INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

E.360.7

(05/2002)

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

International routing plan

QoS routing and related traffic engineering methods – Traffic engineering operational requirements

ITU-T Recommendation E.360.7

ITU-T E-SERIES RECOMMENDATIONS

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

INTERNATIONAL OPERATION	
Definitions	E.100–E.103
General provisions concerning Administrations	E.104–E.119
General provisions concerning users	E.120–E.139
Operation of international telephone services	E.140–E.159
Numbering plan of the international telephone service	E.160–E.169
International routing plan	E.170–E.179
Tones in national signalling systems	E.180–E.189
Numbering plan of the international telephone service	E.190–E.199
Maritime mobile service and public land mobile service	E.200–E.229
OPERATIONAL PROVISIONS RELATING TO CHARGING AND ACCOUNTING IN THE INTERNATIONAL TELEPHONE SERVICE	
Charging in the international telephone service	E.230–E.249
Measuring and recording call durations for accounting purposes	E.260–E.269
UTILIZATION OF THE INTERNATIONAL TELEPHONE NETWORK FOR NON-TELEPHONY APPLICATIONS	
General	E.300–E.319
Phototelegraphy	E.320–E.329
ISDN PROVISIONS CONCERNING USERS	E.330–E.349
INTERNATIONAL ROUTING PLAN	**E.350–E.399**
NETWORK MANAGEMENT	
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	
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.649
Traffic engineering for IP-networks	E.650–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	
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 the list of ITU-T Recommendations.

ITU-T Recommendation E.360.7

QoS routing and related traffic engineering methods – Traffic engineering operational requirements

Summary

The E.360.x series of Recommendations describes, analyzes, and recommends methods which control a network's response to traffic demands and other stimuli, such as link failures or node failures. The functions discussed and recommendations made related to traffic engineering (TE) are consistent with the definition given in the Framework document of the Traffic Engineering Working Group (TEWG) within the Internet Engineering Task Force (IETF):

Internet Traffic Engineering is concerned with the performance optimization of operational networks. It encompasses the measurement, modeling, characterization, and control of Internet traffic, and the application of techniques to achieve specific performance objectives, including the reliable and expeditious movement of traffic through the network, the efficient utilization of network resources, and the planning of network capacity.

The methods addressed in the E.360.x series include call and connection routing, QoS resource management, routing table management, dynamic transport routing, capacity management, and operational requirements. Some of the methods proposed herein are also addressed in or are closely related to those proposed in ITU-T Recs E.170 to E.179 and E.350 to E.353 for routing, E.410 to E.419 for network management and E.490 to E.780 for other traffic engineering issues.

The recommended methods are meant to apply to IP-based, ATM-based, and TDM-based networks, as well as the interworking between these network technologies. Essentially, all of the methods recommended are already widely applied in operational networks worldwide, particularly in PSTN networks employing TDM-based technology. However, these methods are shown to be extensible to packet-based technologies, that is, to IP-based and ATM-based technologies, and it is important that networks which evolve to employ these packet technologies have a sound foundation of methods to apply. Hence, it is the intent that the methods recommended in this series of Recommendations be used as a basis for requirements for specific methods, and, as needed, for protocol development in IP-based, ATM-based, and TDM-based networks to implement the methods.

The methods encompassed in this Recommendation include traffic management through control of routing functions, which include QoS resource management. Results of analysis models are presented which illustrate the tradeoffs between various approaches. Based on the results of these studies, as well as established practice and experience, methods are recommended for consideration in network evolution to IP-based, ATM-based, and/or TDM-based technologies.

Source

ITU-T Recommendation E.360.7 was prepared by ITU-T Study Group 2 (2001-2004) and approved under the WTSA Resolution 1 procedure on 16 May 2002.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. 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 Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 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

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. 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, ITU 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

Introduction

As discussed in the E.360.x series of Recommendations, Figure 1/E.360.1 illustrates a model for network routing and network management and design. The central box represents the network, which can have various configurations, and the traffic routing tables and transport routing tables within the network. Routing tables describe the route choices from an originating node to a terminating node for a connection request for a particular service. Hierarchical, nonhierarchical, fixed, and dynamic routing tables have all been discussed in these series of Recommendations. Routing tables are used for a multiplicity of services on the telecommunications network, such as an MPLS/TE-based network used for illustration in this Recommendation.

Traffic engineering functions include traffic management, capacity management, and network planning. Figure 1/E.360.1 illustrates these functions as interacting feedback loops around the network. The input driving the network is a noisy traffic load, consisting of predictable average demand components added to unknown forecast error and other load variation components. The feedback controls function to regulate the service provided by the network through traffic management controls, capacity adjustments, and routing adjustments. Traffic management provides monitoring of network performance through collection and display of real-time traffic and performance data and allows traffic management controls such as code blocks, connection request gapping, and reroute controls to be inserted when circumstances warrant. Capacity management includes capacity forecasting, daily and weekly performance monitoring, and short-term network adjustment. Forecasting operates over a multiyear forecast interval and drives network capacity expansion. Daily and weekly performance monitoring identify any service problems in the network. If service problems are detected, short-term network adjustment can include routing table updates and, if necessary, short-term capacity additions to alleviate service problems. Updated routing tables are sent to the switching systems either directly or via an automated routing update system. Short-term capacity additions are the exception, and most capacity changes are normally forecasted, planned, scheduled and managed over a period of months, or a year or more. Network design embedded in capacity management includes routing design and capacity design. Network planning includes longer-term node planning and transport network planning, which operates over a horizon of months to years to plan and implement new node and transport capacity.

