International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.1222

(11/2007)

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT-GENERATION NETWORKS

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

---

Traffic control and congestion control in Ethernet-based networks

ITU-T Recommendation Y.1222

ITU-T

International Telecommunication Union

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT-GENERATION NETWORKS

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
NEXT GENERATION NETWORKS	
Frameworks and functional architecture models	Y.2000–Y.2099
Quality of Service and performance	Y.2100–Y.2199
Service aspects: Service capabilities and service architecture	Y.2200–Y.2249
Service aspects: Interoperability of services and networks in NGN	Y.2250–Y.2299
Numbering, naming and addressing	Y.2300–Y.2399
Network management	Y.2400–Y.2499
Network control architectures and protocols	Y.2500–Y.2599
Security	Y.2700–Y.2799
Generalized mobility	Y.2800–Y.2899

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

ITU-T Recommendation Y.1222

Traffic control and congestion control in Ethernet-based networks

Summary

ITU-T Recommendation Y.1222 provides a general description of and procedures for traffic control and congestion control in Ethernet-based networks. It describes the concepts of the traffic contract between a user and the network. It specifies the Ethernet transfer capabilities (ETCs) including, for each ETC, the service model, the associated traffic patterns and conformance definition for an Ethernet flow that are observable at any point in the network.

Source

ITU-T Recommendation Y.1222 was approved on 13 November 2007 by ITU-T Study Group 12 (2005-2008) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). 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.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

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, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

CONTENTS

ITU-T Recommendation Y.1222

Traffic control and congestion control in Ethernet-based networks

1 Scope

This Recommendation describes traffic control and congestion control procedures for Ethernet-based networks.

Traffic control refers to all network actions aiming to meet the negotiated performance objectives and negotiated QoS commitments in an Ethernet-based network, and to avoid congested conditions. Congestion control refers to all network actions to minimize the intensity, spread and duration of congestion.

This Recommendation provides a general description of and procedures for traffic control and congestion control. It describes the concepts of the traffic contract between a user and the network. It specifies the Ethernet transfer capabilities (ETCs) including, for each ETC, the service model, the associated traffic patterns and conformance definition for an Ethernet flow that are observable at any point in the network.

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. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T I.371] ITU-T Recommendation I.371 (2004), *Traffic control and congestion control in B-ISDN*.

[ITU-T Y.1221] ITU-T Recommendation Y.1221 (2002), *Traffic control and congestion control in IP-based networks*.

[Metro Ethernet] Metro Ethernet Forum Technical Specification 10.1 (2006), *Ethernet Services Attributes Phase 2*.

[IEEE 802.3] IEEE 802.3 (2005), *IEEE Standard for information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications*.

3 Terminology

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

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

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

3.3 Ethernet flow: An Ethernet flow at a given interface is defined as the occurrence at that interface of the set of Ethernet frames which match a given classification (see clause 7.1.2).

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

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

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

3.7 overload: Overload is defined as a state of network elements in which buffer overflow results in frame discard for flows with no QoS commitments.

3.8 packet classification: The process of distinguishing Ethernet frames for the purpose of applying appropriate traffic control and congestion control mechanisms is called Ethernet frame classification (see clause 7.1.1).

3.9 traffic contract: For a given Ethernet flow, the selected Ethernet transfer capability (see clause 8), the traffic descriptor at a given interface and the QoS class define the traffic contract at that interface. See clause 7.1.5.

3.10 traffic descriptor: A traffic descriptor is the set of traffic parameters that is used to capture the traffic characteristics of an Ethernet flow at a given standardized interface as part of the traffic contract. See clause 7.1.4.

3.11 traffic parameter: A traffic parameter describes one aspect of a flow. See clause 7.1.3.

