An Overview of Wireless LAN Standards IEEE 802.11 and IEEE 802.11e

Jahanzeb Farooq, Bilal Rauf

Department of Computing Science
Umeå University
Sweden

Jahanzeb Farooq, 2006
(tipputhegreat@hotmail.com)

Chapter 1

Introduction to IEEE 802.11

1.1 Introduction

In 1997, IEEE (Institute of Electrical and Electronics Engineers) released the 802.11 Wireless Local Area Network (WLAN) standard [1]. As the name suggests, it belongs to the group of popular IEEE 802.x standards, e.g., IEEE 802.3 Ethernet [2] and IEEE 802.5 Token Ring [3]. IEEE 802.11 defines Media Access Control (MAC) and Physical (PHY) layers specifications for wireless LANs. Three different Physical layer specifications were defined, namely, Frequency Hopping Spread Spectrum (FHSS), Direct Sequence Spread Spectrum (DSSS) and Infrared (IR), with the maximum data transmission rate of up to 2 Mbps. The DSSS and FHSS Physical layers operated in the license free 2.4 GHz ISM (Industrial, Scientific and Medical) band.

With the passage of time, while the original MAC remained intact, the technology continued evolving with the new Physical layer specifications.

In 1999, IEEE introduced two enhanced Physical layer specifications 802.11b [4] and 802.11a [5] with data transmission rates of up to 11 and 54 Mbps, respectively. 802.11b is also based on DSSS and operates in the 2.4 GHz band, and, 802.11a is based on OFDM (Orthogonal Frequency Division Multiplexing) and operates in the 5 GHz band. In 2003, IEEE released 802.11g [6] that extended 802.11b Physical layer to support data transmission rates of up to 54 Mbps in the 2.4 GHz band.

IEEE 802.11 gained immense popularity due to its cost effectiveness and easy deployment. Today IEEE 802.11 hotspots[1] are available at offices, campuses, airports, hotels, public transport stations and residential places, making it one of the most widely deployed wireless network technologies in the world.

IEEE 802.11 defines two different architectures, BSS (Basic Service Set) and IBSS (Independent Basic Service Set). In a Basic Service Set, number of wireless stations, called STAs[2], are associated to an AP (Access Point). All communications take place through the AP. In an Independent Basic Service Set, STAs can communicate directly to each other, providing that they are within each other's transmission range. This form of architecture is facilitated to form a wireless ad-hoc network in absence of any network infrastructure.

Several BSS can be connected together via a Distribution System (DS) to form an extended network, called Extended Service Set (ESS). Figure 1.1 illustrates the archi-

[1]The area where IEEE 802.11 access is available is referred to as *hotspot*.

[2]For simplicity, we prefer to use the word *station* instead of *STA* throughout the report.

tecture of IEEE 802.11 BSS.

IEEE 802.11 MAC defines two different access mechanisms, the mandatory Distributed Coordination Function (DCF) which provides distributed channel access based on CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance), and the optional Point Coordination Function (PCF) which provides centrally controlled channel access through polling.

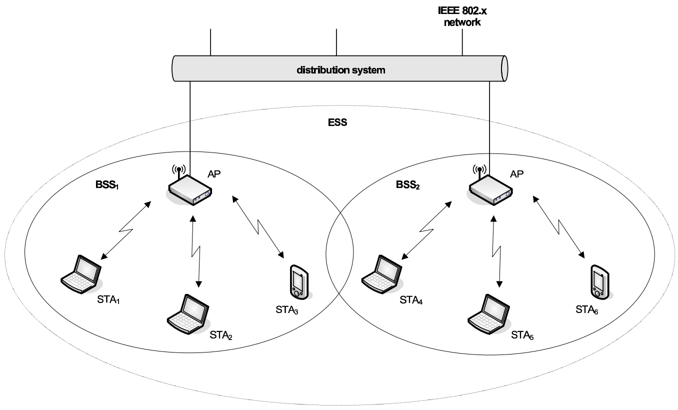

FIGURE 1.1: ARCHITECTURE OF IEEE 802.11 NETWORK.

In DCF, all stations contend for the access to the medium, in distributive manner, based on the CSMA/CA protocol. For this reason the access mechanisms is also referred to as contention-based channel access. In PCF, a Point Coordinator (PC), which is most often collocated in AP, controls the medium access based on the polling scheme, such that the PC polls individual stations to grant access to the medium based on their requirements. As in PCF stations do not contend for the medium and instead the medium access is controlled centrally, the access mechanism is sometimes referred to as contention-free channel access.

Only DCF is explained in the next section as it is the basis for the Enhanced Distributed Channel Access (EDCA) introduced in IEEE 802.11e, which we focus in this work.

1.2 DCF (Distributed Coordination Function)

DCF is the basic access mechanism of 802.11 and is based on Carrier Sense Multiple Access (CSMA). CSMA works as listen-before-talk, i.e., before transmitting a frame, the station senses the medium (carrier sensing). If the medium is found idle at least for DIFS (DCF Inter-Frame Space) time period, the station starts transmission, and other stations wait until medium becomes idle again at least for DIFS time period.

As the destination station successfully receives a frame, it acknowledges by sending back an ACK frame after SIFS (Short Inter-Frame Space) time period. Figure 1.2 illustrates the mechanism.

SIFS is the shortest of the three Inter-Frame Spaces (IFS) defined in IEEE 802.11 to control the access to the medium. IFS relationships can be seen in Figure 1.3 (other parts of figure are explained in next section). Subsequent frame transmissions are separated by these inter-frame spaces depending on the priority of the frame exchange sequence,

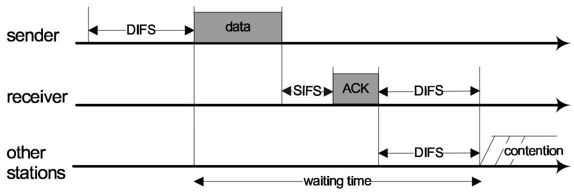

Figure 1.2: DCF basic access mechanism.

i.e., higher the priority of the frame exchange sequence, shorter is the inter-frame space used between the frames. The SIFS between the data and ACK frames, as seen in Figure 1.2, therefore prevents other stations to transmit at the same time the receiver transmits the ACK frame and thereby resulting in transmission failure, because other stations have to wait for the DIFS time prior to start transmission which is longer than SIFS. Thus, in this way, a station transmitting the ACK frame is given priority over the stations trying to transmit data frames.

The second shortest inter-frame space, PIFS (PCF Inter-Frame Space), is used by AP in the PCF, the optional access mechanism of IEEE 802.11, in which PC/AP centrally controls the access to the medium by polling individual stations. In PCF, PC/AP is given priority over ordinary stations such that it has to wait PIFS instead of longer DIFS prior to transmitting a frame.

The values of inter-frame spaces are dependent on the underlying Physical layer (PHY) and are defined in relation to a slot time. Slot time is derived from propagation delay, transmitter delay, and other PHY dependent parameters [7]. PIFS consists of a SIFS plus one slot time and DIFS consists of a SIFS plus two slot times.

