INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

E.360.3

(05/2002)

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

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

International routing plan

QoS routing and related traffic engineering methods – QoS resource management methods

ITU-T Recommendation E.360.3

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.360.3

QoS routing and related traffic engineering methods – QoS resource management methods

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, modelling, 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.360.3 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.

© ITU 2003

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

CONTENTS

Introduction

QoS resource management (sometimes called QoS routing) functions include class-of-service identification, routing table derivation, connection admission, bandwidth allocation, bandwidth protection, bandwidth reservation, priority routing, priority queuing, and other related resource management functions. A broad classification of these functions is as follows:

1) connection-level controls: provide the required connection-level GoS objectives in an economical way;

2) packet-level controls: ensure the required packet-level GoS objectives.

Connection-level controls are discussed in clauses 5 to 8, and packet level controls are discussed in clause 9.

QoS resource management methods have been applied successfully in TDM-based networks [A98], and are being extended to IP-based and ATM-based networks. In an illustrative QoS resource management method, bandwidth is allocated in discrete changes to each of several virtual networks (VNETs), which are each assigned a priority corresponding to either high-priority key services, normal-priority services, or best-effort low-priority services. Examples of services within these VNET categories include:

- high-priority key services such as defense voice communication;
- normal-priority services such as constant rate, interactive, delay-sensitive voice; variable rate, interactive, delay-sensitive IP-telephony; and variable rate, non-interactive, non-delay-sensitive WWW file transfer; and
- low-priority best-effort services such as variable rate, non-interactive, non-delay-sensitive voice mail, email, and file transfer.

Bandwidth changes in VNET bandwidth capacity can be determined by edge nodes on a per-flow (per-connection) basis, or based on an overall aggregated bandwidth demand for VNET capacity (not on a per-connection demand basis). In the latter case of per-VNET bandwidth allocation, based on the aggregated bandwidth demand, edge nodes make periodic discrete changes in bandwidth allocation, that is, either increase or decrease bandwidth, such as on the constraint-based routing label switched paths (CRLSPs) constituting the VNET bandwidth capacity.

In the illustrative QoS resource management method, which we assume is MPLS-based, the bandwidth allocation control for each VNET CRLSP is based on estimated bandwidth needs, bandwidth use, and status of links in the CRLSP. The edge node, or originating node (ON), determines when VNET bandwidth needs to be increased or decreased on a CRLSP, and uses an illustrative MPLS CRLSP bandwidth modification procedure to execute needed bandwidth allocation changes on VNET CRLSPs. In the bandwidth allocation procedure the constraint-based routing label distribution protocol (CRLDP) [J00] or the resource reservation protocol (RSVP-TE) [AGBLSS00] could be used, for example, to specify appropriate parameters in the label request message:

a) to request bandwidth allocation changes on each link in the CRLSP; and

b) to determine if link bandwidth can be allocated on each link in the CRLSP.

If a link bandwidth allocation is not allowed, a notification message with an illustrative crankback parameter allows the ON to search out possible bandwidth allocation on another CRLSP. In particular, we illustrate an optional depth-of-search (DoS) parameter in the label request message to control the bandwidth allocation on individual links in a CRLSP. In addition, we illustrate an optional modify parameter in the label request message to allow dynamic modification of the assigned traffic parameters (such as peak data rate, committed data rate, etc.) of an already existing CRLSP. Finally, we illustrate a crankback parameter in the notification message to allow an edge node to search out additional alternate CRLSPs when a given CRLSP cannot accommodate a bandwidth request.

QoS resource management therefore can be applied on a per-flow (or per-call or per-connection-request) basis, or can be beneficially applied to traffic trunks (also known as "bandwidth pipes" or "virtual trunking") in the form of CRLSPs in IP-based networks or SVPs in ATM-based networks.

QoS resource management provides integration of services on a shared network, for many classes-of-service such as:

- CBR services including voice, 64-, 384-, and 1536-kbit/s N-ISDN switched digital data, international switched transit, priority defense communication, virtual private network, 800/free-phone, fiber preferred, and other services.
- Real-time VBR services including IP-telephony, compressed video, and other services.
- Non-real-time VBR services including WWW file transfer, credit card check, and other services.
- UBR services including voice mail, email, file transfer, and other services.

We now illustrate the principles of QoS resource management, which includes integration of many traffic classes, as discussed above.

ITU-T Recommendation E.360.3

QoS routing and related traffic engineering methods – QoS resource management methods

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 (TWG) within the Internet Engineering Task Force (IETF):

Internet Traffic Engineering is concerned with the performance optimization of operational networks. It encompasses the measurement, modelling, 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 Class-of-service identification, policy-based routing table derivation and QoS resource management steps

QoS resource management functions include class-of-service identification, routing table derivation, connection admission, bandwidth allocation, bandwidth protection, bandwidth reservation, priority routing, and priority queuing. In this clause we discuss class-of-service identification and routing-table derivation.

5.1 Class-of-Service identification

QoS resource management entails identifying class-of-service and class-of-service parameters, which may include, for example:

- service identity (SI);
- virtual network (VNET);
- link capability (LC); and
- QoS and traffic threshold parameters.

The SI describes the actual service associated with the call. The VNET describes the bandwidth allocation and routing table parameters to be used by the call. The LC describes the link hardware capabilities such as fiber, radio, satellite, and digital circuit multiplexing equipment (DCME), that the call should require, prefer, or avoid. The combination of SI, VNET, and LC constitute the class-of-service, which together with the network node number is used to access routing table data.

In addition to controlling bandwidth allocation, the QoS resource management procedures can check end-to-end transfer delay, delay variation, and transmission quality considerations such as loss, echo, and noise, as discussed in clause 10 below.

Determination of class-of-service begins with translation at the originating node. The number or name is translated to determine the routing address of the destination node. If multiple ingress/egress routing is used, multiple destination node addresses are derived for the call. Other data derived from call information, such as link characteristics, Q.931 message information elements, Information Interchange digits, and network control point routing information, are used to derive the class-of-service for the call.

5.2 Policy-based routing table derivation

Class-of-service parameters are derived through application of policy-based routing. Policy-based routing involves the application of rules applied to input parameters to derive a routing table and its associated parameters. Input parameters for applying policy-based rules to derive SI, VNET, and LC could include numbering plan, type of origination/destination network, and type of service. Policy-based routing rules may then be applied to the derived SI, VNET, and LC to derive the routing table and associated parameters.

Hence, policy-based routing rules are used in SI derivation which, for example, uses the type of origin, type of destination, signalling service type and dialed number/name service type to derive the SI. The type of origin can be derived normally from the type of incoming link to the connected network domain, connecting either to a directly connected (also known as nodal) customer equipment location, a switched access local exchange carrier, or an international carrier location.

Similarly, based on the dialed numbering plan, the type of destination network is derived and can be a directly connected (nodal) customer location if a private numbering plan is used (for example, within a virtual private network), a switched access customer location if a National Numbering Plan (NNP) number is used to the destination, or an international customer location, if the international E.164 numbering plan is used. Signalling service type is derived based on bearer capability within signalling messages, information digits in dialed digit codes, numbering plan or other signalling information, and can indicate long-distance service (LDS), virtual private network (VPN) service, ISDN switched digital service (SDS), and other service types. Finally, dialed number service type is derived based on special dialed number codes such as 800 numbers or 900 numbers and can indicate 800 (FreePhone) service, 900 (Mass-Announcement) service, and other service types. Type of origin, type of destination, signalling service type and dialed number service type are then all used to derive the SI.

The following are examples of the use of policy-based routing rules to derive class-of-service parameters. A long-distance service (LDS) SI, for example, is derived from the following information:

1) The type of origination network is a switched access local exchange carrier, because the call originates from a local exchange carrier node.
2) The type of destination network is a switched access local exchange carrier, based on the NNP dialed number.
3) The signalling service type is long-distance service, based on the national numbering plan (NNP).
4) The dialed number service type is not used to distinguish long-distance service SI.

An 800 (FreePhone) service SI, for example, is derived from similar information, except that:

– The dialed number service type is based on the 800 dialed "freephone" number to distinguish the 800 service SI.

A VPN service SI, for example, is derived from similar information, except that:

– The signalling service type is based on the originating customer having access to VPN intelligent network (IN)-based services to derive the VPN service SI.

A service identity mapping table uses the above four inputs to derive the service identity. This policy-based routing table is changeable by administrative updates, in which new service information can be defined without software modifications to the node processing. From the SI and bearer-service capability the SI/bearer-service-to-virtual network mapping table is used to derive the VNET.

Table A.1/E.360.2 illustrates the VNET mapping table. Here the SIs are mapped to individual virtual networks. Routing parameters for priority or key services are discussed further in the clauses below.

Link capability selection allows calls to be routed on links that have the particular characteristics required by these calls. A call can require, prefer, or avoid a set of link characteristics such as fiber transmission, radio transmission, satellite transmission, or compressed voice transmission. Link capability requirements for the call can be determined by the SI of the call, or by other information derived from the signalling message or from the routing number. The routing logic allows the call to skip those links that have undesired characteristics and to seek a best match for the requirements of the call.

5.3 QoS resource management steps

The illustrative QoS resource management method consists of the following steps:

At the ON, the destination node (DN), SI, VNET, and QoS resource management information are determined through the number/name translation database and other service information available at the ON.

The DN and QoS resource management information are used to access the appropriate VNET and routing table between the ON and DN.

The connection request is set up over the first available path in the routing table with the required transmission resource selected based on the QoS resource management data.

In the first step, the ON translates the dialed digits to determine the address of the DN. If multiple ingress/egress routing is used, multiple destination node addresses are derived for the connection request. Other data derived from connection request information includes link characteristics, Q.931 message information elements, information interchange (II) digits, and service control point (SCP) routing information, and are used to derive the QoS resource management parameters (SI, VNET, LC, and QoS/traffic thresholds). SI describes the actual service associated with the connection request: VNET describes the bandwidth allocation and routing table parameters to be used by the connection request, and the LC describes the link characteristics including fiber, radio, satellite, and voice compression that the connection request should require, prefer, or avoid. Each connection request is classified by its SI. A connection request for an individual service is allocated an equivalent bandwidth equal to EQBW and routed on a particular VNET. For CBR services the equivalent bandwidth EQBW is equal to the average or sustainable bit rate. For VBR services the equivalent bandwidth EQBW is a function of the sustainable bit rate, peak bit rate, and perhaps other parameters. For example, EQBW equals 64 kbit/s of bandwidth for CBR voice connections, 64 kbit/s of bandwidth for CBR ISDN switched digital 64-kbit/s connections, and 384-kbit/s of bandwidth for CBR ISDN switched digital 384-kbit/s connections. (Equivalent bandwidth and connection admission control are further discussed in clause 9.)

