In this chapter we discuss the QoS and performance of 802.11 MAC Layer protocol. Furthermore, we discuss about the provisions made in admission controller algorithms for association of STAs to improve the QoS in a WLAN.
Chapter 4: QoS and Admission Controller

4.1 Introduction

The IEEE 802.11 WLAN MAC Layer operates in two modes for accessing the wireless medium: the compulsory Distributed Coordination Function (DCF) and Point Coordination Function (PCF) optionally in the presence of Access Point. The working principle of DCF and PCF has been explained in chapter 2. From the literature review it is clear that DCF and PCF only designed for best effort traffic. But as the popularity of wireless LAN increasing the demand for improving QoS for real-time traffic like video and voice data has also a serious concern. As discussed in chapter 2 various mechanisms have been proposed for enhancing the QoS. The basic approach for DCF access method is to provide priorities and service differentiation mechanisms for different types of traffics. Moreover, the IEEE 802.11e Task Group has finalized the standardization of a QoS conscious MAC protocol with proper service differentiation mechanisms. In the standard of IEEE 802.11e [16 18], the MAC protocol Hybrid Coordination Function (HCF) being developed to manage the channel access to the wireless medium. Two different modes are defined, Firstly the Enhanced Distributed Channel Access (EDCA) [16 18], and secondly the controlled based access mechanism (HCF Controlled Channel Access - HCCA). The HCCA facilitates QoS stations (QSTAs) having prioritized channel access. And on the other
hand EDCA provides prioritized carrier sense multiple access with collision avoidance (CSMA/CA) method and it is compulsory in IEEE 802.11e. The term EDCA is a new name given from the originally named protocol EDCF (*Enhanced DCF*). The HCF divides the time into a contention and a contention-free slot and assigns it to either HCCA or EDCA respectively. Different traffic categories having different QoS requirements are distinguished and tackled using one of these functions or may be with both of these functions.

Under Hybrid Coordination Function, the primary point to authorize the right for accessing the wireless medium is the TXOP [5 16 18]. A TXOP is defined by a beginning time and a defined maximum length. The accessing to the wireless medium in the presence of EDCA function is referred to as contention period (CP) because all traffic categories compete over others to win the EDCA transmission opportunities alias TXOPs. Stations receiving TXOPs are allowed to transmit their frames to the access point. Each traffic category do have a different maximum length for TXOPs and it makes sure that packets from low traffic category don’t hold the medium for long time.

In certain situations the QoS requirements by STAs are not satisfied by EDCA. The hybrid coordinator may initiate its own packets and thus the HCCA TXOPs are assigned to other STAs so that other STAs...
using EDCA cannot access to the wireless media after sensing the channel to be idle. All stations are relayed regarding to the starting and the end of the use of HCCA.

In the following sections we discuss the basics IEEE 802.11e mechanisms with HCCA [17 18] and EDCA [18]. We also explain how admission controllers have been deployed along side EDCA to provide guaranteed QoS in such WLAN.

4.2 EDCA

In IEEE 802.11e standard the compulsory MAC protocol is EDCA. It is the enhanced version of the legacy MAC protocol function DCF of the 802.11 standard to support the QoS requirements by QSTAs. The EDCA mechanism is designed to offer prioritized or differentiated channel access. EDCA adopts the CSMA/CA mechanism of 802.11 channel access. The operation is based on contending for the wireless medium using a backoff procedure. EDCA uses different inter framing spaces as waiting intervals of different length, also called Arbitrary Interframe Spaces (AIFSs), and differentiated backoff time called Contention Windows (CWs), according to the priority of the traffic category or Access Category (AC). AIFS is the amount of time a station would sense the channel to be idle and the length of CW is used for the value of
backoff counter. These different lengths of the AIFS values represents access probabilities for different category of traffics based on their priorities. By assigning different waiting intervals for different traffic classes it is possible to have differentiated service and thus QoS can be supported. More over to avoid collisions EDCA additionally deploys a collision avoidance mechanism using a two-way message passing, called RTS/CTS (Request to Send/Clear to send), and explained in earlier sections. This technique handles to some degree the serious hidden station problem.