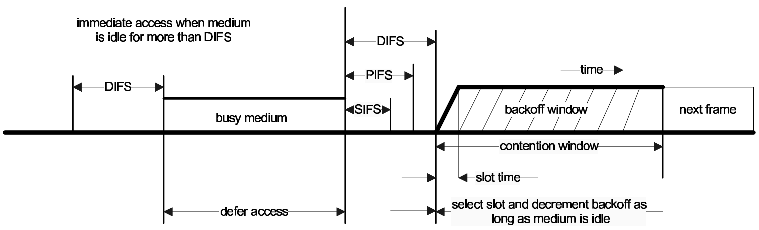

Figure 1.3: Medium access and IFS relationships.

Two types of carrier sensing are used to determine whether the medium is idle or busy. With Physical Carrier Sensing, the wireless channel is sensed itself at the Physical layer. On the other hand, Virtual Carrier Sensing is used at the MAC layer, such that as a station receives a frame that is not directed to it, it examines the *duration* field in the frame header, which specifies the time required to transmit the frame and to receive the ACK frame in response, and then defers the access to the medium for that particular period of time. The process is described in more detail in Section 1.2.3.

1.2.1 Collision Avoidance and Backoff Procedure

The above scenario may lead to collisions if two or more stations sense the medium idle and try to transmit at the same time. In order to avoid such collisions, station has to wait an additional time period prior to transmitting if the medium is sensed busy in the DIFS period, or, if the medium was busy just before the station started waiting the DIFS period. In these situations, the station defers access until the medium becomes idle, and chooses a random backoff value, which specifies the time period, measured in time slots, the station has to wait in addition to the DIFS after the medium becomes idle. This additional random delay in form of backoff helps to avoid collisions, otherwise all stations would try to transmit as soon as medium becomes idle for the DIFS period. This mechanism is called Collision Avoidance (CA), and thus the whole access mechanism is referred to as CSMA/CA.

The reason of using Carrier Sensing with Collision Avoidance instead of Collision Detection (CSMA/CD) used in wired networks, e.g., IEEE 802.3 Ethernet [2], is lack of collision detection capabilities in wireless networks. In wired networks, transceiver has the ability of receiving and transmitting simultaneously and therefore is able to detect collisions, but in wireless networks, the stations very often do not have the ability of simultaneous operation. Even if the station has the ability of receiving while transmitting, the fundamental characteristics of wireless communication do not allow it to detect other signals. Compared to the wired communication where the signal strength in the wire/cable does not drop below an acceptable level[3] and thus makes it possible for the sender to detect the colliding signal, in wireless communication it is not possible because in free space the strength of signal decreases proportionally to the square of distance to the sender. Moreover, various types of interferences and noises, and fading further attenuate the signal strength, thereby making it difficult for the sender to detect other signals in presence of its own signal, because the strength of its own signal is several magnitudes higher than the strength of the signal being detected [7, p. 70].

After choosing the backoff value, as the medium is sensed idle at least for DIFS time period, the station starts decrementing its backoff timer by one for each time slot. If the medium becomes busy during this backoff process, the station backoffs, i.e., it pauses its backoff timer. The backoff timer is then resumed as soon as the medium is sensed idle for the DIFS period again. The station is allowed to transmit as the backoff timer reaches zero.

While a new station has to choose a new backoff value, the station which attempted first continues to count down its paused backoff timer instead of choosing a new one. Thus, in this way a station that attempted first and thus waited longer is given advantage over a station that attempted after it, because it only has to wait for its remaining backoff time.

The random backoff value is uniformly chosen[4] from the interval [0, CW], called the Contention Window. At the first transmission attempt, CW is set to the minimum Contention Window size, CWmin. After each unsuccessful transmission, CW is doubled, actually increased exponentially, using the equation (2 x (CW + 1) - 1), until it reaches the maximum Contention Window size, CWmax. The values of CWmin and CWmax are dependent on the underlying Physical layer (PHY). For the most commonly used Physical layer, DSSS PHY, the values are 31 and 1023 respectively, in which case the Contention Window size increases in the form of 31, 63, 127, 255, 511, 1023, as shown in Figure 1.4. For this reason, the mechanism is also referred to as exponential backoff.

An unsuccessful transmission (in other words, a collision) is determined if the sender station does not receive ACK frame within a specified ACK timeout period. After the

[3] The strength of the signal in wire is directly proportional to the length of the wire. Thus if the length of wire stays within standardized limit, more or less same signal strength can be assumed all over the wire [7, p. 70].

[4] A better choice would be to use the word *draw* or *get* instead of *choose* because the backoff value is not explicitly chosen according to desire and instead is drawn randomly.

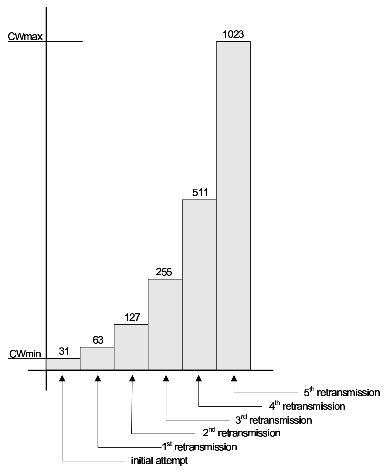

Figure 1.4: Exponential increase of Contention Window.

ACK timeout period, the station assumes that a collision has occurred and enters into the backoff period again after waiting for medium to be idle for DIFS, such that the new backoff value is chosen from the doubled CW. With the doubled CW size, the probability of a bigger random backoff value is higher, which reduces the probability of the stations colliding again.

DCF specifies a retransmit limit (also referred to as retry limit), i.e., the number of times a frame can be retransmitted. If an unsuccessful transmission is determined after reaching the retransmit limit, the frame is dropped.

The CW size is reset to CWmin after each successful transmission. The backoff mechanism is also used after a successful transmission before sending the next frame, i.e., if the sender station has another frame to send just after receiving ACK frame for the previous frame, it waits the medium to be idle for the DIFS time and chooses a new backoff value. This is referred to as *post backoff*, as it is done after the transmission not before. The post backoff ensures that there is at least one backoff interval between two consecutive transmissions. It allows other stations to decrement their backoff timers and thereby to get access to the medium.

The backoff procedure shows that the scenario described earlier, i.e., station transmitting immediately without waiting the backoff time, is rare, and only happens in the situations when, the time the frame arrives at MAC layer, the last post backoff has already been finished, i.e., the transmission queue is empty, and the medium has been idle longer than the DIFS time period. Only then the frame can be transmitted immediately after the DIFS time period. The frames immediately following this frame have to be transmitted after backoff, until the transmission queue becomes empty again.

The Collision Avoidance mechanism does not totally eliminate the risk of collisions. Collisions may still occur if the backoff timers for two or more stations reach zero at the same time, or if two or more stations accidentally get the same backoff values.

The latter also indicates that the probability of collision is inversely proportional to the size of Contention Window, i.e., smaller the Contention Window size, greater the rate of collisions and vice versa. Although a bigger Contention Window size reduces probability of collisions, it results in higher delays and inefficient bandwidth utilization.

An example of DCF operation follows next for further explanation of the backoff procedure and collision. It also highlights the important role of DIFS and CWmin and

CWmax parameters[5] in the access mechanism process.

