Project	**IEEE 802.16 Broadband Wireless Access Working Group <http://ieee802.org/16>**	
Title	**Corrections to QoS parameter text**	
Date Submitted	**2004-11-18**	
Source(s)	Jeff Mandin Streetwaves Networking Amatzia 5 Jerusalem, Israel	Voice: 972-50-724-587 Fax: 972-50-724-587 mailto:jeff@streetwaves-networks.com
	Joel Demarty Ambroise Popper Sequans	ambroise@sequans.com, joel@sequans.com
	Itzikk Kitroser Runcom Technologies	itzikk@runcom.co.il
Re:	Corrigenda call for contributions	
Abstract	Corrections to QoS Parameter Text	
Purpose	Adoption and inclusion in 802.16d corrigenda	
Notice	This document has been prepared to assist IEEE 802.16. It is offered as a basis for discussion and is not binding on the contributing individual(s) or organization(s). The material in this document is subject to change in form and content after further study. The contributor(s) reserve(s) the right to add, amend or withdraw material contained herein.	
Release	The contributor grants a free, irrevocable license to the IEEE to incorporate material contained in this contribution, and any modifications thereof, in the creation of an IEEE Standards publication; to copyright in the IEEE's name any IEEE Standards publication even though it may include portions of this contribution; and at the IEEE's sole discretion to permit others to reproduce in whole or in part the resulting IEEE Standards publication. The contributor also acknowledges and accepts that this contribution may be made public by IEEE 802.16.	
Patent Policy and Procedures	The contributor is familiar with the IEEE 802.16 Patent Policy and Procedures <http://ieee802.org/16/ipr/patents/policy.html>, including the statement "IEEE standards may include the known use of patent(s), including patent applications, provided the IEEE receives assurance from the patent holder or applicant with respect to patents essential for compliance with both mandatory and optional portions of the standard." Early disclosure to the Working Group of patent information that might be relevant to the standard is essential to reduce the possibility for delays in the development process and increase the likelihood that the draft publication will be approved for publication. Please notify the Chair <mailto:chair@wirelessman.org> as early as possible, in written or electronic form, if patented technology (or technology under patent application) might be incorporated into a draft standard being developed within the IEEE 802.16 Working Group. The Chair will disclose this notification via the IEEE 802.16 web site <http://ieee802.org/16/ipr/patents/notices>.	

Corrections to QoS parameters text

Jeff Mandin
Streetwaves Networking

1 Justifications for changes

1. Maximum Sustained Traffic Rate

There is ambiguity regarding at what point (ie. where in the logical stack) the traffic rate should be measured for the purpose of limiting the rate (ingress to BS or SS? before PHS? After PHS?).

There is ambiguity regarding which MAC overhead is not counted (ie. we should emphasize that MAC headers are included when they are part of the SDU eg. with 802.3-style convergence sublayer).

The statement that the instantaneous rate of the service is limited by the rate of the ingress port is plain wrong. Similarly, the statement that policing needs to be done on the uplink only and not the downlink is nonsensical, and conflicts with main body of the QoS text (in section 6).

There is ambiguity regarding what is done with SDUs that exceed the Maximum Sustained Traffic Rate (ie we should say that "you drop them or queue them, it's up to the vendor").

There is inconsistency with the main QoS text in 6.3.5.2.1 which states that this parameter is used to calculate the time between grants for UL-UGS service (ie. we should say that here).

2. Maximum Traffic Burst

The intent of the description of this parameter is not clear. Moreover, it is not mentioned in the main QoS text at all.

3. Minimum Reserved Traffic Rate

The statement that the service flow's less-than-reserved-rate available data "should be transmitted as soon as possible" is nonsensical and should be deleted.

The parameter description only addresses uplink scheduling, even though the main text applies the parameter to downlink scheduling. For consistency, the parameter description must apply to both uplink and downlink.

4. Maximum Delay

There is nowhere in the document where a parameter is imposed that applies outside the MAC or PHY. Hence the measuring of Maximum Latency begins at "reception on a network interface" is an error (probably a vestige from DOCSIS) and should be corrected to refer to ingress at the CS.

5. Applicable direction of Scheduling Services

The QoS framework applies to both uplink and downlink service flows. However, there are several occurrence of the text (probably vestige from DOCSIS) that imply that QoS applies only to uplink service flows. Also, section 6.3.5.2 called "Uplink request/grant scheduling" defines the various scheduling services which is misleading.

