Project	IEEE 802.16 Broadband Wireless Access Working Group <**http://ieee802.org/16**>	
Title	Optimizing the uplink Bandwidth Request Latency	
Date Submitted	2005-11-15	
Source(s)	Muthaiah Venkatachalam Intel Corporation	Voice: 503-712-5583 Email: muthaiah.venkatachalam@intel.com
Re:	Call for comments	
Abstract		
Purpose	This document proposes a method to improve the uplink channel access latency	
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>.	

Optimizing the Bandwidth Request Latency

Muthaiah Venkatachalam, Intel Corporation

1 Introduction

This contribution suggests a mechanism to optimize the uplink bandwidth request latency to be included in [4].

Optimizing the UL bandwidth request latency (and hence the UL response time) for 802.16 [1] is key, if 802.16 has to support some of the key applications like Push-to-talk (PTT) and even some of the elastic data applications like TCP.

802.16 is a centrally scheduled network, which means that the bandwidth grants for the MS to access the UL has to come from the BS.

There are several QoS classes to enable the UL access [2] – depending upon the nature of the application. On the one extreme there is the unsolicited grant service (UGS) where the BS provides a continuous stream of grants to the MS without the MS asking for each and every grant in particular. This works great for real time and streaming applications, however, this will be very wasteful of bandwidth for things like bursty channel access by the MS, as in the case of a bursty TCP connection.

Hence applications that are bursty in nature will need to use a different QoS mechanism like Best Effort (BE) or nrtPS or rtPs. In the case of BE/nrtPS/rtPS, there are four fundamental ways to request bandwidth in the 802.16 specification:

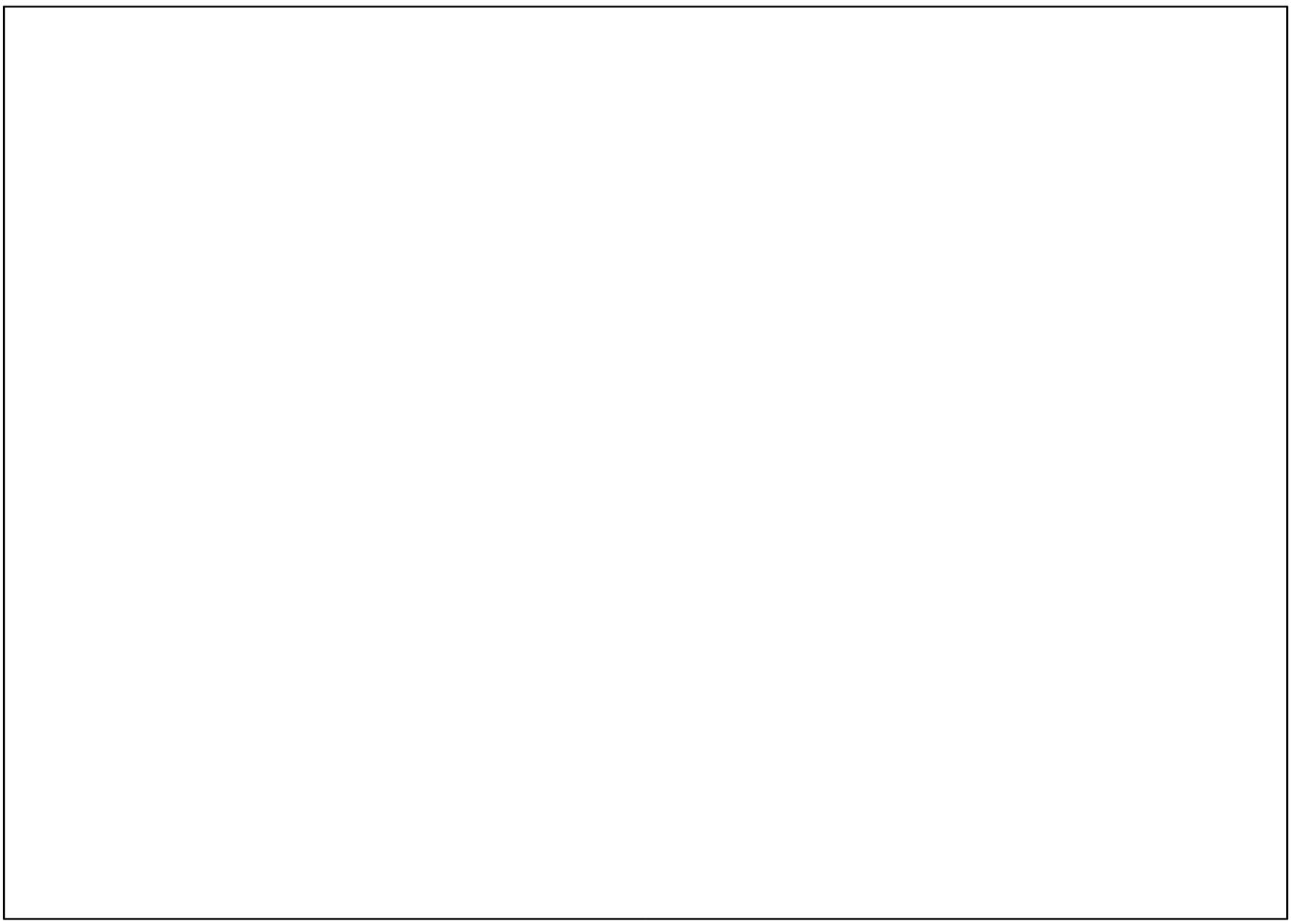
a) use contention bandwidth request mechanism – ie the MS contends for bandwidth on the contention channel and if the request goes through to the BS without any collisions (if there are collisions, the MS tries again), the BS allocates an “alloc-IE” for this MS and then the MS uses the alloc-IE to request bandwidth and the BS allocates the requested bandwidth on this alloc-IE by using a BW-REQ header. The BW-REQ header contains the exact amount of bandwidth needed (incremental or aggregate) and the BS responds to this header by actually allocating the amount of bandwidth needed

