The Impact of Transmit Buffer on EDCF with Frame-Bursting Option for Wireless Networks

S. Selvakennedy
School of Information Technologies
Madsen Bldg. F09, University of Sydney
NSW 2006, Australia.
skennedy@it.usyd.edu.au

Abstract

Wireless LANs are increasingly being used for inelastic applications. Currently, there is little support for quality of service in the IEEE 802.11 MAC protocol, and IEEE task group E is working on the 802.11e MAC extension. Enhanced distributed coordination function (EDCF) is a contention-based scheme of the 802.11e proposal. To allow a station to transmit more than one frame from a single contention, an optional feature known as controlled frame-bursting (CFB) is introduced in its proposal. In this paper, we performed a detailed evaluation and comparison of the EDCF protocol with the CFB option to quantify its performance gain. The impact of the MAC transmit buffer size is also incorporated. Accordingly, we have proposed a suitable approach to guide the configuration of the burst duration limit. It is demonstrated that an optimized CFB configuration allows the MAC protocol to achieve 30% more capacity than the basic EDCF scheme.

Keywords: 802.11e, MAC, contention-based, EDCF, frame-bursting, QoS.

1. Introduction

Wireless local area networks (WLANs) are seeing more deployment as a supplementary technology to the third generation (3G) cellular networks in the provision of data services. Since the standardization of WLAN protocol stack by the IEEE 802.11 subcommittee in 1999, the deployment of these networks is occurring at a greater pace. The original standard provided data rates up to 2 Mbps at the unlicensed ISM (2.4 GHz) band. Later, the IEEE 802.11 working group published several enhanced physical layer (PHY) specifications, namely 802.11b PHY with data rate up to 11 Mbps in the ISM band, 802.11a PHY that can achieve a data rate up to 54 Mbps using orthogonal frequency division multiplexing (OFDM) in the 5 GHz unlicensed national information infrastructure (UNII) band and 802.11g PHY with similar rate as 802.11a but works at the ISM band. All these PHYs have the same IEEE 802.11 MAC protocol specified for the channel access control. With the standards supporting relatively higher data rates, WLANs became widely installed at homes, corporate building, campuses, airports and hot-spots.

As these networks become more ubiquitous, the mobile users are increasingly using common multimedia applications predominantly used on desktop computers with wired access to the networks. When there exists various traffic mix with differing requirements on these networks, there is a serious need to provide service differentiation known as quality of service (QoS) to enable inelastic traffic types like telephony calls, to be treated with urgency against bursty traffic type of web browsing-like applications in the face of shared bandwidth contention. For the wired networks, IETF has defined the integrated services (IntServ) and differentiated services (DiffServ) models for IP networks with QoS requirements. For the wireless networks, the QoS provision is more critical as the wireless bandwidth is significantly more limited. Wireless medium is also subject to fast changes in signal-to-noise ratio, which affects the bit error rate experienced by the wireless stations. Thus, it is difficult to provide hard QoS guarantees. Instead, the wireless networks should minimally provide soft QoS guarantees for the support of these multimedia applications.

There is a significant number of research efforts to address the issues of providing QoS in WLANs prior to IEEE 802.11 standardization and thereafter. MAC protocols could generally be categorized into centralized or distributed type. Centralized protocols, such as time-division multiple access, reservation-based schemes and polling schemes, enable more
stringent QoS guarantees through a single point of coordination. The coordinator (known as base station or access point) is able to carry out admission control, scheduling and channel access control, and thus, is more suited to infrastructure type WLANs (versus ad hoc type). Some examples of centralized protocols are point coordination function (PCF) of 802.11, hybrid coordination function (HCF) of the upcoming 802.11e [1] where both employs polling mechanisms, HIPERLAN/2 of European Telecommunications Standards Institute (ETSI) and many wireless asynchronous transfer mode (ATM) proposals. The acceptance of these protocols has generally been lukewarm due to high overhead, high cost/complexity and limited scalability. On the contrary, the distributed protocols are simpler to implement and require only limited overhead. They also have applications in domains where infrastructure mode is infeasible or difficult to be built like in battlefields temporary festive or business venues and historic sites.