In 6.2 to 6.5, we focus on the steps involved in traffic management of the MPLS/TE-based network (see 6.2), capacity forecasting in the MPLS/TE-based network (see 6.3), daily and weekly performance monitoring (see 6.4), and short-term network adjustment in the MPLS/TE-based network (see 6.5). For each of these three topics, we illustrate the steps involved with examples.

Monitoring of traffic and performance data is a critical issue for traffic management, capacity forecasting, daily and weekly performance monitoring, and short-term network adjustment. This topic is receiving attention in IP-based networks [FGLRR99] where traffic and performance data has been somewhat lacking, in contrast to TDM-based networks where such TE monitoring data has been developed to a sophisticated standard over a period of time [A98]. The discussions in this Recommendation intend to point out the kind and frequency of TE traffic and performance data required to support each function.

See ITU-T Recs E.490 to E.504 on general principles and requirements for traffic measurements and GOS monitoring; E.505 and E.743 on measurements for SS7 signalling networks; E.745 on measurements for ATM networks and E.506 to E.508 on forecasting methods.

ITU-T Recommendation E.360.7

QoS routing and related traffic engineering methods – Traffic engineering operational requirements

1 Scope

The E.360.x series of Recommendations describes, analyzes, and recommends methods which control a network's response to traffic demands and other stimuli, such as link failures or node failures. The functions discussed and recommendations made related to traffic engineering (TE) are consistent with the definitions given in the Framework document of the Traffic Engineering Working Group (TEWG) within the Internet Engineering Task Force (IETF):

Internet Traffic Engineering is concerned with the performance optimization of operational networks. It encompasses the measurement, modeling, characterization, and control of Internet traffic, and the application of techniques to achieve specific performance objectives, including the reliable and expeditious movement of traffic through the network, the efficient utilization of network resources, and the planning of network capacity.

The methods addressed in the E.360.x series include call and connection routing, QoS resource management, routing table management, dynamic transport routing, capacity management, and operational requirements. Some of the methods proposed herein are also addressed in or are closely related to those proposed in ITU-T Recs E.170 to E.179 and E.350 to E.353 for routing, E.410 to E.419 for network management and E.490 to E.780 for other traffic engineering issues.

The recommended methods are meant to apply to IP-based, ATM-based, and TDM-based networks, as well as the interworking between these network technologies. Essentially, all of the methods recommended are already widely applied in operational networks worldwide, particularly in PSTN networks employing TDM-based technology. However, these methods are shown to be extensible to packet-based technologies, that is, to IP-based and ATM-based technologies, and it is important that networks which evolve to employ these packet technologies have a sound foundation of methods to apply. Hence, it is the intent that the methods recommended in this series of Recommendations be used as a basis for requirements for specific methods and, as needed, for protocol development in IP-based, ATM-based, and TDM-based networks, to implement the methods.

Hence the methods encompassed in this series of Recommendations include:

- traffic management through control of routing functions, which include call routing (number/name translation to routing address), connection routing, QoS resource management, routing table management, and dynamic transport routing.
- capacity management through control of network design, including routing design.
- operational requirements for traffic management and capacity management, including forecasting, performance monitoring, and short-term network adjustment.

Results of analysis models are presented which illustrate the tradeoffs between various approaches. Based on the results of these studies as well as established practice and experience, TE methods are recommended for consideration in network evolution to IP-based, ATM-based, and/or TDM-based technologies.

2 References

See clause 2/E.360.1.

3 Definitions

See clause 3/E.360.1.

4 Abbreviations

See clause 4/E.360.1.

5 Traffic management

In this clause we concentrate on the surveillance and control of the MPLS/TE-based network. We also discuss the interactions of traffic managers with other work centres responsible for MPLS/TE-based network operation. Traffic management functions should be performed at a centralized work centre, and be supported by centralized traffic management operations functions (TMOF), perhaps embedded in a centralized bandwidth-broker processor (here denoted TMOF-BBP). A functional block diagram of TMOF-BBP is illustrated in Figure 1.

5.1 Real-time performance monitoring

The surveillance of the MPLS/TE-based network should be performed through monitoring the highest bandwidth-overflow/delay-count node-pair, preferably on a geographical display, which is normally monitored at all times. This display should be used in the auto-update mode, which means that every five minutes TMOF-BBP automatically updates the exceptions shown on the map itself, and displays the node pairs with the highest bandwidth overflow/delay count. TMOF-BBP also should have displays that show the high bandwidth-overflow/delay-percent pairs within threshold values.