1.2.2 Example of DCF Operation

Figure 1.5 illustrates an example scenario to explain the operation of DCF with backoff procedure. Here three stations are contending for the medium. The time Station 3 gets a frame to send, as the medium is idle and there is no ongoing backoff process, the station starts its transmission immediately after sensing the medium idle for the DIFS time period. The transmission is indicated by *busy* in the figure, and includes the complete frame exchange sequence of data + SIFS + ACK, as seen in Figure 1.2. Stations 1 and 2 arrive meanwhile and try to sense the medium idle for DIFS time period, but after sensing it busy, they defer the access.

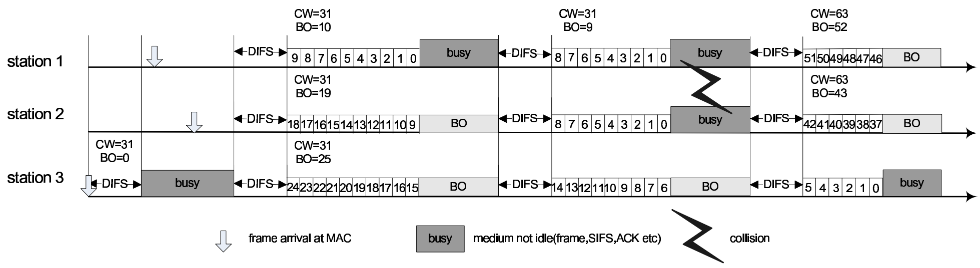

FIGURE 1.5: DCF ACCESS MECHANISM WITH BACKOFF PROCEDURE.

After the transmission finishes and the medium becomes idle, all three stations wait it idle for the DIFS time period and then choose the random backoff values. Here it is assumed that the backoff values chosen by stations 1, 2 and 3 are 10, 19 and 25, respectively. The backoff performed by the Station 3 immediately after the transmission of first frame is referred to as post backoff, as described earlier. All three stations then invoke the backoff procedure by decrementing their backoff timers with each idle time slot. The station with the smallest backoff value, Station 1, is able to count its timer down to zero first, and thus wins the access to the medium. As Station 1 starts transmission, stations 2 and 3 pause their backoff timers, indicated by BO in the figure. After the medium becomes idle again, all three stations wait for the DIFS and start backoff procedure again. While Station 1 chooses a new backoff value of 10, stations 2 and 3 resume their paused backoff timers. As the remaining backoff time of Station 2 is also 10, both Station 1 and Station 2 start transmitting at the same time after counting their timers down to zero. This leads to a collision. As stations 1 and 2 start transmission, Station 3 senses the medium busy and pauses its backoff timer. Stations 1 and 2, being unaware of the collision (because of the reason described in Section 1.2.1), wait for the ACK frames from the receiver stations. As no ACK is received within the ACK timeout period, both stations assume that a collision has occurred, and thus, after waiting for the DIFS time period, double their Contention Windows and choose new backoff values. As the remaining backoff time of Station 3 is smaller compared to the new backoff values of stations 1 and 2, Station 3 is able to transmit first after waiting for the backoff time, and stations 1 and 2 pause their backoff timers.

[5] As these parameters are used to contend for the medium, they are sometimes referred to as *contention parameters*.

1.2.3 RTS/CTS Mechanism

An additional mechanism, RTS/CTS, is defined to solve the hidden terminal problem found in wireless networks that use CSMA. With RTS/CTS, the sender and receiver perform a handshake mechanism by exchanging RTS (Request To Send) and CTS (Clear To Send) control frames. The procedure is shown in Figure 1.6. After waiting the DIFS time, prior to transmit the data frame, the sender sends a RTS frame to the receiver, and the receiver responds with a CTS frame after waiting a SIFS time. The CTS frame indicates that the handshake is successful and ensures that the medium has been reserved for the particular sender and receiver for the transmission.

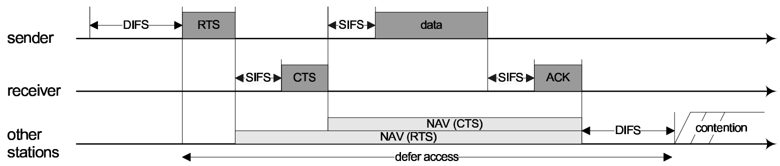

Figure 1.6: Frame exchange sequence with RTS/CTS mechanism.

RTS/CTS uses Virtual Carrier Sensing, such that the RTS and CTS frames include the duration of the complete frame exchange sequence, inclusive of SIFS and ACK. All stations within the receiving range around the sender and receiver set their NAV (Network Allocation Vector) after receiving the RTS and CTS frames, and thus are informed that they have to wait until the current transmission finishes. Here an important thing to note is that the sets of stations receiving RTS and CTS frames can be different. NAV is a timer and is decremented in the similar way the backoff timer is decremented. The station is allowed to transmit after its NAV reaches to zero.

Collisions can only occur at the beginning when the RTS frame is transmitted, as two or more stations may start transmitting, either RTS or data frames, at the same time. Such collisions are determined if the sender does not receive the CTS frame within a specified CTS timeout period. In that case, the sender transmits the RTS frame again. As the size of the RTS frame is significantly smaller compared to that of data frame, the RTS/CTS mechanism provides a mean of fast recovery from collisions, because the sender becomes aware of failure and may retransmit more quickly compared to the case when long data frame is transmitted and failure is determined after ACK timeout.

As seen in Figure 1.6, here again the SIFS intervals between RTS, CTS and data frames prevent other stations to transmit and thereby interrupting the transmission.

RTS/CTS mechanism shall be used for large data frames. Using it for small data frames may result in significant overhead causing inefficient capacity utilization and higher delays.

1.3 Summary

IEEE released the 802.11 Wireless Local Area Network (WLAN) standard in 1997. While the original standard supported maximum data transmission rate of 2 Mbps, in 1999, its enhanced versions 802.11b and 802.11a increased data rates to 11 and 54 Mbps, respectively. IEEE 802.11 MAC defines two different access mechanisms. The mandatory Distributed Coordination Function (DCF) and an optional Point Coordination Function (PCF). Most of the 802.11 installations today deploy DCF, and PCF is hardly implemented due to its complex design.

DCF is based on Carrier Sense Multiple Access (CSMA). Before transmitting a frame, the station senses the medium, and if the medium is found idle at least for DIFS (DCF Inter-Frame Space) time period, the station starts transmission. Otherwise, if the medium is found busy during the DIFS period, the station defers access and chooses a random backoff value that specifies the additional time it has to wait after the medium becomes idle again. As the medium becomes idle for the DIFS time period again, the station starts decrementing its backoff time. If medium becomes busy during this backoff process, the station pauses the backoff timer, and resumes it as the medium becomes idle for the DIFS period again. The station is allowed to transmit as the backoff timer reaches to zero. The additional random backoff time helps avoiding collision, which is defined as the situation when two or more stations transmit at the same time and result in unsuccessful transmissions. As the destination station receives the frame, it acknowledges by sending back an ACK frame after SIFS (Short Inter-Frame Space) time period. SIFS, like DIFS, is one of the three inter-frame spaces (IFS) defined to control the medium access.

The random backoff value is uniformly chosen in the range (0, CW), where CW is called the Contention Window. CW is initialized to the minimum size CWmin and doubled after each unsuccessful transmission, until it reaches the maximum Contention Window size, CWmax. CW is reset to CWmin after every successful transmission.