Some proposals prior to 802.11 standardization were Aloha, CSMA, MACA [2] and MACAW [3]. After the standardization of 802.11 protocol stack that was based on a variant of CSMA, there were a number of proposals to extend distributed coordination function (DCF) to support service differentiation by varying random backoff periods, contention windows or both. Some of these main proposals were Distributed Fair Sharing (DFS) [4], Blackburst [5] and EDCF of IEEE 802.11e (recently renamed as Enhanced Distribution Channel Access). All these distributed protocols are contention-based schemes and thus, need to address the contention and collision problems efficiently.

The suitability of the proposed schemes in 802.11e and their enhancements had already been investigated [6]-[11]. It was shown in [8] that EDCF works well for service differentiation and priority access. EDCF was demonstrated to assure better performance for high priority classes but at the cost of lower priority classes [9], and HCF was able to assure better QoS support when the medium is heavily loaded [10]. However, HCF involves state at the access point (AP) making it a less robust protocol. EDCF was also shown to support simultaneously a large number of voice and video flows in hot-spots and home network scenarios [11]. The authors here have also suggested to make HCF optional to speed up the standardization process to avoid vendors from providing proprietary QoS solutions.

Within the current IEEE 802.11e proposal, there is also an option to use the controlled frame-bursting (CFB) mechanism to achieve better medium utilization through reduced collisions [1]. Once a wireless station (WSTA) contends and wins access to the channel, the CFB option would enable it to send more than one data frame without further contention during the current transmission opportunity. The WSTA is allowed to transmit as many frames as permitted within a certain limit specified in the MAC MIB. Since uncontrolled bursting may increase the frame delay variation, judicious use of the feature is necessary. It was reported that the use of CFB might enhance the performance and achieve better utilization [12].

In this paper, we investigate the effectiveness of the CFB feature in detail. The effect of the bursting mechanism is studied in a network with various mix of multimedia traffic and different offered load levels. It is also expected that the MAC transmit buffer size to have a critical effect on the CFB mechanism. As such, the impact of the MAC buffer size is also investigated. Based on these investigations, we also propose a suitable guide for the configuration of the burst duration limit to achieve an optimized EDCF protocol performance. The detailed investigation is carried out through simulation.

The remainder of this paper is organized as follows. In next section, we describe the basic EDCF scheme and the CFB mechanism in detail. Our simulation methodology and the adopted network scenario are described in Section 3. Section 4 presents the relevant performance metrics for each traffic class with its tolerable limits. Subsequently, the simulation results from the system experimentation are analyzed. The paper concludes by summarizing main findings in the final section, while highlighting extensions to this work.

2. The enhanced DCF (EDCF) scheme

The upcoming IEEE 802.11e MAC protocol is conceived as a compatible extension of the previous IEEE 802.11 MAC. It includes both a contention-based scheme known as EDCF and a more complete polling-based scheme known as HCF.

EDCF is governed through a distributed mechanism very similar to the existing DCF that relies on CSMA/CA protocol. EDCF adds prioritization by allowing different traffic classes to be mapped to access categories (ACs). As specified in the upcoming standard, EDCF supports up to eight traffic classes. One or more traffic classes can be mapped to each AC. Each AC has a separate queue and its associated contention and backoff values. In DCF, the backoff counter of a WSTA only commences when the wireless medium has been idle for at least a DCF interframe space (DIFS) time interval of fixed length. In EDCF, the interval can be different for each AC and
is designated as Arbitration IFS (AIFS). An AIFS may be equal or greater than a DIFS.