In the second step, the SI value is used to derive the VNET. In the multi-service, QoS resource management network, bandwidth is allocated to individual VNETs which is protected as needed but otherwise shared. Under normal non-blocking/delay network conditions, all services fully share all available bandwidth. When blocking/delay occurs for VNET i, bandwidth reservation acts to prohibit alternate-routed traffic and traffic from other VNETs from seizing the allocated capacity for VNET i. Associated with each VNET are average bandwidth (BWavg) and maximum bandwidth (BWmax) parameters to govern bandwidth allocation and protection, which are discussed further in clause 6. As discussed, LC selection allows connection requests to be routed on specific transmission links that have the particular characteristics required by a connection request.

In the third step, the VNET routing table determines which network capacity is allowed to be selected for each connection request. In using the VNET routing table to select network capacity, the ON selects a first choice path based on the routing table selection rules. Whether or not bandwidth can be allocated to the connection request on the first choice path is determined by the QoS resource management rules given below. If a first choice path cannot be accessed, the ON may then try alternate paths determined by FR, TDR, SDR, or EDR path selection rules outlined in ITU-T Rec. E.360.2. Whether or not bandwidth can be allocated to the connection request on the alternate path again is determined by the QoS resource management rules now described.

6 Dynamic bandwidth allocation, protection, and reservation principles

QoS resource management functions include class-of-service identification, routing table derivation, connection admission, bandwidth allocation, bandwidth protection, bandwidth reservation, priority routing and priority queuing. In this clause we discuss connection admission, bandwidth allocation, bandwidth protection, and bandwidth reservation.

This clause specifies the resource allocation controls and priority mechanisms, and the information needed to support them. In the illustrative QoS resource management method, the connection/bandwidth-allocation admission control for each link in the path is performed based on the status of the link. The ON may select any path for which the first link is allowed according to QoS resource management criteria. If a subsequent link is not allowed, then a release with crankback/bandwidth-not-available is used to return to the ON and select an alternate path. This use of an EDR path selection, which entails the use of the release with crankback/bandwidth-not-available mechanism to search for an available path, is an alternative to SDR path selection which may entail flooding of frequently changing link state parameters such as available-cell-rate. The tradeoffs between EDR with crankback and SDR with link-state flooding are further discussed in ITU-T Rec. E.360.6. In particular, when EDR path selection with crankback is used in lieu of SDR path selection with link-state flooding, the reduction in the frequency of such link-state parameter flooding allows for larger peer group sizes. This is because link-state flooding can consume substantial processor and link resources in terms of message processing by the processors, and link bandwidth consumed by messages on the links.

Two cases of QoS resource management are considered in this Recommendation: per-virtual-network (per-VNET) management and per-flow management. In the per-VNET method, such as illustrated for IP-based MPLS networks, aggregated LSP bandwidth is managed to meet the overall bandwidth requirements of VNET service needs. Individual flows are allocated bandwidth within the CRLSPs accordingly, as CRLSP bandwidth is available. In the per-flow method, bandwidth is allocated to each individual flow, such as in SVC set-up in an ATM-based network, from the overall pool of bandwidth, as the total pool bandwidth is available. A fundamental principle applied in these bandwidth allocation methods is the use of bandwidth reservation techniques. We first review bandwidth reservation principles and then discuss per-VNET and per-flow QoS resource allocation.

Bandwidth reservation (the TDM-network terminology is "trunk reservation") gives preference to the preferred traffic by allowing it to seize any idle bandwidth in a link, while allowing the non-preferred traffic to only seize bandwidth if there is a minimum level of idle bandwidth available, where the minimum-bandwidth threshold is called the reservation level. P. J. Burke [Bur61] first analyzed bandwidth reservation behavior from the solution of the birth-death equations for the bandwidth reservation model. Burke's model showed the relative lost-traffic level for preferred traffic, which is not subject to bandwidth reservation restrictions, as compared to non-preferred traffic, which is subject to the restrictions. Figure 1 illustrates the percent lost traffic of preferred and non-preferred traffic on a typical link with 10 percent traffic overload. It is seen that the preferred traffic lost traffic is near zero, whereas the non-preferred lost traffic is much higher, and this situation is maintained across a wide variation in the percentage of the preferred traffic load. Hence, bandwidth reservation protection is robust to traffic variations and provides significant dynamic protection of particular streams of traffic.

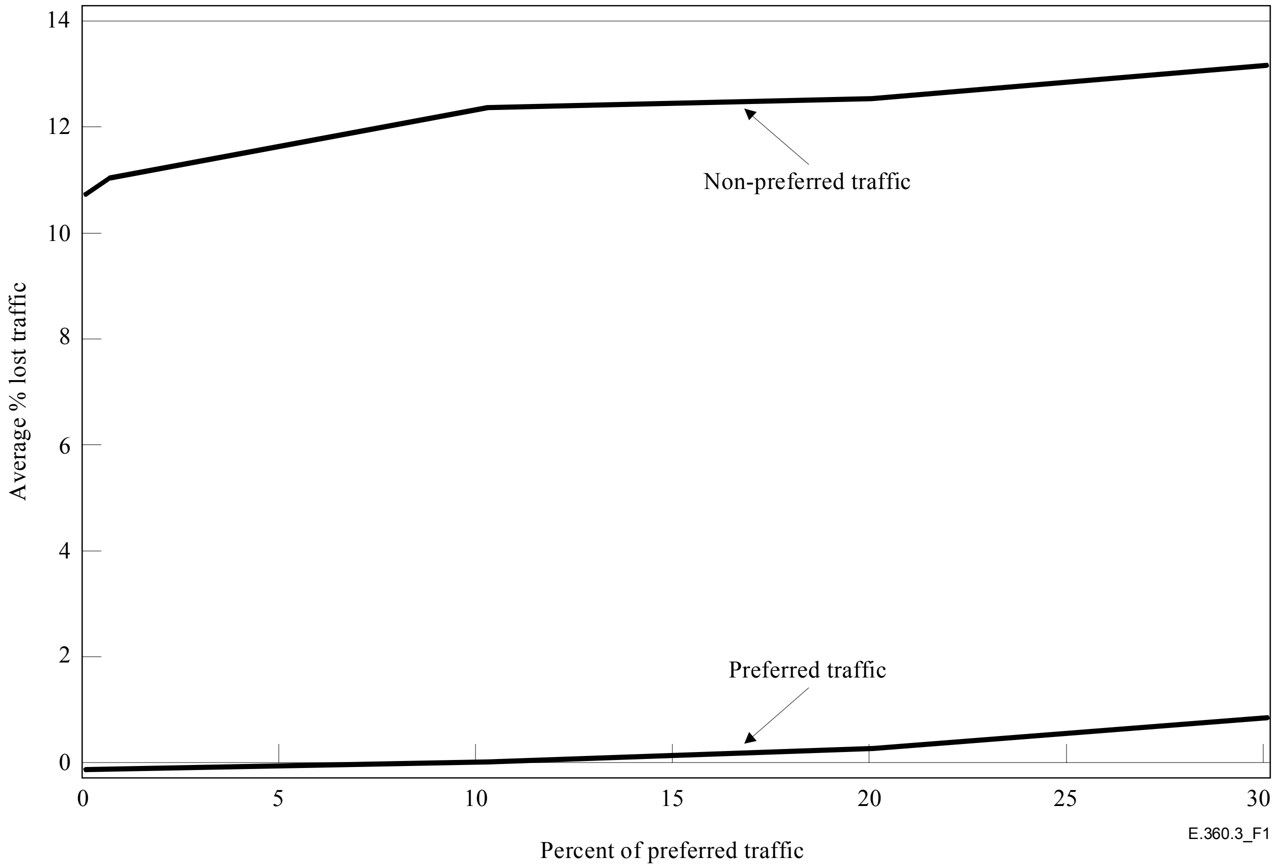


Figure 1/E.360.3 – Dynamic bandwidth reservation performance under 10% overload

Bandwidth reservation is a crucial technique used in nonhierarchical networks to prevent "instability," which can severely reduce throughput in periods of congestion, perhaps by as much as 50 percent of the traffic-carrying capacity of a network [E.525]. The phenomenon of instability has an interesting mathematical solution to network flow equations which has been presented in several studies [NaM73], [Kru82], [Aki84]. It is shown in these studies that nonhierarchical networks exhibit two stable states, or bistability, under congestion and that networks can transition between these stable states in a network congestion condition that has been demonstrated in simulation studies. A simple explanation of how this bistable phenomenon arises is that, under congestion, a network is often not able to complete a connection request on the primary shortest path which consists, in this example, of a single link. If alternate routing is allowed, such as on longer, multiple-link paths, which are assumed, in this example, to consist of two links, then the connection request might be completed on a two-link path selected from among a large number of two-link path choices, only one of which needs sufficient idle bandwidth on both links to be used to route the connection. Because this two-link connection now occupies resources that could perhaps otherwise be used to complete two one-link connections, this is a less efficient use of network resources under congestion. In the event that a large fraction of all connections cannot complete on the direct link but instead occupy two-link paths, the total network throughput capacity is reduced by one-half because most connections take twice the resources needed. This is one stable state; that is, most or all connections use two links. The other stable state is that most or all connections use one link, which is the desired condition.

Bandwidth reservation is used to prevent this unstable behavior by organizing traffic in such a way that the preferred traffic on a link corresponds to the direct traffic on the primary, shortest path, and that the non-preferred traffic, subjected to bandwidth reservation restrictions as described above, corresponds to the alternate-routed traffic on longer paths. In this way the alternate-routed traffic is inhibited from selecting longer alternate paths when sufficient idle trunk capacity is not available on all links of an alternate-routed connection, which is the likely condition under network and link congestion. Mathematically, the studies of bistable network behavior have shown that bandwidth reservation used in this manner to favor primary shortest connections eliminates the bistability problem in nonhierarchical networks and allows such networks to maintain efficient utilization under congestion by favoring connections completed on the shortest path. For this reason, dynamic bandwidth reservation is universally applied in nonhierarchical TDM-based networks [E.529], and often in hierarchical networks [Mum76].