4 Abbreviations

This Recommendation uses the following abbreviations:

B	Bucket size of a token bucket
Bc	Bucket size of committed token bucket
BCB	Backbone Core Bridge
BE	Best Effort
Be	Bucket size of excess token bucket
BEB	Backbone Edge Bridge
B-TAG	Backbone VLAN TAG
CBS	Committed Burst Size
CF	Coupling Flag
CM	Colour Mode
CIR	Committed Information Rate
C-TAG	Custom VLAN TAG
DBW	Dedicated Bandwidth
EBS	Excess Burst Size
EIR	Excess Information Rate
ETC	Ethernet Transfer capability
GBRA	Generic Byte Rate Algorithm
I-TAG	Service instance TAG
LAN	Local Area Network

M	Maximum allowed frame size
MAC	Media Access Control
MEF	Metro Ethernet Forum
PBN	Provider Bridged Network
PBBN	Provider Backbone Bridged Network
PC	Parameter Control
PVID	Port VLAN ID
QoS	Quality of Service
R	Rate of a token bucket
Rc	Rate of committed token bucket
Re	Rate of excess token bucket
SBW	Statistical Bandwidth
S-TAG	Service VLAN TAG
TB	Token Bucket
TC	Transfer Capability
VID	VLAN Identifier

5 Convention

None.

6 Introduction

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

In Ethernet-based networks, congestion is defined as a state of network elements (e.g., 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 Ethernet frame 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. This Recommendation also specifies a set of Ethernet transfer capabilities (ETCs) including, for each ETC, the service model, the associated traffic patterns and conformance definition for an Ethernet flow that are observable at any point in the network.

7 Traffic parameters and descriptors

7.1 Definitions

7.1.1 Ethernet frame classification

For the purpose of Ethernet traffic control and congestion control, it is essential that not all Ethernet frames are treated the same way but differently, depending on the objectives and on the commitments made. Therefore, it is useful to classify Ethernet frames into different categories.

Ethernet frame can be classified based on information in the Ethernet frame header fields, such as priority, VID, source MAC address and destination MAC address. The level of detail in the classification may be different, which depends on its intended use. Usually, the classification can be done as follows:

– In order to recognize Ethernet frames which are with a given priority in a given VLAN, the priority field and VID field in the corresponding TAG field (C-TAG in IEEE 802.1Q LAN, S-TAG in 802.1ad PBN, I-TAG for BEB and B-TAG for BCB in 802.1ah PBBN) are used in the classification.

– In order to recognize Ethernet frames which flow from a given source MAC address to a given destination MAC address in a given VLAN with a given priority, the source and destination address fields, the priority field and VID field in the corresponding TAG field (C-TAG in IEEE 802.1Q LAN, S-TAG in 802.1ad PBN, I-TAG for BEB and B-TAG for BCB in 802.1ah PBBN) are used in the classification.

For the 802.1D LAN, the VID corresponds to the PVID, and the priority corresponds to the port default priority.

In some conditions, according to different applications, there may be other modes for the classification.

7.1.2 Ethernet flow

An Ethernet flow at a given interface is defined as the occurrence at that interface of the set of Ethernet frames which match a given classification (see clause 7.1.1).

7.1.3 Traffic parameter

A traffic parameter is a specification of a particular traffic aspect. It may be qualitative or quantitative. Traffic parameters may, for example, describe peak bit rate, average bit rate, the average or maximum frame size, etc.

7.1.4 Traffic descriptor

A traffic descriptor is the set of traffic parameters which can be used to capture the traffic characteristics of an Ethernet flow at a given standardized interface as part of the traffic contract (see clause 7.1.5).

7.1.5 Traffic contract

For a given Ethernet flow, the selected Ethernet transfer capability (see clause 8), the traffic descriptor at a given interface (see clause 7.1.4) and the associated QoS class define the traffic contract at that interface.

7.2 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 clause 9.1.3.

7.3 Traffic parameter specifications

7.3.1 Reference configuration

No specific reference configuration is required as this Recommendation is applicable at any point where an Ethernet flow can be observed or monitored. Future editions of this Recommendation may include examples of configurations where the functions described in clause 9 can be applied.

7.3.2 Traffic parameter description

The following traffic parameters may be used in the traffic descriptor.

7.3.2.1 Maximum allowed frame size

The maximum allowed frame size *M* is expressed in bytes. It is a mandatory traffic parameter for each ETC.