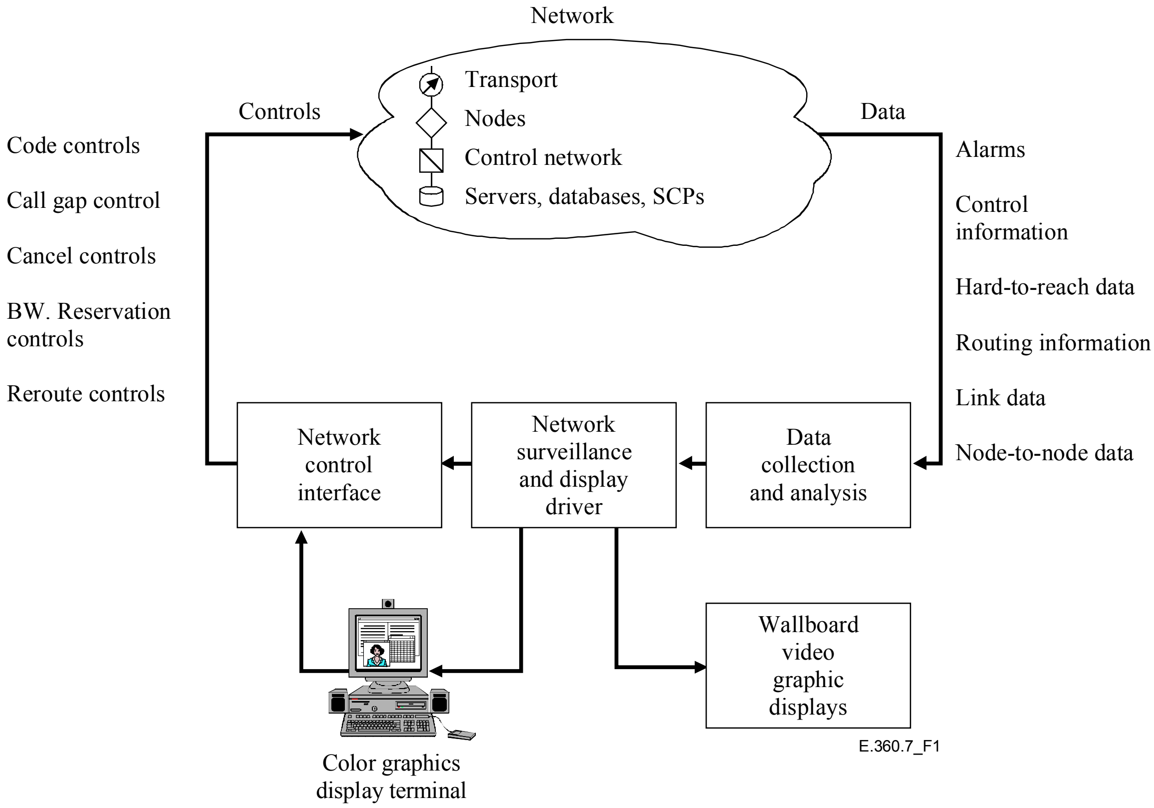

Figure 1/E.360.7 – Traffic management operations functions within bandwidth-broker processor

Traffic managers are most concerned with what connection requests can be rerouted and, therefore, want to know the location of the heaviest concentrations of blocked call routing attempts. For that purpose, overflow/delay percentages can be misleading. From a service revenue standpoint, the difference between 1 percent and 10 percent blocking/delay on a node pair may favor concentration on the 1 percent blocking/delay situation, because there are more connection requests to reroute. TMOF-BBP should also display all the exceptions that there are with the auto threshold display, which displays everything exceeding the present threshold, for example, either 1 percent bandwidth-overflow/delay or 1 or more blocked connection requests, in 5 minutes. In the latter case, this display then shows the total blocked connection requests and not just the highest pairs.

For peak-day operation, or operation on a high day (such as a Monday after Thanksgiving), traffic managers should work back and forth between the auto threshold display and the highest blocked-connection-count pair display. They can spend most of their time with the auto threshold display, where they can see everything that is being blocked. Then, when traffic managers want to concentrate on clearing out some particular problem, they can look at the highest blocked-connection-count pair display, an additional feature of which is that it allows the traffic manager to see the effectiveness of controls.

The traffic manager can recognize certain patterns from the surveillance data. For example, a focused overload on a particular city/node such as caused by a flooding situation discussed further in 6.3, 6.4, and 6.5. The typical traffic pattern under a focused overload is that most locations show heavy overflow/delay into and out of the focus-overload node. Under such circumstances, the display should show the bandwidth overflow/delay percent for any node pair in the MPLS/TE-based network that exceeds 1 percent bandwidth overflow/delay percent.

One of the other things traffic managers should be able to see with TMOF-BBP using the highest bandwidth-overflow/delay-count pair display is a node failure. Transport failures should also show on the displays, but the resulting display pattern depends on the failure itself.

5.2 Network control

The MPLS/TE-based network needs automatic controls built into the node processing and also has automatic and manual controls that can be activated from TMOF-BBP. We first describe the required controls and what they do, and then we discuss how the MPLS/TE-based traffic managers work with these controls. Two protective automatic traffic management controls are required in the MPLS/TE-based network: dynamic overload control (DOC), which responds to node congestion, and dynamic bandwidth reservation (DBR), which responds to link congestion. DOC is an identical concept to Automatic Congestion Control (ACC), as defined in ITU-T Rec. E.412. DOC and DBR should be selective in the sense that they control traffic destined for hard-to-reach points more stringently than other traffic.

The complexity of MPLS/TE networks makes it necessary to place more emphasis on fully automatic controls that are reliable and robust and do not depend on manual administration. DOC and DBR should respond automatically within the node software program. For DBR, the automatic response can be coupled, for example, with two bandwidth reservation threshold levels, represented by the amount of idle bandwidth on an MPLS/TE-based link. DBR bandwidth reservation levels should be automatic functions of the link size.

