INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.1221

(03/2002)

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE AND INTERNET PROTOCOL ASPECTS

Internet protocol aspects – Architecture, access, network capabilities and resource management

Traffic control and congestion control in IP based networks

ITU-T Recommendation Y.1221

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE AND INTERNET PROTOCOL ASPECTS

GLOBAL INFORMATION INFRASTRUCTURE	
General	Y.100–Y.199
Services, applications and middleware	Y.200–Y.299
Network aspects	Y.300–Y.399
Interfaces and protocols	Y.400–Y.499
Numbering, addressing and naming	Y.500–Y.599
Operation, administration and maintenance	Y.600–Y.699
Security	Y.700–Y.799
Performances	Y.800–Y.899
INTERNET PROTOCOL ASPECTS	
General	Y.1000–Y.1099
Services and applications	Y.1100–Y.1199
Architecture, access, network capabilities and resource management	**Y.1200–Y.1299**
Transport	Y.1300–Y.1399
Interworking	Y.1400–Y.1499
Quality of service and network performance	Y.1500–Y.1599
Signalling	Y.1600–Y.1699
Operation, administration and maintenance	Y.1700–Y.1799
Charging	Y.1800–Y.1899

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

ITU-T Recommendation Y.1221

Traffic control and congestion control in IP-based networks

Summary

This Recommendation provides a general description as well as objectives and procedures for traffic control and congestion control for IP-based networks. In particular, it describes the concepts of the traffic contract between a user and the network. It specifies the IP transfer capabilities (IPTCs) including, for each IPTC, the service model, the associated traffic patterns and conformance definition.

Source

ITU-T Recommendation Y.1221 was prepared by ITU-T Study Group 13 (2001-2004) and approved under the WTSA Resolution 1 procedure on 16 March 2002.

FOREWORD

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

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

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

NOTE

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

INTELLECTUAL PROPERTY RIGHTS

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

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

CONTENTS

Page

ITU-T Recommendation Y.1221

Traffic control and congestion control in IP-based networks

1 Scope

This Recommendation describes traffic control and congestion control procedures for IP-based networks. Such control procedures are necessary in order to support services with QoS, where the QoS is negotiated between a user and the network.

Traffic control refers to all network actions aiming to meet the negotiated performance objectives in an IP-based network and to allow the avoidance of congested conditions.

Congestion control refers to all network actions to minimize the intensity, spread and duration of congestion.

This Recommendation provides a general description as well as objectives and procedures for traffic control and congestion control. In particular, it describes the concepts of the traffic contract between a user and the network. It specifies the IP transfer capabilities (IPTCs) including, for each IPTC, the service model, the associated traffic patterns and conformance definition.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published.

2.1 Normative references

- ITU-T Recommendation I.371 (2000), *Traffic control and congestion control in B-ISDN.*
- ITU-T Recommendation Y.1241 (2001), *Support of IP-based services using IP Transfer Capabilities*.
- ITU-T Recommendation Y.1540 (1999), *Internet protocol communication service – IP packet transfer and availability performance parameters*.
- ITU-T Recommendation Y.1541 (2002), Network performance objectives for IP-based services.

2.2 Informative references

- IETF RFC 791 (1981), *Internet Protocol – DARPA Internet Program – Protocol Specification*.
- IETF RFC 1633 (1994), *Integrated Services in the Internet Architecture: an Overview*.
- IETF RFC 2210 (1997), *The use of RSVP with IETF Integrated Services*.
- IETF RFC 2211(1997), *Specification of the Controlled-Load Network Element Service*.
- IETF RFC 2212 (1997), *Specification of Guaranteed Quality of Services*.
- IETF RFC 2698 (1999), *A Two Rate Three Color Marker*.
- IETF RFC 2475 (1998), *An Architecture for Differentiated Services*.
- IETF RFC 2481 (1999), *A Proposal to Add Explicit Congestion Notification (ECN) to IP*.
- IETF RFC 2597 (1999), *Assured Forwarding PHB Group*.
- IETF RFC 2598 (1999), *An Expedited Forwarding PHB*.

3 Abbreviations and terminology

3.1 Abbreviations

This Recommendation uses the following abbreviations.

B	Bucket size of a token bucket
BE	Best Effort
Bp	Bucket size of peak token bucket
Bs	Bucket size of sustainable token bucket
DBW	Dedicated Bandwidth IP Transfer Capability
DS	Differentiated Services (IP header field)
GBRA	Generic Byte Rate Algorithm
GCRA	Generic Cell Rate Algorithm
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPDV	IP Delay Variation
IPLR	IP Loss Ratio
IPTC	IP Transfer Capability
IPTD	IP Transfer Delay
LR	Line Rate
M	Maximum allowed packet size
N	Size (in byte) of an IP packet
PC	Parameter Control
QoS	Quality of Service
R	Rate of a token bucket
Rp	Rate of peak token bucket
Rs	Rate of sustainable token bucket
SBW	Statistical Bandwidth IP Transfer Capability
TB	Token Bucket
TC	Transfer Capability
ToS	Type of Service (IP header field, see DS)

3.2 Terminology

For the purposes of this Recommendation, the following terminology applies.

3.2.1 conformance: Conformance is the application of one or more criteria, at a given standardized interface, to a flow.

3.2.2 congestion: Congestion is defined as a state of network elements (e.g. router, switches) in which the network is not able to meet the network performance objectives and the negotiated QoS commitments for the already established flow.

3.2.3 IP flow: An IP flow at a given interface is defined as the occurrence at that interface of the set of IP packets which match a given classification (see 5.1.2).

3.2.4 IP traffic control: Refers to network actions aiming to meet the network performance objectives and negotiated QoS commitments.

3.2.5 IP congestion control: Refers to all network actions to minimize the intensity, spread and duration of congestion.

3.2.6 IP transfer capability: An IP transfer capability is a set of network capabilities provided by IP based network to transfer IP-flows.

3.2.7 overload: Overload is defined as a state of network element in which buffer overflow results in packet discard for flows with no QoS commitments.