7.3.2.2 GBRA or token bucket[1]

The generic byte rate algorithm (GBRA) or token bucket (TB) is used to characterize a rate and the associated burstiness. Their concepts are equivalent and use the following set of two parameters:

– the rate *R* expressed in byte/s;
– the burstiness 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 Ethernet transfer capabilities (see clause 8) recognizes an excess TB (with parameters excess rate *Re* and excess bucket size *Be*) and a committed TB (with parameters committed rate *Rc* and committed bucket size *Bc*).

7.3.2.3 Bandwidth profiles service attributes

A bandwidth profile is a method of characterizing service frames for the purpose of rate enforcement or policing. In order to describe the bandwidth profile, the following parameters are used.

– **Committed information rate** (*CIR*) expressed as bits per second. CIR **MUST** be ≥ 0.
– **Committed burst size** (*CBS*) expressed as bytes. When CIR > 0, CBS **MUST** be greater than or equal to the largest maximum transmission unit size among all of the EVCs that the bandwidth profile applies to.
– **Excess information rate** (*EIR*) expressed as bits per second. EIR **MUST** be ≥ 0.
– **Excess burst size** (*EBS*) expressed as bytes. When EIR > 0, EBS **MUST** be greater than or equal to the largest maximum transmission unit size among all of the EVCs that the bandwidth profile applies to.
– **Coupling flag** (*CF*) **MUST** have only one of two possible values, 0 or 1.

[1] The detailed description of GBRA and token bucket can be found in Annex A of [ITU-T Y.1221].

– **Colour Mode** (*CM*) **MUST** have only one of two possible values, "colour-blind" and "colour-aware".

Bandwidth profile also can be used to describe the traffic capability of a flow. There are relationships between these bandwidth profile parameters and the TB parameters mentioned above. Committed rate *Rc* equals *CIR* divided by 8, and committed bucket size *Bc* equals *CBS*. Excess rate *Re* can equal *EIR* divided by 8 or be calculated according to *EIR* and associated committed bucket situation that is chosen by *CF*, and excess bucket size *Be* equals *EBS*.

8 Ethernet transfer capabilities

Ethernet transfer capabilities are network capabilities provided by Ethernet-based networks to transfer Ethernet frames. An Ethernet transfer capability is a set of traffic parameters and procedures that is intended to support an Ethernet service model and a range of associated QoS classes. Each Ethernet transfer capability specified in terms of a service model, a traffic descriptor, a conformance definition and associated QoS commitments are defined. An Ethernet transfer capability is supported by a set of traffic control and congestion control functions.

Ethernet-based networks should be able to provide multiple transfer capabilities to offer different QoS classes for different applications and to optimize the usage of network resources.

In this Recommendation, three types of transfer capability are defined. They are consistent with the specification in clause 7.11 of [Metro Ethernet].

8.1 Dedicated bandwidth (DBW) transfer capability

8.1.1 Description

The dedicated bandwidth (DBW) transfer capability is intended to support applications with stringent loss and delay requirements. It is used to forward the guaranteed and timely delivery of Ethernet frames.

8.1.2 Service model

The DBW ETC provides a specified committed rate (*Rc*) for real-time applications with limited burst duration.

The commitment made by the network is that all conforming Ethernet frames are assured of the negotiated Ethernet QoS, while the non-conforming frames should be discarded by the network.

The DBW ETC can be associated with specified loss commitments (Ethernet frame loss ratio, FLR) and specified delay commitments (Ethernet frame transfer delay and Ethernet frame delay variation).

8.1.3 Traffic descriptor

The traffic descriptor consists of the following parameters:

– The committed rate *Rc* and committed bucket size *Bc* as specified in clause 7.3.2.2.

– The maximum allowed frame size *M* as specified in clause 7.3.2.1.

8.1.4 Conformance definition

An Ethernet frame is conforming if its arrival conforms to the following two parts:

– the arrival is conforming to the GBRA(*Rc*,*Bc*);

– the actual frame length is shorter than the maximum allowed frame size *M*.

The GBRA is updated for conforming frames only.

8.1.5 QoS commitments

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

If all Ethernet frames are conforming, the QoS commitments apply to all frames. The DBW user should expect that (possibly all) non-conforming frames be discarded by the network. If not all frames are conforming, the network may choose to commit QoS only to the conforming frames.

