# PERFORMANCE ANALYSIS OF THE POLLING SCHEME IN IEEE 802.16

Alexey Vinel
State University of Aerospace Instrumentation, Saint-Petersburg, Russia
Dirk Staehle, Rastin Pries
Department of Distributed Systems, University of Wuerzburg, Germany
Qiang Ni
Electronics and Computer Engineering, Brunel University, London, UK
E-mail: VINEL@IEEE.ORG

In this paper, a method for the computation of the mean delay for the request transmission in IEEE 802.16 broadband network is presented. It is assumed, that unicast polling is used and the base station cyclically polls the subscriber stations. Performance of the polling is compared to the performance of the binary exponential backoff random multiple access algorithm in terms of mean delay for the request transmission. The analysis is conducted with an error-prone wireless channel.

## 1. Introduction

Broadband wireless access (BWA) has gained great attention recently. In 1999 a working group called IEEE 802.16 was set up by the IEEE 802 committee to develop a new standard for BWA applications. Later, another industrial association, namely worldwide interoperability for microwave access (WiMAX) forum [1], was formed to promote the 802.16 standards by defining the interoperability specifications between 802.16 products from different vendors. In October 2001 the first IEEE 802.16 standard [2] was completed. It addressed radio frequency bands from 10 to 66 GHz, and thus line of sight (LOS) is required between a base station (BS) and subscriber stations (SSs). In January 2003 an amendment called IEEE 802.16a was ratified by operating the physical (PHY) layer at lower frequency bands from 2 to 11 GHz and thus allowing the possibility of non-line-of-sight (NLOS) operation. The new standard 802.16-2004 [3] was published in October 2004. It is actually an amalgamation of 802.16 and 802.16a, which specifies interoperable air interfaces from 2 to 66 GHz with a common medium access control (MAC) layer. Finally, the 802.16e standard [4] was also ratified in December 2005 by allowing the upgrade from fixed BWA systems to mobile service provision up to vehicular speeds.

The IEEE 802.16 MAC protocol supports two operational modes: a mandatory point-to-multipoint (PMP) mode and an optional mesh mode. In the PMP mode, a centralized BS controls all the communications between SSs and the BS using an antenna sector, whereas, in the mesh mode, SSs can also serve as routers and cooperate access control in a distributed manner. In this article, we focus on the centralized PMP mode. In a downlink subframe of the PMP mode, the BS transmits a burst of MAC protocol data units (PDUs) using Time Division Multiplexing (TDM) mechanism; while in an uplink subframe of the PMP mode, an SS transmits a burst of MAC PDUs to the BS using the Time Division Multiple Access (TDMA) technique.

Resource management and allocation mechanisms are crucial to guarantee QoS performance in 802.16 WiMAX networks. Under a centralized PMP architecture of 802.16, multiple SSs share a common uplink to the BS on a demand basis. This means that if an SS needs some amounts of bandwidth, it requests from the BS by transferring a request message. On accepting the request from a SS, the BS scheduler should determine and grant it a transmission opportunity in time slots by using some scheduling algorithms, which should take into account the requirements from all authorized SSs and the available channel resources. These grants are made based on the negotiated QoS agreements between the BS and SSs. Two main methods are suggested in the WiMAX standard to offer transmission opportunities for SSs to send their bandwidth request (BW-REQ) messages: centralized polling and contention-based random access. In the first case, each SS station is only allowed to send its request when it is polled by the BS, where in the latter one all SS stations contend to obtain transmission opportunities for sending requests by using some contention resolution mechanisms. For the high priority traffic, only polling can be used, while the bandwidth for the transmission of the lower priority traffic can be requested either by means of random access or in dedicated slots in contention-free manner. Thus, on the one hand it is important to study performance of polling itself as well as to compare the performance of different BW-REQ delivery methods. In [5] and [6] random access performance of IEEE 802.16 has been analyzed thoroughly and very simplified analysis has been used for the investigation of polling in [6]. Moreover, the analysis has been conducted for the error-free channel. In this paper, we investigate polling mechanisms of WiMAX by considering error-prone channel with arbitrary request arrival rates using Markov-process model. For the general theoretical discussions of different polling systems we refer the reader, for instance, to [7], [8] and [9]. Here, without losing the generality of the problem, a Bernoulli request arrival process is chosen with finite number of stations. We limit our analysis to the investigation of BW-REQ delay performance during the reservation process, where actual data packet transmission was not included in this work. This allows us to first focus our attention on improving the efficiency of the BW-REQ algorithms as it is a fundamental component for the complete analysis of the 802.16 MAC protocols.