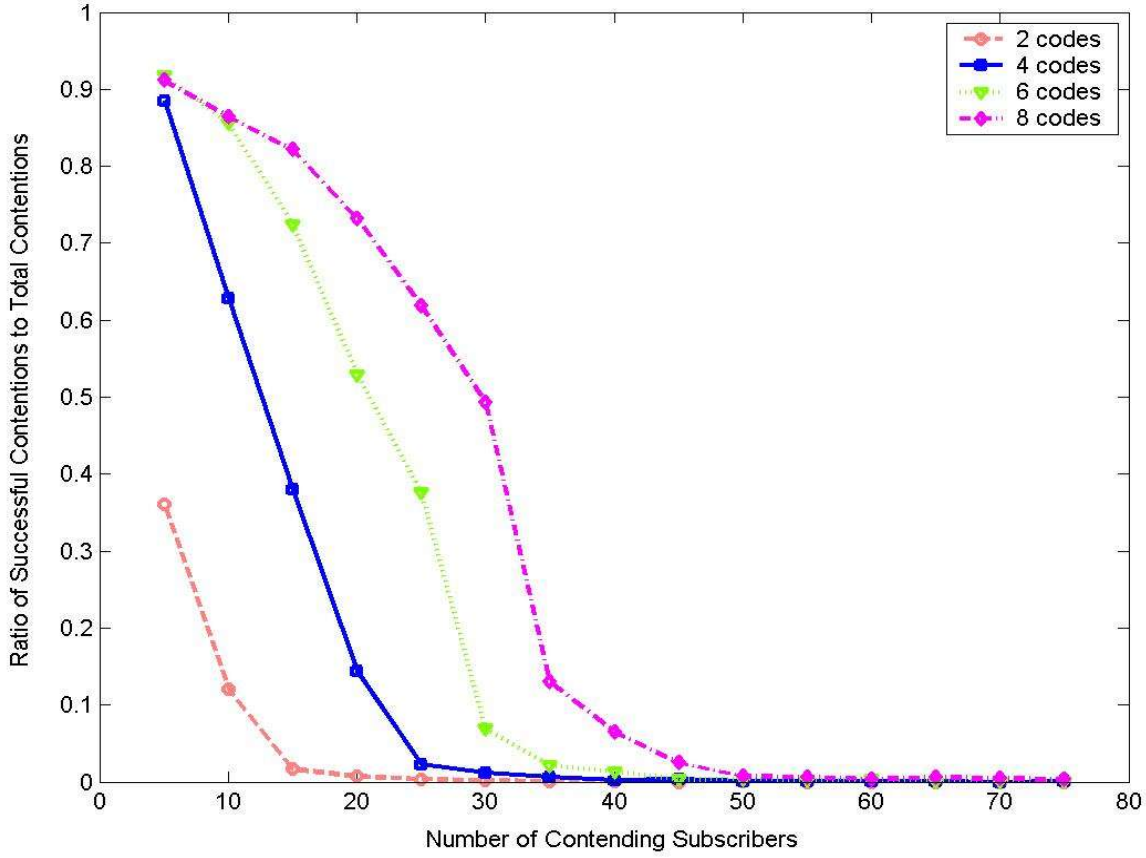
b) grant management subheader - if the MS has an ongoing UGS CID for a different application, the MS can set a “poll ME” bit in the Grant management subheader of the UGS connection, and when the BS receives this PM bit, it will understand that the MS needs more bandwidth for the current or new application and subsequently allocates some extra bandwidth for the MS to send its BW-REQ header. The rest of the steps are as in step a). If the MS doesn’t have an UGS connection, it can request the amount of bandwidth needed via a piggy back request for that particular CID.