There are differences in how and when bandwidth reservation is applied, however, such as whether the bandwidth reservation for connections routed on the primary path is in place at all times or whether it is dynamically triggered to be used only under network or link congestion. This is a complex network throughput trade-off issue, because bandwidth reservation can lead to some loss in throughput under normal, low-congestion conditions. This loss in throughput arises because if bandwidth is reserved for connections on the primary path, but these connection requests do not arrive, then the capacity is needlessly reserved when it might be used to complete alternate-routed traffic that might otherwise be blocked. However, under network congestion, the use of bandwidth reservation is critical to preventing network instability, as explained above [E.525], [E.529], [E.731].

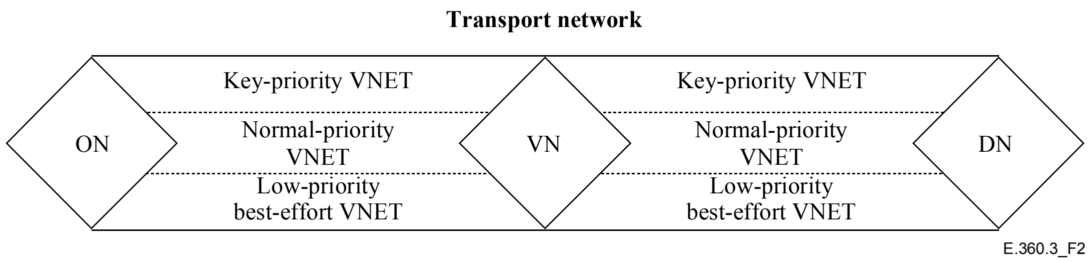
It is beneficial for bandwidth reservation techniques be included in IP-based and ATM-based routing methods, in order to ensure the efficient use of network resources especially under congestion conditions. Currently recommended path-selection methods, such as methods for optimized multipath for traffic engineering in IP-based MPLS networks [V99], or path selection in ATM-based PNNI networks [ATM960055], give no guidance on the necessity for using bandwidth-reservation techniques. Such guidance is essential for acceptable network performance [E.737].

Examples are given in this Recommendation for dynamically triggered bandwidth reservation techniques, where bandwidth reservation is triggered only under network congestion. Such methods are shown to be effective in striking a balance between protecting network resources under congestion and ensuring that resources are available for sharing when conditions permit. In clause 6 the phenomenon of network instability is illustrated through simulation studies, and the effectiveness of bandwidth reservation in eliminating the instability is demonstrated. Bandwidth reservation is also shown to be an effective technique to share bandwidth capacity among services integrated on a primary path, where the reservation in this case is invoked to prefer link capacity on the primary path for one particular class-of-service as opposed to another class-of-service when network and link congestion are encountered. These two aspects of bandwidth reservation, that is, for avoiding instability and for sharing bandwidth capacity among services, are illustrated in clauses 6, 7, and 8.

7 Per-virtual-network bandwidth allocation, protection, and reservation

Through the use of bandwidth allocation, reservation, and congestion control techniques, QoS resource management can provide good network performance under normal and abnormal operating conditions for all services sharing the integrated network. Such methods have been analyzed in practice for TDM-based networks [A98], and in modelling studies for IP-based networks [ACFM99]. In this Recommendation these IP-based QoS resource management methods are described. However, the intention here is to illustrate the general principles of QoS resource management and not to recommend a specific implementation.

As illustrated in Figure 2, in the multi-service, QoS resource management network, bandwidth is allocated to the individual VNETs (high-priority key services VNETs, normal-priority services VNETs, and best-effort low-priority services VNETs).



- Distributed method applied on a per-virtual-network basis.
- ON allocates bandwidth to each virtual-network (VNET) based on demand.
- For VNET bandwidth increase:
  - ON decides link-bandwidth-modification threshold (Pi) based on:
    - bandwidth-in-progress (BWIP);
    - routing priority (key, normal, best-effort);
    - bandwidth allocation BWavg;
    - first/alternate choice path.
  - ON launches a CRLDP label request message with explicit route, modify-flag, traffic parameters, & threshold Pi (carried in setup priority).
- VNs keep local link state of idle link bandwidth (ILBW), including lightly loaded (LL), heavily loaded (HL), reserved (R), & busy (B).
- VNs compare link state to Pi threshold.
- VNs send crankback/bandwidth-not-available notification message to ILSR if Pi threshold not met.

ON Originating node
DN Destination node
VN Via node
Node

Figure 2/E.360.3 – Virtual Network (VNET) bandwidth management

This allocated bandwidth is protected by bandwidth reservation methods, as needed, but otherwise shared. Each ON monitors VNET bandwidth use on each VNET CRLSP, and determines when VNET CRLSP bandwidth needs to be increased or decreased. Bandwidth changes in VNET bandwidth capacity are determined by ONs based on an overall aggregated bandwidth demand for VNET capacity (not on a per-connection demand basis). Based on the aggregated bandwidth demand, these ONs make periodic discrete changes in bandwidth allocation, that is, either increase or decrease bandwidth on the CRLSPs constituting the VNET bandwidth capacity. For example, if connection requests are made for VNET CRLSP bandwidth that exceeds the current CRLSP bandwidth allocation, the ON initiates a bandwidth modification request on the appropriate CRLSP(s). For example, this bandwidth modification request may entail increasing the current CRLSP bandwidth allocation by a discrete increment of bandwidth denoted here as delta-bandwidth (DBW). DBW is a large enough bandwidth change so that modification requests are made relatively infrequently. Also, the ON periodically monitors CRLSP bandwidth use, such as once each minute, and, if bandwidth use falls below the current CRLSP allocation, the ON initiates a bandwidth modification request to decrease the CRLSP bandwidth allocation by a unit of bandwidth such as DBW.

In making a VNET bandwidth allocation modification, the ON determines the QoS resource management parameters including the VNET priority (key, normal, or best-effort), VNET bandwidth-in-use, VNET bandwidth allocation thresholds, and whether the CRLSP is a first choice CRLSP or alternate CRLSP. These parameters are used to access a VNET depth-of-search (DoS) table to determine a DoS load state threshold (Pi), or the "depth" to which network capacity can be allocated for the VNET bandwidth modification request. In using the DoS threshold to allocate VNET bandwidth capacity, the ON selects a first choice CRLSP based on the routing table selection rules.

Path selection in this IP network illustration may use open shortest path first (OSPF) for intra-domain routing. In OSPF-based layer 3 routing, as illustrated in Figure 3, ON A determines a list of shortest paths by using, for example, Dijkstra's algorithm.

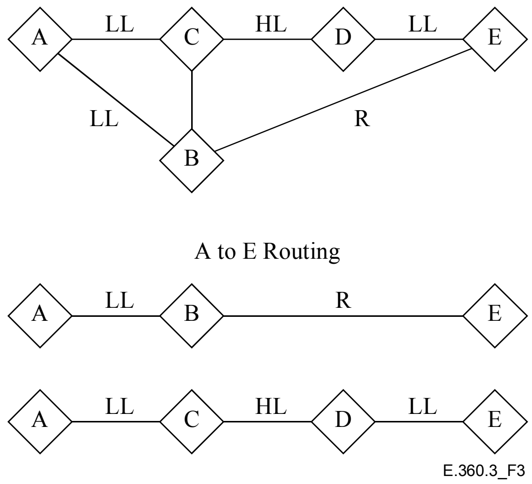


Figure 3/E.360.3 – Label switched path selection for bandwidth modification request

This path list could be determined based on administrative weights of each link, which are communicated to all nodes within the autonomous system (AS) domain. These administrative weights may be set, for example, to [1 + epsilon × distance], where epsilon is a factor giving a relatively smaller weight to the distance in comparison to the hop count. The ON selects a path from the list based on, for example, FR, TDR, SDR, or EDR path selection, as discussed in ITU-T Rec. E.360.2.

For example, in using the first CRLSP A-B-E in Figure 3, ON A sends an MPLS label request message to VN B, which in turn forwards the label request message to DN E. VN B and DN E are passed in the explicit routing (ER) parameter contained in the label request message. Each node in the CRLSP reads the ER information, and passes the label request message to the next node listed in the ER parameter. If the first path is blocked at any of the links in the path, an MPLS notification message with a crankback parameter is returned to ON A, which can then attempt the next path. If FR is used, then this path is the next path in the shortest path list, for example path A-C-D-E. If TDR is used, then the next path is the next path in the routing table for the current time period. If SDR is used, OSPF implements a distributed method of flooding link status information, which is triggered either periodically and/or by crossing load state threshold values. This method of distributing link status information can be resource intensive and may not be any more efficient than simpler path selection methods such as EDR. If EDR is used, then the next path is the last successful path, and if that path is unsuccessful another alternate path is searched out according to the EDR path selection method.

Hence, in using the selected CRLSP, the ON sends the explicit route, the requested traffic parameters (peak data rate, committed data rate, etc.), a DoS-parameter, and a modify-parameter in the MPLS label request message to each VN and the DN in the selected CRLSP. Whether or not bandwidth can be allocated to the bandwidth modification request on the first choice CRLSP is determined by each VN applying the QoS resource management rules. These rules entail that the VN determine the CRLSP link states, based on bandwidth use and bandwidth available, and compare the link load state to the DoS threshold Pi sent in the MPLS signalling parameters, as further explained below. If the first choice CRLSP cannot admit the bandwidth change, a VN or DN returns control to the ON through the use of the crankback-parameter in the MPLS notification message. At that point, the ON may then try an alternate CRLSP. Whether or not bandwidth can be allocated to the bandwidth modification request on the alternate path again is determined by the use of the DoS threshold compared to the CRLSP link load state at each VN. Priority queuing is used during the time the CRLSP is established, and at each link the queuing discipline is maintained such that the packets are given priority according to the VNET traffic priority.

Hence, determination of the CRLSP link load states is necessary for QoS resource management to select network capacity on either the first choice CRLSP, or alternate CRLSPs. Four link load states are distinguished: lightly loaded (LL), heavily loaded (HL), reserved (R), and busy (B). Management of CRLSP capacity uses the link state model and the DoS model to determine if a bandwidth modification request can be accepted on a given CRLSP. The allowed DoS load state threshold Pi determines if a bandwidth modification request can be accepted on a given link to an available bandwidth "depth." In setting up the bandwidth modification request, the ON encodes the DoS load state threshold allowed on each link in the DoS-parameter Pi which is carried in the MPLS label request. If a CRLSP link is encountered at a VN in which the idle link bandwidth and link load state are below the allowed DoS load state threshold Pi, then the VN sends an MPLS notification message with the crankback-parameter to the ON, which can then route the bandwidth modification request to an alternate CRLSP choice. For example, in Figure 3, CRLSP A-B-E may be the first path tried where link A-B is in the LL state and link B-E is in the R state. If the DoS load state allowed is Pi = HL or better, then the CRLSP bandwidth modification request in the MPLS label request message is routed on link A-B but will not be admitted on link B-E, wherein the CRLSP bandwidth modification request will be cranked back in the MPLS notification message to the originating node A to try alternate CRLSP A-C-D-E. Here, the CRLSP bandwidth modification request succeeds since all links have a state of HL or better.