3.2.8 packet classification: The process of distinguishing IP packets for the purpose of applying appropriate traffic control and congestion control mechanisms is called IP packet classification (see 5.1.1).

3.2.9 traffic contract: For a given IP flow, the selected IP transfer capability (see clause 6), the traffic descriptor at a given interface and the QoS class (see ITU-T Rec. Y.1541) define the traffic contract at that interface. See 5.1.5.

3.2.10 traffic descriptor: A traffic descriptor is the set of traffic parameters that is used to capture the traffic characteristics of an IP flow at a given standardized interface as part of the traffic contract. See 5.1.4.

3.2.11 traffic parameter: A traffic parameter describes one aspect of a flow. See 5.1.3.

4 Introduction

The primary role of traffic control and congestion control procedures is to protect the IP network and the traffic entering the network in order to achieve the IP network performance objectives and QoS commitments. Traffic and congestion control allows the use of IP network resources to be optimized.

In IP-based networks, congestion is defined as a state of network elements (e.g. routers, switches) in which the network is not able to meet the network performance objectives and the negotiated QoS commitments for the established flow. Congestion is to be distinguished from the state where buffer overflow causes IP packet loss, but the negotiated quality of service is still met. For services with no QoS commitments such as best effort services, the term overload is used rather than congestion.

This Recommendation defines a set of traffic control and congestion control capabilities. It may be appropriate to consider additional sets of such capabilities, for which additional traffic control mechanisms will be used to achieve increased network efficiency.

It should be noted that traffic control procedures apply to IP packet flows. For some IP flows, the network commits to meet QoS objectives, assuming the user generated IP packet flow conforms to a traffic contract. For IP flows without a traffic contract between the user and the network, the network may provide a best effort service.

5 Traffic parameters and descriptors

5.1 Definitions

5.1.1 IP packet classification

Typically, IP packets from many different sources may cross a given interface to many different destinations. For the purpose of IP traffic control and congestion control, it is essential that not all IP packets are treated the same way but differently, depending on the objectives and on the commitments made. Therefore, it is useful to distinguish IP packets into relevant categories. The process of distinguishing IP packets for the purpose of applying appropriate traffic control and congestion control mechanisms is called IP packet classification. IP packet classification is based on information in the following IP header fields: source address, destination address [RFC 791] and TOS/DS field [RFC 791/RFC 2474]. The level of detail in the classification may differ, depending on its intended use. The following examples illustrate this:

- Example 1 – In order to recognize IP packets which flow from a given source address to a given destination address, the entire length of the source and destination address fields is used in the classification.
- Example 2 – In order to recognize IP packets which flow from a given source address to a given destination address with a given service level, the entire length of the source and destination address fields and the DS field is used in the classification.
- Example 3 – In order to recognize IP packets which flow from a source in one given subnet to a destination in another given subnet, the respective subnet masks are applied to the source and destination address fields.

NOTE – In this version of this Recommendation, only the packet classification for IP version 4 has been addressed. Classification for IPv6 and MPLS is for further study.

5.1.2 IP flow and IP subflow

An IP flow at a given interface is defined as the occurrence at that interface of the set of IP packets which match a given classification (see 5.1.1). An IP flow may consist of packets from a single application session, or it may be an aggregation comprising the combined traffic from a number of application sessions. When a classification may be subdivided into different sub-classifications (separate or overlapping), different IP subflows may be recognized in the corresponding IP flow.

5.1.3 Traffic parameter

A traffic parameter describes one aspect of a flow. It may be qualitative or quantitative. A traffic parameter may for example describe the peak bit rate, the peak packet rate, the average bit rate, the average or maximum packet size, the (average of maximum) burst length of a flow, etc.

5.1.4 Traffic descriptor

A traffic descriptor is the set of traffic parameters that is used to capture the traffic characteristics of an IP flow at a given standardized interface as part of the traffic contract (see 5.1.5).

5.1.5 Traffic contract

For a given IP flow, the selected IP transfer capability (see clause 6), the traffic descriptor at a given interface (see 5.1.4) and the QoS class (see ITU-T Rec. Y.1541) define the traffic contract at that interface.

5.2 Requirements on traffic parameters and traffic descriptors

Any traffic parameter to be involved in a traffic descriptor should:

- have the same interpretation on both sides of an interface;
- be meaningful in resource allocation schemes to meet network performance requirements;
- be enforceable by the parameter control (PC), see 7.1.3.

5.3 Traffic parameter specifications

5.3.1 Reference configuration

For further study.

5.3.2 Traffic parameter description

The following traffic parameters are defined which may be used in the traffic descriptor.

5.3.2.1 Maximum allowed packet size

The maximum allowed packet size *M* is expressed in byte. It is a mandatory traffic parameter for each IPTC.

5.3.2.2 GBRA or Token Bucket

The Generic Byte Rate Algorithm (GBRA) or Token Bucket (TB), as described in Annex A, is used to characterize a rate and the associated burstiness. Both concepts are equivalent and use the following set of two parameters:

- the rate *R* expressed in byte/s;
- the bucket size *B* expressed in byte.

The traffic descriptor may contain zero or more token buckets (with the respective values of *R* and *B*).

The current set of IP transfer capabilities (see clause 6) recognizes a peak TB (with parameters peak rate *Rp* and peak bucket size *Bp*) and a sustainable TB (with parameters sustainable rate *Rs* and sustainable bucket size *Bs*).

6 IP transfer capabilities

An IP transfer capability is a set of network capabilities provided by IP-based networks to transfer IP packets. For each IP transfer capability, the service model, traffic descriptor, conformance definition and any QoS commitments are defined. An IP transfer capability is supported by a set of traffic control and congestion control functions.

In order to offer multiple classes of QoS to multiple applications and to optimize the usage of network resources, IP-based networks should be capable of providing multiple transfer capabilities.

Three IP transfer capabilities are defined:

- Dedicated Bandwidth (DBW) IP transfer capability.
- Statistical Bandwidth (SBW) IP transfer capability.
- Best-effort (BE) IP transfer capability

This set of IP transfer capabilities are based on current IP service models and this set may be extended in the future.

