}, T$$

$$C_{1k} \geq C_{2k} \geq 0, \quad k = 1, 2, \text{c}, K,$$

where the suffix t denotes the hour. If $w_{k,t} \equiv 1$, then b_t is the design grade of service for hour t.

The implied cost methodology can be used to solve the problem D-9 with the optimization jointly over $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$. However, practically speaking the circuit reservation parameters are chosen to ensure that the network is robust in the statistical sense. Moreover, sensible values of the circuit reservation parameters can be set fairly easily [KEY], and in what follows it is assumed that the circuit reservation parameters are fixed.

D.3.2 Use of shadow prices

If C_{1k} is regarded as a continuous parameter, then the Kuhn-Tucker conditions necessary for a local optimum are:

$$\lambda_t \left\{W(\nu_t, \boldsymbol{C}_1, \boldsymbol{C}_2) - (1 - b_t) \sum_{k=1,\ldots K} w_{kt}\nu_{kt}\right\} = 0 \quad \text{(D-10)}$$

$$\lambda_t \geq 0 \quad \text{(D-11)}$$

$$C_{1k}\left\{A_k - \sum_t \lambda_t \frac{dW_t}{dC_{1k}}\right\} = 0 \quad \text{(D-12)}$$

$$A_k - \sum_t \lambda_t \frac{dW_t}{dC_{1k}} \geq 0 \qquad \text{(D-13)}$$

$$C_{1k} \geq 0 \qquad \text{(D-14)}$$

If there is just a single hour and if $C_{1k} > 0$ then at a local optimum for a single hour,

$$R_k = \frac{d_{1k}}{A_k} \text{ a constant (the multiplier } \lambda \text{) where } d_{1k} = \frac{dW}{dC_{1k}} \qquad \text{(D-15)}$$

Thus with integral C_{1k} it is natural to augment the circuit groups for which R_k is large and decrement those for which it is small.

The introduction of multiple hours complicates matters. For the purpose of illustration, consider the case of $t = 2$. If one of the constraints is not active at the optimum then we are in the former situation. However if both are active then if $C_{1k} > 0$ then:

$$A_k - \sum_t \lambda_t \frac{dW}{dC_{1k}} = 0 \qquad \text{(D-16)}$$

The Lagrange multipliers λ_t measure the change in the objective function with respect to the constraints. A practical solution is to start with these equal for all hours and use the sum of the ratios $(R_{k1} + R_{k2})/2$ as a guide to capacity allocation.

This does not guarantee stream k's call blocking objective and it might be decided to bound the acceptable objective,

$$L_{\max} \leq b_{\max} \text{ a constant with } b_{\max} \geq b_t \text{ for all } t \qquad \text{(D-17)}$$

A natural way to proceed is to optimize the network with this constraint absent. If the conditions are violated for certain streams, then the worths of these streams can be set to one and the rest to zero. The network can then be augmented using the new shadow prices until these constraints are met.

ITU-T RECOMMENDATIONS SERIES

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