An additional mechanism RTS/CTS is defined that allows the sender and receiver to handshake by exchanging RTS (Request To Send) and CTS (Clear To Send) frames prior to transmitting the data frame. In this way, the medium is reserved such that all stations which received RTS or CTS frames defer the access until the transmission is finished.

Chapter 2

Introduction to IEEE 802.11e

2.1 Introduction

IEEE is currently working on a new standard, called IEEE 802.11e [8], which is an enhanced version of the legacy[1] 802.11 MAC in order to support quality of service (QoS). IEEE 802.11e is in standardization process and the final draft has been released.

IEEE 802.11e supports quality of service by introducing priority mechanism. All types of data traffic are not treated equally as it is done in the original standard, instead, 802.11e supports service differentiation by assigning data traffic with different priorities based on their QoS requirements. Furthermore, four different Access Categories (AC) have been defined each for data traffic of a different priority. Access to the medium is then granted based on the priorities of data traffic, such that each frame with a particular priority is mapped to an Access Category, and service differentiation is realized by using a different set of contention parameters to contend for the medium, for each AC.

In IEEE 802.11e, the AP and STA that provides QoS services are referred to as QAP (QoS Access Point) and QSTA[2] (QoS Station) respectively, and the BSS they are operating in is called QBSS (QoS Basic Service Set).

IEEE 802.11e introduces a new coordination function, called Hybrid Coordination Function (HCF), to provide QoS support. Subsequent sections describe HCF together with the detailed description of its service differentiation mechanism.

2.2 HCF (Hybrid Coordination Function)

IEEE 802.11e defines a new coordination function called Hybrid Coordination Function (HCF). HCF is a centralized coordination function that combines the aspects of DCF and PCF with enhanced QoS mechanisms to provide service differentiation. HCF provides both distributed and centrally controlled channel access mechanisms similar to DCF and PCF in the original standard. The distributed, contention-based channel access mechanism of HCF is called Enhanced Distributed Channel Access (EDCA), and the centrally controlled, contention-free channel access mechanism is called HCF Controlled Channel Access (HCCA).

IEEE 802.11e introduces Transmission Opportunity (TXOP), defined as the time period during which a QSTA has the right to transmit. In other words, in 802.11e

[1] The word *legacy* is often used to refer to the original 802.11 standard.

[2] For simplicity, most of the times in this report we use the word *station* for both *STA* and *QSTA*.

when a station gets access to the medium, it is said to be granted the TXOP. TXOP is characterized by a starting time and a maximum duration, called TXOP Limit. As a QSTA gets the TXOP, it can then start transmitting frames such that the transmission duration does not exceed the TXOP limit. TXOP Limit is specified by the QAP.

The next section describes EDCA, the distributed access mechanism of HCF. The detailed functionality of the centrally controlled access mechanism HCCA is beyond the scope of this report as we focus on the EDCA.

2.3 EDCA (Enhanced Distributed Channel Access)

The EDCA provides differentiated, distributed access to the medium using different priorities for different types of data traffic. The detailed description of the components and operation of EDCA is presented next.

2.3.1 Access Categories (ACs)

EDCA defines four Access Categories (ACs) for different types of data traffic, and service differentiation is introduced such that for each AC, a different set of parameters is used to contend for the medium. These parameters are referred to as EDCA parameters and are described in the next subsection.

Frames from different types of data traffic are mapped into different ACs depending on the QoS requirements of the traffic/application the frames belong to. The four Access Categories are named AC_BK, AC_BE, AC_VI and AC_VO, for Background, Best Effort, Video and Voice data traffic, respectively, where AC_BK has the lowest and AC_VO has the highest priority.

Each frame from the higher layer arrives at the MAC layer along with a priority value. This priority value is referred to as User Priority (UP) and assigned according to the type of application/traffic the frame belongs to. There are eight different priority values ranging from 0 to 7.

Priority	User Priority (UP)	Access Category (AC)	Designation
Lowest	1	AC_BK	Background
.	2	AC_BK	Background
.	0	AC_BE	Best Effort
.	3	AC_BE	Best Effort
.	4	AC_VI	Video
.	5	AC_VI	Video
.	6	AC_VO	Voice
Highest	7	AC_VO	Voice

TABLE 2.1: USER PRIORITY (UP) TO ACCESS CATEGORY (AC) MAPPINGS.

How to assign such a priority to each frame is a higher layer implementation issue. Interestingly, the draft standard does not specify anything how such a priority is assigned at the higher layers. Generally, it can be assigned by application generating the traffic, or by the user using the application. The latter solution indicates that every application has to be updated in order to be compatible with 802.11e. Another possibility would be to adaptively assign the priority at the Application layer, based on the traffic characteristics e.g., data rate, packet[3] interval, packet size etc. How this priority will travel through different layers down to the MAC layer further indicates that modifications at the higher layers would be inevitable.

At the MAC layer, a frame with a particular UP is further mapped to an AC. ACs are derived from the UPs as illustrated in Table 2.1.

2.3.2 EDCAF (Enhanced Distributed Channel Access Function)

Every station maintains four transmit queues one for each AC, and four independent EDCAFs (Enhanced Distributed Channel Access Function), one for each queue, as illustrated in Figure 2.1. EDCAF is an enhanced version of DCF, and contends for the medium on the same principles of CSMA/CA and backoff, but based on the parameters specific to the AC it is contending for[4]. Next section discusses these parameters, referred to as EDCA parameters .

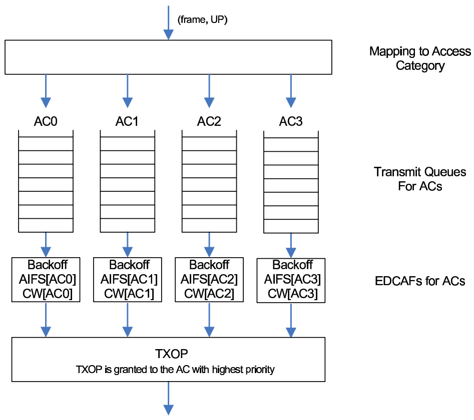

Figure 2.1: Four ACs, each with its own queue, AIFS, CW and backoff timer.

2.3.2.1 EDCA Parameters

An EDCAF contends for medium based on the following parameters associated to an AC:

- AIFS - The time period the medium is sensed idle before the transmission or backoff is started.
- CWmin, CWmax - Size of Contention Window used for backoff.
- TXOP Limit - The maximum duration of the transmission after the medium is acquired.

[3] A packet at the MAC layer is generally referred to as *frame*. Therefore sometimes the words *packet* and *frame* are used interchangeably.

[4] AC and EDCAF are sometimes referred interchangeably because of very close relationship between them.

The values of EDCA parameters[5] are different for different ACs. The higher priority ACs wait a small AIFS time period while the lower priority ACs have to wait a longer AIFS time before they can access the medium. The size of Contention Window varies such that the higher priority ACs choose backoff values from a smaller Contention Window compared to the lower priority ACs. TXOP Limit is also set in a way that the higher priority ACs get the access to the medium for longer durations. Basically, the higher the priority of an AC, the smaller the AIFS, CWmin and CWmax, and larger the TXOP Limit. As the values of EDCA parameters are AC specific, they are sometimes referred to as AIFS[AC], CWmin[AC], CWmax[AC] and TXOP Limit[AC].