6.1 Dedicated Bandwidth (DBW) transfer capability

6.1.1 Description

The Dedicated Bandwidth (DBW) transfer capability is intended to support applications with stringent delay requirements. It aims to support the guaranteed and timely delivery of IP packets along the end-to-end path of the network.

The DBW transfer capability strives for compatibility with the Guaranteed Service [RFC 2212] and the end-to-end services based on the Expedited Forwarding per-hop behaviour [RFC 2598].

6.1.2 Service model

The DBW IPTC can be used by applications that characterize the traffic with a single token bucket.

The commitment made by the network is that the negotiated IP QoS is assured to all IP packets when all packets are conforming to the conformance tests. The DBW user should expect that (possibly all) non-conforming packets be discarded by the network.

The DBW capability can be associated with specified loss commitments (IP Loss Ratio, IPLR) and specified delay commitments (IP Transfer Delay, IPTD and IP Delay Variation, IPDV), (see ITU-T Rec. Y.1541).

The network does not fragment packets. In addition, the network commits to attempt, as long as possible (e.g. until there is a need to reroute the flow), to maintain packet sequence integrity.

6.1.3 Traffic descriptor

The Traffic Descriptor consists in the following parameters:

- The peak rate *Rp* and peak bucket size *Bp* as specified in 5.3.2.2.
- The maximum allowed packet size *M* as specified in 5.3.2.2.

6.1.4 Conformance definition

An IP packet is conforming if the arrival conforms to the following two parts:

- the arrival is conforming to the GBRA(*Rp,Bp*);
- the actual packet length does not exceed the maximum allowed packet size *M*.

The GBRA is updated for conforming packets only.

6.1.5 QoS commitments

The DBW capability may be associated with specified loss commitments and specified delay commitments.

If all packets are conforming, the QoS commitments apply to all IP packets. The DBW user should expect that (possibly all) non-conforming packets be discarded by the network. If not all packets are conforming, the network may choose to commit QoS to some of the packets, for example to a volume of packets that is conforming.

6.2 Statistical Bandwidth (SBW) transfer capability

6.2.1 Description

The Statistical Bandwidth (SBW) transfer capability is intended to support applications, which do not have stringent delay requirements. It aims to support the guaranteed delivery of IP packets along the end-to-end path of the network.

The SBW transfer capability strives for compatibility with the Controlled-Load Network Element Service [RFC 2211] and the end-to-end services based on the Assured Forwarding per-hop behaviour [RFC 2597].

6.2.2 Service model

The SBW transfer capability provides a specified sustainable rate (*Rs*) for non-real time applications with limited burst duration with the expectation that traffic in excess of GBRA(*Rs,Bs*) will be delivered within the limits of available resources.

The following two examples describe the commitment the SBW user will receive:

- If the user sends conforming packets at a constant rate that is less than or equal to the *Rs*, then the commitments is that all these packets are delivered across the network, corresponding to the associated QoS class.
- If the user has not sent packets for a long time and the user sends conforming packets in a burst with a duration that does not exceed the bound set by GBRA(*Rs*,*Bs*), then the commitments is that all these packets are delivered across the network, corresponding to the associated QoS class.

The SBW capability also allows the user to send conforming packets in excess of the GBRA(*Rs*,*Bs*), but traffic that exceeds this bound will only be delivered within the limits of available resources.

The SBW capability may be associated with a specified packet loss commitment.

The network does not fragment packets. In addition, the network commits to attempt, as long as possible (e.g. until there is a need to reroute the flow), to maintain packet sequence integrity.

6.2.3 Traffic descriptor

The traffic descriptor consists in:

- The peak rate *Rp* and the peak bucket size *Bp* as specified in 5.3.2.2;
- The sustainable rate *Rs* and the sustainable token bucket size *Bs* as specified in 5.3.2.2;
- The maximum allowed packet size *M* as specified in 5.3.2.1.

6.2.4 Conformance definition

An IP packet is conforming if the arrival conforms to the following three parts:

- the arrival is conforming to the peak GBRA(*Rp*,*Bp*);
- the arrival is conforming to the sustainable GBRA(*Rs*,*Bs*);
- the actual packet length does not exceed the maximum packet size *M*.

The GBRAs are updated in coordinated mode (see Annex B) for conforming packets only.

6.2.5 QoS commitments

The SBW capability may be associated with specified loss commitments.

If all packets are conforming, the QoS commitments apply to all packets. Otherwise, the QoS commitments apply to a number of bytes in conforming packets. Non-conforming traffic will be delivered within the limits of available resources.

6.3 Best effort (BE) transfer capability

6.3.1 Description

The best effort IP transfer capability is intended to support applications which do not have stringent loss or delay requirements.

6.3.2 Service model

The service model for the best effort (BE) IPTC requires that available resources be used for forwarding packets of best effort flows. Even though there are no QoS commitments specified, the expectation is that packets be delivered provided that sufficient resources are available.

6.3.3 Traffic descriptor

- The maximum allowed packet size *M*.

6.3.4 Conformance definition

Conformance to the maximum packet size *M* is required.

6.3.5 QoS commitments

There is no absolute QoS requirement for this transfer capability.

7 Functions for traffic control, congestion control, and overload treatment

Traffic control refers to a set of functions that control the flow of packets via a series of functions such as admission control, network resource management, traffic parameter control. The main objective of traffic control is to satisfy user requirements such as Quality of Service while still supporting efficient network utilization.

As opposed to traffic engineering, traffic control is accomplished in a short time scale. Therefore, a well-established and automated mechanism is to be provided to control the flow of traffic into the network and out of the network.

Under normal operation, i.e. when no network failures occur, functions referred to as traffic control functions in this Recommendation are intended to avoid network congestion.

However, congestion may occur, e.g. because of misfunctioning of traffic control functions caused by unpredictable statistical fluctuations of traffic flows or of network failures. Therefore, additionally, functions referred to as congestion control functions in this Recommendation are intended to react to network congestion in order to minimize its intensity, spread and duration.