7.1 Per-VNET bandwidth allocation/reservation – Meshed network case

For purposes of bandwidth allocation reservation, two approaches are illustrated: one applicable to meshed network topologies and the other applicable to sparse topologies. In meshed networks, a greater number of logical links results in less traffic carried per link, and functions such as bandwidth reservation need to be more carefully controlled than in a sparse network. In a sparse network, the traffic is concentrated on much larger, and many fewer logical links, and here bandwidth reservation does not have to be as carefully managed. Hence, in the meshed network case, functions such as automatically triggering of bandwidth reservation on and off, dependent on the link/network congestion level, are beneficial to use. In the sparse network case, however, the complexity of such automatic triggering is not essential and bandwidth reservation may be permanently enabled without performance degradation.

Here we discuss a meshed network example of bandwidth allocation/reservation and in 7.2 we discuss the sparse network case.

The DoS load state threshold is a function of bandwidth-in-progress, VNET priority, and bandwidth allocation thresholds, as follows:

Table 1/E.360.3 – Determination of Depth-of-Search (DoS) load state threshold (Per-VNET Bandwidth Allocation, Meshed Network)

Load state allowed$_i$	Key priority VNET	Normal priority VNET		Best effort priority VNET
		First choice CRLSP	Alternate CRLSP	
R	if $BWIP_i \le 2 \times BWmax_i$	If $BWIP_i \le BWavg_i$	Not Allowed	Note
HL	if $BWIP_i \le 2 \times BWmax_i$	If $BWIP_i \le BWmax_i$	if $BWIP_i \le BWavg_i$	Note
LL	All $BWIP_i$	All $BWIP_i$	All $BWIP_i$	Note

where: $BWIP_i$ = bandwidth-in-progress on VNET i $BWavg_i$ = minimum guaranteed bandwidth required for VNET i to carry the average offered bandwidth load $BWmax_i$ = the bandwidth required for VNET i to meet the blocking/delay probability grade-of-service objective for CRLSP bandwidth allocation requests = $1.1 \times BWavg_i$ NOTE – CRLSPs for the best effort priority VNET are allocated zero bandwidth; Diffserv queuing admits best effort packets only if there is available bandwidth on a link.

Note that BWIP, BWavg, and BWmax are specified per ON-DN pair, and that the QoS resource management method provides for a key priority VNET, a normal priority VNET, and a best effort VNET. Key services admitted by an ON on the key VNET are given higher priority routing treatment by allowing greater path selection DoS than normal services admitted on the normal VNET. Best effort services admitted on the best effort VNET are given lower priority routing treatment by allowing lesser path selection DoS than normal. Note that these designations of key, normal, and best effort are connection level priorities, whereas packet-level priorities are discussed in clause 9. The quantities $BWavg_i$ are computed periodically, such as every week w, and can be exponentially averaged over a several week period, as follows:

$$BWavg_i(w) = .5 \times BWavg_i(w-1) + .5 \times [\, BWIPavg_i(w) + BWOVavg_i(w) \,]$$

$BWIPavg_i$ = average bandwidth-in-progress across a load set period on VNET i

$BWOVavg_i$ = average bandwidth allocation request rejected (or overflow) across a load set period on VNET i

where all variables are specified per ON-DN pair, and where $BWIP_i$ and $BWOV_i$ are averaged across various load set periods, such as morning, afternoon, and evening averages for weekday, Saturday and Sunday, to obtain $BWIPavg_i$ and $BWOVavg_i$.

Table 2/E.360.3 – Determination of link load state (meshed network)

Link load state		Condition
Busy	B	$ILBW_k < DBW$
Reserved	R	$ILBW_k \le Rthr_k$
Heavily Loaded	HL	$Rthr_k < ILBW_k \le HLthr_k$
Lightly Loaded	LL	$HLthr_k < ILBW_k$

where:

$ILBW_k$ = idle link bandwidth on link k

DBW = delta bandwidth requirement for a bandwidth allocation request

$Rthr_k$ = reservation bandwidth threshold for link k

= $N \times .05 \times TBW_k$ for bandwidth reservation level N

$HLthr_k$ = heavily loaded bandwidth threshold for link k

= $Rthr_k + .05 \times TRBW_k$

$TRBW_k$ = the total bandwidth required on link k to meet the blocking/delay probability grade-of-service objective for bandwidth allocation requests on their first choice CRLSP

QoS resource management implements bandwidth reservation logic to favor connections routed on the first choice CRLSP in situations of link congestion. If link congestion (or blocking/delay) is detected, bandwidth reservation is immediately triggered and the reservation level N is set for the link according to the level of link congestion. In this manner, bandwidth allocation requests attempting to alternate-path over a congested link are subject to bandwidth reservation, and the first choice CRLSP requests are favored for that link. At the same time, the LL and HL link state thresholds are raised accordingly in order to accommodate the reserved bandwidth capacity N for the VNET. Figure 4 illustrates bandwidth allocation and the mechanisms by which bandwidth is protected through bandwidth reservation. Under normal bandwidth allocation demands bandwidth is fully shared, but under overloaded bandwidth allocation demands, bandwidth is protected through the reservation mechanisms wherein each VNET can use its allocated bandwidth. Under failure, however, the reservation mechanisms operate to give the key VNET its allocated bandwidth before the normal priority VNET gets its bandwidth allocation. As noted in Table 1, the best effort low-priority VNET is not allocated bandwidth nor is bandwidth reserved for the best effort VNET. Further illustrations are given in clause 7 of the robustness of dynamic bandwidth reservation in protecting the preferred bandwidth requests across wide variations in traffic conditions.

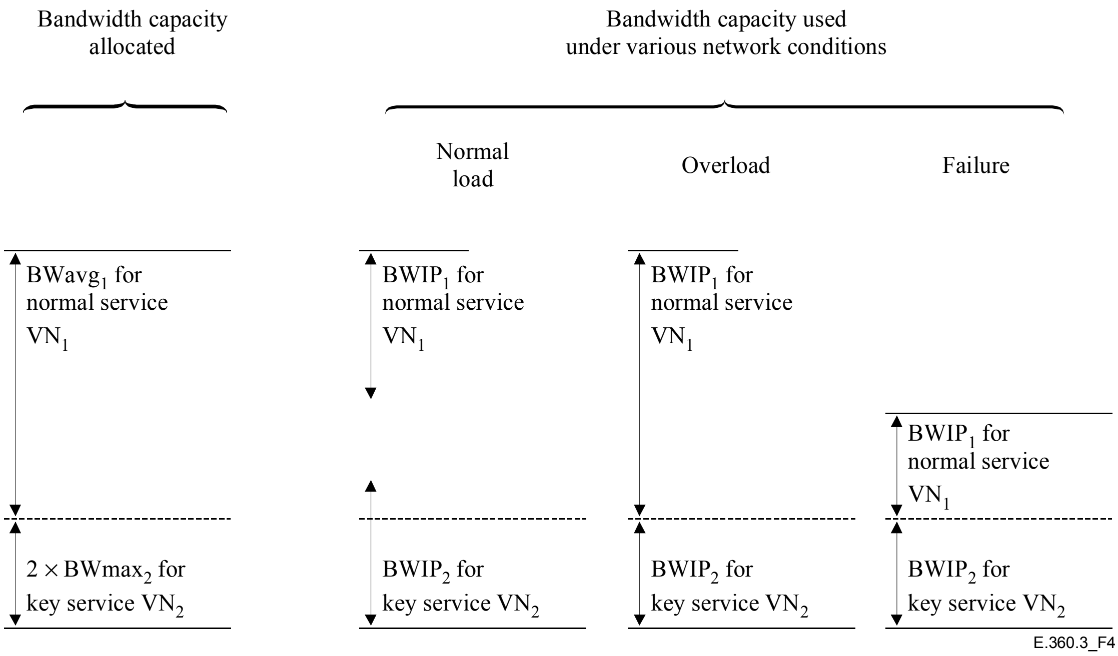


Figure 4/E.360.3 – Bandwidth allocation, protection and priority routing

The reservation level N (for example, N may have 1 of 4 levels), is calculated for each link k based on the link blocking/delay level of bandwidth allocation requests. The link blocking/delay level is equal to the total requested but rejected (or overflow) link bandwidth allocation (measured in total bandwidth), divided by the total requested link bandwidth allocation, over the last periodic update interval, which is, for example, every three minutes. That is:

$BWOV_k$ = total requested bandwidth allocation rejected (or overflow) on link k

$BWOF_k$ = total requested or offered bandwidth allocation on link k

LBL_k = link blocking/delay level on link k

$= BWOV_k/BWOF_k$

If LBL_k exceeds a threshold value, the reservation level N is calculated accordingly. The reserved bandwidth and link states are calculated based on the total link bandwidth required on link k, $TRBW_k$, which is computed on-line, for example every 1-minute interval m, and approximated as follows:

$TRBW_k(m) = .5 \times TRBW_k(m-1) + .5 \times [\ 1.1 \times TBWIP_k(m) + TBWOV_k(m)]$

$TBWIP_k$ = sum of the bandwidth in progress ($BWIP_i$) for all VNETs i for bandwidth requests on their first choice CRLSP over link k

$TBWOV_k$ = sum of bandwidth overflow ($BWOV_i$) for all VNETs i for bandwidth requests on their first choice CRLSP over link k

Therefore, the reservation level and load state boundary thresholds are proportional to the estimated required bandwidth load which means that the bandwidth reserved, and the bandwidth required to constitute a lightly loaded link, rise and fall with the bandwidth load as, intuitively, they should.

