

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.248.11

(03/2013)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS
Infrastructure of audiovisual services – Communication
procedures

Gateway control protocol: Media gateway overload control package

Recommendation ITU-T H.248.11



International
Telecommunication
Union

Documents
Document Sharing Hub

ITU-T H-SERIES RECOMMENDATIONS
AUDIOVISUAL AND MULTIMEDIA SYSTEMS

CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100–H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	
General	H.200–H.219
Transmission multiplexing and synchronization	H.220–H.229
Systems aspects	H.230–H.239
Communication procedures	H.240–H.259
Coding of moving video	H.260–H.279
Related systems aspects	H.280–H.299
Systems and terminal equipment for audiovisual services	H.300–H.349
Directory services architecture for audiovisual and multimedia services	H.350–H.359
Quality of service architecture for audiovisual and multimedia services	H.360–H.369
Supplementary services for multimedia	H.450–H.499
MOBILITY AND COLLABORATION PROCEDURES	
Overview of Mobility and Collaboration, definitions, protocols and procedures	H.500–H.509
Mobility for H-Series multimedia systems and services	H.510–H.519
Mobile multimedia collaboration applications and services	H.520–H.529
Security for mobile multimedia systems and services	H.530–H.539
Security for mobile multimedia collaboration applications and services	H.540–H.549
Mobility interworking procedures	H.550–H.559
Mobile multimedia collaboration inter-working procedures	H.560–H.569
BROADBAND, TRIPLE-PLAY AND ADVANCED MULTIMEDIA SERVICES	
Broadband multimedia services over VDSL	H.610–H.619
Advanced multimedia services and applications	H.620–H.629
Ubiquitous sensor network applications and Internet of Things	H.640–H.649
IPTV MULTIMEDIA SERVICES AND APPLICATIONS FOR IPTV	
General aspects	H.700–H.719
IPTV terminal devices	H.720–H.729
IPTV middleware	H.730–H.739
IPTV application event handling	H.740–H.749
IPTV metadata	H.750–H.759
IPTV multimedia application frameworks	H.760–H.769
IPTV service discovery up to consumption	H.770–H.779
Digital Signage	H.780–H.789

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

Recommendation ITU-T H.248.11

Gateway control protocol: Media gateway overload control package

Summary

Recommendation ITU-T H.248.11 describes a package for media gateway (MG) overload control for use with the ITU-T H.248.1 gateway control protocol. It serves to protect an MG from processing overload that prevents the timely execution of ITU-T H.248.1 transactions.

In summary, in this Recommendation, overload protection is achieved as follows:

- 1) An MG (or virtual MG) detects that it is in overload and notifies its media gateway controller (MGC) of that fact whenever it receives an ADD command.
- 2) The MGC adaptively throttles the rate at which it sets up calls using that MG (or virtual MG) to maximize the MG's effective throughput whilst bounding its response times. It does this by throttling the rate at which transactions that set up new calls or that new call legs are sent to the overloaded MG, so as to cause the rate of overload notifications the MGC receives from the overloaded MG (or virtual MG) to converge to a suitably low level.

A separate instance of the overload control shall be initiated at an MGC for each of its dependent MGs (or virtual MGs) that is overloaded. These separate instances should run independently (that is, they do not explicitly exchange information). Their overload control parameters shall be separately configurable, for example, by means of a proprietary management interface, or the use of SNMP to invoke configuration functions.

The most general overload scenario the control can handle is where one or more MGCs are jointly overloading a single MG that has several virtual MGs (virtual MG 'i' interacting only with MGC 'i'). The control does not need to know how many MGCs are causing the MG to be overloaded, nor what the MG capacity is. Informative reference [b-Ofitel] provides a full explanation of one way this can be achieved, and informative reference [b-Whitehead] provides further material on designing overload controls.

The overload control is largely specified by saying how it shall behave, but not how it should be implemented to achieve that behaviour. This has two important consequences.

As a first consequence, part of the package (see clause 8.5) defines a set of overload scenarios, and any fully-compliant implementation of the package must automatically (i.e., without the need for operator intervention to adjust parameter values from one overload scenario to another) satisfy all the requirements for each of the scenarios.

As a second consequence, not all configurable parameters can be known to this package since they depend upon specific implementations of the control. Nevertheless, there is a requirement that the implementation shall provide a means by which an operator can change all parameters which affect the performance of the control. See clause 9 for the management requirements associated with this package. It is expected that they will be realized by a proprietary management interface, or the use of SNMP.

This revision adds a note indicating a way for a media gateway to ensure prompt delivery of overload indications.

History

Edition	Recommendation	Approval	Study Group
1.0	ITU-T H.248.11	2002-11-29	16
1.1	ITU-T H.248.11 (2002) Cor. 1	2008-06-13	16
2.0	ITU-T H.248.11	2013-03-16	16

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/>.

© ITU 2013

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

Table of Contents

		Page
1	Scope	1
2	References.....	1
3	Definitions	1
4	Abbreviations and acronyms	3
5	Overload Control package	3
	5.1 Properties.....	3
	5.2 Events	4
6	Signals	4
7	Statistics.....	4
8	Procedures	4
	8.1 Actions at overloaded MG (or virtual MG).....	4
	8.2 Actions at an MGC.....	5
	8.3 Bounding MG response times in the steady-state	9
	8.4 Bounding offered rates to MG during initial overload transient	10
	8.5 Range of overload scenarios.....	10
	8.6 Normalization of the control loop	10
9	Management requirements for ITU-T H.248.11.....	11
	9.1 An approximate performance analysis of leaky bucket restrictors	11
	9.2 Configuration of leaky bucket restrictors at MGC	15
	9.3 Configuration of proprietary parameters relating to overload detection at an MG	16
	9.4 Configuration of proprietary parameters relating to control activation at an MGC	16
	9.5 Configuration of proprietary parameters relating to adaptation of admitted rate at an MGC	16
	9.6 Configuration of control termination at an MGC.....	16
	9.7 MGC statistics	17
	Bibliography.....	18

Recommendation ITU-T H.248.11

Gateway control protocol: Media gateway overload control package

1 Scope

This Recommendation describes a package for the ITU-T H.248.1 gateway protocol related to media gateway overload control. With the root termination implementing this package, a gateway is expected to report MG_Overload events to a media gateway controller.