## 2. Overview of IEEE 802.16 BW-REQ Mechanisms

The WiMAX MAC layer is connection-oriented, and it is designed with QoS support by allowing bandwidth reservation and flexible implementation of resource scheduling/admission control mechanisms. All services are mapped to connections. Any application from upper-layer first has to establish a connection with the BS. The BS then assigns each connection with a unique connection ID (CID). This mechanism applies to all services, including inherently connectionless services, in order to provide a mechanism for requesting bandwidth, associating QoS and traffic parameters, transporting and routing data, and other actions associated with the services. Both Time Division Duplexing (TDD) and Frequency Division Duplexing (FDD) modes are supported in WiMAX. In the TDD case, each MAC frame includes a downlink sub-frame followed by an uplink sub-frame; where in the FDD case, the uplink sub-frame could be slightly delayed with respect to the downlink sub-frame so that the SSs can receive necessary information about the uplink channel access from the downlink.

Under the PMP architecture, all the transmissions between the BS and SSs are coordinated by the BS. In this article, we only focus on the TDMA/TDD transmission mode, where similar analysis can also be applied to FDD configuration. The TDMA/TDD frame consists of a downlink sub-frame for transmission from the BS to SSs and an uplink sub-frame for transmissions in the reverse direction. The Tx/Rx transition gap (TTG) and the Rx/Tx transition gap (RTG) are specified between the downlink and uplink sub-frames, and between the uplink and following downlink sub-frames in the next frame duration respectively to allow SS terminals to turn around from reception to transmission and vice versa. In the downlink sub-frame, both the Downlink MAP (DL-MAP) and Uplink MAP (UL-MAP) messages are transmitted, which comprise the bandwidth allocations for data transmission in both downlink and uplink directions, respectively. Moreover, the lengths of uplink and downlink sub-frames are determined dynamically by the BS and are broadcast to the SSs through UL-MAP and DL-MAP messages at the beginning of each frame. Therefore, each SS knows when and how long to receive from and transmit data to the BS. This functionality considers that most Internet applications have more downstream traffic than upstream (known as bandwidth asymmetry) and the bandwidth allocated to each direction can be tuned dynamically to match the traffic in the corresponding direction.

The uplink sub-frame contains transmission opportunities scheduled for the purpose of sending BW-REQ messages, in which BW-REQ messages can be transmitted, which serves for SSs to indicate to the BS that they need UL bandwidth allocation. The BS controls both the number of transmission opportunities for BW-REQ and data packet transmission through the UL-MAP message.

A BW-REQ can be issued either in a stand-alone request, or in an uplink data packet as a piggyback request. Note that the capability of piggyback request is optional. In order to determine which SS is allowed to transmit its BW-REQ from multiple candidates, two main methods are suggested in the standard, contention-based random access and contention-free based polling. In both schemes, no explicit acknowledgement (ACK) frame is sent back to indicate whether a BW-REQ message is successfully transmitted, or distorted (possibly due to channel noise or collision), or how much bandwidth the SS is granted. If a grant is not given within a special timeout - T16 3, the SS should determine that BW-REQ was corrupted, and then start contention resolution process. On the other hand, on receiving a grant within the timeout, the SS will stop contention resolution process and uses the allocated bandwidth for uplink transmission of data packets, or to piggyback additional request if necessary. Furthermore, the SS might know how much bandwidth awarded by observing the following grant from the BS. Due to different scheduling algorithms at a BS, a grant may be given at any time.