7.2 Per-VNET bandwidth allocation/reservation – Sparse network case

Here we discuss a sparse network example of bandwidth allocation/reservation. For the sparse network case of bandwidth reservation, a simpler method is illustrated which takes advantage of the concentration of traffic onto fewer, higher capacity backbone links. A small, fixed level of bandwidth reservation is used and permanently enabled on each link, as follows:

The DoS load state threshold again is a function of bandwidth-in-progress, VNET priority, and bandwidth allocation thresholds, however, only the reserved (R) and non-reserved (NR) states are used, as follows:

Table 3/E.360.3 – Determination of Depth-of-Search (DoS) load state threshold (Per-VNET bandwidth allocation, sparse network)

Load state allowed_i	Key priority VNET	Normal priority VNET		Best effort priority VNET
		First choice CRLSP	Alternate CRLSP	
R	If $BWIP_i \leq 2 \times BWmax_i$	If $BWIP_i \leq BWavg_i$	Not Allowed	Note
NR	If $2 \times BWmax_i < BWIP_i$	If $BWavg_i < BWIP_i$	If $BWavg_i < BWIP_i$	Note

where:

$BWIP_i$ = bandwidth-in-progress on VNET i

$BWavg_i$ = minimum guaranteed bandwidth required for VNET i to carry the average offered bandwidth load

$BWmax_i$ = the bandwidth required for VNET i to meet the blocking/delay probability grade-of-service objective for CRLSP bandwidth allocation requests

= $1.1 \times BWavg_i$

NOTE – CRLSPs for the best effort priority VNET are allocated zero bandwidth; Diffserv queuing admits best effort packets only if there is available bandwidth on a link.

The corresponding load state table for the sparse network case is as follows:

Table 4/E.360.3 – Determination of link load state (Sparse network)

Link load state		Condition
Busy	B	$ILBW_k < DBW$
Reserved	R	$ILBW_k - RBWr_k < DBW$
Not Reserved	NR	$DBW \leq ILBW_k - RBWr_k$

where:

$ILBW_k$ = idle link bandwidth on link k

DBW = delta bandwidth requirement for a bandwidth allocation request

$RBWr_k$ = reserved bandwidth for link k

= $.01 \times TLBW_k$

$TLBW_k$ = the total link bandwidth on link k

Note that reservation level is fixed and not dependent on any link blocking level (LBL) calculation or total required bandwidth (TRBW) calculation. Therefore, LBL and TRBW monitoring are not required in this example bandwidth allocation/protection method.

8 Per-flow bandwidth allocation, protection, and reservation

Per-flow QoS resource management methods have been applied successfully in TDM-based networks where bandwidth allocation is determined by edge nodes based on bandwidth demand for each connection request. Based on the bandwidth demand, these edge nodes make changes in bandwidth allocation using, for example, an SVC-based QoS resource management approach illustrated in this clause. Again, the determination of the link load states is used for QoS resource management in order to select network capacity on either the first choice path or alternate paths. Also the allowed DoS load state threshold determines if an individual connection request can be admitted on a given link to an available bandwidth "depth". In setting up each connection request, the ON encodes the DoS load state threshold allowed on each link in the connection-setup IE. If a link is encountered at a VN in which the idle link bandwidth and link load state are below the allowed DoS load state threshold, then the VN sends a crankback/bandwidth-not-available IE to the ON, which can then route the connection request to an alternate path choice. For example, in Figure 3, path A-B-E may be the first path tried where link A-B is in the LL state and link B-E is in the R state. If the DoS load state allowed is HL or better, then the connection request is routed on link A-B but will not be admitted on link B-E, wherein the connection request will be cranked back to the originating node A to try alternate path A-C-D-E. Here the connection request succeeds since all links have a state of HL or better.

8.1 Per-flow bandwidth allocation/reservation – Meshed network case

Here again, two approaches are illustrated for bandwidth allocation reservation: one applicable to meshed network topologies and the other applicable to sparse topologies. In meshed networks, a greater number of links results in less traffic carried per link, and functions, such as bandwidth reservation, need to be more carefully controlled than in a sparse network. In a sparse network the traffic is concentrated on much larger, and many fewer links, and here bandwidth reservation does not have to be as carefully management (such as automatically triggering bandwidth reservation on and off, dependent on the link/network congestion level).

Here we discuss a meshed network example of bandwidth allocation/reservation, and in 8.2 we discuss the sparse network case.

The illustrative DoS load state threshold is a function of bandwidth-in-progress, service priority, and bandwidth allocation thresholds, as follows:

Table 5/E.360.3 – Determination of Depth-of-Search (DoS) load state threshold (per-flow bandwidth allocation, meshed network)

Load state allowed$_i$	Key service	Normal service		Best effort service
		First choice path	Alternate path	
R	If $BWIP_i \le 2 \times BWmax_i$	If $BWIP_i \le BWavg_i$	Not Allowed	Not Allowed
HL	If $BWIP_i \le 2 \times BWmax_i$	If $BWIP_i \le BWmax_i$	If $BWIP_i \le BWavg_i$	Not Allowed
LL	All $BWIP_i$	All $BWIP_i$	All $BWIP_i$	All $BWIP_i$

where:

$BWIP_i$ = bandwidth-in-progress on VNET i

$BWavg_i$ = minimum guaranteed bandwidth required for VNET i to carry the average offered bandwidth load

$BWmax_i$ = the bandwidth required for VNET i to meet the blocking/delay probability grade-of-service objective

$= 1.1 \times BWavg_i$

Note that all parameters are specified per ON-DN pair, and that the QoS resource management method provides for key service and best effort service. Key services are given higher priority routing treatment by allowing greater path selection DoS than normal services. Best effort services are given lower priority routing treatment by allowing lesser path selection DoS than normal. The quantities $BWavg_i$ are computed periodically, such as every week w, and can be exponentially averaged over a several-week period, as follows:

$$BWavg_i(w) = .5 \times BWavg_i(w-1) + .5 \times [\,BWIPavg_i(w) + BWOVavg_i(w)]$$

$$BWIPavg_i = \text{average bandwidth-in-progress across a load set period on VNET } i$$

$$BWOVavg_i = \text{average bandwidth overflow across a load set period}$$

where $BWIP_i$ and $BWOV_i$ are averaged across various load set periods, such as morning, afternoon, and evening averages for weekday, Saturday, and Sunday, to obtain $BWIPavg_i$ and $BWOVavg_i$. Illustrative values of the thresholds to determine link load states are given in Table 2.

The illustrative QoS resource management method implements bandwidth reservation logic to favor connections routed on the first choice path in situations of link congestion. If link blocking/delay is detected, bandwidth reservation is immediately triggered and the reservation level N is set for the link according to the level of link congestion. In this manner, traffic attempting to alternate-route over a congested link is subject to bandwidth reservation, and the first choice path traffic is favored for that link. At the same time, the LL and HL link state thresholds are raised accordingly in order to accommodate the reserved bandwidth capacity for the VNET. The reservation level N (for example, N may have 1 of 4 levels), is calculated for each link k based on the link blocking/delay level and the estimated link traffic. The link blocking/delay level is equal to the equivalent bandwidth overflow count divided by the equivalent bandwidth peg count over the last periodic update interval, which is typically three minutes. That is:

$$BWOV_k = \text{equivalent bandwidth overflow count on link } k$$

$$BWPC_k = \text{equivalent bandwidth peg count on link } k$$

$$LBL_k = \text{link blocking/delay level on link } k$$

$$= BWOV_k / BWPC_k$$

If LBL_k exceeds a threshold value, the reservation level N is calculated accordingly. The reserved bandwidth and link states are calculated based on the total link bandwidth required on link k, TBW_k, which is computed on-line, for example every 1-minute interval m, and approximated as follows:

$$TBW_k(m) = .5 \times TBW_k(m-1) + .5 \times [1.1 \times TBWIP_k(m) + TBWOV_k(m)]$$

$TBWIP_k$ = sum of the bandwidth in progress ($BWIP_i$) for all VNETs i for connections on their first choice path over link k

$TBWOV_k$ = sum of bandwidth overflow ($BWOV_i$) for all VNETs i for connections on their first choice path over link k

Therefore, the reservation level and load state boundary thresholds are proportional to the estimated required bandwidth traffic load which means that the bandwidth reserved, and the bandwidth required to constitute a lightly loaded link, rise and fall with the traffic load, as, intuitively, they should.

8.2 Per-Flow bandwidth allocation/reservation – Sparse network case

Here we discuss a sparse network example of bandwidth allocation/reservation. For the sparse network case of bandwidth reservation, a simpler method is illustrated which takes advantage of the concentration of traffic onto fewer, higher capacity backbone links. A small, fixed level of bandwidth reservation is used on each link, as follows:

The DoS load state threshold again is a function of bandwidth-in-progress, VNET priority, and bandwidth allocation thresholds: however, only the reserved (R) and non-reserved (NR) states are used, as follows:

Table 6/E.360.3 – Determination of Depth-of-Search (DoS) load state threshold (per-flow bandwidth allocation, sparse network)

Load state allowed$_i$	Key priority VNET	Normal priority VNET		Best effort priority VNET
		First choice path	Alternate path	
R	If $BWIP_i \le 2 \times BWmax_i$	If $BWIP_i \le BWavg_i$	Not Allowed	Note
NR	If $2 \times BWmax_i < BWIP_i$	If $BWavg_i < BWIP_i$	If $BWavg_i < BWIP_i$	Note

where:

$BWIP_i$ = bandwidth-in-progress on VNET i

$BWavg_i$ = minimum guaranteed bandwidth required for VNET i to carry the average offered bandwidth load

$BWmax_i$ = the bandwidth required for VNET i to meet the blocking/delay probability grade-of-service objective

= $1.1 \times BWavg_i$

NOTE – CRLSPs for the best effort priority VNET are allocated zero bandwidth; Diffserv queuing admits best effort packets only if there is available bandwidth on a link.

The corresponding load state table for the sparse network case is as follows:

Table 7/E.360.3 – Determination of link load state (sparse network)

Link load state		Condition
Busy	B	$ILBW_k < EQBW$
Reserved	R	$ILBW_k - RBWr_k < EQBW$
Not Reserved	NR	$EQBW \le ILBW_k - RBWr_k$

where:

$ILBW_k$ = idle link bandwidth on link k

EQBW = equivalent bandwidth requirement for a bandwidth allocation request

$RBWr_k$ = reserved bandwidth for link k

= $.01 \times TLBW_k$

$TLBW_k$ = the total link bandwidth on link k

Note that reservation level is fixed and not dependent on any link blocking level (LBL) calculation or total required bandwidth (TRBW) calculation. Therefore LBL and TRBW monitoring are not required in this example.

9 Packet-level traffic control

Traffic controls may be distinguished according to whether their function is to enable quality of service guarantees at packet level (e.g. packet loss ratio) or at connection level (e.g. connection blocking probability). Connection-level controls are already covered in clauses 5 to 8.