Each AC contends for medium access as a separate DCF instance within a WSTA using its own contention parameters, AIFS, minimum contention window (\(CW_{\text{min}}\)) and maximum contention window (\(CW_{\text{max}}\)). When the medium is sensed idle for an AIFS period, a WSTA waits for an additional random period known as a backoff period. This period is computed from the contention window (CW) value. CW is initially set to the \(CW_{\text{min}}\). The backoff period (in multiple of time slots) is determined uniformly from the interval \([0, CW]\). When the medium is found busy before the backoff timer expires, the timer is frozen and the WSTA has to abstain from any transmission for the duration of the current usage. When the medium becomes idle again, this WSTA restarts the frozen timer. When the timer reaches zero, the frame is transmitted. If an external collision is detected, CW is increased by a factor known as persistence factor (PF) as follows:

\[
CW_i = \min(PF \times CW_{i-1}, CW_{\text{max}})
\]

where \(i\) is the number of attempts to transmit the frame. PF can be varied too for each AC to provide further differentiation. In 802.11 MAC, its value defaults to two representing a binary exponential backoff scheme.

ACs with smaller values for these contention parameters will generally experience lower mean waiting delay; thus, getting higher priority to the medium. For each WSTA, multiple ACs may set off in parallel leading to internal virtual collision. Virtual collisions between ACs within a WSTA are resolved such that the MAC frames from the higher priority AC receives the transmission opportunity (TXOP) and the lower priority AC(s) backs off without counting the collision like an external collision, i.e. does not affect its CW value.

2.1. The CFB mechanism

In IEEE 802.11e, there is also the possibility to use the option of CFB mechanism to enhance the performance, and achieve improved channel utilization. Through this option, once a WSTA wins the TXOP to the wireless medium under normal contention, it is allowed to transmit more than one frame without further contention of the channel during its current TXOP. This operation is illustrated in Fig. 1. After obtaining access to the channel, the WSTA is allowed to transmit as many frames from the same AC queue as available provided that the total duration does not exceed a certain limit bounded in the dot11EDCFTableLimit (shortened here as TXOPLimit for simplicity) MIB variable, which is set based on either the most recently received QoS Parameter Set element of a beacon frame from the AP in an infrastructure setting or pre-configured in the WSTA’s MIB in an ad hoc setting. To ensure that no other WSTA interrupts the burst, the interframe space between the reception of an acknowledgement and the subsequent transmission of the next data frame in the burst is a short IFS (SIFS) interval.

![Diagram of EDCF controlled frame bursting mechanism](image)

Even though it is felt that CFB would increase the medium utilization, the selection of an appropriate value for the TXOPLimit is necessary especially when considering delay-sensitive flows being present in the networks. At any time, when a TXOP is won, the maximum number of frames that could be transmitted depends on the current number of buffered frames in the MAC transmit buffer and the physical transmission rate. The TXOPLimit should be directly proportional to the buffer size (B) and inversely proportional to the physical data rate (R). Thus, the TXOPLimit should be configured as:

\[
\text{TXOPLimit} = \alpha \times \frac{B}{R}, \text{ where } 0 < \alpha \leq 1.
\]

When \(\alpha\) is one, the limit is set to the maximum data transfer at the physical rate. For \(\alpha\) less than one, the limit represents the time to transmit a fraction of its buffer occupancy.

When a relatively large burst limit (relative to the buffer size) is used, a WSTA may continue holding the channel for a long time resulting in larger delay variations for sensitive flows like interactive voice and video in other WSTAs. Also, setting the value to a relatively small value allowing only the transmission of
a few frames in a burst will reduce the CFB option to the basic EDCF operation.

To evaluate and verify the choice of the TXOPLimit value, simulation experiments are performed. The adopted simulation scenario and traffic characteristics for the simulation study is discussed further.

3. Simulation scenario

In order to evaluate the performance of the EDCF MAC protocol, a discrete-event simulator has been developed. The simulator is built using C++ and models a WLAN with an 802.11a PHY and 802.11e MAC layers. The simulator is developed using the object-oriented approach. Its overall software architecture is given in Fig. 2 as a simplified class diagram. The symbols used on the diagram represent the standard UML notations.