The following MGC/MG characteristics are required for this package to work effectively:

- 1) The applications consist of individual point-to-point calls between two end users, and of calls involving several additional call legs (for example, a conference call where an MG hosts a conference bridge).
- 2) Point-to-point call set-up using an MG typically involves no more than two ADD commands requested by its MGC.

An example of the applications covered by such an MGC is real-time conversational voice transported using TDM, ATM, IP or frame relay.

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 H.248.1] Recommendation ITU-T H.248.1 (2013), *Gateway control protocol: Version 3*.
- [ITU-T Q.543] Recommendation ITU-T Q.543 (1993), *Digital exchange performance design objectives*.
- [ITU-T Q.714] Recommendation ITU-T Q.714 (2001), *Signalling connection control part procedures*.

3 Definitions

This Recommendation defines the following terms:

3.1 call: In this package, "Call" means a generic term, meaningful only to MGCs, related to the establishment of a path to carry user data between end users via an MG. Normally a qualifier is necessary to make clear the aspect being considered, e.g., call attempt.

3.2 call attempt: In this package, "Call Attempt" means an attempt to set up a call.

3.3 load: In this package, "Load" means the total number of "call attempts" in a given interval of time (i.e., offered load).

This definition is motivated by alignment with performance objectives of [ITU-T Q.543].

3.4 MG overload and MG capacity: In this package, "MG Overload" means that the MG (or virtual MG) is close to being unable to respond to MGC transactions in a sufficiently timely manner to avoid the calling customer abandoning the call in set up.

This definition is motivated by the need to maximize the calls/second handled by the MG subject to keeping most MG response times short enough to prevent customers from abandoning a call before it is established end-to-end between the calling and the called customers.

In this package, "MG Capacity" means the calling rate (calls/second) at which the MG detects it is in MG Overload. Thus the MG capacity depends on the specific sequence of ITU-T H.248.1 commands (ADD, MODIFY, etc.) which implement the calls, as well as on the processing power of the MG.

The precise way MG overload is detected is likely to be implementation-specific, since different suppliers' switches are likely to have different architectures. Overload detection could, for example, be based on the use of processor occupancy thresholds, or queue thresholds or internal delay thresholds. Each overload detection scheme would have an associated set of configurable parameters, which however cannot in general be known to this package. Despite this, an operator will need to be able to configure an MG's overload detection scheme to ensure suitably short response times under overload. It is therefore necessary for the MG/MGC implementation to provide a means for an operator to configure such parameters via, for example, a proprietary management interface, or the use of SNMP, not the ITU-T H.248.1 interface. Clause 9 gives the management requirements for this package.

There may be MG resources (e.g., codecs, ATM bandwidth, tone generators) whose congestion, although preventing a call from being set up, does not prevent timely response of the MG to MGC transactions. Congestion of those resources must not cause MGC overload control (as defined in this package) to be invoked. The reasons for this are that:

- a) congestion of such resources does not (by definition) prevent timely response to MGC transactions; and
- b) it is probably more appropriate to flag such congestion via existing ITU-T H.248.1 error messages in order to trigger a capacity upgrade of the congested resource.

3.5 leaky bucket: This package requires the MGC to restrict the rate it admits calls to an MG which has reported that it is congested. Any one of the following three types of leaky bucket is acceptable for use as a call restrictor with this package, provided its implementation can be configured to cover the required range of admitted rates implied by the set of overload control scenarios given in clause 8.5. See clause 9 for an approximate performance analysis.

Type 1 leaky bucket: This is a count which decreases by a configurable LeakAmount every LeakInterval (subject to not falling below 0), and which increases by a configurable SplashAmount at each call arrival (subject to not exceeding the configurable MaximumFill count). A call which arrives to find the count less than or equal to the {MaximumFill – SplashAmount} is admitted (and the count is incremented by SplashAmount); otherwise, the call is rejected and the count is not incremented. The maximum sustained admitted rate is approximately (LeakAmount/SplashAmount)/LeakInterval provided that LeakAmount ≤ MaximumFill. The LeakInterval is adaptively changed by the overload control.

Type 2 leaky bucket: This is a count that is decremented by (Now – TimeOfLastDecrement) × LeakAmount/LeakInterval at each call arrival instant (subject to not falling below 0). It is then incremented by SplashAmount if the count is less than or equal to the {MaximumFill – SplashAmount} and the call is admitted; otherwise, the call is rejected and the count is not incremented. The maximum sustained admitted rate is (LeakAmount/SplashAmount)/LeakInterval provided that LeakAmount ≤ MaximumFill. The LeakInterval is adaptively changed by the overload control.

Type 3 leaky bucket: This is a count that is decremented by LeakAmount every LeakInterval (subject to not falling below 0). LeakInterval is an operator-configurable fixed period; LeakAmount is adaptively changed by the overload control. A call which arrives to find the count less than or equal to the {MaximumFill – SplashAmount} is admitted (and the count is incremented by