Overload treatment applies only to traffic flows with no QoS commitments. It refers to a set of functions in the network that can detect and reduce the amount of overload.

7.1 Traffic control functions

The following functions are identified for traffic control:

i) Network resource management.

ii) Admission control.

iii) Parameter control.

iv) Packet marking.

v) Traffic shaping.

vi) Packet scheduling.

7.1.1 Network resource management

Network resource management refers to a set of policies and rules for allocating the network resources such as the bandwidth and buffer spaces in a switch or router.

7.1.2 Admission control

Admission control refers to the policies of the network to admit commitments to a new IP flow or to refuse the commitments when the demand for the network resources such as the bandwidth and buffer spaces exceeds the available capacity in the network.

7.1.3 Parameter control

Parameter control is a set of policies that monitors and controls that the traffic contract is not exceeded.

7.1.4 Packet marking

When an IP flow is found to be non-conforming to one or more aspects of the traffic contract, the non-conforming packets may be marked by modifying the relevant bits in the ToS/DS field of the IP header.

7.1.5 Traffic shaping

Traffic shaping is an action by the network to modify the traffic characteristics of the flow such that the flow becomes more suitable for the network. One example is shaping to the peak rate of a flow.

7.1.6 Packet scheduling

Packet scheduling is a function of the network with two different objectives. One objective is to bind the queuing delay for a flow, which has stringent delay requirements. Another objective is to divide an available resource (e.g. bandwidth) over different flows in a network-specific manner.

7.2 Functions for congestion control and overload treatment

The following functions are identified for congestion control and overload treatment:

i) Packet discard control.

ii) Explicit Congestion Notification.

7.2.1 Packet discard control

Packet discard control is used to discard marked traffic in a congested situation. Another application of packet discard control is, in a congested situation, to discard packets for which no stringent QoS commitments apply.

7.2.2 Explicit Congestion Notification

Explicit Congestion Notification is an optional method with which the network may inform the receiver of existing network congestion. The receiver can use this information to take appropriate action, e.g. to inform the sender to reduce its transmission rate.

8 Methods and tools for IP traffic engineering

For further study.

Annex A

The Token Bucket and the Generic Byte Rate Algorithm (GBRA)

This annex provides several algorithms, which are used to define packet conformance of an IP flow with respect to a rate *R* (in byte per second) assuming a bucket size *B*. The algorithms are considered equivalent in the sense that they will identify the same packets to be conforming.

Traffic metering is a basic component of packet networks that provide various levels of quality of service, such as throughput guarantees and upper bounds on delay and loss. An IP packet stream must conform to a predefined traffic profile in order to assure that the network resources allocated to that stream can provide the expected service level, and to assure that this stream does not cause other network traffic to experience an unacceptable service level.

Starting point is the Token Bucket. This algorithm is considered equivalent to the continuous state token bucket. This algorithm in turn is then shown to be equivalent to the Generic Byte Rate Algorithm (GBRA). The GBRA defines the characteristics that relate an information rate parameter to a burst tolerance parameter. Together, these two traffic parameters and their interrelationship form the basis for a generalized traffic profile conformance test. A secondary goal of GBRA is to avoid referencing one or more implementation options on the basis that it is not central to the conformance concept and that it is a source of unnecessary bias. GBRA specifies the fundamental aspects of a conformance test to a single profile, and as such does not include other related policing functions such as multi-stage metering, packet coloring, packet marking, and packet discard policies.

A.1 The token bucket

This algorithm is described as follows (compare also e.g. RFC 2698):

The token bucket has two fixed parameters:

- The token bucket rate R (in bytes per second).
- The token bucket size B (in bytes).

The token bucket uses the variable:

- The token count Tc (in bytes) which is initially (at the arrival at time ta of the first packet of the flow) full, i.e. the token count Tc(ta) = B.

Thereafter, the token count Tc is incremented by one every 1/R seconds up to B.

When a packet of size N bytes arrives at time ta, the following happens:

- If Tc(ta)–N < 0, the packet is non-conforming; else
- the packet is conforming and Tc is decremented by N.

NOTE – The arrival time is considered to be when the last bit of the packet arrives.

A.2 The continuous-state token bucket

To be more precise and unambiguous, a continuous-state token bucket is proposed. The continuous-state token bucket is considered to be equivalent to the token bucket. The continuous-state token bucket has the advantage that it is easily implementable.

The continuous-state token bucket has two fixed parameters per IP flow:

- The token bucket rate R (in bytes per second) for the flow.
- The token bucket size B (in bytes) for the flow.

The continuous-state token bucket uses the following variables:

- The token count Tc (in bytes) of the flow.
- LCT is the last conformance time of the flow (in seconds).

Initially (at time ta of the arrival of the first packet of the flow):

- Tc = B
- LCT = ta

At arrival of a packet with size N (bytes) at time ta:

```
Tc' = Tc + R*(ta – LCT)
If Tc' < N
    Then    packet is not conforming
    Else    packet is conforming
            Tc = min(Tc', B) – N
```

For the continuous-state token bucket, the variable Tc and LCT are only modified at packet arrival.

NOTE – A packet can never be conforming if N > B.

A.3 The Generic Byte Rate Algorithm

For the generic byte rate algorithm, there are two variants: one is the virtual scheduling byte rate algorithm and one is the continuous-state leaky bucket byte rate algorithm.

For both algorithms, the fixed parameters per IP flow are:

- The increment I in units of seconds per byte. I is related to R by I = 1/R
- The Limit L in seconds. L is related to R and B by L = B/R.

A.4 The virtual scheduling byte rate algorithm

The virtual scheduling byte rate algorithm uses the following variable:

- TAT in units of seconds is the theoretical arrival time of the next packet.

Initially (at time ta of the arrival of the first packet of the flow):

- TAT = ta

The theoretical arrival time (TAT) is updated only after a packet is considered conforming by the algorithm.