In a connection-oriented network, each connection request is specified by a traffic descriptor, delay variation tolerance and QoS requirements. The source traffic descriptor is a list of traffic parameters which should:

a) be understandable and conformance should be possible;

b) be used in resource allocation meeting network performance requirements; and

c) be enforceable by the usage parameter control and network parameter control.

The traffic parameters may relate explicitly to connection traffic characteristics such as the peak data rate or implicitly define these characteristics by reference to a service type. End-to-end packet level QoS criteria use the following performance parameters:

a) transfer delay;

b) delay variation;

c) loss ratio.

End-to-end performance objectives relevant to traffic engineering are as follows:

a) maximum end-to-end queuing delay;

b) mean queuing delay;

c) packet loss ratio.

These performance objectives must be apportioned to the various network elements contributing to the performance degradation of a given connection so that the end-to-end QoS criteria are satisfied.

When the establishment of a new connection is requested, the network must decide if it has sufficient resources to accept it without infringing packet level GoS requirements for all established connections, as well as the new connection. This is the function of connection admission control (CAC) which determines if a link or path connection is capable or not of handling the requested connection. This decision can sometimes be made by allocating resources to specific connections or groups of connections and refusing new requests when insufficient resources are available. Note that the allocation is generally logical: no particular physical resources are attributed to a specific connection. The resources in question are typically bandwidth and buffer space. It is assumed that resources are allocated independently for each link or path connection with a separate decision made for each transmission direction of a connection. A connection will be established only if resources are available on every link of its path, in both directions. Admission control could be applied for peak rate allocation wherein a network operator may choose to apply an overbooking factor. It is possible to base admission control on an equivalent bandwidth: connection i is attributed an equivalent bandwdith EQBW_i and connections are admitted while $\Sigma\ \mathrm{EQBW}_i < c$, where c is the link bandwidth. If the only performance requirement consists in guaranteeing a minimum bandwidth, EQBW_i may be set equal to this minimum bandwidth. (As discussed in ITU-T Rec. E.736, when there is more than one packet-level priority for different services, there should be multiple constraints for the different equivalent bandwidths, with one constraint for each priority level. In practice, when applying this model to a network with key, normal, and best-effort priorities, the following simplification can be made. Assuming that key-priority traffic does not consume a major portion of the link bandwidth and zero equivalent bandwidth for best-effort traffic, then the set of three constraints for the three priority levels can be reduced to a single constraint, $\Sigma\ \mathrm{EQBW}_i < c$.)

Packet level traffic control encompasses the control procedures which allow packet level GoS objectives to be fulfilled [E.736]. Once a flow is admitted through the connection admission control (CAC) functions, packet level control:

a) ensures, through traffic policing (e.g. usage parameter control), that the user in fact emits traffic in conformity with the declared traffic parameters;

b) ensures, through packet priority and queue management, that the network provides the differentiation of quality of service according to service requirements.

If the connection is accepted, there is an implicitly defined traffic contract whereby the network operator provides the requested quality of service on condition that the user emits traffic in conformity with the declared traffic descriptor; this is the role of usage parameter control. When more than one network is involved in a connection, it is also incumbent on each network to verify that the traffic it receives from the neighboring network conforms; this is network parameter control. One of the requirements on traffic parameters is that they be enforceable by the usage parameter control and network parameter control. This has led to a definition of traffic parameters: peak rate, sustainable rate and intrinsic burst tolerance allowing user conformance to be determined by the generic leaky bucket algorithm.

Users, or networks, may introduce supplementary packet delays to shape the characteristics of a given flow. By smoothing packet rate variations, shaping generally allows an increase in the utilization of network resources leading to greater multiplexing gains. On the other hand, shaping may introduce non-negligible delays and a part of the end-to-end GoS objective must be allocated to the shaper. Shaping may be performed by the user to ensure compliance with declared traffic parameters and delay variation tolerance. The network operator may employ shaping at the network entrance, within the network, or at the network egress (to meet constraints on output traffic characteristics). Shaping is an option for users and networks. A particular example of shaping is the reduction of delay variation by means of packet spacing. The spacer tries to produce a packet stream with a time between consecutive packets at least equal to the inverse of the peak data rate by imposing a variable delay on each packet.

We now discuss priority queuing as an illustrative traffic scheduling method, and further assume that a traffic policing function is employed, as discussed above, such as a leaky-bucket model to determine out-of-contract traffic behavior, and appropriately mark packets for possible dropping under congestion. These scheduling and policing mechanisms compliment the connection admission mechanisms, described in the previous clauses to appropriately allocate bandwidth on links in the network.

Note that priority queuing is used as an illustrative scheduling mechanism, whereas other methods may be used. DiffServ does not require that a particular queuing mechanism be used to achieve EF, AF, etc. QoS. Therefore, the queuing implementation used for DiffServ could be weighted fair queuing (WFQ), priority queuing (PQ), or other queuing mechanism, depending on the choice in the implementation. In the analysis, PQ is used for illustration, however, the same or comparable results would be obtained with WFQ or other queuing mechanisms.

In addition to the QoS bandwidth management procedure for bandwidth allocation requests, a QoS priority of service queuing capability is used during the time connections are established on each of the three VNETs. At each link, a queuing discipline is maintained such that the packets being served are given priority in the following order: key VNET services, normal VNET services, and best effort VNET services. Following the MPLS CRLSP bandwidth allocation setup and the application of QoS resource management rules, the priority of service parameter and label parameter need to be sent in each IP packet, as illustrated in Figure 5. The priority of service parameter may be included in the type of service (ToS), or differentiated services (DiffServ) [RFC2475], [LDVKCH00], [ST98], parameter already in the IP packet header. Another possible alternative is that the priority of service parameter can be associated with the MPLS label appended to the IP packet [LDVKCH00]. In either case, from the priority of service parameters, the IP node can determine the QoS treatment based on the QoS resource management (priority queuing) rules for key VNET packets, normal VNET packets, and best effort VNET packets. From the label parameter, the IP node can determine the next node to route the IP packet to, as defined by the MPLS protocol. In this way, the backbone nodes can have a very simple per-packet processing implementation to implement QoS resource management and MPLS routing.

IP payload	IP header (contains ToS/DIFFSERV QoS parameter)	LDP label (contains MPLS routing parameters)

E.360.3_F5

DIFFSERV	Differentiated services
IP	Internet protocol
LDP	Label distribution protocol
MPLS	Multiprotocol label switching
QoS	Quality of service
ToS	Type of service

Figure 5/E.360.3 – IP packet structure under MPLS switching

10 Other QoS resource management constraints

Other QoS routing constraints are taken into account in the QoS resource management and route selection methods in addition to bandwidth allocation, bandwidth protection and priority routing. These include end-to-end transfer delay, delay variation [G99a], and transmission quality considerations such as loss, echo, and noise [D99], [G99a], [G99b]. Additionally, link capability (LC) selection allows connection requests to be routed on specific transmission media that have the particular characteristics required by these connection requests. In general, a connection request can require, prefer, or avoid a set of transmission characteristics such as fiber optic or radio transmission, satellite or terrestrial transmission, or compressed or uncompressed transmission. The routing table logic allows the connection request to skip links that have undesired characteristics and to seek a best match for the requirements of the connection request. For any SI, a set of LC selection preferences is specified for the connection request. LC selection preferences can override the normal order of selection of paths. If a LC characteristic is required, then any path with a link that does not have that characteristic is skipped. If a characteristic is preferred, paths having all links with that characteristic are used first. Paths having links without the preferred characteristic will be used next. A LC preference is set for the presence or absence of a characteristic. For example, if fiberoptic transmission is required, then only paths with links having Fiberoptic = Yes are used. If we prefer the presence of fiberoptic transmission, then paths having all links with Fiberoptic = Yes are used first, then paths having some links with Fiberoptic = No.

11 Interdomain QoS resource management

In current practice, interdomain routing protocols generally do not incorporate standardized path selection or per class-of-service QoS resource management. For example, in IP-based networks BGP [RL00] is used for interdomain routing but does not incorporate per class-of-service resource allocation as described in this clause. Also, MPLS techniques have not yet been addressed for interdomain applications. Extensions to interdomain routing methods discussed in this clause can, therefore, be considered to extend the call routing and connection routing concepts to routing between network domains.

Interdomain routing can also apply class-of-service routing concepts described in clause 5 and increased routing flexibility for interdomain routing. Principles discussed in clause 5 for class-of-service derivation and policy-based routing table derivation also apply in the case of interdomain QoS resource management. As described in ITU-T Rec. E.360.2, interdomain routing works synergistically with multiple ingress/egress routing and alternate routing through transit domains. Interdomain routing can use link status information in combination with call completion history to select paths and also use dynamic bandwidth reservation techniques, as discussed in clauses 6 to 8.

Interdomain routing can use the virtual network concept that enables service integration by allocating bandwidth for services and using dynamic bandwidth reservation controls. These virtual network concepts have been described in this Recommendation, and can be extended directly to interdomain routing. For example, the links connected to the originating domain gateway nodes, such as links OGN1-DGN1, OGN2-DGN1, OGN1-VGN1, OGN1-VGN2, and OGN2-VGN2 in Figure 5/E.360.2, can define VNET bandwidth allocation, protection, reservation, and routing methods, exactly as discussed in clauses 6 to 8. In that way, bandwidth can be fully shared among virtual networks in the absence of congestion. When a certain virtual network encounters congestion, bandwidth is reserved to ensure that the virtual network reaches its allocated bandwidth. Interdomain routing can employ class-of-service routing capabilities including key service protection, directional flow control, link selection capability, automatically updated time-variable bandwidth allocation, and alternate routing capability through the use of overflow paths and control parameters such as interdomain routing load set periods. Link selection capability allows specific link characteristics, such as fiber transmission, to be preferentially selected. Thereby, interdomain routing can improve performance and reduce the cost of the interdomain network with flexible routing capabilities, such as described in ITU-T Rec. E.360.2.

Similar to intradomain routing, interdomain routing may include the following steps for call establishment:

– At the originating gateway node (OGN), the destination gateway node (DGN), SI, VNET, and QoS resource management information are determined through the number/name translation database and other service information available at the OGN.
– The DGN and QoS resource management information are used to access the appropriate VNET and routing table between the OGN and DGN.
– The connection request is set up over the first available path in the routing table with the required transmission resource selected based on the QoS resource management data.