In the case of random access, an SS transmits a BW-REQ during a predefined contention period and a random backoff mechanism is used to resolve contention among the BW-REQ PDUs from multiple SSs. The mandatory method of random access-based contention resolution mechanism used in WiMAX, is based on a truncated binary exponential backoff (BEB) scheme ([5],[6]).

When polling-based BW-REQ allocation is chosen, the BS shall maintain a list of registered SSs and poll them according to this list. Each SS is only allowed to transmit the BW-REQ message after it is polled. Actually, the poll schedule information for polling-based BW-REQ is carried by the UL-MAP and UCD in the downlink sub-frames. Note that scheduling algorithms for polling are vendor-dependent, and are not specified in the standard. One may choose a simple round-robin scheduler to poll each SS sequentially in the polling list, but other priority-based polling mechanisms might also be used for BW-REQ scheduling if different QoS levels are required by different SSs. Furthermore, the polling allocation can be issued to a group of SSs. Allocation schedules to groups are also indicated in UL-MAP and UCD. This grouping mechanism is particularly important when available bandwidth is insufficient for a BS to individually poll many inactive SSs, and thus only those active groups of SSs should be polled in multicast groups or a broadcast poll may be used to save the resource usage. Certain CIDs are reserved for multicast groups and broadcast messages as specified in the standard.

## 3. System Model for Studying Subscriber Stations Polling

Let us consider a PMP system in which there is one BS, and the total number of SS stations is $n$. Each SS station has a buffer sufficient to store exactly one request. A station, which has a request in the considered moment of time is referred to as "active", otherwise it is called a "non-active" one. According to the WiMAX standard [3], each SS may potentially establish several connections with different negotiated QoS parameters with the BS, and a BW-REQ should be issued per-connection based. In this work, we assume that each SS has only one connection at a given time. In the case of multiple connections per SS, n is referred to the total number of connections in the system.

During one frame duration, each "non-active" SS generates a request with a probability $p = \lambda/n$, where $\lambda$ is the mean number of the requests generated by the system in that frame if all SSs are "non-active". This new request is put into the buffer, and will be transmitted no earlier than a next frame transmission. Since only "non-active" stations can generate a request, the actual requests arrival rate in a frame can sometimes be lower than $\lambda$ depending on the system load.

We assume, that all SSs are numbered and each slot is consequently assigned to all the SSs. Taking into account, that for most of the practical cases $n >> K$, complete polling cycle takes $f = n / K$ frames (for simplicity, we assume, that $K$ is the divisor of $n$ and all frames have cyclic numeration from 1 to $f$). Information about the slots assignment in the current frame is provided by the BS to SSs by its.

In this paper, we focus our analysis on uplink BW-REQ transmission. The transmission of data packets in both directions is ignored. The time duration of each frame is fixed, in which $K$ BW-REQ slots are included. The duration of a slot corresponds to the time needed for a BW-REQ transmission, which is PHY-layer dependent.

Error-prone channel is considered in our model. When channel is error-prone, BW-REQ message may be corrupted due to poor channel conditions, e.g., path-loss, multipath fading, thermal noise or interference from other emitting sources nearby. This additional damaging effect should be modelled in order to investigate effective solutions as the actual channels are normally noisy. We assume that the wireless channel is a Gaussian one, in which each bit has the same bit error probability, and bit errors are identically and independently distributed (i.i.d.) over the whole BW-REQ frame. While Gaussian channel assumption is not realistic, it is widely used due to its simplicity. The consideration of other sophisticated channel models will be our future work. In this work, pe denotes the probability of a request corrupted by channel noise.

The delay performance of the system is defined as the time interval (measured in frames) from the moment of issuing the BW-REQ until the moment, when the station knows that the request has been successfully transmitted. We denote D the mean of this random variable. The value of mean delay for a particular BW-REQ mechanism is a significant performance metric indicating its efficiency, by which we choose to compare different BW-REQ mechanisms.

## 4. Analytical Modeling of the Simplest Polling Scheme

First let us notice that in the framework of our model, the request delay is not independent of the request generation probability and cannot be written out explicitly. This section provides the detailed mathematical analysis of the reservation process.