DOC and DBR are not strictly link-dependent but should also depend on the node pair to which a controlled connection request belongs. A connection request offered to an overloaded via node should either be canceled at the originating node or advanced to an alternate via node, depending on the destination of the call. DBR should differentiate between primary (shortest) path and alternate path connection requests.

DOC and DBR should also use a simplified method of obtaining hard-to-reach control selectivity. In the MPLS/TE-based network, hard-to-reach codes can be detected by the terminating node, which then communicates them to the originating nodes and via nodes. Because the terminating node is the only exit point from the MPLS/TE-based network, the originating node should treat a hard-to-reach code detected by a terminating node as hard to reach on all MPLS/TE-based links.

DOC should normally be permanently enabled on all links. DBR should automatically be enabled by an originating node on all links when that originating node senses general network congestion. DBR is particularly important in the MPLS/TE-based network because it minimizes the use of less efficient alternate path connections and maximizes useful network throughput during overloads. The automatic enabling mechanism for DBR ensures its proper activation without manual intervention. DOC and DBR should automatically determine whether to subject a controlled connection request to a cancel, or skip control. In the cancel mode, affected connection requests are blocked from the network whereas, in the skip mode, such connection requests skip over the controlled link to an alternate link. DOC and DBR should be completely automatic controls. Capabilities such as automatic enabling of DBR, the automatic skip/cancel mechanism and the DBR one-link/two-link traffic differentiation, adapt these controls to the MPLS/TE-based network and make them robust and powerful automatic controls.

Code-blocking controls block connection requests to a particular destination code. These controls are particularly useful in the case of focused overloads, especially if the connection requests are blocked at or near their origination. Code blocking controls need not block all calls, unless the destination node is completely disabled through natural disaster or equipment failure. Nodes equipped with code-blocking controls can typically control a percentage of the connection requests to a particular code. The controlled E.164 name (dialed number code), for example, may be NPA, NXX, NPA-NXX, or NPA-NXX-XXXX, when in the latter case one specific customer is the target of a focused overload.

A call-gapping control, illustrated in Figure 2, is typically used by network managers in a focused connection request overload, such as sometimes occurs with radio call-in give-away contests.

Call gapping allows one connection request for a controlled code, or set of codes, to be accepted into the network by each node once every x seconds, and connection requests arriving after the accepted connection request are rejected for the next x seconds. In this way, call gapping throttles the connection requests and prevents the overload of the network to a particular focal point.

An expansive control is also required. Reroute controls should be able to modify routes by inserting additional paths at the beginning, middle, or end of a path sequence. Such reroutes should be inserted manually, or automatically, through TMOF-BBP. When a reroute is active on a node pair, DBR should be prevented on that node pair from going into the cancel mode, even if the overflow/delay is heavy enough on a particular node pair to trigger the DBR cancel mode. Hence, if a reroute is active, connection requests should have a chance to use the reroute paths and not be blocked prematurely by the DBR cancel mode.

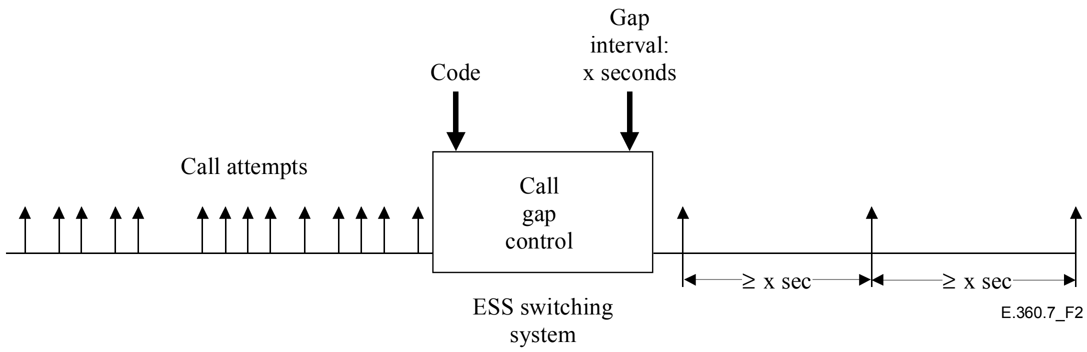

Figure 2/E.360.7 – Call gap control

In the MPLS/TE-based network, a display should be used to graphically represent the controls in effect. Depending on the control in place, either a certain shape or a certain color should tell traffic managers which control is implemented. Traffic managers should be able to tell if a particular control at a node is the only control on that node. Different symbols should be used for the node, depending on the controls that are in effect.

5.3 Work centre functions

Work centre functions are outlined in various clauses of this Recommendation as guidance reflecting best current practices. Implementation choices are, of course, left to the discretion of each network operator.

5.3.1 Automatic controls

The MPLS/TE-based network requires automatic controls, as described above, and if there is spare capacity, traffic managers can decide to reroute. In the example focus-overload situation, the links are occupied sufficiently, and there is often no network capacity available for reroutes. The DBR control is normally active at the time. In order to get connection requests out of focus-overload-node, traffic managers must sometimes manually disable the DBR control at the focus-overload-node. This gave preference to connection requests going out of the focus-overload-node. Thereby, the focus-overload-node gets much better completion of outgoing connection requests than will the other nodes at completing calls into the focus-overload node. This control results in using the link capacity more efficiently.