Thus, basically the main difference between DCF and EDCAF is that EDCAF uses AC specific parameters AIFS[AC], CWmin[AC] and CWmax[AC] instead of using fixed values DIFS, CWmin, and CWmax.

EDCA parameters are periodically advertised by the QAP. QAP can adapt these parameters dynamically depending on the network conditions. The draft standard specifies default values of EDCAF parameters if not advertised by the QAP.

A brief description of each of the EDCA parameters and its role in providing service differentiation is presented in next section.

AC	**CWmin**	**CWmax**	**AIFSN**	**TXOP Limit**	
				FHSS	**DSSS**
AC_BK	CWmin	CWmax	7	0	0
AC_BE	CWmin	CWmax	3	0	0
AC_VI	(CWmin+1)/2-1	CWmin	2	6.016ms	3.008ms
AC_VO	(CWmin+1)/4-1	(CW+1)/2-1	2	3.264ms	1.504ms

TABLE 2.2: DEFAULT EDCA PARAMETER VALUES.

AIFS (Arbitration Inter-Frame Space) - The minimum time period for which the medium must be sensed idle before an EDCAF/station may start transmission or backoff is not the fixed value DIFS, as it is in DCF, but is a variable value, AIFS, that depends on the AC for which the EDCAF is contending for. AIFS is derived from the following equation:

$$\text{AIFS} = \text{AIFSN x aSlotTime} + \text{aSIFSTime},$$

where aSlotTime is the slot time, aSIFSTime is the SIFS time period and AIFSN (Arbitration Inter-Frame Space Number) is used to determine the length of the AIFS. AIFSN specifies the number of time slots in addition to the SIFS time period the AIFS consists of. Different AIFSN values are used for different ACs such that the high priority ACs use smaller values compared to the low priority ACs. The minimum possible value of AIFSN is 2. As a DIFS is equal to 2 x aSlotTime + aSIFSTime, it shows that the minimum length of AIFS is same as of DIFS. For QAP operating in HCCA, the minimum possible value of AIFSN is 1, which makes it equal to PIFS as PIFS is 1 x aSlotTime + aSIFSTime.

The default AIFSN values for all four ACs can be seen in Table 2.2. Figure 2.2 further explains how priority is given to different ACs based on the AIFS time periods.

[5]The terms *EDCA parameters*, *AC parameters*, *QoS parameters*, and *contention parameters* are used interchangeably to refer to these parameters throughout the report.

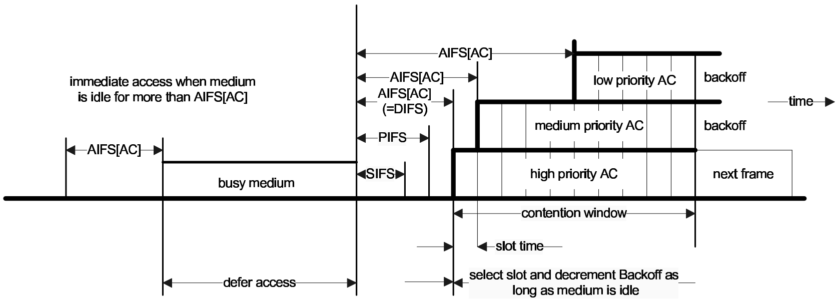

Figure 2.2: Prioritization based on AIFS.

The smaller AIFSN value for a higher priority AC explains that the corresponding EDCAF has to wait shorter time period before it can start transmission or counting down its backoff timer compared to the EDCAF for a low priority AC. In this way, the higher priority ACs are guaranteed greater share of the bandwidth. Moreover, smaller AIFS lengths ensure that the higher priority ACs will not suffer from long delays, which are vary critical for the delay-sensitive applications/traffics, as discussed in Section **??**. The lower priority ACs may suffer from longer delays because of the larger AIFS durations they have to wait, but since these ACs are designed for delay-tolerant applications/traffics, certain amount of delays do not degrade their performance beyond an acceptable limit.

CWmin and CWmax - The minimum and maximum Contention Window size limits are not fixed as it is in DCF, but are variable depending on the AC. The higher priority ACs have smaller CWmin and CWmax values compared to lower priority ACs. The default values of CWmin and CWmax parameters for each of the four ACs are presented in Table 2.2. The Contention Window parameters specific for the physical layers are presented in Table 2.3.

	FHSS	**DSSS**
CWmin	15	31
CWmax	1023	1023

Table 2.3: Contention Window parameters for different Physical layers.

A smaller Contention Window for an AC will cause the corresponding EDCAF to choose smaller random backoff values, and thereby waiting shorter time period in addition to AIFS as the medium becomes idle. It gives such an AC priority over the AC with a larger Contention Window, which results in larger backoff values and thereby longer delays.

As seen in Table 2.2, for the commonly used Physical layer DSSS, the CWmin values for lower priority ACs, AC_BE and AC_BK, are same as it is for the legacy 802.11 DCF, but these values for higher priority ACs, AC_VO and AC_VI, are as small as one half or quarter of those of the lower priority ACs. This results in smaller backoff values for the high priority ACs and thereby shorter medium access delays. The negative aspect of small Contention Window sizes for higher priority ACs is, that they suffer from higher number of collisions. The reason is, as described in Section 1.2.1, that the probability of choosing the same backoff values or counting the backoff timers to zero at the same time increases with the decreasing size of Contention Windows.

CWmax values for high priority ACs are also set such that they are equal or less than the CWmin values for the lower priority ACs, i.e., Contention Windows are non-overlapping. This shows that after doubling the Contention Window size in case of an unsuccessful transmission, i.e., collision, its size still remains smaller than the CWmin size of lower priority ACs. Furthermore, it also indicates that while a low priority AC has to double its CW size after each unsuccessful transmission, until it reaches the CWmax, and with higher probability, has to choose a bigger backoff value for each retransmission, the Contention Window size of a high priority AC becomes constant after fewer retransmissions, allowing it to consistently choose smaller backoff values and thereby winning access to the medium. In this way, high priority ACs are given consistent and greater share of the bandwidth in the situations when the network has become congested. On the other hand, this may severely degrade the performance of the low priority ACs since they might not be able to decrement their backoff timers because of the smaller post backoff durations of the higher priority ACs. The situation is further explained with an example later in this section.

As it can be seen in Table 2.2, the default values for CWmin and CWmax for both AC_BE and AC_BK are same, but priority is given to AC_BE over AC_BK by assigning it a much smaller AIFSN, i.e., 3 compared to 7, indicating that AC_BK has to wait four additional slots prior to starting transmission or backoff procedure. It also shows that AC_BK suffers from much high delays compared to the other ACs.

TXOP (Transmission Opportunity) - As described above, TXOP is the time duration an EDCAF may transmit after winning access to the medium. TXOP is characterized by a maximum duration, called TXOP Limit. As an EDCAF gets the TXOP, it can then start transmitting frames such that the transmission duration does not exceed the TXOP Limit. The transmission duration covers the whole frame exchange sequence, including the intermediate SIFS periods and ACKs, and the RTS and CTS frames if RTS/CTS mechanism is used.