EDCA primarily access to the wireless medium in a differentiated and distributed manner based on the 8 different UPs (User Priorities). Any packet destined from higher layers to the MAC layer is given a UP value. The entire higher layer UPs are mapped into some ACs (Access Categories) as shown in figure 4.2. The UP to AC mapping procedure can be found in [18]. As shown in figure 4.2, there are four transmission queues being implemented in a QSTA and each of these queues supports one AC and each of this AC functions like a single DCF entity, contending for the wireless medium access, and also each access category initiates its backoff procedure by its own. In EDCA AC-0 is having the lowest priority also known as best effort traffic and AC-3 is the highest priority of all in the real time traffic category. Therefore to
have service differentiation EDCA defines the following parameters in each AC\([k]\) where \(0 \leq k \leq 3\):

- Minimum CW size is \(CW_{\text{min}}[k]\),
- Maximum CW size is \(CW_{\text{max}}[k]\),
- Arbitration inter frame space is \(AIFS[k]\),
- Arbitration inter frame space number is \(AIFSN[k]\).

The relation between \(AIFS[k]\) and \(AIFSN[k]\) is defined as follows:

\[
AIFS[k] = AIFSN[k] \times \text{slotTime} + SIFS
\]

*Here* \(\text{slotTime}\) *is the amount of time defined in the setting of each PHY layer setting*. The backoff procedure in each AC \(k\) is selected in the range of \([0, CW[k]]\), where \(CW[k]\) represents the current CW in the particular AC \(k\). At the very first attempt, the value of \(CW[k]\) is assigned to the value of \(CW_{\text{min}}[k]\). When successive transmissions fail because of collisions, the value of \(CW[k]\) increases up to the maximum value of \(CW_{\text{max}}[k]\) in binary exponential manner. In EDCA standard, the smaller the \(AIFSN[AC]\) and \(CW_{\text{min}}[AC]\), the higher the probability of winning the contention with the other ACs. The smaller the \(AIFSN[AC]\) and \(CW_{\text{min}}[AC]\), the higher the probability of winning the contention with...
the other ACs. If a particular AC has smaller AIFS value and shorter CW values than traffics from that category gets more access to the wireless medium. In EDCA standard, an AC of real-time category has always smaller and shorter values of AIFS and CW size than that of other ACs of non real-time category.

Table 4.1 shows the default values of the channel access parameters defined in EDCA for the four ACs (BK = background, BE = best effort, VI = video, VO = voice). These parameters are not fixed: in each beacon frame, the access point (AP) broadcasts the values chosen for each AC. Indeed, these values may also be dynamically adjusted according to network conditions when more than one AC of the same station expire its backoff counter, a virtual collision occurs, and the packet having highest-priority among the colliding ones is selected for actual transmission on the radio channel. Figure 4.1 shows the IEEE 802.11e channel access method.
The QAP broadcasts the current EDCA parameter-set such as $CW_{\text{min}}[k]$, $CW_{\text{max}}[k]$ and $AIFS_N[k]$ in a special field of the beacon frames. In a situation when more than one AC of the same station expire its backoff counter, a virtual collision occurs, and the highest-priority
packet among the colliding ones is selected for actual transmission on the radio channel and the other ACs increase their CW values and goes for backoff. Although, for a QSTA an AC of real-time category gets always more opportunities for accessing the channel than that of non real-time traffic, but collisions frequently occurs when there exist more QSTAs with real time traffic and STAs with non real time traffic. So in certain situations traffics from real time category may not be satisfied with its QoS requirement. Therefore, it is necessary to protect real-time traffic from other real-time traffic and non real-time traffic.

Figure 4.2: Implementation model for EDCA
4.3 EDCA TXOPs

In EDCA standard, two types of TXOPS [18 19] are defined, firstly initiation of the EDCA TXOP and transmission of multiple frames within an EDCA TXOP. The initiation of the TXOP takes place whenever the EDCA norms allow the access to the wireless channel. And secondly transmission of multiple frames within the TXOP occurs whenever an EDCAF keeps the authority for accessing the wireless channel in a way to complete a frame exchange sequence. The values of the TXOP limit duration are broadcasted by the QAP within the EDCA Parameter in Beacon frame and Probe Response frames from the QAP. The value of a TXOP limit 0 is an indication that a single MSDU additionally with a possible RTS/CTS exchange may be transmitted at any rate for each TXOP. The non-AP QSTAs will ensure that the length of TXOPs derived using the EDCA norms should not exceed the TXOP limit. The TXOP duration is defined as the time amount during which the QSTA retaining uninterrupted authorization of the wireless channel, and it also comprises the time required for transmitting frames as an immediate reply to the QSTA holding the TXOP transmission. A QSTA fragments MSDU so that while transmitting the first MPDU within the TXOP does not result the TXOP limit exceeding at the PHY rate considered for the
initial transmission attempt of that MPDU. In particular the maximum TXOP limit may be exceeded if lower PHY rate is used.