![Figure 2. The simplified class diagram of the simulator](image)

The single active entity in the real system, i.e. a WSTA (represented by the Node class), also represents the most complicated object in this simulator. Since a WSTA may implement more than one ACs, this is represented as the Node class’s composition relationship with multiple AC objects. The Channel class represents the spectrum resource in a wireless network. Here, the channel is assumed to be error-free. The other passive objects in the system are made the subclass of the Frame class. The resulting architecture is highly cohesive and loosely coupled enabling easy extension for many other future works.

The adopted topology comprises a number of WSTAs configured in an ad hoc setting as shown in Fig. 3. All WSTAs are located within a basic service set (i.e. a cell) such that every WSTA is able to detect a transmission from any other WSTA, and they are static in the simulations. It is also assumed that the wireless medium is error-free. The adopted IEEE 802.11a PHY parameters are presented in Table 1.

![Figure 3. The single-BSS ad hoc network topology](image)

<table>
<thead>
<tr>
<th>Table 1. 802.11a PHY parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
</tr>
<tr>
<td>aSlotTime</td>
</tr>
<tr>
<td>SIFS</td>
</tr>
<tr>
<td>Channel bit rate</td>
</tr>
<tr>
<td>Propagation delay</td>
</tr>
</tbody>
</table>

Our simulations involve a few different traffic sources to represent the typical traffic mix on the network. Voice signals are known to have a two-state ON/OFF behavior, where talkspurts are followed by silence periods. The G.729 speech codec [13] has been selected to model the voice calls, with voice packets of 60 bytes generated every 20 ms during a talkspurt, which corresponds to 24 kbps bit rate. The talkspurt and silence duration times are exponentially distributed with a mean value of 1 sec and 1.35 sec, respectively [14]. As for the video sources, we have considered low-quality video application with 500 kbps data rate, which emulates an MPEG downlink streaming service. We assume that these flows are generated by constant bit rate sources. A video packet length of 1500 bytes has been selected. Finally, it is assumed that the bursty data source, emulating an FTP application, follow a Poisson arrival with packet size of 1000 bytes and bit rate of 200 kbps. Two AC queues are used in each WSTA to enqueue voice and FTP packets or video and FTP packets. It is assumed that there is equal number of such WSTAs in the network.

These sources’ parameters and their corresponding AC’s contention parameter values are summarized in Table 2. The contention parameter values are chosen to ensure sufficient differentiation among the ACs to minimize collision, and were validated through
simulations in [9]. To emulate the increase of system load, we gradually increase the number of WSTAs.

Table 2. MAC and source traffic parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Voice</th>
<th>Video</th>
<th>Bursty Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIFS (SIFS+*aSlotTime)</td>
<td>2</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>CWmin/CWmax</td>
<td>5/200</td>
<td>15/511</td>
<td>31/1023</td>
</tr>
<tr>
<td>Packet size (bytes)</td>
<td>60</td>
<td>1500</td>
<td>1000</td>
</tr>
<tr>
<td>Packet Interval (ms)</td>
<td>20</td>
<td>24</td>
<td>-</td>
</tr>
<tr>
<td>Bit rate (kbps)</td>
<td>24</td>
<td>500</td>
<td>200</td>
</tr>
</tbody>
</table>

Based on the adopted simulation scenario, simulations are performed and results are presented and discussed in the following section.

4. Performance evaluation

To evaluate the performance of the CFB mechanism, the following metrics are used:
- **Mean transfer delay**: the average time taken to transmit a packet from the time it is generated to the time the source receives a successful acknowledgement.
- **Delay jitter**: the variation of frame transfer delay.
- **Throughput**: the offered load that is actually delivered to the destination, which excludes the MAC and PHY overheads.
- **Packet loss ratio**: the fraction of discarded to generated packets.