At arrival of a packet with size N (bytes) at time ta:

```
If ta < TAT + N*I – L
    Then    packet is not conforming
    Else    packet is conforming
            TAT = max (ta, TAT) + N*I
```

A.5 The continuous-state leaky bucket byte rate algorithm

The continuous-state leaky bucket byte rate algorithm uses the following variables:

- LCT is the last conformance time of the flow (in seconds).
- X is the Leaky Bucket counter (in seconds).

Initially (at time ta of the arrival of the first packet of the flow):

```
–    LCT = ta
–    X = 0
```

At arrival of a packet with size N (bytes) at time ta:

```
X' = X – (ta – LCT)
If X' + N*I > L
    Then    packet is not conforming
    Else    packet is conforming
X = max(0, X') + N*I
```

NOTE – Equivalence of the algorithms:

- The continuous-state leaky bucket byte rate algorithm is equivalent to the continuous-state token bucket: Just set Tc = (L–X)*R, Tc' = (L–X')*R and use B = L*R.
- The virtual scheduling byte rate algorithm is equivalent to the continuous-state leaky bucket byte rate algorithm: Just observe that after the two algorithms have been executed, one has TAT = X + LCT.

NOTE – Relationship between the GBRA and the GCRA in ITU-T Rec. I.371:

In case of ATM, the cell size is N = 53 bytes and N*I becomes equal to T which is the inverse of the cell rate. If one sets L = T + τ where τ is the tolerance of the GCRA, then one sees that the GCRA is a special case of the GBRA (see Annex A/I.371).

Annex B

Behaviour of two coordinated Generic Byte Rate Algorithms

The conformance definition of the Statistical Bandwidth (SBW) transfer capability is based, besides the test of the maximum allowed packet size, on two instances of the Generic Byte Rate Algorithm (GBRA) (see Annex A) where the two instances operate in a coordinated mode. Coordinated mode means the states of the GBRA are updated if and only if an arriving cell is conforming to all relevant instances of the GBRA. An IP packet passing this test is conforming when it is at least conforming to both the peak GBRA(*Rp*,*Bp*) and the sustainable GBRA(*Rs*,*Bs*).

Figure B.1 shows the behaviour that results from two instances of the GBRA operated in a coordinated mode.

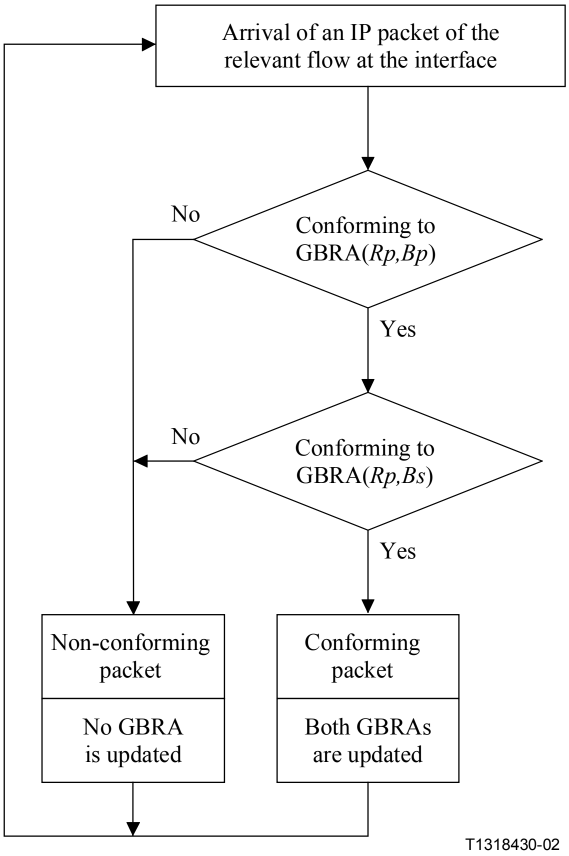
Arrival of an IP packet of the relevant flow at the interface
No
Conforming to GBRA(*Rp,Bp*)
Yes
No
Conforming to GBRA(*Rp,Bs*)
Yes
Non-conforming packet
No GBRA is updated
Conforming packet
Both GBRAs are updated
T1318430-02

Figure B.1/Y.1221 – Behaviour of two co-ordinated Generic Byte Rate Algorithms.

Appendix I

Illustration of Token Bucket behaviour

This appendix contains informative text.

Introduction: Token bucket behaviour

A rate, expressed in byte/s, may be used to partially characterize a flow. The value of the rate may be dependent on the period over which it is measured. For example, a source, which sends 50 bytes every 1 ms, is different from a source, which sends 500 bytes every 10 ms, though their long-term average rates are the same. To allow this difference to be reflected, a token bucket (see Annex A) with bucket rate R and a bucket size B is used to characterize the flow. In practice, real flows may not show the same recurrent pattern as in the examples above. Yet, the token bucket characterization may be used to characterize the 'worst case behaviour' of all possible flows of which all packets are conforming to a token bucket with given parameters (R,B). This will be further described and illustrated in the following.

I.1 Characterization of a flow with a single rate and a single token bucket

A flow is said to be confined or upper bounded by a token bucket with rate R and token bucket size B if all packets of the flow conform to the token bucket algorithm with bucket rate R and bucket size B (see Annex A).