The rules for selecting the interdomain primary path and alternate paths for a call can be governed by the availability of primary path bandwidth, node-to-node congestion, and link capability, as described in clauses 6 to 8. The path sequence consists of the primary shortest path, lightly loaded alternate paths, heavily loaded alternate paths, and reserved alternate paths, where these load states are further refined by combining link load state information with path congestion state information. Interdomain alternate paths, which include nodes in the originating domain and terminating domain, are selected before alternate paths, which include transit domain nodes, are selected. As described in clauses 7 and 8, greater path selection depth is allowed if congestion is detected to the destination network domain, because more alternate path choices serve to reduce the congestion. During periods of no congestion, capacity not needed by one virtual network is made available to other virtual networks that are experiencing loads above their allocation.

The gateway node, for example, may automatically compute the bandwidth allocations once a week and may use a different allocation for various load set periods, for example each of 36 two-hour load set periods: 12 weekday, 12 Saturday, and 12 Sunday. The allocation of the bandwidth can be based on a rolling average of the traffic load for each of the virtual networks, to each destination node, in each of the load set periods. Under normal network conditions in which there is no congestion, all virtual networks fully share all available capacity. Under call overload, however, link bandwidth is reserved to ensure that each virtual network gets the amount of bandwidth allotted. This dynamic bandwidth reservation during times of overload results in network performance that is analogous to having the link bandwidth allocation between the two nodes dedicated for each VNET.

12 Conclusions/recommendations

The conclusions/recommendations reached in this Recommendation are as follows:

- QoS resource management is recommended and is shown to be effective in achieving connection-level and packet-level GoS objectives, as well as key service, normal service, and best effort service differentiation.
- Admission control is recommended and is the basis that allows for applying most of the other controls described in this Recommendation.
- Bandwidth reservation is recommended and is critical to the stable and efficient performance of TE methods in a network, and to ensure the proper operation of multiservice bandwidth allocation, protection and priority treatment.
- Per-VNET bandwidth allocation is recommended and is essentially equivalent to per-flow bandwidth allocation in network performance and efficiency. Because of the much lower routing table management overhead requirements, as discussed and modeled in ITU-T Rec. E.360.4, per-VNET bandwidth allocation is preferred to per-flow allocation.
- Both MPLS QoS and bandwidth management and DiffServ priority queuing management are recommended and are important for ensuring that multiservice network performance objectives are met under a range of network conditions. Both mechanisms operate together to ensure QoS resource allocation mechanisms (bandwidth allocation, protection, and priority queuing) are achieved.

Annex A

Modelling of traffic engineering methods

In this annex, we again use the full-scale national network model developed in ITU-T Rec. E.360.2 to study various TE scenarios and tradeoffs. The 135-node national model is illustrated in Figure A.1/E.360.2, the multiservice traffic demand model is summarized in Table A.1/E.360.2, and the cost model is summarized in Table A.2/E.360.2.

A.1 Performance of bandwidth reservation methods

As discussed in clauses 6, 7, and 8, dynamic bandwidth reservation can be used to favor one category of traffic over another category of traffic. A simple example of the use of this method is to reserve bandwidth in order to prefer traffic on the shorter primary paths over traffic using longer alternate paths. This is most efficiently done by using a method which reserves bandwidth only when congestion exists on links in the network. We now give illustrations of this method, and compare the performance of a network in which bandwidth reservation is used under congestion to the case when bandwidth reservation is not used.

In the example, traffic is first routed on the shortest path, and then allowed to alternate route on longer paths if the primary path in not available. In the case where bandwidth reservation is used, five percent of the link bandwidth is reserved for traffic on the primary path when congestion is present on the link.

Table A.1 illustrates the performance of bandwidth reservation methods for a high-day network load pattern. This is the case for multilink path routing being used to set up per-flow CRLSPs in a sparse network topology.

Table A.1/E.360.3 – Performance of dynamic bandwidth reservation methods for CRLSP setup – percent lost/delayed traffic under overload (per-flow multilink path routing in sparse network topology; 135-node multi-service network model)

OverloadFactor	Without bandwidth reservation	With bandwidth reservation
7	11.94	3.86
8	22.85	9.66
10	37.74	24.78

We can see from the results of Table A.1 that performance improves when bandwidth reservation is used. The reason for the poor performance without bandwidth reservation is due to the lack of reserved capacity to favor traffic routed on the more direct primary paths under network congestion conditions. Without bandwidth reservation nonhierarchical networks can exhibit unstable behavior in which essentially all connections are established on longer alternate paths as opposed to shorter primary paths, which greatly reduces network throughput and increases network congestion [Aki84], [Kru82], [NaM73]. If we add the bandwidth reservation mechanism, then performance of the network is greatly improved.

Another example is given in Table A.2, where 2-link SDR is used in a meshed network topology. In this case, the average business day loads for a 65-node national network model were inflated uniformly by 30 percent [A98]. The Table A.2 gives the average hourly lost traffic due to blocking of connection admissions in hours 2, 3, and 5, which correspond to the two early morning busy hours and the afternoon busy hour.

Table A.2/E.360.3 – Performance of dynamic bandwidth reservation methods – percent lost traffic under 30% overload (per-flow 2-link SDR in meshed network topology; 65-node network model)

Hour	Without bandwidth reservation	With bandwidth reservation
2	12.19	0.22
3	22.38	0.18
5	18.90	0.24

Again, we can see from the results of Table A.2 that performance dramatically improves when bandwidth reservation is used. A clear instability arises when bandwidth reservation is not used, because, under congestion, a network state in which virtually all traffic occupies 2 links instead of 1 link is predominant. When bandwidth reservation is used, flows are much more likely to be routed on a 1-link path, because the bandwidth reservation mechanism makes it less likely that a 2-link path can be found in which both links have idle capacity in excess of the reservation level.

A performance comparison is given in Table A.3 for a single link failure in a 135-node design averaged over 5 network busy hours, for the case without bandwidth reservation, and with bandwidth reservation. Clearly, the use of bandwidth reservation protects the performance of each virtual network class-of-service category.

Table A.3/E.360.3 – Performance of dynamic bandwidth reservation methods – percent lost/delayed traffic under 50% general overload (multilink STT-EDR; 135-node network model)

Virtual network	Without bandwidth reservation	With bandwidth reservation
Business-voice	2.42	0.00
Consumer-voice	2.33	0.02
INTL-out	2.46	1.33
INTL-in (key)	2.56	0.00
Key voice	2.41	0.00
64-kbit/s ISDN data	2.37	0.10
64-kbit/s ISDN data (key)	2.04	0.00
384-kbit/s ISDN data	12.87	0.00
VBR-RT voice	1.25	0.07
VBR-NRT MM	1.90	0.01
UBR MM	24.95	11.15

A.2 Multiservice network performance: Per-VNET vs. Per-flow bandwidth allocation

Here we use the 135-node model to compare the per-virtual-network methods of QoS resource management, as described in clause 7, and the per-flow methods described in clause 8. We look at these two cases in Figure A.1, which illustrates the case of per-virtual-network CRLSP bandwidth allocation the case of per-flow CRLSP bandwidth allocation. The two figures compare the performance in terms of lost or delayed traffic under a focused overload scenario on the Oakbrook (OKBK), IL node (such as might occur, for example, with a radio call-in give-away offer). The size of the focused overload is varied from the normal load (1 × case) to a 10 times overload of the traffic to OKBK (10 × case). Here a fixed routing (FR) CRLSP bandwidth allocation is used for both the per-flow CRLSP bandwidth allocation case and the per-virtual-network bandwidth allocation case. The results show that the per-flow and per-virtual-network bandwidth allocation performance is similar; however, the improved performance of the key priority traffic and normal priority traffic in relation to the best-effort priority traffic is clearly evident.

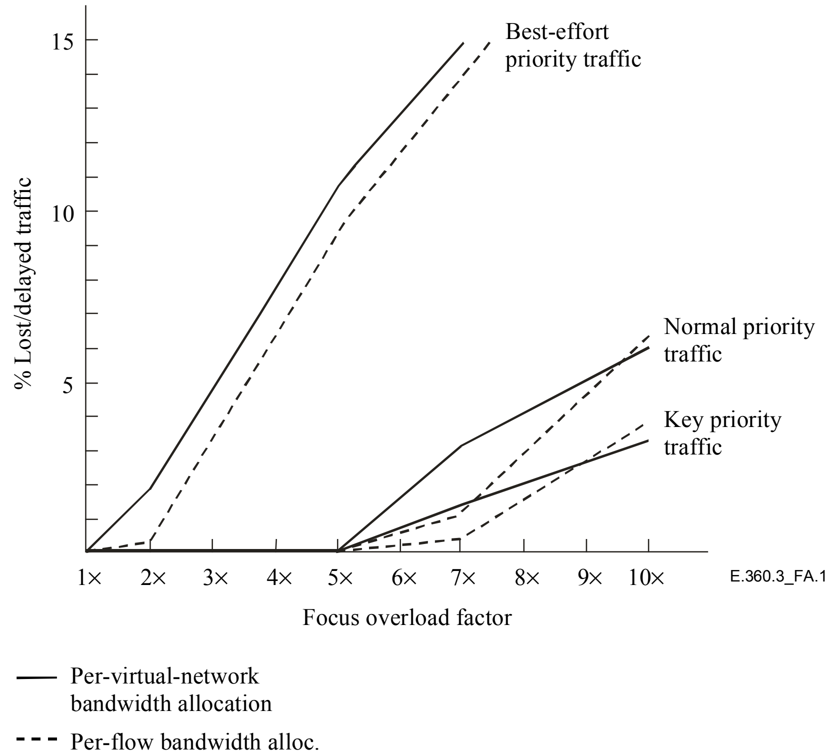


Figure A.1/E.360.3 – Performance under focused overload on OKBR node

The performance analyses for overloads and failures for the per-flow and per-virtual-network bandwidth allocation are now examined in which event dependent routing (EDR), with success-to-the-top (STT) path selection, are used. Again, the simulations include call admission control with QoS resource management in which we distinguish the key services, normal services, and best-effort services, as indicated in the tables below. Table A.4 gives performance results for a 30% general overload, Table A.5 gives performance results for a six-times overload on a single network node, and Table A.6 gives performance results for a single transport link failure. Performance analysis results show that the multilink STT-EDR per-flow bandwidth allocation and per-virtual-network bandwidth allocation options perform similarly under overloads and failures.