Table 2.2 shows the default TXOP limits for different ACs. A non zero value of TXOP Limit indicates that the EDCAF may transmit multiple frames in a TXOP, provided that the transmission duration does not exceed the TXOP Limit and the frames belong to the same AC. This is then referred to as Contention Free Bursting (CFB). The consecutive frame transmissions in a TXOP are then separated by SIFS time periods instead of AIFS plus the post backoff periods, as illustrated in Figure 2.3. It is important to note that the multiple frame transmission is granted to EDCAF (or AC) and not to the station, i.e., it is only allowed for the transmission of frames of the same AC as of the frame for which the TXOP was obtained.

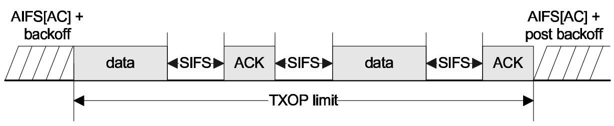

FIGURE 2.3: CONTENTION FREE BURSTING (CFB).

If RTS/CTS mechanism is used with CFB, then the RTS CTS frames handshake is done only once, before the first frame, instead of for every frame in the CFB.

The TXOP Limit of zero indicates that CFB is disabled, and thus only one frame, in addition to RTS/CTS if enabled, can be transmitted in a TXOP. In that case, if there is a risk that the transmission duration of the first frame may exceed the TXOP limit, then the frame should be fragmented.

As it can be seen in Table 2.2, the default values of TXOP limits for the low priority ACs, AC_BK and AC_BE, are zero, indicating that CFB is disabled for these ACs. For high priority ACs, CFB allows to seize the medium for certain amount of time periods, which results in significantly reduced delay. However, too large TXOP limits for high priority ACs may result in higher delays for the low priority ACs.

Thus, service differentiation is introduced through TXOP limits by allowing higher priority ACs to gain continuous access to the medium for longer time periods compared to the lower priority ACs.

In the case when CFB is enabled, the virtual carrier sensing is applied such that the *Duration* field in frame header is set to the remaining duration of the whole TXOP and thus all stations receiving the frame set their NAVs for the duration of whole TXOP instead of that of one frame, i.e., first frame in TXOP, plus the intermediate SIFS times and ACK.

2.3.3 Example of EDCA Operation

Besides the different AIFS, CWmin, CWmax and TXOP limit values for different ACs, the rest of medium access mechanism is same as in DCF, i.e., as the medium becomes idle at least for AIFS time period, the EDCAF chooses a random backoff value from its Contention Window and starts decreasing its backoff timer. The EDCAF can start transmission as its backoff timer reaches to zero.

Figure 2.4 shows an example of EDCA operation to further explain how individual EDCAFs in a station contend for the medium, assuming that all four EDCAFs have frames to transmit.

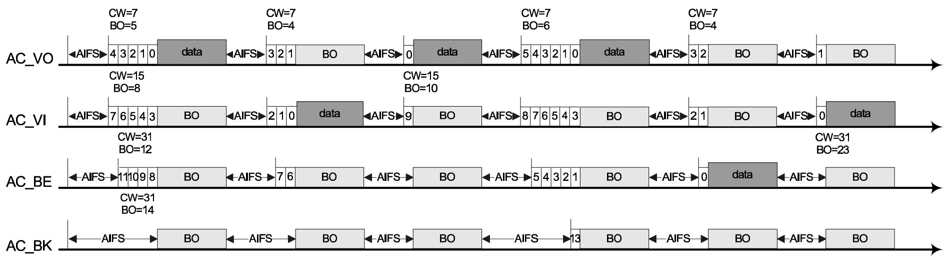

Figure 2.4: EDCA access mechanism.

Figure 2.4 shows all four EDCAFs inside a station contending for the medium. The current Contention Window size and backoff value for each of the AC/EDCAF is also shown. As the EDCAFs for higher priority ACs, AC_VO and AC_VI, have to wait smaller AIFS time periods, they starts counting their backoff timers down prior to EDCAFs for lower priority ACs, AC_BK and AC_BE. The figure assumes the default EDCA parameter values. As the default AIFSN values for AC_VO and AC_VI are same, i.e., 2, the corresponding EDCAFs start to count their backoff timers down at the same time. EDCAFs for AC_BK and AC_BE have to wait some additional slots because of their longer AIFS time periods. As a high priority AC has smaller minimum and maximum Contention Window limits compared to a lower priority AC, EDCAF for a high priority AC most of the time gets smaller backoff values and thus has to wait less time, as shown in the example.

As an EDCAF gets access to the medium, others pause their backoff timers, and continue to count down as soon as the medium becomes idle again for the AIFS time period. Thus, at a certain time, a lower priority EDCAF will have smaller backoff value because while the higher priority EDCAF has to choose a new backoff value for every next frame, the lower priority EDCAF continues to count down its paused backoff timer. This avoids the starvation of the lower priority ACs in a similar way it is avoided for different stations in the original standard, as described in Section 1.2.1.

The example clearly shows the role of EDCA parameters in achieving the service differentiation, i.e., a higher priority AC gets the larger share of the bandwidth by transmitting more frames compared to a lower priority AC. It also shows that the EDCAF for lowest priority AC, AC_BK, is starved. The reason is that since it has to sense the medium to be idle for a much longer AIFS time period, most of the time it is unable to decrement its backoff timer because another EDCAF starts transmitting and hence the medium becomes busy before its AIFS is finished. The other big reason is the smaller Contention Window sizes for high priority ACs, which with high likelihood results in shorter post backoff durations. As seen in Figure 2.4, the EDCAF for AC_BK is able to decrement its backoff timer either if there are no pending frames for high priority ACs, or, if the EDCAFs for AC_VO and AC_VI get post backoff values greater than 5 and the EDCAF for AC_BE gets post backoff value greater than 4.

Internal Collisions **-** As shown in Figure 2.4, as the four EDCAFs at the AC transmit queues behave like virtual stations inside the real station such that each EDCAF contends for the medium independently of other EDCAFs, there exist two levels of contention, internal contention among different EDCAFs/ACs inside the same station, and external contention among different stations. This may result into a situation where more than one EDCAF in the same station count their backoff timers to zero and try to transmit at the same time. This leads to a situation referred to as *internal collision* or *virtual collision*. In such situation, the access to the medium is granted to the EDCAF for the highest priority AC among the colliding EDCAFs, and the lower priority colliding EDCAF doubles its Contention Window and backoffs, just as if an external collision occurred. Figure 2.5 shows an example of internal collision.

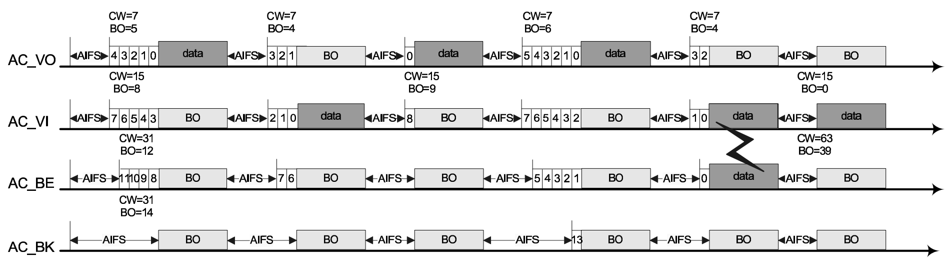

FIGURE 2.5: EDCA ACCESS MECHANISM AND INTERNAL COLLISION.