Traffic managers should be able to manually enable or inhibit DBR, and also inhibit the skip/cancel mechanism for both DBR and DOC. Traffic managers should monitor DOC controls very closely because they indicate switching congestion or failure. Therefore, DOC activations should be investigated much more thoroughly, and more quickly, than DBR activations which are frequently triggered by normal heavy traffic.

5.3.2 Code controls

Code controls are used to cancel connection requests for very hard-to-reach codes. Code control is used when connection requests cannot complete to a point in the network, or there is isolation. For example, traffic managers should use code controls for a focus overload situation, such as caused by an earthquake, in which there can be isolation. Normal hard-to-reach traffic caused by heavy calling volumes will be blocked by the DBR control, as described above.

Traffic managers should use data on hard-to-reach codes in certain situations for problem analysis. For example, if there is a problem in a particular area, one of the early things traffic managers should look at is the hard-to-reach data to see if they can identify one code, or many codes, that are hard to reach, and if they are from one location, or several locations.

5.3.3 Reroute controls

Traffic managers should sometimes use manual reroute even when an automatic reroute capability is there. Reroutes are used primarily for transport failures or heavy traffic surges, such as traffic on heavier than normal days, where the surge is above the normal capabilities of the network to handle the load. Those are the two prime reasons for rerouting. Traffic managers do not usually reroute into a disaster area.

5.3.4 Peak-day control

Peak-day routing in the MPLS/TE-based network should involve using the primary (shortest) path (CRLSP) as the only engineered path and then the remaining available paths as alternate paths all subject to DBR controls. The effectiveness of the additional alternate paths and reroute capabilities depends very much on the peak day itself. The greater the peak-day traffic, the less effective the alternate paths are. That is, on the higher peak days, such as Christmas and Mother's Day, the network is filled with connections mostly on shortest paths. On lower peak days, such as Easter or Father's Day, the use of alternate paths and rerouting capabilities are more effective. This is because the peaks, although they are high and have an abnormal traffic pattern, are not as high as on Christmas or Mother's Day. So, on these days, there is additional capacity to complete connection requests on the alternate paths. Reroute paths are particularly available in the early morning and late evening. Depending on the peak day, at times there is also a lull in the afternoon, and TMOF-BBP should normally be able to find reroute paths that are available.

5.4 Traffic management on peak days

A typical peak-day routing method uses the shortest path between node pairs as the only engineered path, followed by alternate paths protected by DBR. This method is more effective during the lighter peak days such as Thanksgiving, Easter, and Father's Day. With the lighter loads, when the network is not fully saturated, there is a much better chance of using the alternate paths. However, when we enter the network busy hour, or combination of busy hours, with a peak load over most of the network, the routing method at that point drops back to shortest-path routing because of the effect of bandwidth reservation. At other times, the alternate paths are very effective in completing calls.

5.5 Interfaces to other work centres

The main interaction traffic managers have is with the capacity managers. Traffic managers notify capacity managers of conditions in the network, that are affecting the data that they use in making decisions, as to whether or not to add capacity. Examples are transport failures and node failures that would distort traffic data. A node congestion signal can trigger DOC; DOC cancels all traffic destined to a node while the node congestion is active. All connection requests to the failed node are reflected as overflow connection requests for the duration of the node congestion condition. This can be a considerable amount of canceled traffic. The capacity manager notifies traffic managers of the new link capacity requirements that they are trying to get installed, but that are delayed. Traffic managers can then expect to see congestion on a daily basis, or several times a week, until the capacity is added. This type of information is passed back and forth on a weekly or perhaps daily basis.

6 Capacity management – Forecasting

In this clause we concentrate on the forecasting of MPLS/TE-based node-to-node loads and the sizing of the MPLS/TE-based network. We also discuss the interactions of network forecasters with other work centres responsible for MPLS/TE-based network operations.

Network forecasting functions should be performed from a capacity administration centre and supported by network forecasting operations functions integrated into the BBP (NFOF-BBP). A functional block diagram of NFOF-BBP is illustrated in Figure 3. In the following two clauses we discuss the steps involved in each functional block.

6.1 Load Forecasting

6.1.1 Configuration database functions

As illustrated in Figure 3, the configuration database is used in the forecasting function, and within this database are defined various specific components of the network itself, for example: backbone nodes, access nodes, transport points of presence, buildings, manholes, microwave towers, and other facilities.