Let us consider the operation of some arbitrary SS. Without loosing the generality, let us assume, that it is polled in the frames having number 1, then its operation can be modeled by means of two-dimensional discrete-time Markov chain

$$\{F(t), S(t)\}, \quad (1)$$

where states can change each other on the frames bounds, $F(t)$ – number of a frame, which starts at moment $t$, and $S(t)$ – state of considered SS (0 – "non-active", 1 – "active"). Then SS operation can be modeled by means of ergodic chain (Figure 1), having the following transition probabilities:

$$\begin{aligned}
&P\{F(t+1) = (i+1)\bmod(f+1)+1, S(t+1) = 0 \mid F(t) = i, S(t) = 0\} = 1-p, \\
&\qquad i \in [1, f]; \\
&P\{F(t+1) = (k+1)\bmod(f+1)+1, S(t+1) = 1 \mid F(t) = k, S(t) = 0\} = p, \\
&\qquad \text{for } k \in [1, f]; \\
&P\{F(t+1) = (j+1)\bmod(f+1)+1, S(t+1) = 1 \mid F(t) = j, S(t) = 1\} = 1, \text{ for} \\
&\qquad j \in [2, f]; \\
&P\{F(t+1) = 2, S(t+1) = 0 \mid F(t) = 1, S(t) = 1\} = 1-p_e \\
&P\{F(t+1) = 2, S(t+1) = 1 \mid F(t) = 1, S(t) = 1\} = p_e,
\end{aligned} \quad (2)$$

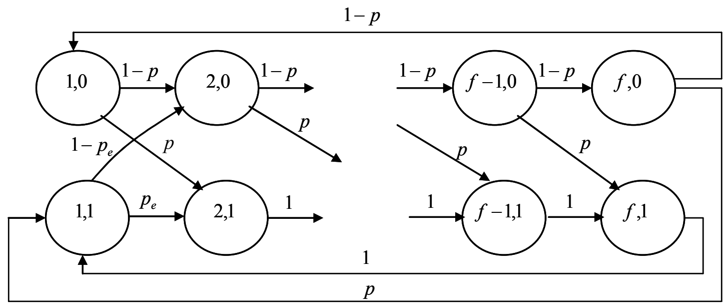


*Figure 1. Markovian polling model*

Below are some comments for (2). First line describes the probabilities of events, which corresponds to non-active state of a subscriber, second one – transitions to active, third one – waiting till the beginning of next frame with number 1, forth one – transmission in this frame and the last one is the request distortion by noise. Denote $p_{i,j}$ – stationary probability corresponding to the state $\{F(t)=i, S(t)=j\}$, then following system of equations holds:

$$\begin{cases} p_{1,0} = p_{f,0}(1-p) \\ p_{2,0} = p_{1,0}(1-p) + p_{1,1}(1-p_e) \\ p_{3,0} = p_{2,0}(1-p) \\ \dots \\ p_{f,0} = p_{f-1,0}(1-p) \\ \\ p_{1,1} = p_{f,0}p + p_{f,1} \\ p_{2,1} = p_{1,0}p + p_{1,1}p_e \\ p_{3,1} = p_{2,0}p + p_{2,1} \\ \dots \\ p_{f,1} = p_{f-1,0}p + p_{f-1,1} \\ \\ \sum_{i=1}^{f} p_{i,0} + \sum_{j=1}^{f} p_{j,1} = 1 \end{cases} \quad (3)$$

From equations from 2 to $f$ of system (3) it is easy to obtain

$$p_{i,0} = [p_{f,0}(1-p)^2 + p_{1,1}(1-p_e)](1-p)^{i-2}, \text{ for } i \in [2, f], \quad (4)$$

and, from equations having numbers from $f+1$ to $2f$ of system (3) and, taking into account the first equation from (3), we have

$$p_{j,1} = p\sum_{i=1}^{j-1} p_{i,0} + p_{1,1}p_e = p[p_{1,0} + \sum_{i=2}^{j-1} p_{i,0}] + p_{1,1}p_e, \text{ for } j \in [2, f]. \quad (5)$$