As described above, after the internal collision, out of the colliding EDCAFs the EDCAF for lower priority AC, AC_BE, doubles its Contention Window and chooses a new backoff value, and the EDCAF for higher priority AC, AC_VI, starts the transmission without any backoff. This explains that the traffic of higher priority AC does not suffer from additional delays after the occurrence of internal collisions, although it may starve the lower priority AC even more, i.e., it will take long time for EDCAF for AC_BE to count its new backoff of 39 down to zero. This condition may be far worse if EDCAF for AC_BK collides; it will hardly be able to decrement its backoff timer after higher priority EDCAFs will have transmitted dozens of frames.

An external collision occurs if backoff timers of the EDCAFs at two or more stations reach zero at the same time, or, if the EDCAFs at two or more stations accidentally get the same backoff values and win access to the medium. Similar to the original standard, after the external collision the colliding EDCAFs increase (double) their Contention Windows and choose new backoff values, and the rest of the EDCAFs retain their paused backoff timers. Figure 2.6 shows an example of external collision where EDCAFs for AC_VO and AC_VI in two different stations count their backoff timers down to zero and try to transmit at the same time. After determining the collision, both colliding EDCAFs double their Contention Windows and start decrementing the newly chosen backoff values while other EDCAFs in both stations continue to decrement their paused backoff timers.

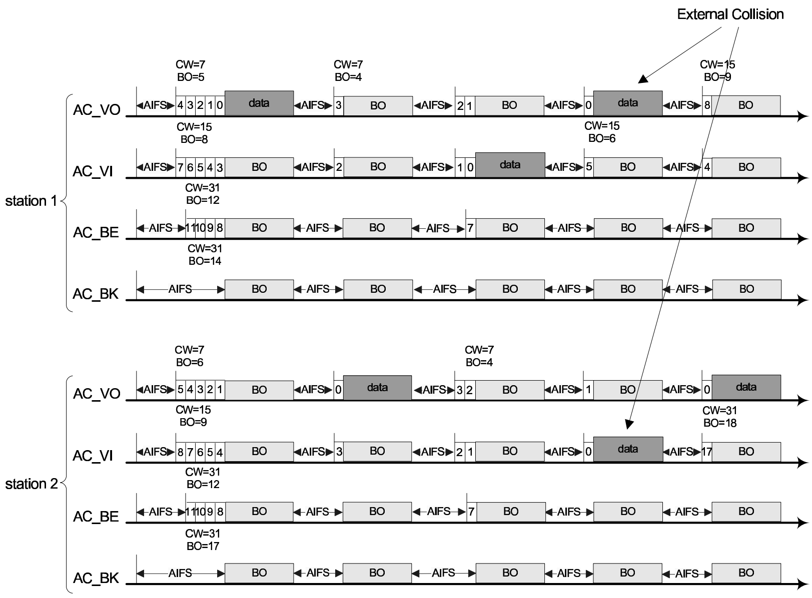

Figure 2.6: EDCA access mechanism and external collision.

2.4 Architecture and Important Frame Formats

Together with HCF and its two access mechanisms EDCA and HCCA, IEEE 802.11e also includes the two coordination functions from the original 802.11, DCF and PCF, in order to provide backward compatibility. Figure 2.7 illustrates the architecture of 802.11e MAC.

For backward compatibility, a QSTA can also operate in a non-QoS BSS (nQBSS) by associating itself to a non-QoS AP (nQAP), in case a QAP is not available. On the other hand, a non-QoS STA (nQSTA) may also associate with a QAP in a QBSS, such that it operates just like an ordinary STA in 802.11 and the transmissions from QAP to nQSTA do not use frame formats specific for QoS services.

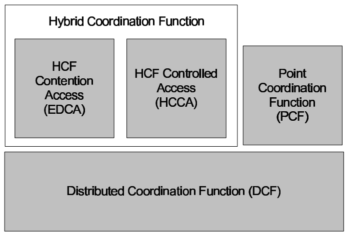

FIGURE 2.7: IEEE 802.11E MAC ARCHITECTURE.

The centrally controlled, contention-free channel access mechanism of HCF, i.e., HCCA, uses a centralized coordinator called HC (Hybrid Controller), which is collocated in QAP. HC operates concurrently with the EDCA just like in the legacy 802.11, i.e., a Contention Free Period (CFP) is followed by a Contention Period (CP), such that the EDCA operates in CP while HC operates both in CP and CFP. This is in contrast with legacy 802.11 where PC can only operate in CFP. It indicates that HC is capable of polling QSTAs both in CP and CFP, and explains why it is referred to as Hybrid Controller.

The one bit *QoS subfield* in *Frame Control* field of MAC header indicates whether the station is acting as a QSTA or an nQSTA. The field is set to 1 if the station is QSTAs, and 0 otherwise. Figure 2.8 shows the QoS subfield in the MAC header.

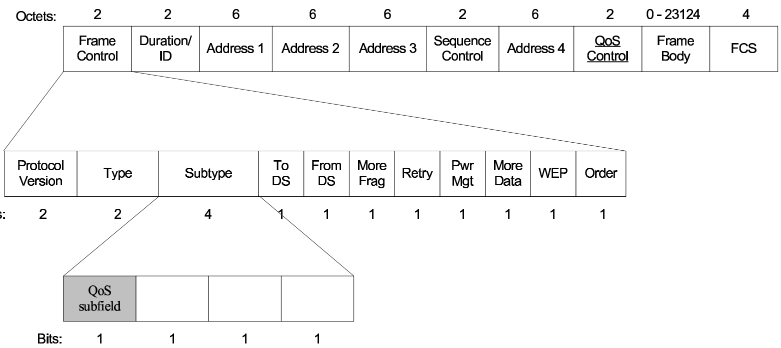

FIGURE 2.8: MAC DATA FRAME HEADER AND QOS SUBFIELD.

At MAC layer, each frame is assigned a priority in the form of a traffic identifier (TID). *TID* field in the newly added *QoS Control* field in the MAC header contains this TID value. UP of the frame is then determined based on this TID value, such that when QAP receives a frame from a QSTA, it gets the UP value from *TID* field, ranging from 0 to 7. Figure 2.9 illustrates the *QoS Control* and *TID* fields in the MAC header.

The priority value in *TID* field is supported only if the station has its *QoS subfield* in the *Frame Control* field set to 1, i.e., the station is associated with a QAP and thus working as a QSTA. If no QAP is available and a QSTA is associated with an ordinary AP, i.e., nQAP, then the QSTA is functioning just like an ordinary STA, which is indicated by setting the *QoS subfield* to 0. In that case, the TID value is meaningless and

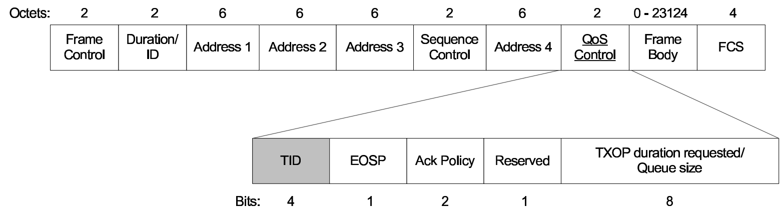

Figure 2.9: TID field in QoS Control field.

all frames from the station are treated as frames with priority of *Contention*, indicating that they shall be transmitted without any priority, as it is done in the DCF. Similarly, if an ordinary station, i.e., STA or nQSTA, is associated with a QAP, all frames from the station are treated as frames with priority of 0.