2 Text Changes

[Modify 6.3.5]

6.3.5 Scheduling services

Scheduling services represent the data handling mechanisms supported by the MAC scheduler for data transport on a connection. Each connection is associated with a single data service. Each data service is associated with a set of QoS parameters which quantify aspects of its behavior. These parameters are managed using the DSA and DSC message dialogs. Four services (11.13.11) are supported: Unsolicited Grant Service (UGS), Real-time Polling Service (rtPS), Non-real-time Polling Service (nrtPS), and Best Effort (BE). The following text provides a brief description of each of the supported scheduling services, including the mandatory QoS parameters that shall be included in the service flow definition when the scheduling service is enabled for a service flow. A detailed description of each QoS parameter is provided in 11.13.

The UGS is designed to ~~support~~ transport in the uplink direction real-time data streams consisting of fixed-size data packets issued at periodic intervals, such as T1/E1 and Voice over IP without silence suppression. The mandatory QoS service flow parameters for this scheduling service are Maximum Sustained Traffic Rate (11.13.6), Maximum Latency (11.13.14), Tolerated Jitter (11.13.13), and Request/Transmission Policy (11.13.12). If present, the Minimum Reserved Traffic Rate parameter (11.13.8) shall have the same value as the Maximum Sustained Traffic Rate parameter.

The rtPS is designed to ~~support~~ transport in the uplink direction real-time data streams consisting of variable-sized data packets that are issued at periodic intervals, such as moving pictures experts group (MPEG) video. The mandatory QoS service flow parameters for this scheduling service are Minimum Reserved Traffic Rate (11.13.8), Maximum Sustained Traffic Rate (11.13.6), Maximum Latency (11.13.14), and Request/Transmission Policy (11.13.12).

The nrtPS is designed to ~~support~~ transport in the uplink direction delay-tolerant data streams consisting of variable-sized data packets for which a minimum data rate is required, such as FTP. The mandatory QoS service flow parameters for this scheduling service are Minimum Reserved Traffic Rate (11.13.8), Maximum Sustained Traffic Rate (11.13.6), Traffic Priority (11.13.5), and Request/Transmission Policy (11.13.12).

The BE service is designed to support data streams for which no minimum service level is required and therefore may be handled on a space-available basis. The mandatory QoS service flow parameters for this scheduling service are Maximum Sustained Traffic Rate (11.13.6), Traffic Priority (11.13.5), and Request/ Transmission Policy (11.13.12).

[Renumber 6.3.5.2.1 to 6.3.5.3 and modify]

~~6.3.5.2.1~~ **6.3.5.3 UGS**

The UGS is designed to support real-time uplink service flows that generate fixed size data packets on a periodic basis, such as T1/E1 and Voice over IP without silence suppression. The service offers fixed size grants on a real-time periodic basis, which eliminate the overhead and latency of SS requests and assure that grants are available to meet the flow's real-time needs. The BS shall provide Data Grant Burst IEs to the SS at periodic intervals based upon the Maximum Sustained Traffic Rate of the service flow. The size of these grants shall be sufficient to hold the fixed length data associated with the service flow (with associated generic MAC header and Grant management subheader) but may be larger at the discretion of the BS scheduler. In order for this service to work correctly, the Request/Transmission Policy (see 11.13.12) setting shall be such that the SS is prohibited from using any contention request opportunities for this connection. The key service IEs are the Maximum Sustained Traffic, Maximum Latency, the Tolerated Jitter, and the Request/Transmission Policy. If present, the Minimum Reserved Traffic Rate parameter shall have the same value as the Maximum Sustained Traffic Rate parameter.

[Renumber 6.3.5.2.2 to 6.3.5.4 and modify]

~~6.3.5.2.2~~ **6.3.5.4 rtPS**

The rtPS is designed to support real-time uplink service flows that generate variable size data packets on a periodic basis, such as moving pictures experts group (MPEG) video. The service offers real-time, periodic, unicast request opportunities, which meet the flow's real-time needs and allow the SS to specify the size of the desired grant. This service requires more request overhead than UGS, but supports variable grant sizes for optimum data transport efficiency.

The BS shall provide periodic unicast request opportunities. In order for this service to work correctly, the Request/Transmission Policy setting (see 11.13.12) shall be such that the SS is prohibited from using any contention request opportunities for that connection. The BS may issue unicast request opportunities as prescribed by this service even if prior requests are currently unfulfilled. This results in the SS using only unicast request opportunities in order to obtain uplink transmission opportunities (the SS could still use unsolicited Data Grant Burst Types for uplink transmission as well). All other bits of the Request/Transmission Policy are irrelevant to the fundamental operation of this scheduling service and should be set according to network policy. The key service IEs are the Maximum Sustained Traffic Rate, the Minimum Reserved Traffic Rate, the Maximum Latency and the Request/Transmission Policy.