The main performance metrics for voice traffic are frame transfer delay, jitter and packet loss [15]. In order to preserve the user-perceived voice quality, commonly accepted maximum values in an end-to-end connection over an IP network are 150 ms for the one-way delay, several milliseconds for delay jitter and 3% packet loss. However, since a WLAN is likely to represent only the first or last hop of an end-to-end connection, we have instead chosen 20 ms as a maximum acceptable value, as suggested in [11]. As for the video traffic, we have adopted 25 ms as the maximum tolerable one-way delay [16]. Throughput degradation has more serious consequence on the video performance when adaptive algorithms and jitter buffers are employed.

The above metrics are evaluated for different network sizes, $N$. In order to see the impact of MAC transmit buffer size on the protocol performance, we have adopted three different buffer sizes, namely 32 KB, 64 KB and 128 KB. In the forthcoming graphs, these MAC buffer sizes are shown in the legend as 32, 64 and 128, respectively. For the initial investigation, we have assumed that the TXOPLimit value is fixed at 10 ms.

Figures 4 and 5 depict the network performance in terms of mean delay and jitter for the three different buffer sizes. Consistent results are obtained for both metrics. It is observed that the smallest buffer size ensures the minimal average delay and jitter for the applications. The difference becomes more evident when the network operates at heavy loads. As more loads are introduced, nodes with bigger buffer have more queued packets to be transmitted. Subsequently, these packets wait longer periods to gain media access resulting in longer transfer delays and jitters. All the considered source types exhibited similar behavior. For example, considering the voice application, it experienced an increase of more than 70% in both delay as well as jitter when buffer size is increased from 32 KB to 128 KB for the network of 100 nodes.

![Figure 4. Mean frame transfer delay for different MAC transmit buffer sizes](image)

![Figure 5. Delay jitter for different MAC transmit buffer sizes](image)

In Fig. 6, the throughput performance is plotted against network size for the different buffer sizes. Contrary to the above delay performances, the nodes with smaller buffer size exhibit lower throughput especially when the network has more than 50 nodes. This behavior is due to the increased contentions at higher loads, which results in packets waiting longer periods to gain access. As the rate of packet transfer decreases, the throughput growth slows and saturates at
a certain level. This effect is more pronounced for the smallest buffer size. Video sources enjoy the highest throughput due to their data rates and access priority. Consistent results to supplement the throughput behavior is shown in Fig. 7. It is seen that packet loss probability is minimal when limited nodes exist in the network. As the network grows, the packets tend to wait longer, and newly arriving packets may be dropped as the buffer space runs out. This is more evident for nodes with smaller transmit buffer as depicted. However, as seen earlier, bigger buffers tend to increase mean delay and jitter values, which may be intolerable for certain inelastic applications like voice and video. A single buffer size to fit all cases may not be possible. Instead, we may use separate queue thresholds for each AC, and any incoming packets exceeding its queue threshold may be proactively dropped, as its deadline might not be met. Such a move might reduce the transmission of useless packets, while achieving a fairer utilization among the nodes.

In the further study, we chose to use 64 KB as the buffer size, and vary the CFB TXOPLimit to see its impact on the network performance. As highlighted earlier, the bursting limit is varied with relation to the buffer size. We have accordingly selected 0.5 ms, 5 ms and 10 ms representing transmission of packets occupying 5%, 50% and 100% of the buffer size, respectively. The following plots show the chosen limits as part of their legend.

