Project	**IEEE 802.16 Broadband Wireless Access Working Group <http://ieee802.org/16>**
Title	**Classes of Service Definition for 802.16e**
Date Submitted	**10-01-2003**
Source(s)	Vladimir Yanover Leonid Shousterman Ronnie Shechter Alvarion Ltd. 11/5 Shtern Str. Herzlya, Israel Voice: +972-36457834 Fax: +972-36456222 mailto:vladimir.yanover@alvarion.com mailto::leonid.shousterman@alvarion.com mailto:ronnie.shechter@alvarion.com
Re:	Request for contributions for 802.16e Global Service Flows Ad Hoc Group, 19 Sep 2003; the proposal is a revision of "Classes of Services Definition" section of IEEE 802.16.1mc-00/03 "802.16 BWA Air Interface Medium Access Control. Proposal for Standard"
Abstract	The document contains suggestions for definition of Classes of Service for 802.16e from the prospect of 1) types of QoS contracts 2) set of parameters for each contract 3) expected system behavior
Purpose	The document is contributed to be discussed in 802.16e Global Service Flows Ad Hoc Group
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 text 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 (Version 1.0) <http://ieee802.org/16/ipr/patents/policy.html>, including the statement "IEEE standards may include the known use of patent(s), including patent applications, if there is technical justification in the opinion of the standards-developing committee and provided the IEEE receives assurance from the patent holder that it will license applicants under reasonable terms and conditions for the purpose of implementing 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:r.b.marks@ieee.org> as early as possible, in written or electronic form, of any patents (granted or under application) that may cover technology that is under consideration by or has been approved by IEEE 802.16. The Chair will disclose this notification via the IEEE 802.16 web site <http://ieee802.org/16/ipr/patents/notices>.

Classes of Service Definition for 802.16e

Vladimir Yanover, Leonid Shousterman, Ronnie Shechter
Alvarion Ltd.

1. References

[1] IEEE P802.16/D5-2001, IEEE Draft Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed Broadband Wireless Access Systems, 2001-10-18
[2] IEEE P802.16c/D4-2002, Draft Amendment to IEEE Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed Broadband Wireless Access Systems – Detailed System Profiles for 10-66 GHz
[3] IEEE P802.16a/D7-2002, Draft Amendment to IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems - Medium Access Control Modifications and Additional Physical Layer Specifications for 2-11 GHz, 2002-11-17
[4] P802.16d/D3-2003 Draft Amendment to IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems - Amendment 3: Detailed System Pro?les for 2-11 GHz, 2003-06-03
[5] IEEE P802.16-2003/D0, IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems, Consolidated Draft (2003-08-01)
[6] IEEE 802.16e-03/07r3 TGe Working Document (2003-08-12)
[7] IEEE 802.16.1mc-00/03 “802.16 BWA Air Interface Medium Access Control. Proposal for Standard” by Vladimir Yanover and Leonid Shousterman

2. The Document’s Goal

The document’s goal is to propose a definition of Classes of Service for 802.16e from the prospect of 1) types of QoS contracts 2) set of parameters for each contract 3) expected system behavior. The proposal is a revision of “Classes of Services Definition” section of IEEE 802.16.1mc-00/03 “802.16 BWA Air Interface Medium Access Control. Proposal for Standard” by Vladimir Yanover and Leonid Shousterman.

3. Incentive for Development of Data Delivery Service QoS Specification

There is no ***direct*** specification of data delivery services in QoS terms in 802.16 MAC, except UGS service in 6.2.5 “Uplink scheduling service”. Note that UGS service is for UL only. Other definitions in that section refer to request / grant process but not data delivery. For example, definition of Real-Time Polling Service (rtPS) in 6.2.5.2: ”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**.” Note that BS is not explicitly prescribed to react to the request.

Reason for that lays in the scope of the standard, that does not include specifications related to scheduling.

On the contrary, 802.16e needs tools for description of data delivery services in the terms of QoS parameters (rate, delay etc.). Classes of Services shall be used then for quantization of possible types of data delivery services.