[Renumber 6.3.5.2.3 to 6.3.5.5 and modify]

~~6.3.5.2.3~~ **6.3.5.5 nrtPS**

The nrtPS offers unicast polls on a regular basis, which assures that the uplink service flow receives request opportunities even during network congestion. The BS typically polls nrtPS CIDs on an interval on the order of one second or less.

The BS shall provide timely unicast request opportunities. In order for this service to work correctly, the Request/Transmission Policy setting (see 11.13.12) shall be set such that the SS is allowed to use contention request opportunities. This results in the SS using contention request opportunities as well as unicast request opportunities and unsolicited Data Grant Burst Types. All other bits of the Request/Transmission Policy are

irrelevant to the fundamental operation of this scheduling service and should be set according to network policy.

[Renumber 6.3.5.2.4 to 6.3.5.6 and modify]

~~6.3.5.2.4~~ 6.3.5.6 BE service

The intent of the BE service is to provide efficient service for best effort traffic in uplink and downlink. In the case of an uplink service flow, in order for this service to work correctly, the Request/Transmission Policy setting shall be set such that the SS is allowed to use contention request opportunities. This results in the SS using contention request opportunities as well as unicast request opportunities and unsolicited Data Grant Burst Types. All other bits of the Request/Transmission Policy are irrelevant to the fundamental operation of this scheduling service and should be set according to network policy. In the case of a downlink service flow, the Request/Transmission Policy parameter is irrelevant and shall be omitted.

[Modify 11.13.7 as follows:]

11.13.7 Maximum sustained traffic rate

This parameter defines the peak information rate of the service. The rate is expressed in bits per second and pertains to the SDUs at the input to the Convergence Sublayer ~~system~~. Hence ~~Explicitly~~, this parameter does not include 802.16 MAC overhead such as MAC headers or CRCs. ~~This parameter does not limit the instantaneous rate of the service since this is governed by the physical attributes of the ingress port. However,~~ at the BS and SS ~~in the uplink direction~~, the service shall be policed to conform to this parameter, on the average, over time. ~~At the BS in the downlink direction, it may be assumed that the service was already policed at the ingress to the network and the BS is not required to do additional policing~~. If this parameter is omitted or set to zero, then there is no explicitly mandated maximum rate. This field specifies only a bound, not a guarantee that the rate is available.

The algorithm for measuring whether a flow exceeds its Maximum Sustained Traffic Rate ~~for policing to this parameter~~ is left to vendor differentiation and is outside the scope of the standard - however the algorithm should ensure that within a vendor-defined time interval T, the amount of data for the SF that is forwarded to the MAC CPS should not exceed [(T * Maximum Sustained Traffic Rate) + Maximum Traffic Burst]. SDUs deemed to exceed the Maximum Sustained Traffic Rate may be delayed, dropped, or fragmented according to the discretion of the vendor.

[replace section 11.13.8]

This parameter describes the maximum burst for the service flow ie. the maximum number of bytes by which the SF is allowed to exceed its Maximum Sustained Traffic Rate.

[Modify section 11.13.9 as follows:]

11.13.9 Minimum reserved traffic rate

This parameter specifies the minimum rate reserved for this service flow. The rate is expressed in bits per second and specifies a guaranteed ~~the minimum~~ amount of SDU data to be transported on behalf of the service flow over the air interface when averaged over time. The specified rate shall only be honored when sufficient data is available for scheduling. ~~When insufficient data exists, the requirement imposed by this parameter shall be satisfied by assuring the the available data is transmitted as soon as possible.~~

~~The BS shall be able to satisfy bandwidth requests for a service flow up to its Minimum Reserved Traffic Rate.~~ If less data ~~bandwidth~~ than the ~~its~~ Minimum Reserved Traffic Rate is available ~~requested~~ for a service flow, the BS may reallocate the excess reserved bandwidth for other purposes. The aggregate Minimum Reserved Traffic Rate of all service flows may exceed the amount of available bandwidth. The data for this parameter is measured before the application of any header suppresion. ~~The value of this parameter is calculated from the byte following the MAC header HCS to the end of the MAC PDU payload.~~ If this parameter is omitted, then it defaults to a value of 0 bits per second (i.e., no bandwidth is reserved for the flow).

[Modify 11.13.14 as follows:]

11.13.14 Maximum latency

The value of this parameter specifies the maximum latency between the ingress of a packet to the Convergence Sublayer ~~reception of a packet by the BS or SS on its network interface~~ and the forwarding of the SDU ~~packet~~ to its RF Interface.

If defined, this parameter represents a service commitment (or admission criteria) at the BS or SS and shall be guaranteed by the BS or SS. A BS or SS does not have to meet this service commitment for service flows that exceed their minimum reserved rate.