c) Steal bandwidth from other ongoing QoS connections like UGS/rtPS/exrtPS to issue a bandwidth request PDU.

d) Utilize the polling grants by the BS to request bandwidth via BW-REQ-PDU. (only for nrtPS/nrtPs QoS)

Option a): Simulation studies (See figure-1) have shown that option a) can have a lot of latency for bandwidth request, since it is dependent on the contention request getting through without collisions; or under heavy load - several retries for the request to get through; and even after the request getting through, 2 more cycles of messaging for the MS to actually get the bandwidth. From figure-1, it can be seen that under high load conditions, the average number of retries (given by 1/P(success)) can be of the order of 100s of times or more. The resulting latency can be very detrimental to end to end TCP performance and associated round trip times.





(GMH), instead of just the grant management sub header of the UGS connection as it stands today. The poll ME bit can be included without an increase in the size of the GMH, by utilizing an existing reserved bit in the header. The newly proposed generic MAC header looks like in fig 1.

Field	Size (bits)
HT	1
EC	1
Type	6
ESF	1
CI	1
EKS	2
PM (was reserved)	**1**
LEN	11
CID	16
HCS	8
TOTAL	48

Figure-2: New proposed 802.16 Generic MAC header to reduce the UL latency.

Also the processing at the MS and BS for this is simple enough that no significant processing cycles are needed to do this in both MS and BS.

With this new feature, one usage model will could be as follows: A MS can request to be polled for incremental bandwidth for any of its CIDs by setting the PM bit on ANY MAC PDU that it transmits on any of its CIDs. Consider a scenario, where the MS has two CIDs - CID1 and CID2, which happens to be BE type, and all its data sessions are transmitted on these CIDs. CID1 has ongoing data session(s) (eg: HTTP connection or a File transfer etc). CID2 is trying to start up a data session. For the initiation of a new data session on CID2, the MS

can get quick access to the channel using our proposal, by setting the PM bit in the GMH of a MAC PDU that is transmitted by active data session(s) on CID1. It cannot use the Grant management subheader and associated bandwidth request piggybacking in this case.

Post initial access, for regular data transfer, the MS can get the required bandwidth for any CID1/CID2 in a more optimal/quicker fashion, by using the poll me bit in the GMH of the MAC PDUs being transmitted on any of its CIDs. If there is a temporary lapse in transmission on CID1, due to the bursty nature of traffic, the MS can quickly request bandwidth on CID1 at a later point by setting the GMH Poll ME bit on PDUs being transmitted on CID2 at that point in time.

2 Proposed Text Changes

Section 6.3.2.1 –

Include Poll-ME (PM) bit in table 4 based on Figure-2 of this contribution

Section 6.3.2.1.1 –

Include PM bit in Figure 19 based on Figure-2 of this contribution

Include the following description of the PM bit in table 5.

Name	Length (bits)	Description
PM	1	Poll ME bit. Can be set by the MS when transmitting an UL MAC PDU on its UL grant.

Section 6.3.5.2 –

Table 112: Add “polling using the PM bit in GMH” under the polling column for all the QoS classes

Add a new section 6.3.6.7 with the following text:

Another method of requesting UL bandwidth is to use the Poll ME bit in the GMH. Any MS transmitting any MAC PDU on the UL on any service class and on any CID can use this bit. When the BS receives this PDU with the PM bit set, it knows that the MS requires more bandwidth. The BS may then use this information to schedule some grant for the MS to request bandwidth, in a future frame. The MS may utilize this grant by sending a BW-REQ PDU to the BS, indicating the exact amount of bandwidth needed.

3 References

[1] IEEE P802.16e/D12 Draft standard

[2] IEEE 802.16-2004 Standard

[3] IEEE P802.16-2004/Cor1/D5 Draft Standard

[4] IEEE P802.16g baseline document http://ieee802.org/16/netman/docs/80216g-05_008r1.pdf