Such a description will be stored in centrally located databases (ASA servers [6]) and then used each time when a MSS considers a handover to another BS and – after the handover – when creating connections to/from the MSS.

Such a definition is assumed referring to *a **set** of service flows* and may, for example, include a requirement to provide service for two service flows: one with up to 128 Kbps at DL and another 64 Kbps at UL, both with max latency = 20 ms – all together referred to as "SILVER-1" Class of Service (name given by network operator or group of operators). Then, prior to the HO, the MSS may ask a candidate BS whether addition of one MSS "SILVER-1" service is available etc. Note that in [1] meaning of "Class of Service" term is different.

"QoS Profile" means the set of QoS-related Service Flow parameters. The above example is described by the following table (see definitions of QoS Profiles types in 5):

Table 1. Example of service with specification of QoS parameters

Service that fits Class of Service "SILVER-1"	
DL Service Flow with SFID = 0100 and QoS Profile - Type = NRT-VR - Minimum Reserved Traffic Rate = 128 Kbps - Maximum Sustained Traffic Rate = 256 Kbps - Maximum Latency = 20 ms	UL Service Flow with SFID = 0101 and QoS Profile - Type = NRT-VR - Minimum Reserved Traffic Rate = 64 Kbps - Maximum Sustained Traffic Rate = 128 Kbps - Maximum Latency = 20 ms

4. Correction in Definition of Parameters

4.1. Minimum Reserved Traffic Rate

This definition should be detailed similarly to the definition of Maximum Sustained Traffic Rate ([1] , 11.4.8.7). A new element is time base T. This value is assumed to be much greater than the system latency ("Maximum Latency" parameter; if not specified, default = 1 sec).

"Minimum Reserved Traffic Rate = R (bits/sec) with time base T(sec)" means the following. Let S denote additional demand accumulated at the MAC SAP of the transmitter during an arbitrary time interval of the length T. Then the amount of data forwarded to PHY (in bits) during this interval should be not less than min{S, R * T}.

4.2. Minimum Tolerable Traffic Rate

"Minimum Tolerable Traffic Rate = R (bits/sec) with time base T(sec)" means the following. Let S denote additional demand accumulated at the MAC SAP of the

transmitter during an arbitrary time interval of the length T. Then the amount of data forwarded at the receiver to CS (in bits) during this interval should be not less than min{S, R * T}.

In the case of DL connection, Minimum Tolerable Traffic Rate should be monitored by the BS to make decisions on rate change and/or deletion of the connection in the case SDUs' loss ratio is too high. So SS has to measure and report certain parameters to BS. See possible solution in 7.

5. Types of QoS Profile

5.1. Service Flow QoS Profile Type parameter

The following are possible values for "Type of QoS profile" parameter

Type	Symbolic Name	Meaning
0	CG	Continuing Grant service
1	CG-AD	Continuing Grant with Activity Detection service
2	RT-VR	Real Time – Variable Rate service
3	NRT-VR	Non-Real Time – Variable Rate service
4	BE	Best Efforts service

See below detailed definitions for the services.

5.2. Continuing Grant (CG)

This type of QoS Profile is designed to support real-time services generating fixed-size data packets on a periodic basis.

The following are the parameters of the profile:

Table 2. Continuing Grant Service Parameters

Parameter	Meaning
SDU Size	According to [1] , 11.4.8.16. The BS should grant to the connection resources sufficient for transferring SDU of the given size (it means that all MAC and PHY overheads should be taken into account)
Maximum Latency	As specified in [1] , 11.4.8.14
Minimum Tolerable Traffic Rate	Optional parameter described in 4.2. This value should not be higher than the corresponding Minimum Reserved Traffic Rate value = (SDU Size)*8/(Maximum Latency). The difference between these two values reflects SDUs' loss rate. Default = 0.

Description of service

Each SDU arrived to the transmitter's MAC SAP, should be transmitted within time interval T = Maximum Latency.

5.3. Continuing Grant with Activity Detection (CG-AD)

This QoS Profile is designed to support real-time applications generating fixed-size data packets on a periodic basis, which may become inactive for long periods of time.