It particular when multiple frames are transmitted within the TXOP using acknowledgment mechanisms (other than Normal Ack), a protective method must be deployed (such as RTS/CTS). On the other hand a QAP may broadcast or multicast frames without considering any protection method.

4.3.1 EDCA Backoff procedure

EDCA maintains a backoff variable CW[AC], which initially set to the value of the parameter CWmin[AC]. When a frame is successfully transmitted by a particular EDCAF, following by successful reception of a CTS, the successful reception of an ACK, the successful reception of a BlockAck or ACK frame, or the transmission of a multicast frame or a frame with No Ack policy, CW[AC] will be reset to CWmin[AC].

The EDCA backoff procedure is initiated for an EDCAF when any of the following situations arises:

- When a frame from a particular access category is ready for transmission and the medium is sensed to be busy informed by
either physical carrier sensing or virtual Carrier sensing, and the backoff timer has a value of zero in that particular AC.

- When the final transmission by the TXOP holder initiated during the TXOP for that AC was successful.

- When a transmission of a frame fails from a particular AC due to failing receiving a CTS in a reply of an RTS, failing to receive an ACK frame or a failing to receive a BlockAck or ACK frame in reply to a BlockAckReq frame.

- When the transmissions collide internally within the same QSTA due to the higher priority of an AC, in other words two or more EDCA frames within the same QSTA are allotted the TXOP simultaneously.

When the backoff is initiated due to the first condition then the value of CW[AC] is left unchanged. If backoff procedure is initiated due to second condition then the value of CW[AC] must be reset to CWmin[AC].
4.4 HCCA

The HCF (hybrid coordination function) controlled channel access (HCCA) works a lot like PCF. However, in contrast to PCF, in which the interval between two beacon frames is divided into two periods of CFP [19] and CP [19], the HCCA allows for CFPs being initiated at almost any time during a CP. This kind of CFP is called a Controlled Access Phase (CAP) in 802.11e. A CAP is initiated by the AP whenever it wants to send a frame to a station or receive a frame from a station in a contention-free manner. In fact, the CFP is a CAP too. During a CAP, the Hybrid Coordinator (HC) -- which is also the AP controls the access to the medium? During the CP, all stations function in EDCA. The other difference with the PCF is that Traffic Class (TC) and Traffic Streams (TS) are defined. This means that the HC is not limited to per-station queuing and can provide a kind of per-session service. Also, the HC can coordinate these streams or sessions in any fashion it chooses (not just round-robin). Moreover, the stations give info about the lengths of their queues for each Traffic Class (TC). The HC can use this info to give priority to one station over another, or better adjust its scheduling mechanism. Another difference is that stations are given a TXOP: they may send multiple packets in a row, for a given time period selected by
the HC. During the CP, the HC allows stations to send data by sending CF-Poll frames.

HCCA is generally regarded to be the most advanced (and complex) coordination function. With the HCCA, QoS can be configured with great precision. QoS-enabled stations have the ability to request specific transmission parameters (data rate, jitter, etc.) which should allow advanced applications like VoIP and video streaming to work more effectively in a Wireless network.

HCCA support is not mandatory for 802.11e QAPs. In fact, few (if any) APs currently available are enabled for HCCA. Implementing the HCCA on end stations uses the existing DCF mechanism for channel access (no change to DCF or EDCA operation is needed). Stations only need to be able to respond to poll messages. On the AP side, a scheduler and queuing mechanism is needed.

4.5 Backoff Procedure

The backoff period of each AC is set based on a uniform distribution over \([0, \text{CW} [\text{AC}]]\). The initial value of the CW is to the minimum value of the CW and the value of the CW is doubled at the time of collision or when transmission fails up. As explained in the earlier sections when the CW values ranges in small size, then the channel access delay is very small for the prioritized channels and...
therefore the particular access category acquires adequate access to the wireless media. When a particular traffic is admitted, it will be attached with QoS parameters. When more than one traffic categories within the QSTA finishes their backoff timers at the same time virtual collision will come to happen. The internal scheduler from the station’s solves the virtual collision. EDCA support QoS; however, it cannot guarantee strict QoS required by real-time services without proper network control mechanisms.