Table A.4/E.360.3 – Performance of per-flow and per-virtual-network bandwidth allocation – percent lost/delayed traffic under 30% general overload (single-area flat network topology; multilink STT-EDR routing; 135-node network model)

Virtual network	Per-flow bandwidth allocation	Per-virtual-network bandwidth allocation
Business-voice	0.00	0.00
Consumer-voice	0.00	0.00
INTL-out	0.00	0.00
INTL-in (key)	0.00	0.00
Key voice	0.00	0.00
64-kbit/s SDS	0.00	0.00
64-kbit/s ISDN data (key)	0.00	0.00
384-kbit/s ISDN data	0.00	0.00
VBR-RT voice	0.00	0.00
VBR-NRT MM	0.00	0.00
UBR MM	4.15	3.94

Table A.5/E.360.3 – Performance of Per-flow and per-virtual-network bandwidth allocation – percent lost/delayed traffic under 6× focused overload on OKBK (single-area flat network topology; multilink STT-EDR routing; 135-node network model)

Virtual network	Per-flow bandwidth allocation	Per-virtual-network bandwidth allocation
Business-voice	0.00	0.01
Consumer-voice	0.00	0.01
INTL-out	0.00	0.01
INTL-in (key)	0.00	0.00
Key voice	0.00	0.00
64-kbit/s ISDN data	0.00	0.00
64-kbit/s ISDN data (key)	0.00	0.00
384-kbit/s ISDN data	0.00	0.00
VBR-RT voice	0.00	0.00
VBR-NRT MM	0.00	0.01
UBR MM	12.46	12.30

Table A.6/E/360.3 – Performance of per-flow and per-virtual-network bandwidth allocation – percent lost/delayed traffic under failure on CHCG-NYCM link (single-area flat network topology; multilink STT-EDR routing; 135-node network model)

Virtual network	Per-flow bandwidth allocation	Per-virtual-network bandwidth allocation
Business-voice	0.00	0.00
Consumer-voice	0.00	0.00
INTL-out	0.00	0.00
INTL-in (key)	0.00	0.00
Key voice	0.00	0.00
64-kbit/s ISDN data	0.00	0.00
64-kbit/s ISDN data (key)	0.00	0.00
384-kbit/s ISDN data	0.00	0.00
VBR-RT voice	0.00	0.00
VBR-NRT MM	0.00	0.00
UBR MM	0.17	0.17

A.3 Multiservice network performance: single-area flat topology vs. Multi-area 2-level hierarchical flat topology

We also investigate the performance of hierarchical network designs, which represent the topological configuration to be expected with multi-area (or multi-autonomous-system (multi-AS), or multi-domain) networks. In Figure A.5/E.360.2 we show the model considered, which consists of 135 edge nodes each homed onto one of 21 backbone nodes. Typically, the edge nodes may be grouped into separate areas or autonomous systems, and the backbone nodes into another area or autonomous system. Within each area a flat routing topology exists, however between edge areas and the backbone area a hierarchical routing relationship exists. This routing hierarchy is modeled for both the per-flow and per-virtual-network bandwidth allocation examples, and the results are given in Tables A.7 to A.9 for the 30% general overload, 6-times focused overload, and link failure examples, respectively. We can see that the performance of the hierarchical network case is substantially worse than the flat network model, which models a single area or autonomous system consisting of 135 nodes.

Table A.7/E.360.3 – Performance of multi-area 2-level hierarchical network topology – percent lost/delayed traffic under 30% general overload per-flow and per-virtual-network bandwidth allocation (multilink STT-EDR routing; 135-node network model)

Virtual network	Per-flow bandwidth allocation	Per-virtual-network bandwidth allocation
Business-voice	0.00	0.00
Consumer-voice	0.00	0.00
INTL-out	0.00	0.00
INTL-in (key)	0.00	0.00
Key voice	0.00	0.00
64-SDS	0.00	0.00
64-kbit/s ISDN data (key)	0.00	0.00
384-kbit/s ISDN data	0.00	0.00
VBR-RT voice	0.00	0.00
VBR-NRT MM	0.00	0.00
UBR MM	9.88	9.06

Table A.8/E.360.3 – Performance of multi-area 2-level hierarchical network topology – percent lost/delayed traffic under 6× focused overload on OKBK per-flow and per-virtual-network bandwidth allocation (multilink STT-EDR routing; 135-node network model)

Virtual network	Per-flow bandwidth allocation	Per-virtual-network bandwidth allocation
Business-voice	1.64	1.70
Consumer-voice	2.27	2.22
INTL-out	1.11	0.89
INTL-in (key)	0.00	0.00
Key voice	0.00	0.00
64-kbit/s ISDN data	0.40	0.27
64-kbit/s ISDN data (key)	0.00	0.00
384-kbit/s ISDN data	0.00	0.00
VBR-RT voice	0.94	0.93
VBR-NRT MM	1.85	1.80
UBR MM	12.86	12.88

Table A.9/E.360.3 – Performance of multi-area 2-level hierarchical network topology – percent lost/delayed traffic under failure on CHCG-NYCM link per-flow and per-virtual-network bandwidth allocation (multilink STT-EDR routing; 135-node network model)

Virtual network	Per-flow bandwidth allocation	Per-virtual-network bandwidth allocation
Business-voice	0.00	0.00
Consumer-voice	0.00	0.00
INTL-out	0.00	0.00
INTL-in (key)	0.00	0.00
Key voice	0.00	0.00
64-kbit/s ISDN data	0.00	0.00
64-kbit/s ISDN data (key)	0.00	0.00
384-kbit/s ISDN data	0.00	0.00
VBR-RT voice	0.00	0.00
VBR-NRT MM	0.00	0.00
UBR MM	1.22	1.38

A.4 Multiservice network performance: Need for MPLS and DiffServ

We illustrate the operation of MPLS and DiffServ in the multiservice network model with some examples. First, suppose there is 10 Mbit/s of normal-priority traffic and 10 Mbit/s of best-effort priority traffic being carried in the network between node A and node B. Best-effort traffic is treated in the model as UBR traffic and is not allocated any bandwidth. Hence, the best-effort traffic does not get any CRLSP bandwidth allocation, and is not treated as MPLS forward equivalence class (FEC) traffic. As such, the best-effort traffic would be routed by the interior gateway protocol, or IGP, such as OSPF. Hence, the best-effort traffic cannot be denied bandwidth allocation as a means to throttle back such traffic at the edge router, which can be done with the normal-priority and key-priority traffic (i.e. normal and key traffic could be denied bandwidth allocation). The only way that the best-effort traffic gets dropped/lost is to drop it at the queues, therefore it is essential that the traffic that is allocated bandwidth on the CRLSPs have higher priority at the queues than the best-effort traffic. Therefore, in the model, the three classes of traffic get these DiffServ markings: best-effort get no-DiffServ marking, which ensures that it will get best-effort priority queuing treatment. Normal-priority traffic gets the assured forwarding (AF) DiffServ marking, which is a middle priority level of queuing treatment, and key-priority traffic gets the expedited forwarding (EF) DiffServ marking, which is the highest priority queuing level.

Now suppose that there is 30 Mbit/s of bandwidth available between A and B and that all the normal-priority and best-effort traffic is getting through. Now suppose that the traffic for both the normal-priority and best-effort traffic increases to 20 Mbit/s. The normal-priority traffic requests, and gets, a CRLSP bandwidth allocation increase to 20 Mbit/s on the A to B CRLSP. However, the best-effort traffic, since it has no CRLSP assigned and therefore no bandwidth allocation, is just sent into the network at 20 Mbit/s. Since there is only 30 Mbit/s of bandwidth available from A to B, the network must drop 10 Mbit/s of best-effort traffic in order to leave room for the 20 Mbit/s of normal-priroity traffic. The way this is done in the model is through the queuing mechanisms governed by the DiffServ priority settings on each category of traffic. Through the DiffServ marking, the queuing mechanisms in the model discard about 10 Mbit/s of the best-effort traffic at the priority queues. If the DiffServ markings were not used, then the normal-priority and best-effort traffic would compete equally on the first-in/first-out (FIFO) queues, and perhaps 15 Mbit/s of each would get through, which is not the desired situation.

Taking this example further, if the normal-priority and best-effort traffic both increase to 40 Mbit/s, then the normal-priority traffic tries to get a CRLSP bandwidth allocation increase to 40 Mbit/s. However, the most it can get is 30 Mbit/s, so 10 Mbit/s is denied for the normal-priority traffic in the MPLS constraint-based routing procedure. By having the DiffServ markings of AF on the normal-priority traffic and none on the best-effort traffic, essentially all the best-effort traffic is dropped at the queues since the normal-priority traffic is allocated and gets the full 30 Mbit/s of A to B bandwidth. If there were no DiffServ markings, then again perhaps 15 Mbit/s of both normal-priority and best-effort get through. Or, in this case, perhaps a greater amount of best-effort traffic is carried than normal-priority traffic, since 40 Mbit/s of best-effort traffic is sent into the network and only 30 Mbit/s of normal-priority traffic is sent into the network, and the FIFO queues will receive more best-effort pressure than normal-priority pressure.

Some general observations on the operation of MPLS and DiffServ in the multiservice TE models include the following:

1) In a multiservice network environment, with best-effort priority traffic (WWW traffic, email, …), normal-priority traffic (CBR voice, IP-telephony voice, switched digital service, …), and key-priority traffic (800-gold, incoming international, ...) sharing the same network, MPLS bandwidth allocation plus DiffServ/priority-queuing are both needed. In the models, the normal-priority and key-priority traffic use MPLS to receive bandwidth allocation, while the best-effort traffic gets no bandwidth allocation. Under congestion (e.g. from overloads or failures), the DiffServ/priority-queuing mechanisms push out the best-effort priority traffic at the queues so that the normal-priority and key-priority traffic can get through on the MPLS-allocated CRLSP bandwidth.
2) In a multiservice network where the normal-priority and key-priority traffic use MPLS to receive bandwidth allocation, and there is no best-effort priority traffic, the MPLS bandwidth allocation more-or-less assures that the queues will not overflow. In this case, DiffServ/priority-queuing can still be used to ensure packet-level performance.
3) As bandwidth gets more and more plentiful/lower-cost, the point at which the MPLS and DiffServ mechanisms have a significant effect under traffic overload goes to a higher and higher threshold. For example, the models show that the overload factor at which congestion occurs gets larger as the bandwidth modules get larger (i.e. OC3 to OC12 to OC48 to OC192, etc.). However, the congestion point will always be reached with failures and/or large-enough overloads necessitating the MPLS/DiffServ mechanisms.

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