The *Queue size* field in *QoS Control* field of the frame header, as seen in Figure 2.9, specifies the total number of frames of the particular priority/TID the station have in its AC transmit queue, excluding the current frame.

TXOP are obtained both in EDCA and HCCA, such that the former is referred to as EDCA TXOP, and the later is referred to as HCCA TXOP or Polled TXOP. An EDCA TXOP is obtained as soon as a QSTA wins access to the medium while operating in EDCA. A HCCA TXOP is granted by the HC while operating in HCCA, such that the HC polls individual stations to grant HCCA TXOPs based on their requirements.

A QSTA can specify the intention to transmit multiple frames in a TXOP by setting the *Duration/ID* field in the frame header, such that it also includes the time required to transmit the additional frames. While operating in HCCA, a QSTA can request the TXOP of particular duration by setting the *TXOP duration requested* subfield of *QoS Control* field shown in Figure 2.9. The *TID* field in that case indicates the AC for which the TXOP is being requested. The HC/QAP may then assign a TXOP of the size requested or of a smaller size.

EDCA parameters are defined in *EDCA Parameter Set* element, and are periodically advertised by the QAP in selected frames (beacons). QAP can adapt these parameters dynamically, depending on the network condition. Figure 2.10 shows the structure of EDCA Parameter Set element.

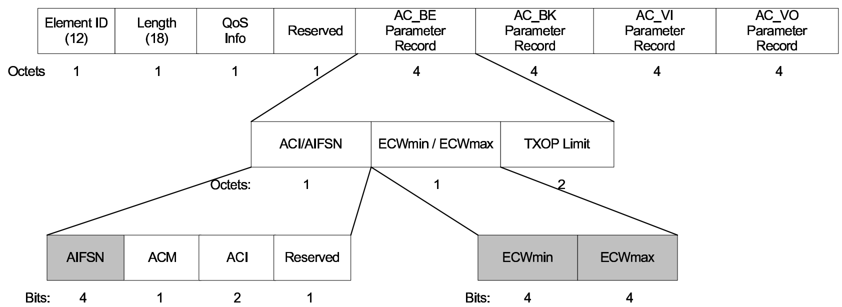

Figure 2.10: EDCA Parameter Set element.

The values of EDCA parameters are specified in the subfields *AIFSN*, *ECWmin*, *ECWmax*, and *TXOP Limit*, in the EDCA Parameter Set element. All QSTAs that receive the EDCA Parameter Set element from QAP update their EDCA parameter

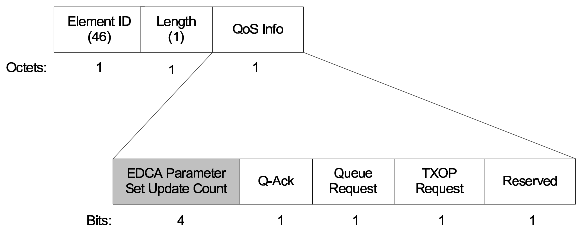

FIGURE 2.11: QoS CAPABILITY ELEMENT AND QoS INFO FIELD.

values and use new values to contend for the medium. The draft standard specifies the default values of EDCA parameters if not advertised by the QAP, as presented in Table 2.2.

Every time QAP updates the EDCA parameters, it increments the value of *EDCA Parameter Set Update Count* field in *QoS Info* field in the *QoS Capability element* sent in selected frames. The QSTAs use this information to confirm that they are using the latest set of EDCA parameters. The structure of QoS Capability element is illustrated in Figure 2.11.

2.5 Summary

IEEE 802.11e is an enhanced version of IEEE 802.11 WLAN standard to support quality of service (QoS). IEEE 802.11e provides support for QoS by introducing service differentiation. Four Access Categories (ACs) are defined to serve different types of data traffic. These Access Categories are named AC_BK, AC_BE, AC_VI and AC_VO, for Background, Best Effort, Video and Voice data traffic types, respectively. At the higher layers, each frame from a traffic stream is assigned a priority value, called User Priority (UP), ranging from 0 to 7. At the MAC layer this priority is mapped to one of the four Access Categories. For each Access Category, an enhanced version of DCF, called Enhanced Distributed Channel Access Function (EDCAF), contends for the medium based on a different set of contention parameters. These parameters are: (1) AIFS - the time period the medium is sensed idle before the transmission or backoff is started, (2) CWmin, CWmax - size of the Contention Window used for backoff, and, (3) TXOP Limit - the maximum duration of TXOP (Transmission Opportunity), which is the time period during which an EDCAF has the right to transmit after the medium is acquired. AIFS (Arbitration Inter-Frame Space) is derived from the equation AIFS = AIFSN x Slot Time + SIFS, where AIFSN (Arbitration Inter-Frame Space Number) specifies the number of time slots in addition to the SIFS time period an AIFS consists of. The values of EDCA parameters are AC specific and are set in a way that a higher priority AC waits a shorter AIFS time period before it can access the medium, chooses backoff values from a smaller Contention Window, and occupies the medium for a longer time period, compared to a lower priority AC. Basically, the higher the priority of an AC, the smaller the AIFS, CWmin and CWmax, and larger the TXOP Limit. These parameters are referred to as EDCA Parameters and are periodically advertised by the AP. The standard specifies the default values for EDCA Parameters if not advertised by the AP.

Besides the different sets of EDCA parameters for different ACs, the rest of medium access mechanism is same as in DCF, i.e., as the medium becomes idle for the AIFS time period, the EDCAF chooses a backoff value and starts decreasing its backoff timer.

Transmission is started as the backoff timer reaches to zero.

As the four EDCAFs inside the station contend for the medium independently of each other, it may result into a situation where more than one EDCAF in the same station count their backoff timers to zero and try to transmit at the same time. This is then referred to as *internal collision* and is resolved such that among the colliding EDCAFs, the EDCAF for higher priority Access Category transmits and the lower priority EDCAF backoffs, just as if an external collision occurred.

A new feature of 802.11e, called Contention Free Bursting (CFB), enables an EDCAF to transmit multiple frames once the medium/TXOP is acquired, without contending for the medium for every frame. The consecutive frame transmissions are separated by SIFS time periods. The transmission duration of CFB is bounded by TXOP Limit.

References

[1] IEEE Std. 802.11, Part 11: *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.* 1997.

[2] IEEE Std. 802.3, Part 3: *Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications*. 1985.

[3] IEEE Std. 802.5, *Token Ring Access Method and Physical Layer Specificationss*. 1985.

[4] IEEE Std. 802.11b, Supplement to Part 11: *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band.* 1999.

[5] IEEE Std. 802.11a, Supplement to Part 11: *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extension in the 5 GHz Band.* 1999.

[6] IEEE Std. 802.11g, Supplement to Part 11: *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band.* 2003.

[7] J. Schiller. *Mobile Communications.* Addison-Wesley, second edition, 2003.

[8] IEEE 802.11e/D13.0, Draft Supplement to Part 11: *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Quality of Service (QoS) Enhancements.* January 2005.