In 802.11e standard by default CW sizes are set to small values for real time access categories and that traffic result in small backoff time. But in situation when plenty of real-time traffics exist, collisions between them frequently occur and their QoS requirements are not satisfied such as throughput, delay or jitter. There are so many schemes for adjusting the CW values dynamically for each category in a way to reduce collisions among real-time traffics. The issue of increasing or decreasing CW values in a should be applied normalized for all ACs because when CW values in a real rime category AC3 or AC2 are updated leaving non-real time traffic category unchanged will result in the collapse of service differentiation defined in IEEE802.11e standard. Therefore when CW values in AC3 or AC2 are updated, the CW values of
other ACs are accordingly updated. For example, when CW values in AC3 are updated, those in AC2, AC1, and AC0 may have to be updated accordingly.

4.6 Admission Control at HC

In IEEE 802.11 standard admission controller may be deployed to administer or regulate the radio resources. To guarantee the QSTAs for accessing the wireless medium for adequate amount of time admission control is very important, in other words admission controller provides enhanced QoS for real time traffic. The administration of the admission controller in a network is done by the HC located in a QAP. In IEEE 802.11e standard, to support the QoS facility, there exist two distinct admission control policies: one for contention-based access and another for controlled access. Admission control [16 18] policy normally, depends on vendor specific implementation of the scheduler, available admission capacity, link conditions, limit on retransmission, and the scheduling specification for a given stream.

4.6.1 Contention-based admission control procedures

In contention based admission control [18] procedure a QSTA may support transmitting frames in a particular AC where admission control
 CHAPTER 4: QoS and Admission Controller Page 89

is mandated; but, if the QSTA does not support that procedure, it may use the EDCA parameters having a lower priority AC, which does not need admission control. QAPs support admission control policies, at least for the minimal degree of advertising that admission is not compulsory on its ACs. The QAP uses the admission control mandatory (ACM) subfields advertised in the EDCA Parameter Set to specify whether admission control is needed for each of the ACs. While the CW minimum and maximum values, AIFS, TXOP limit parameters may be adjusted time to time by the QAP, the ACM bit remains fixed for the duration of the lifetime of the BSS.

When Admission controller is deployed under EDCA a QSTA is assumed to send an ADDTS [18] (Add traffic specification) frame to the AP as shown in figure 4.3. The ADDTS frame contains traffic category and traffic specification (TSPEC) [18] corresponding to the specific application. The TSPEC contains the following information.

- **Mean data rate**: the average bit rate for packet transmission, in bits per second (bps)
Chapter 4: QoS and Admission Controller

- **Delay bound**: the maximum delay allowed to transport a packet across the wireless interface (including queuing delay), in milliseconds;

- **Maximum service interval**: the maximum time allowed between neighbor TXOPs allocated to the same station, in microseconds;

- **Nominal MSDU size**: the nominal size of a packet, in octets;

- **Minimum PHY rate**: the minimum physical bit rate assumed by the scheduler for calculating transmission time, in bps.

The QAP assigns the received UP of the ADDTS Request frame to its appropriate AC following the mapping procedure of UP-to-AC as explained in earlier section. In particular situations a QSTA may send un-admitted traffic for particular ACs and for that the QAP does not need admission control. When a QSTA wants to transmit data without admission control using an AC that emphasizes on admission control, the QSTA uses EDCA parameters that correspond to a relatively lower priority and do not need admission control.
4.6.2 Procedures at the QAP

Upon getting an ADDTS request frame from the QSTA, the QAP responds to the ADDTS Request frame with an ADDTS Response frame. The ADDTS response frame indicates the admission or rejection of current the request from a particular category. The decision-making part or mechanism used by the controller in the QAP is very much a local concern within the QAP. When the QAP accepts a request from a traffic category, then the QAP has also to derive the medium or channel time from the information conveyed in the TSPEC [18] element of the ADDTS Request frame. Any suitable mechanism may be used for deriving the medium time.
The EDCAF has two duration values: admitted_time and used_time. These values are set to 0 at the time of (re)association. A QSTA may consequently explicitly request medium time for an AC which is associated with the specified priority. When a request is accepted by the HC then QSTA recompute the admitted_time for the specified EDCAF as follows:

\[
\text{admitted\_time} = \text{admitted\_time} + \text{dot11EDCA\_AveragingPeriod} \times (\text{medium time of TSPEC}).
\]