Figures 8 and 9 show the mean delay and jitter performance for various TXOPLimits. It is observed in Fig. 8 that mean delay increases significantly for all the cases when $N$ is more than 50 nodes. EDCF with the smallest TXOPLimit duration (i.e. 0.5 ms) exhibits a wide-varying delay especially for larger networks as each TXOP only occupies the medium for a short period. Its delay performance is unpredictable, and behaves more like the basic EDCF scheme without the CFB option. For larger limits, nodes are able to transmit more frames without contentions, while achieving more controlled use of the media. Between 5 ms and 10 ms, there is no significant difference in delay for all the traffic sources. The jitter metric comparison is shown in Fig. 9. Jitter values for all the sources are mostly within the acceptable range for the different burst limits. Thus, the burst duration does not have a critical impact here on the jitter metric for the chosen scenario.
Figure 10 displays the results of throughput metric for different network sizes. Here, it is evident that the CFB option enables the EDCF protocol to achieve significantly higher throughput in an error-free condition. With CFB, the network is able to efficiently utilize the network and achieve higher throughput especially at higher loads. However, any increase of the burst duration beyond 5 ms has only negligible improvement on the throughput. A consistent result is also obtained for the loss probability as shown in Fig. 11. As before, the limited-burst EDCF protocol faces significant frame losses especially when \( N \) goes beyond 50 nodes. Again, there is limited improvement in frame loss probability when the TXOPLimit is increased beyond 5 ms. Thus, the 5-ms limit ensures a more controlled sharing of the media. By prohibiting nodes from over or under occupying the medium, we are able to ensure an optimized performance and controlled collision rate.

![Figure 10. Throughput for different TXOPLimit values](image1)

Figure 10. Throughput for different TXOPLimit values

![Figure 11. Packet loss ratio for different TXOPLimit values](image2)

Figure 11. Packet loss ratio for different TXOPLimit values

To illustrate the significant improvement obtained using the CFB feature, we plotted the overall network throughput against network size for the same TXOPLimits, shown in Fig. 12. It is evident that the bursting option allows the MAC protocol to achieve substantially higher utilization especially in larger networks. For a shared medium access situation, the basic EDCF scheme may only be suitable for small to medium-sized networks. Any BSS with more than 40 nodes will suffer significantly. It is found that the CFB feature utilizing a 5-ms limit achieves a throughput increase of about 50% than the limited bursting option of 0.5 ms when the channel is error-free and absent of hidden/exposed terminal problems. This feature allows the MAC protocol to realize more than 60% of the physical data rate. Setting the limit beyond 5 ms does not result in any further improvement.

![Figure 12. Network throughput for different TXOPLimit values](image3)

Figure 12. Network throughput for different TXOPLimit values

Therefore, it is conclusive that employing the frame bursting option is indeed beneficial especially in terms of EDCF scalability and its support for various traffic mix. It is also evident that proper configuration of the TXOPLimit variable is necessary for an optimized operation of the EDCF protocol. We found that setting its value to about 50% of the MAC transmit buffer size achieves the best trade-off in terms of delay, delay variation, throughput and packet loss. Furthermore, the optimized CFB configuration allows the MAC protocol to realize 30% more capacity of the available physical data rate.

5. Conclusion and future work

Distributed QoS support for 802.11 networks is seeing a lot of interests from the researchers as the technology becomes more accessible to the general users. As the users begin to use various multimedia applications, these wireless networks support of such applications is crucial as their bandwidths are more limited. EDCF is one such protocol proposed to address the QoS issue in the upcoming IEEE 802.11e standard.

It is clearly demonstrated here and elsewhere that the basic EDCF operation fails to scale well. It was proposed in the standard that EDCF may use the frame bursting technique to increase utilization when a station wins a contention. In this paper, we have shown that the frame-bursting option can indeed be used to achieve improved performance and network utilization.
However, it is shown that for an optimized operation, the proper configuration of the TXOPLimit variable is crucial and should be closely associated with the MAC transmit buffer size. A limit proportional to at least 50% of buffer occupancy and not larger than 100% should be utilized. It is also shown that a bigger buffer size may not necessarily improve the inelastic applications’ performance. It may be helpful to adopt a separate queue threshold for each traffic class to guide a proactive packet discard to realize a performance trade-off.

As a future work, we would choose to configure the burst limit dynamically according to the current network usage. A bandwidth estimation algorithm might be used to estimate each node’s bandwidth share, which in turn would be used to estimate and control its burst duration dynamically.

6. References