Substituting (4) into (5) we can simplify to

$$p_{j,1} = p_{f,0}(1-p) + p_{1,1} - p_{j,0} = \tag{6}$$

$$= p_{f,0}(1-p) + p_{1,1} - [p_{f,0}(1-p)^2 + p_{1,1}(1-p_e)](1-p)^{j-2},$$

what, substituting $j = f$, and, applying equation number $f+1$ from (3), after algebraic simplification, leads to

$$p_{f,1} = \frac{p_{f,0}}{1-p_e}[\frac{1}{(1-p)^{f-2}} - (1-p-pp_e+p^2)]$$

and

$$p_{1,1} = \frac{p_{f,0}}{1-p_e}[\frac{1}{(1-p)^{f-2}} - (1-p-pp_e+p^2)] + p_{f,0}p = \tag{7}$$

$$= \frac{p_{f,0}}{1-p_e}[\frac{1}{(1-p)^{f-2}} - (1-p)^2].$$

Finally, using normalization condition (the last equation from (3)), and substituting into it equations (4) – (7), as well as applying the first equation from (3), we obtain

$$p_{f,0} = \frac{(1-p_e)(1-p)^{f-2}}{f[1+(1-p)^{f-1}(p-p_e)]},$$

what, taking into account (4) – (7) allows us to compute the stationary distribution

$$p_{1,0} = \frac{(1-p_e)(1-p)^{f-1}}{f(1+(1-p)^{f-1}(p-p_e))},$$

$$p_{i,0} = \frac{(1-p_e)(1-p)^{i-2}}{f(1+(1-p)^{f-1}(p-p_e))},\ 2 \le i \le f;$$

$$p_{1,1} = \frac{1-(1-p)^f}{f(1+(1-p)^{f-1}(p-p_e))},$$

$$p_{j,1} = \frac{1+(1-p)^{f-1}(p-p_e)-(1-p_e)(1-p)^{j-2}}{f(1+(1-p)^{f-1}(p-p_e))},\ 2 \le j \le f.$$

Thus, mean delay D equals to:

$$D(p,p_e) = 0.5 + \sum_{i=1}^{f} \frac{p_{f-i+1,0}}{\sum_{j=1}^{f} p_{j,0}} i = \tag{8}$$

$$= 0.5 + \frac{(1-p)^f + f(1-p)^{f-1}p^2 + fp - 1}{p(1-(1-p)^f)} + \frac{p_e f}{1-p_e},$$

From (7), applying Lopital rule, we see, that

$$\lim_{p\to 0}\left[\frac{(1-p)^f + f(1-p)^{f-1}p^2 + fp - 1}{p(1-(1-p)^f)} + \frac{p_e f}{1-p_e}\right] = \frac{f+1}{2} + \frac{p_e f}{1-p_e},$$

So in the error-free channel for $p \to 0$ mean delay equals to $f/2+1$, and for "saturation" case, when $p=1$, mean delay equals to $D(1) = f - 1/2$. The heuristic explanation of this fact is the following. In "saturation" case, approximately complete polling cycle is needed to transmit the request, because new request is generated as soon as the previous one has been sent.

## 5. Numerical results

Let us consider first example scenario with parameters $n = 30$, $K = 5$ and compute the mean delay for the BW-REQ transmission arrival rate values per slot $np/K$ (Figure 2) for error-free channel ( $p_e = 0$ ). Analogous values are computed for the binary exponential backoff algorithm by means of method from 6. First of all, notice, that analytical model results (lines) match the simulation ones (points). One can see, that for this example it is better to switch from random access to polling, when the probability $p$ of SS to generate a request for the frame duration is larger than 0.4. Let us compare the performance of binary exponential backoff and cyclic polling in the noisy channel conditions (Figure 3). An interesting observation, that crossing-point between the curves moves left, when the noise probability $p_e$ increases. This means, that for the noisy channel (for instance, when $p_e$=0.3) it is reasonable to switch to the polling when $p$ becomes larger than approximately 0.3 (and not 0.4 as for the ideal channel conditions).