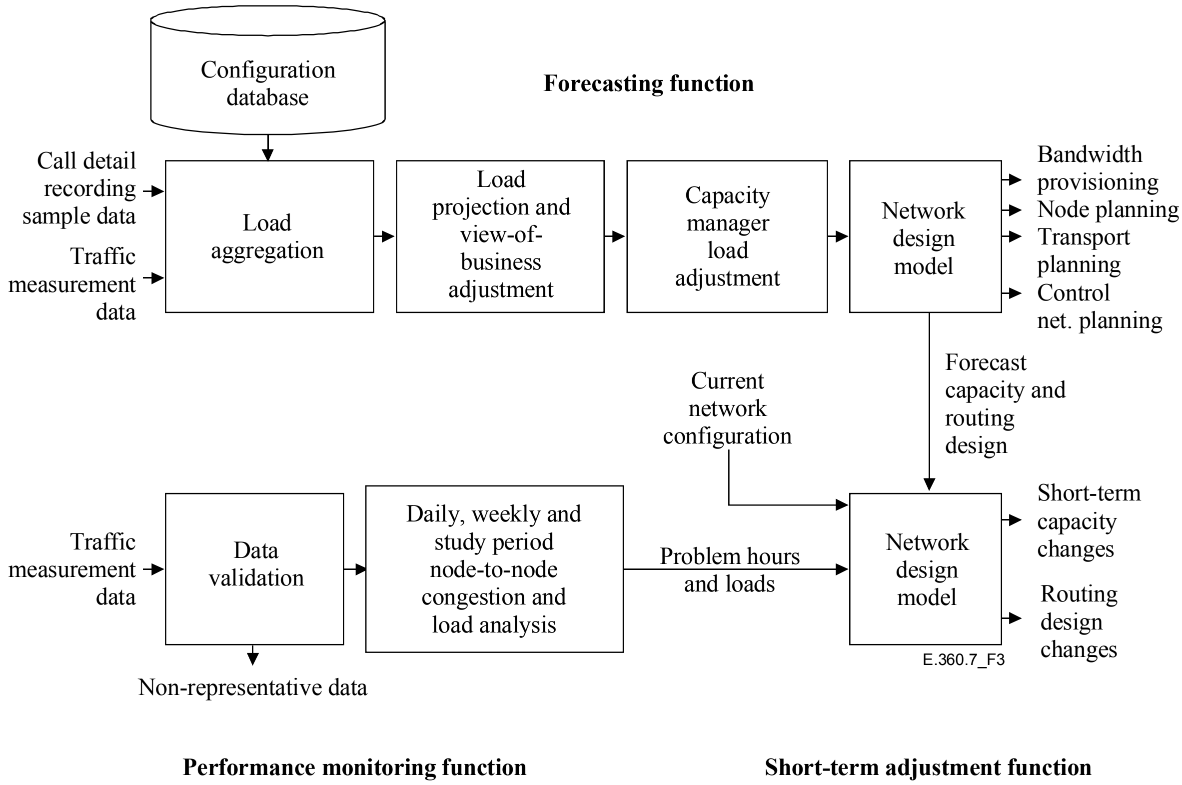

Figure 3/E.360.7 – Capacity management functions within bandwidth-broker processor

Forecasters maintain configuration data for designing and forecasting the MPLS/TE-based network. Included in the data for each backbone node and access node, for example, are the number/name translation capabilities, equipment type, type of signalling, homing arrangement, international routing capabilities, operator service routing capabilities, and billing/recording capabilities. When a forecast cycle is started, which is normally each month, the first step is to extract the relevant pieces of information from the configuration database that are necessary to drive network forecasting operations functions (NFOF-BBP). One of information items indicates the date of the forecast view; this is when the configuration files were frozen, which then represents the network structure at the time the forecast is generated.

6.1.2 Load aggregation, basing and projection functions

NFOF-BBP should process data from a centralized message database, which represents a record of all connection requests placed on the network, over four study periods within each year, for example, March, May, August, and November, each a 20-day study period. Based on the centralized data, a sampling method can be used, for example, a 5 percent sample of recorded connections for 20 days. Forecasters can then equate that 5 percent, 20-day sample to one average business day. The load information then consists of messages and traffic load by study period. In the load aggregation step, NFOF-BBP may apply nonconversation time factors to equate the traffic load, obtained from billed traffic load, to the actual holding time traffic load.

The next step in load forecasting is to aggregate all of the access-node-to-access-node loads up to the backbone node-pair level. This produces the backbone-node-to-backbone-node traffic item sets. These node-to-node traffic item sets are then routed to the candidate links. NFOF-BBP should then project those aggregated loads into the future, using smoothing techniques to compare the current measured data with the previously projected data, and to determine an optimal estimate of the base and projected loads. The result is the initially projected loads that are ready for forecaster adjustments and business/econometric adjustments.

6.1.3 Load adjustment cycle and view of business adjustment cycle

Once NFOF-BBP smoothes and projects the data, forecasters can then enter a load adjustment cycle. This should be an online process that has the capability to go into the projected load file for all the forecast periods for all the years and apply forecaster-established thresholds to those loads. For example, if the forecaster requests to see any projected load that has deviated more than 15 percent from what it was projected to be in the last forecast cycle, a load analysis module in NFOF-BBP should search through all the node pairs that the forecaster is responsible for, sort out the ones that exceed the thresholds, and print them on a display. The forecaster then has the option to change the projected loads or accept them.

After the adjustment cycle is complete, and the forecasters have adjusted the loads to account for missing data, erroneous data, more accurate current data, or specifically planned events that cause a change in load, forecasters should then apply the view of the business adjustments. Up to this point, the projection of loads has been based on projection models and network structure changes, as well as the base study period billing data. The view of the business adjustment is intended to adjust the future network loads to compensate for the effects of competition, rate changes, and econometric factors on the growth rate. This econometric adjustment process tries to encompass those factors in an adjustment that is applied to the traffic growth rates. Growth rate adjustments should be made for each business, residence and service category, since econometric effects vary according to service category.