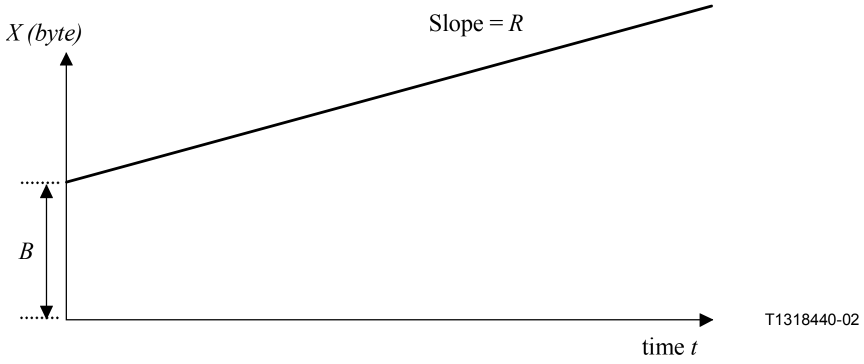

Figure I.1a/Y.1221 – Single rate, single bucket size token bucket – 'worst case' flow

Figure I.1a shows the behaviour of a theoretical worst case flow which conforms to the token bucket algorithm with token bucket size B and rate R and where the line rate is assumed to be infinite. The horizontal axis represents the time since the first packet on that flow, the vertical axis shows the number of bytes X that at maximum could have been sent on the flow and still may pass the token bucket up to a time t. This theoretical worst case flow has sent a burst of B bytes (at infinite rate) followed by a constant flow of 1 byte packets each 1/R second. Note that in a practical IP network, with a minimum packet size larger than 1 byte, a real flow cannot show such a packet rate. It is easily verified that this is the maximum behaviour allowed by the token bucket and that it is indicated by the top line with slope equal to the token bucket rate R, (i.e. $X \leq B + t \times R$).

Another worst case flow may show a recurrent pattern of B bytes (at infinite rate) at regular intervals such that the long term average rate equals the bucket rate R. Figure I.1b indicates the trajectory of this flow as a dashed line. In a practical IP network, a real flow can realize or approximate such behaviour. It is easily verified that, in order to send the second burst of B bytes, the flow should show reduced activity (e.g. be silent) for a sufficient number of tokens to be replenished.

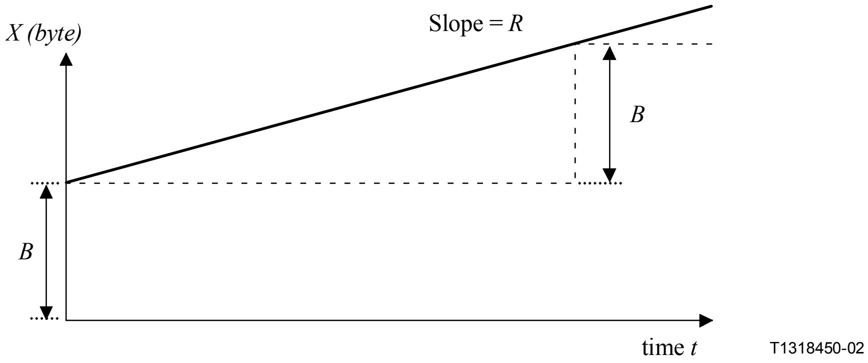

Figure I.1b/Y.1221 – Single rate, single bucket size token bucket – 'worst case bursty' flow.

Figure I.1b also illustrates that the expression $X \leq B + t \times R$ still holds but is not at all moments an accurate representation of the actual number of bytes sent on this flow.

In a third example, a flow shows to be less active for some time and then resumes 'worst case' behaviour similar to the first two examples. This is illustrated in Figure I.1c.

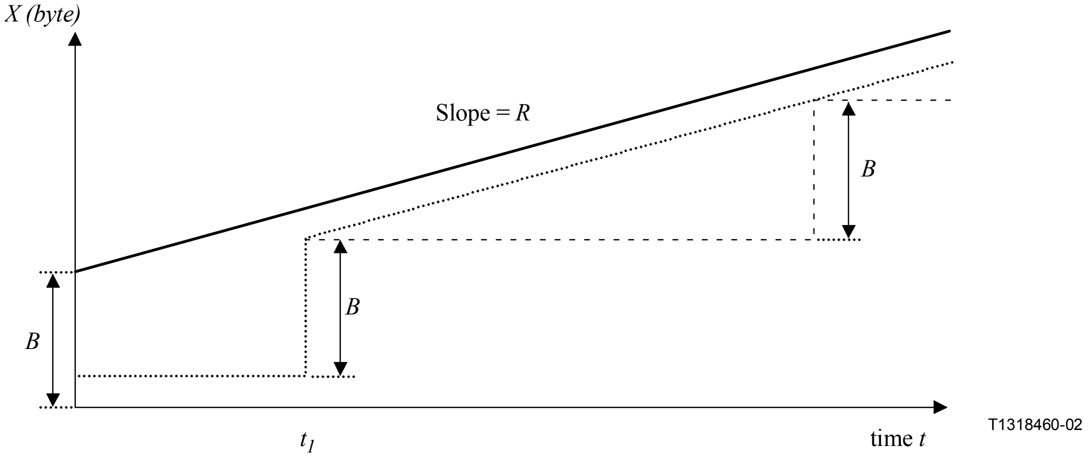

Figure I.1c/Y.1221 – Single rate, single bucket size token bucket – 'worst case' flows after inactivity

The first, theoretical worst case flow resumes maximum activity at time t_1. Due to the modest activity in the pre-t_1 era, the token bucket has been fully replenished to B byte. This means that, at moment t_1, the flow is constrained to sending a burst not larger than B byte (at infinite rate) and may subsequently show the (only theoretically possible) worst case behaviour of sending 1-byte packets each 1/R second. Figure I.1c illustrates the trajectory of this flow as a dotted line.

A second worst case flow, which resumes maximum activity at time t_1, shows a recurrent pattern of B bytes (at infinite rate). This flow also finds itself constrained by the token bucket size B and, when all tokens have been consumed, needs to wait at least B/R s before sending the next B-size packets. Figure I.1c illustrates the trajectory of this flow as a dashed line.

Real flows may show many different characteristics, for example by sending different size packets and sending at different rates during the lifetime of the flow.

It is the token bucket algorithm that confines the flow, at any moment in two ways:

– it confines the flow to a burst size (not exceeding B byte); and

– it confines the flow to a rate (R byte/s) measured over a sufficiently long period (not exceeding B/R s) and disregarding the allowed burst size.