All parameters described in the Table 2 are relevant; one more parameter needed:

Table 3. Continuing Grant with Activity Detection Additional Parameters

Parameter	Meaning
Activity Ramp up Time	Maximum time (sec) allocated for the system to recognize a demand that appears after certain period of inactivity

Description of service

Each SDU arrived to the transmitter's MAC SAP, should be transmitted within time interval T = Maximum Latency. If during time interval A = Activity Ramp up Time no SDUs delivered to MAC SAP, the system enters the idle state, when delivery of SDUs is not guaranteed (and then the SDUs are discarded), and it may take time = A sec to return to active state and thus to start delivering service flow SDUs.

5.4. Real-Time Variable Rate (RT-VR) QoS Profile

This QoS Profile is designed to support real-time data applications with variable bit rates which require guaranteed data rate and delay. The QoS Profile is defined by the following parameters:

Table 4. Real Time Variable rate Parameters

Parameter	Meaning
Maximum Latency	As specified in [1] , 11.4.8.14
Time Base	T as used in 4; should be much larger than Maximum Latency
Minimum Reserved Traffic Rate	As specified in 4.1 with time base T
Minimum Tolerable Traffic Rate	As specified in 4.2. with time base T. This value should not be larger than Minimum Reserved Traffic Rate. The difference between these two values reflects SDUs' loss rate

Description of service

Let S denote the amount of data arrived to the transmitter's MAC SAP, during time interval T = Time Base; R = Minimum Reserved Traffic Rate. Then the amount of data forwarded at the receiver to CS (in bits) during this interval should be not less than min{S, R * T} while any SDU delivered to the CS, should be delayed by the system for not more than D = Maximum Latency.

5.5. Non-Real-Time Variable Rate (NRT-VR)

This QoS profile shall support applications with that require a guaranteed data rate but are insensitive to delays. It is desirable in certain cases to limit the data rate of these services to some maximum rate. The QoS profile is defined by the following parameters defined as in the Table 4 (note absence of Maximum Latency parameter):

- Time Base
- Minimum Reserved Traffic Rate
- Minimum Tolerable Traffic Rate

There are two additional parameters:

Parameter	Meaning
Maximum Sustained Traffic Rate	As specified in [1] , 11.4.8.7
Minimum Tolerable Traffic Rate	Optional parameter described in 4.2. This value should not be larger than Maximum Sustained Traffic Rate. Default = 0.

Description of service
Let S denote the amount of data arrived to the transmitter's MAC SAP, during time interval T = Time Base; R = Minimum Reserved Traffic Rate. Then the amount of data forwarded at the receiver to CS (in bits) during this interval should be not less than min{S, R * T}

5.6. Best Effort (BE)

This QoS profile is designed for services with no rate or delay requirements. The following parameters are applicable:

Parameter	Meaning
Maximum Sustained Traffic Rate	As specified in [1] , 11.4.8.7

6. Encoding of QoS Parameters

TBD

7. Reporting of QoS Related Measurements from SS to BS

7.1. Bandwidth Usage Report

This section includes a suggestion on implementation of new function when SS reports to BS on actual amount of data arrived through certain DL connection. This information can be used for estimation whether requested Minimum Tolerable Traffic Rate value is followed and then for decision on rate change and/or deletion of the connection for relevant QoS profiles in the case SDUs' loss ratio is too high.

The suggestion is to use a short BW Report MAC message similar to Bandwidth Request Header structure ([1] , 6.2.2.1.2) but with new Type value:

HT(1)	EC(1)	Type = 000002	BWRPT msb (8)
BWRPT lsb (8)			CID msb (8)
CID lsb (8)			HCS (8)

The value of BWRPT (bandwidth usage report) field is the total accumulated amount of data, in units of 1024 bytes, correctly received at the connection with given CID. This number wraps around after reached the value 2^16; interval between reports should be small enough to resolve ambiguity of this value.

7.2. Report Request from BS

It is suggested to redefine the second reserved bit in Generic MAC Header ([5] [1] , 6.4.2.1.1) as “BW Report Request”. If it is set to ‘1’ in an arbitrary MAC message sent at DL connection, then it means that a bandwidth usage report from the SS is requested.