6.2 Network design

Given the MPLS/TE-based node-pair loads, adjusted by the forecasters, and also adjusted for econometric projections, the network design model should then be executed by NFOF-BCC based on those traffic loads. The node-to-node loads are estimated for each hourly backbone-node-to-backbone-node traffic load, including the minute-to-minute variability and the day-to-day variation, plus the control parameters. The access-node-to-backbone-node links should also be sized in this step.

A list of all the MPLS/TE-based node pairs should then be sent to the transport planning database, from which is extracted transport information relative to the transport network between the node pairs on that list. Once the information has been processed in the design model, NFOF-BBP should output the MPLS/TE-based forecast report. Once the design model has run for a forecast cycle, the forecast file and routing information should be sent downstream to the provisioning systems, planning systems, and capacity management system, and the capacity manager takes over from there as far as implementing the routing and the link capacity called for in the forecast.

6.3 Work centre functions

Capacity management and forecasting operation should be centralized. Work should be divided on a geographic basis so that the MPLS/TE-based forecaster and capacity manager for a region can work with specific work centres within the region. These work centres include the node planning and implementation organizations and the transport planning and implementation organizations. Their primary interface should be with the system that is responsible for issuing the orders to augment link capacity. Another interface is with the routing organization that processes the routing information coming out of NFOF-BBP.

NFOF-BBP should provide a considerable amount of automation and, as such, people can spend their time on more productive activities. By combining the forecasting job and the capacity management job into one centralized operation, additional efficiencies are achieved from a reduction in fragmentation. By centralizing the operations, this avoids duplication from distributing the operation within regional groups and, with the automation, time need only be spent to clear a problem, or analyze data outliers, rather than to check and verify everything.

This operation requires people who are able to understand and deal with a more complex network, and the network complexity will continue to increase as new technology and services are introduced. Other disciplines can usefully centralize their operations, for example, node planning, transport planning, equipment ordering, and inventory control. With centralized equipment-ordering and inventory control, for example, all equipment required for the network can be bulk ordered and distributed. This leads to a much more efficient use of inventory.

6.4 Interfaces to other work centres

Network forecasters work cooperatively with node planners, transport planners, traffic managers, and capacity managers. With an MPLS/TE network, forecasting, capacity management and traffic management must tie together closely. One way to develop those close relationships is by having centralized, compact work centres. The forecasting process essentially drives all the downstream construction and planning processes for an entire network operation.

7 Capacity management – Daily and weekly performance monitoring

In this clause we concentrate on the analysis of node-to-node capacity management data and the design of the MPLS/TE-based network. Capacity management becomes mandatory at times, as seen from the node-to-node traffic data, when significant congestion problems are extant in the network, or when it is time to implement a new forecast. We discuss the interactions of capacity managers with other work centres responsible for MPLS/TE-based network operation. Capacity management functions should be performed from a capacity administration centre and should be supported by the capacity management operations functions embedded, for example, in the BBP (denoted here as the CMOF-BBP). A functional block diagram of the CMOF-BBP is illustrated within the lower three blocks of Figure 3. In the following clauses we discuss the processes in each functional block.

7.1 Daily congestion analysis functions

A daily congestion summary should be used to give a breakdown of the highest to the lowest node-pair congestion that occurred the preceding day. This is an exception-reporting function in which there should be an ability to change the display threshold. For example, the capacity manager can request to see only node pairs whose congestion level is greater than 10 percent. Capacity managers investigate to find out whether they should exclude these data and, if so, for what reason. One reason for excluding data is to keep them from downstream processing if they are associated with an abnormal network condition. This would prevent designing the network for this type of nonrecurring network condition. In order to find out what the network condition was, capacity managers consult with the traffic managers. If, for example, traffic managers indicate that the data is associated with an abnormal network condition, such as a focused overload due to flooding the night before, then capacity managers may elect to exclude the data.

7.2 Study-week congestion analysis functions

The CMOF-BBP functions should also support weekly congestion analysis. This should normally occur after capacity managers form the weekly average using the previous week's data. The study-week data should then be used in the downstream processing to develop the study-period average. The weekly congestion data are set up basically the same way as the daily congestion data and give the node pairs that had congestion for the week. This study-week congestion analysis function gives another opportunity to review the data to see if there is a need to exclude any weekly data.

7.3 Study-period congestion analysis functions

Once each week, the study-period average should be formed using the most current four weeks of data. The study-period congestion summary gives an idea of the congestion during the most current study period in which node pairs, that experienced average business day average blocking/delay greater than 1 percent, are identified. If congestion is found for a particular node pair in a particular hour, the design model may be exercised to solve the congestion problem. In order to determine whether they should run the design model for that problem hour, capacity managers should first look at the study-period congestion detail data. For the node pair in question, they look at the 24-hours of data to see if there are any other hours for that node pair that should be investigated. Capacity managers should also determine if there is pending capacity addition for the problem node pair.

8 Capacity management – Short-term network adjustment

8.1 Network design functions

There are several features which should be available in the design model. First, capacity managers should be able to select a routing change option. With this option, the design model should make routing table changes to utilize the network capacity that is in place to minimize congestion. The design model should also design the network to the specified grade-of-service objectives. If it cannot meet the objectives with the network capacity in place, it specifies how much capacity to add to which links in order to meet the performance objectives. The routing table update implementation should be automatic from the CMOF-BBP all the way through to the network nodes. An evaluator option of the design model should be available to determine the carried traffic per link, or network efficiency, for every link in the network for the busiest hour.