Given a suitable set of token bucket parameters, the network can rely on these simplified flow characteristics for its resource allocation, instead of requiring the originator of the flow to specify its real (possibly very complicated and possibly very diverse) flow or source characteristics and take these into account. For a given flow, there may be more than one set of token bucket parameter sets (R,B) to characterize the flow. It is the users responsibility to select a suitable parameter set for his flow.

NOTE – It should be noted that, after a sufficiently long period of reduced activity, the expression $X \leq B + t \times R$ still holds, but it is no longer an accurate representation of the maximum number of bytes which can be sent on the flow.

I.2 Characterization of a flow with two rates and two token buckets (*Bp* < *Bs*)

When two rates are defined for a flow, then two token buckets are used. The higher of the two rates is designated as the peak rate (*Rp*), the lower rate is the sustainable rate (*Rs*). When the token bucket size for the peak rate is smaller than the token bucket size for the sustainable rate (as is the case for IETF GS and CLS), then the peak rate represents the short-term throughput. The sustainable rate (*Rs*) represents the long-term throughput.

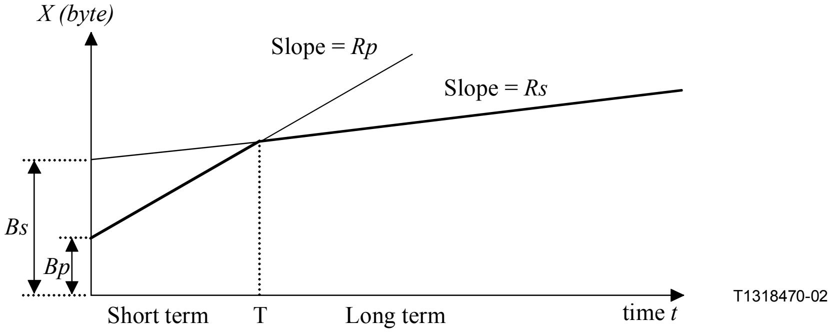

Figure I.2/Y.1221 – Two rates, two token bucket (*Bp* < *Bs*) – 'worst case' flow

Figure I.2 shows the behaviour of a theoretical worst case flow which conforms to two token buckets where one has rate *Rp* with token bucket size *Bp* and the other has rate of *Rs* with token bucket size *Bs* and *Bp* < *Bs*. The line rate is assumed to be infinite. The number of bytes X is the maximum number of bytes that could have been sent on the flow as a function of time t since the first packet arrival and still pass both token buckets. The number of X is determined by the lines whose slopes are equal to the token bucket rates *Rp* and *Rs*. At times before T the higher rate (i.e. *Rp*) token bucket dominates the worst case flow behaviour while for time values exceeding T the lower rate (i.e., *Rs*) token bucket dominates the worst case flow behaviour. These regions are indicated as short term and long term in Figure I.2.

The following relationships define the maximum number for X as function of the time t and are indicated with thick lines in Figure I.2.

$$X \leq Bp + Rp \times t \qquad (0 \leq t \leq T)$$

$$X < Bs + Rs \times t \qquad (t \geq T)$$

Similar to the case described in I.1, a suitable set of token bucket parameters may be used to better characterize some flows than would be possible with a single rate token bucket and the network can rely on these flow characteristics for its resource allocation.

NOTE – Similar to the case described in I.1, it should be noted that, after a sufficiently long period of reduced activity, the expressions above still hold but they are no longer an accurate representation of the maximum number of bytes, which can be sent on the flow.

I.3 Characterization of a flow with two rates and two token buckets (*Bp* > *Bs*)

When two rates are used and the token bucket size for the peak rate is bigger than the token bucket size for the lower rate (as is the case for IETF Diffserv), then the two profile lines that are shown in Figure I.2 do not cross each other any more. The traffic descriptor includes (*Rp*,*Bp*) and (*Rs*,*Bs*) and defines a committed rate token bucket TB(*Rs*,*Bs*), in which *Rs* is the lower rate, that represents the traffic profile below which the network guarantees delivery of the packets. Packets, which conform to this TB, are conforming packets. This is very similar to the case of conforming packets for the single rate characterization in I.1. For the sake of discussion, these conforming packets are also indicated as "green" packets.

The traffic descriptor also defines a peak token bucket TB(*Rp*,*Bp*), in which *Rp* is the higher rate that represents the traffic profile above which the network does not provide any delivery guarantee. This allows to further distinguishing between packets that are not conforming to the sustainable TB.

A packet which is non-conforming to the peak rate TB is called a "red" packet. This is very similar to the case of a non-conforming packet of the single rate characterization in I.1, yet with different token bucket parameter set.

A packet, which does not conform to the lower bucket but does conform to the peak rate bucket, is called "yellow". The network may offer different 'commitments' to these non-conforming (i.e. yellow and red) packets. For example, for yellow packets a lesser commitment may apply then to conforming (green) packets and a better commitment than applies to red packets. Figure I.3 illustrates the lines with slopes *Rs* and *Rp*, which separate three different regions as follows.

$X \leq Bs + Rs \times t$ → conforming region (green)

$X > Bp + Rp \times t$ → non-conforming region (red)

$Bs + Rs \times t < X \leq Bp + Rp \times t$ → non-conforming region (yellow)

NOTE – Clarification on the use of the distinction between the two non-conforming regions (yellow and red) is requested.

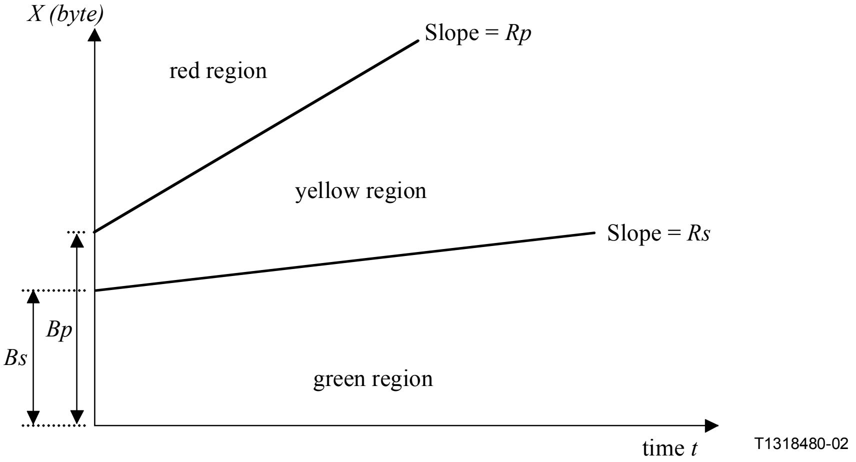

Figure I.3/Y.1221 – Two rates, two token bucket (*Bp* > *Bs*) – 'worst case' flow

Describing the possible characteristics of a 'worst case' flow in this configuration becomes fuzzier.