8.2 Statistical bandwidth (SBW) transfer capability

8.2.1 Description

The statistical bandwidth (SBW) transfer capability is intended to support applications without stringent delay requirements. It aims to support the guaranteed delivery of Ethernet frames.

8.2.2 Service model

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

The commitment made by the network is that all conforming Ethernet frames (not exceed the bound set by GBRA (*Rc*,*Bc*)) are delivered across the network, corresponding to the associated Ethernet QoS class, while the non-conforming frames (exceed the bound set by GBRA (*Rc*,*Bc*)) will only be delivered within the limits of available resources.

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

8.2.3 Traffic descriptor

The traffic descriptor consists of:

– the excess rate *Re* and the excess bucket size *Be* as specified in clause 7.3.2.2;

– the committed rate *Rc* and the committed token bucket size *Bc* as specified in clause 7.3.2.2;

– the maximum allowed frame size *M* as specified in clause 7.3.2.1.

8.2.4 Conformance definition

An Ethernet frame is conforming if its arrival conforms to the following three parts:

– the arrival is conforming to the excess GBRA(*Re*,*Be*);

– the arrival is conforming to the committed GBRA(*Rc*,*Bc*);

– the actual frame length is shorter than the maximum frame size *M*.

The GBRAs are updated in coordinated mode for conforming frames only.

8.2.5 QoS commitments

The SBW capability may be associated with specified loss commitments.

If all Ethernet frames are conforming, the QoS commitments apply to all frames. Otherwise, the QoS commitments apply to the conforming frames. Non-conforming frames will be delivered within the limits of available resources.

8.3 Best effort (BE) transfer capability

8.3.1 Description

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

8.3.2 Service model

The BE transfer capability provides no QoS commitment. Only when sufficient resources are available, frames of best effort flows are forwarded.

8.3.3 Traffic descriptor

The maximum allowed frame size *M*.

8.3.4 Conformance definition

The actual frame length is shorter than the maximum frame size *M*.

8.3.5 QoS commitments

There is no absolute QoS requirement for the best effort transfer capability.

9 Functions for traffic control, congestion control and overload treatment

Traffic control refers to a set of functions that control the flow of frames 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 mis-functioning 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.

9.1 Traffic control functions

The following functions are identified for traffic control:

a) Network resource management.

b) Admission control.

c) Parameter control.

d) Frame marking.

e) Traffic shaping.

f) Frame scheduling.

9.1.1 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.

9.1.2 Admission control

Admission control refers to the policies of the network to admit commitments to a new Ethernet 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.

9.1.3 Parameter control

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

9.1.4 Frame marking

When an Ethernet flow is found to be non-conforming to one or more aspects of the traffic contract, the non-conforming frames may be marked.

9.1.5 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.

9.1.6 Frame scheduling

Frame 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.

9.2 Functions for congestion control and overload treatment

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

a) Frame discard control.

b) Explicit congestion notification.

9.2.1 Frame discard control

Frame discard control is used to discard marked traffic in a congested situation. Another application of frame discard control is, in a congested situation, to discard frames for which no stringent QoS commitments apply.

9.2.2 Explicit congestion notification

The PAUSE operation can be used as explicit congestion notification to inhibit transmission of data frames for a specified period of time (only when IEEE 802.3 LAN operates in full duplex mode). This function is optional.

When congestion occurs, a station will send a PAUSE frame to the corresponding station using the assigned multicast address 01-80-C2-00-00-01. PAUSE frame is a type of MAC control frame whose opcode is 0x0001. The pause_time field in the frame indicates the length of time which it wishes to inhibit data frame transmission. The PAUSE frame is limited between the two stations in the IEEE 802.3 LAN operating in full duplex mode. IEEE 802.1D-conformant bridges will not forward frames sent to this multicast destination address.

After receiving the PAUSE frame, the station will start a pause_timer whose value is set according to the pause_time field in the PAUSE frame received, and stop sending data frames to the corresponding station till the pause_timer expires. If a new PAUSE frame is received before the pause_timer expires, the pause_timer will update its value according to the new PAUSE frame.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
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	Telecommunication management, including TMN and network maintenance
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, open system communications and security
Series Y	**Global information infrastructure, Internet protocol aspects and next-generation networks**
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2008