8.2 Work centre functions

Certain sections of the network should be assigned so that all capacity managers have an equal share of links that they are responsible for. Each capacity manager therefore deals primarily with one region. Capacity managers also need to work with transport planners so that the transport capacity planned for the links under the capacity manager's responsibility is available to the capacity manager. If, on a short-term basis, capacity has to be added to the network, capacity managers find out from the transport planner whether the transport capacity is available. CMOF-BBP is highly automated, and the time the capacity manager spends working with CMOF-BBP system displays should be small compared with other daily responsibilities. One of the most time-consuming work functions is following up on the capacity orders to determine status: Are they in the field? Does the field have them? Do they have the node equipment working? If capacity orders are delayed, the capacity manager is responsible for making sure that the capacity is added to the network as soon as possible. With the normal amount of network activity going on, that is the most time-consuming part of the work centre function.

8.3 Interfaces to other work centres

The capacity manager needs to work with the forecasters to learn of network activity that will affect the MPLS/TE-based network. Of concern are new nodes coming into the network capacity management activity that affects the MPLS/TE-based network. Capacity managers should interact quite frequently with traffic managers to learn of network conditions such as cable cuts, floods, or disasters. Capacity managers detect such activities the next day in the data; the network problem stands out immediately. Before they exclude the data, however, capacity managers need to talk with the traffic managers to find out specifically what the problem was in the network. In some cases, capacity managers will share information with them about something going on that they may not be aware of. For example, capacity managers may be able to see failure events in the data, and they can share this type of information with the traffic managers. Other information capacity managers might share with traffic managers relates to peak days. The next morning, capacity managers are able to give the traffic managers the actual reports and information of the load and congestion experienced in the network.

Capacity managers also work with the data collection work centre. If they miss collecting data from a particular node for a particular day, capacity managers should discuss this with that work centre to get the data into CMOF-BBP. In CMOF-BBP, capacity managers should have some leeway in getting data into the system that may have been missed. So, if data are missed one night on a particular node, the node should be available to be repolled to pull data into CMOF-BBP.

Capacity managers frequently communicate with the routing work centres because there is so much activity going on with routing. For example, capacity mangers work with them to set up the standard numbering/naming plans so that they can access new routing tables when they are entered into the network. Capacity managers also work with the people who are actually doing the capacity order activity on the links. Capacity managers should try to raise the priority on capacity orders if there is a congestion condition, and often a single congestion condition may cause multiple activities in the MPLS/TE network.

9 Comparison of off-line (TDR) versus on-line (SDR/EDR) TE methods

With an on-line (SDR/EDR-based) MPLS/TE network, as compared to an off-line (TDR-based) network, several improvements occur in TE functions. Under TDR-based networks, TMOF-BBP should automatically put in reroutes to solve congestion problems by looking everywhere in the network for additional available capacity, and adding additional alternate paths to the existing preplanned paths, on a five-minute basis. With SDR/EDR-based networks, in contrast, this automatic rerouting function is replaced by real-time examination of all admissible routing choices.

Hence, an important simplification introduced with the SDR/EDR-based networks is that routing tables need not be calculated by the design model, because these are computed in real time by the node or BBP. This leads to simplifications in that the routing tables, computed in TDR-based networks, are no longer needed. Hence, simplifications are introduced into the administration of network routing. With TDR, routing tables must be periodically reoptimized and downloaded into nodes via the CMOF-BBP process. Reoptimizing and changing the routing tables in the TDR-based network represents an automated yet large administrative effort involving perhaps millions of records. This function is simplified in SDR/EDR-based networks since the routing is generated in real time for each connection request and then discarded. Also, because SDR/EDR-based TE adapts to network conditions, less network churn and short-term capacity additions are required. This is one of the operational advantages of SDR/EDR-based MPLS/TE networks, that is, to automatically adapt TE so as to move the traffic load to where capacity is available in the network.

10 Conclusions/recommendations

Conclusions/recommendations reached in this Recommendation include the following:

- Monitoring of traffic and performance data is recommended and is required for traffic management, capacity forecasting, daily and weekly performance monitoring and short-term network adjustment.
- Traffic management is recommended and is required to provide monitoring of network performance through collection and display of real-time traffic and performance data and allow traffic management controls such as code blocks, connection request gapping and reroute controls to be inserted when circumstances warrant.
- Capacity management is recommended and is required for capacity forecasting, daily and weekly performance monitoring and short-term network adjustment.

- Forecasting is recommended and is required to operate over a multiyear forecast interval and drive network capacity expansion.
- Daily and weekly performance monitoring is recommended and is required to identify any service problems in the network. If service problems are detected, short-term network adjustment can include routing table updates and, if necessary, short-term capacity additions to alleviate service problems. Updated routing tables are sent to the switching systems either directly or via an automated routing update system.
- Short-term capacity additions are recommended and are required as needed, but only as an exception, whereas most capacity changes are normally forecasted, planned, scheduled and managed over a period of months, or a year or more.
- Network design, which includes routing design and capacity design, is recommended and is required within the capacity management function.
- Network planning is recommended and is required for longer-term node planning and transport network planning, and operates over a horizon of months to years to plan and implement new node and transport capacity.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of 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	Cable networks and 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 communications
Series Y	Global information infrastructure and Internet protocol aspects
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2003