Similar, to the worst case flows described in I.1, a 'worst case green' flow can be described of which all packets conform to the TB(*Rs*,*Bs*).

Also, a 'worst case yellow' flow can be described of which all packets conform to the TB(*Rp*,*Bp*). Note that a 'worst case yellow' flow consists in not only yellow packets but in a mixture of green and yellow packets.

Finally, a flow which is not conforming to TB(*Rp*,*Bp*) will consist in a mixture of red, yellow and green packets, depending on the severity of exceedence of the respective token buckets.

Appendix II

Relation between IPTC/QoS class and IETF IntServ/DiffServ-specification

The main body of this Recommendation describes two IP transfer capabilities (see 6.1, 6.2) which strive for compatibility with service descriptions and specifications generated by IETF IntServ and DiffServ groups. The intention is that network elements and networks conforming to these specifications, are likely able or adaptable to support the IP transfer capabilities specified in this Recommendation.

This informative appendix lists and motivates the most important differences between the IP transfer capabilities and the IETF counterparts.

II.1 Dedicated BandWidth IPTC associated with a suitable QoS class

The description for the DBW IP Transfer Capability (see 6.1.1) lists that the DBW capability strives for compatibility with the Guaranteed Service (GS) [RFC 2212] and the end-to-end services based on the Expedited Forwarding per-hop behaviour (EF-PHB) [RFC 2598].

Because ITU-T uses a combination of an IPTC and a QoS class, the following assumes that a QoS class with specified loss commitments and specified delay (variation) commitments is selected for a given flow. A flow characterized by the DBW capability in combination with such a QoS class is in the sequel referred to as 'DBW' . Similarly, a flow which uses the GS or a flow using the EF PHB is referred to as 'GS' and 'EF' respectively.

The common property between DBW, GS and EF is that the network commits to transport IP packets with a specified capacity [byte/s] with a low delay and low delay variation.

The following differences are noted.

– The DBW capability allows the network provider to discard (possibly all) non-conforming packets, whereas in GS and EF non-conforming packets are expected to be carried on a best-effort basis. It is expected that the ability to restrict the influx of DBW packets to the agreed rate and characterized by GBRA(*Rp*,*Bp*), greatly simplifies the traffic control in the network (e.g. priority queuing) and minimises queuing delays for all conforming traffic. It is further expected that most applications using DBW (e.g. streaming media) are able to specify suitable parameters and to generate traffic within these bounds (conforming packets).

– The association with a QoS class allows the end-to-end commitments to be a priori (before agreeing on a traffic contract) known to the user. GS allows (e.g. via the RSVP protocol) to inform the user about the maximum queuing delay expected to be experienced on the particular flow path. This information is provided a posteriori (after the reservation has been made). In addition, it has been shown that on a path with many hops the accumulation of per-hop maximum queuing delay values leads to much larger values than the 'maximum delay' (or a suitable quantile thereof) experienced on that path. Consequently, the accumulated information is of little or no practical value to the user. EF does not specify an upper bound to the delay (variation).

– DBW applies (in addition to the maximum packet size) the minimum traffic characterisation with a single set of Token Bucket parameters (*Rp*,*Bp*). GS requires the specification of an additional peak rate parameter, however without the possibility to specify a 1-point packet delay variation tolerance related to that rate. In addition, GS is invoked using two sets of traffic parameters (Tspec and Rspec, which may have different values) plus a so-called 'slack term'.

II.2 Statistical BandWidth IPTC associated with a suitable QoS class

The description for the SBW IP transfer capability (see 6.2.1) lists that the SBW capability strives for compatibility with the Controlled-Load Network Element Service (CL) [RFC 2211] and the end-to-end services based on the Assured Forwarding per-hop behaviour (AF-PHB) [RFC 2597].

Because ITU-T uses a combination of an IPTC and a QoS class, the following assumes that a QoS class with specified loss commitments and possibly unspecified delay (variation) commitments is selected for a given flow. A flow characterized by the SBW capability in combination with such a QoS class is in the sequel referred to as the 'SBW'. Similarly, a flow which uses the CL or a flow using the AF PHB/PSC is referred to as 'CL' and 'AF' respectively.

The common property between SBW, CL and AF is that the network commits to transport IP packets with at least a specified capacity [byte/s] and allows additional (excess) traffic to be transported within the limits of the available network resources.

The following differences are noted:

– The SBW capability and AF is intended to support applications which do not have stringent delay requirements. CL does not commit to quantified delay properties but aims at properties 'under unloaded conditions', i.e. a delay variation which is small compared to the minimum delay (e.g. propagation delay). This property is expected to be difficult to realize in combination with the ability to efficiently support excess traffic beyond and in addition to the guaranteed rate.

– SBW applies (in addition to the maximum packet size) a double set of Token Bucket parameters (*Rp*,*Bp*) and (*Rs*,*Bs*). CL also requires the specification of the peak rate parameter, however without the possibility to specify a 1-point packet delay variation tolerance related to that rate. In addition, CL is invoked using two sets of traffic parameters: Tspec and Rspec, which may have different values.

– Currently, the SBW capability distinguishes conforming and non-conforming packets. AF distinguishes three so-called 'drop precedence', yet is required to support at minimum two different drop probabilities.

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

22157

Printed in Switzerland
Geneva, 2002