A QSTA may decide at any point of time to clear or delete the explicit request. For deleting the explicit admission, QSTA can send a DELTS frame containing the TSID and direction that specify the TSPEC to the QAP. After a QSTA sends or receives the DELTS frame, then it will re-compute the admitted_time for the specified EDCAF as follows:

\[
\text{admitted\_time} = \text{admitted\_time} - \text{dot11EDCA\_AveragingPeriod} \times (\text{medium time of TSPEC}).
\]

4.6.3 Procedure at QSTAs

In order to make such a request, the non-AP QSTA shall transmit a TSPEC element contained in an ADDTS Request frame with the...
following fields specified (i.e., nonzero): Nominal MSDU Size, Mean Data Rate, Minimum PHY Rate, Inactivity Interval, and Surplus Bandwidth Allowance. The Medium Time field is not used in the request frame and shall be set to 0. On receipt of a TSPEC element contained in an ADDTS Response frame indicating that the request has been accepted, the non-AP QSTA shall re-compute the admitted_time for the specified EDCAF as follows:

\[
\text{admitted\_time} = \text{admitted\_time} + \text{dot11EDCAAverage\_Period} \times (\text{medium\_time\_of\_TSPEC}).
\]

The non-AP QSTA may choose to tear down the explicit request at any time. For the teardown of an explicit admission, the non-AP QSTA shall transmit a DELTS frame containing the TSID and direction that specify the TSPEC to the QAP. If the non-AP QSTA sends or receives a DELTS frame, it shall recompute the admitted_time for the specified EDCAF as follows:

\[
\text{admitted\_time} = \text{admitted\_time} - \text{dot11EDCAAverage\_Period} \times (\text{medium\_time\_of\_TSPEC}).
\]
To describe the behavior at the non-AP QSTA, two parameters are defined. The parameter \textit{used\_time} signifies the amount of time used, in units of 32 μs, by the non-AP QSTA in \texttt{dot11EDCA\_AveragingPeriod}. The parameter \textit{admitted\_time} is the medium time allowed by the QAP, in units of 32 μs, in \texttt{dot11EDCA\_AveragingPeriod}. The default values used by non-AP QSTAs for the parameters in the EDCA Parameter Set element are defined in Figure 4.5.

### 4.7 EDCA Parameter Set element

The EDCA Parameter Set element [18] provides information needed by non-AP QSTAs for proper operation of the QoS facility during the CP. The format of the EDCA Parameter Set element is defined in Figure 4.4.

<table>
<thead>
<tr>
<th>Element ID</th>
<th>Length</th>
<th>QoS Info</th>
<th>Reserved</th>
<th>AC_BE Parameter Record</th>
<th>AC_BK Parameter Record</th>
<th>AC_VI Parameter Record</th>
<th>AC_VO Parameter Record</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>18</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Octets: 1 1 1 4 4 4 4

Figure 4.4: EDCA parameter set element
The EDCA Parameter Set element is used by the QAP to establish policy (by changing default MIB attribute values), to change policies when accepting new STAs or new traffic, or to adapt to changes in offered load. The non-AP QSTA uses the recent EDCA parameter set element to update the appropriate MIB values. The QoS Info field contains the EDCA Parameter Set Update Count subfield, which is initially set to have the value 0 and is incremented each time any of the AC parameters changes. This subfield is utilized by non-AP QSTAs to decide whether the EDCA parameter set has changed and requires updating the appropriate MIB attributes. The most recent EDCA parameter set element received by a non-AP QSTA is used to update the appropriate MIB values.

---

**Figure 4.5: Default QoS parameter Set**

<table>
<thead>
<tr>
<th>AC</th>
<th>CWmin</th>
<th>CWmax</th>
<th>AIFS</th>
<th>TXOP limit (802.11b)</th>
<th>TXOP limit (802.11a/g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>aCWmin</td>
<td>aCWmax</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>aCWmin</td>
<td>aCWmax</td>
<td>1</td>
<td>3.0ms</td>
<td>1.5ms</td>
</tr>
<tr>
<td>2</td>
<td>(aCWmin+1)/2-1</td>
<td>aCWmin</td>
<td>1</td>
<td>6.0ms</td>
<td>3.0ms</td>
</tr>
<tr>
<td>3</td>
<td>(aCWmin+1)/4-1</td>
<td>(aCWmin+1)/2-1</td>
<td>1</td>
<td>3.0ms</td>
<td>1.5ms</td>
</tr>
</tbody>
</table>

---