We do not give any "typical" values for $p$ as this probability strongly depends on the SSs scheduling scheme (which is vendor-dependent) as well as traffic type. On the one hand, it seems that polling scheme is preferable in most of the cases. However, note, that if optional "piggybacking" capability to transmit the requests is used, in real system probability $p$ has a rather small value, what is a good scenario for the random access usage.

## 6. Conclusion

We have introduced the model for the SSs unicast polling analysis in IEEE 802.16. Assuming, that each SS can store not more than one request per unit of time (what is a reasonable assumption, since BW-REQ includes actual SS's bandwidth need), we have investigated the influence of the request generation probability and its distortion by noise on the efficiency of polling and random access.

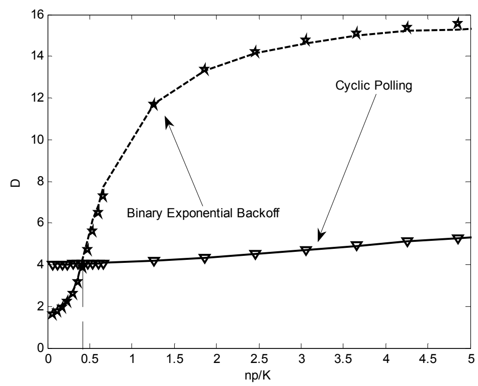


*Figure 2. Random access and cyclic polling for ideal channel conditions*

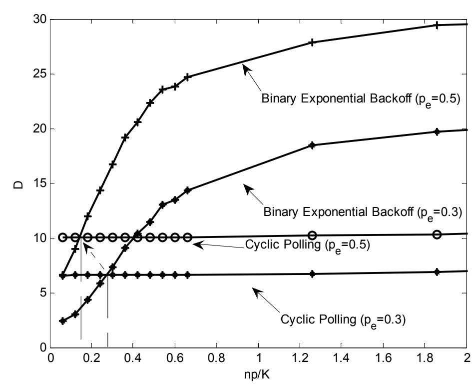


*Figure 3. Random access and cyclic polling for noisy channel conditions*

**Acknowledgement**

This work was funded for Dr. Alexey Vinel by the German Academic Exchange Service (DAAD) jointly with Russian Federal Agency for Education in the framework of "Mikhail Lomonosov 2006/2007" program and was done during his internship at the Department of Distributed Systems, University of Wuerzburg, Germany.

## REFERENCES

1. WiMAX Forum. [Online]. Available: http://www.wimaxforum.org/.
2. IEEE Std. 802.16-2001, IEEE Standard for Local and Metropolitan Area Networks, Part 16: Air Interface for Fixed Broadband Wireless Access Systems, 2001.

3. IEEE Std. 802.16-2004, IEEE Standard for Local and Metropolitan Area Networks, Part 16: Air Interface for Fixed Broadband Wireless Access Systems, Oct 2004.
4. IEEE Std. 802.16e-2005, IEEE standard for local and metropolitan area networks, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems, Dec 2005.
5. Vinel A., Zhang. Y., Lott M., Turlikov A. Performance Analysis of the Random Access in IEEE 802.16 // Proc. of the 16th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications – IEEE PIMRC'05, Berlin, Germany, 2005, pp. 1596–1600.
6. Vinel A., Zhang Y., Ni Q., Lyakhov A. Efficient Request Mechanisms Usage in IEEE 802.16 // Proc. of 49th IEEE Global Telecommunications Conference – GLOBECOM'06, San Francisco, California, USA, 2006.
7. Takagi H. Analysis of polling systems. MIT Press, 1986. 175 p.
8. Borst S.C. Polling systems. Amsterdam: Stichting Mathematisch Centrum, 1996.
9. Vishnevsky V.M., Semenova O.V., Mathematical methods to study the polling systems // Automation and Remote Control, 2006, Vol. 67, No. 2, pp. 173–220.