8 Medium access control

The medium access control (MAC) layer provides, among other things, addressing and channel access control that makes it possible for multiple stations on a network to communicate. IEEE 802.11 is often referred to as wireless Ethernet and, in terms of addressing and channel access, 802.11 is indeed similar to Ethernet, which was standardized as IEEE 802.3. As a member of the IEEE 802 LAN family, IEEE 802.11 makes use of the IEEE 802 48-bit global address space, making it compatible with Ethernet at the link layer. The 802.11 MAC also supports shared access to the wireless medium through a technique called carrier sense multiple access with collision avoidance (CSMA/CA), which is similar to the original (shared medium) Ethernet's carrier sense multiple access with collision detect (CSMA/CD). With both techniques, if the channel is sensed to be "idle," the station is permitted to transmit, but if the channel is sensed to be "busy" then the station defers its transmission. However, the very different media over which Ethernet and 802.11 operate mean that there are some differences.

The Ethernet channel access protocol is essentially to wait for the medium to go "idle," begin transmitting and, if a collision is detected while transmitting, to stop transmitting and begin a random backoff period. It is not feasible for a transmitter to detect a collision while transmitting in a wireless medium; thus the 802.11 channel access protocol attempts to avoid collisions. Once the medium goes "idle," the station waits a random period during which it continues to sense the medium, and if at the end of that period the medium is still "idle," it begins transmitting. The random period reduces the chances of a collision since another station waiting to access the medium would likely choose a different period, hence the collision avoidance aspect of CSMA/CA.

The simple distributed, contention-based access protocol supported by the CSMA/CA technique is the basis for the 802.11 MAC protocol and also where the similarity to Ethernet ends. The wireless medium, being very different from the wired medium, necessitates a number of additional features:

- The wireless medium is prone to errors and benefits significantly from having a low latency, link level error recovery mechanism.
- In a wireless medium not all stations can "hear" all other stations. Some stations may "hear" the station on one end of an exchange but not the station at the far end (the hidden node problem).
- The data rate that a channel can support is affected greatly by distance and other environmental effects. Also, channel conditions may change with time due to station

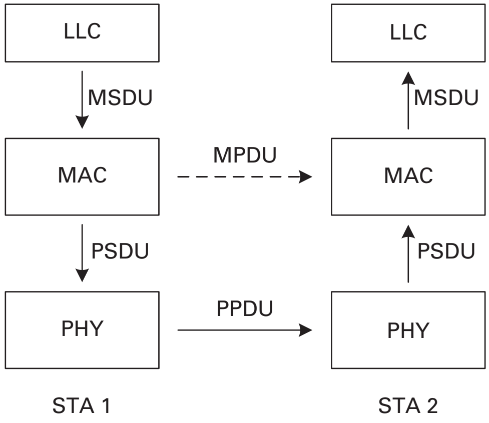

Figure 8.1 Protocol layering and messaging.

mobility or environmental changes. Stations need to continually adjust the data rate at which they exchange information to optimize throughput.

- Stations, often being mobile, need management mechanisms for associating with and disassociating from WLANs as they change location.

This chapter provides an overview of the 802.11 MAC prior to the enhancements introduced in 802.11n. After some background information on protocol layering, there is a brief overview of 802.11 management functions. This is followed by a more detailed overview of the channel access and data transfer aspects.

8.1 Protocol layering

Some basic concepts regarding protocol layering and messaging and illustrated in Figure 8.1 are needed in order to understand the MAC functionality. In this layered model, each entity, PHY and MAC, offers services to the entity in the layer immediately above it and user data is transferred between the layers as a service data unit (SDU). The MAC receives data from the logical link control (LLC) layer, and delivers data to the LLC layer through the MAC SDU (MSDU). The PHY receives data from the MAC and delivers data to the MAC in a PHY SDU[1] (PSDU).

A protocol is the means by which entities in the layered model exchange data and control information with their peer entities. This exchange takes place through protocol data units (PDUs). The MAC exchanges MAC PDUs (MPDUs) with its peer and the PHY exchanges PHY PDUs with its peer.

Another commonly used term in the 802.11 standard is station or STA, which refers to the MAC and PHY in the context of the device that incorporates these entities. In practice

[1] The IEEE 802.11 standard uses the term PLCP SDU instead of PHY SDU, where PLCP is the physical layer convergence procedure, a sublayer at the top of the PHY. The terms are equivalent.

this is the network adaptor in a laptop or the communication subsystem in a mobile phone. An AP (access point) is a station with additional functions related to managing an infrastructure BSS and providing access to the distribution system.

8.2 Finding, joining, and leaving a BSS

As described in Chapter 1, the BSS is the basic building block of an 802.11 WLAN. IEEE 802.11 defines two types of BSS, the independent BSS (IBSS), which is an ad-hoc association of stations that communicate with one another directly, and the infrastructure BSS, which is anchored by an AP that may be connected to a distribution system (DS) and through which the majority of the data transfer takes place, both station to station and station to DS. The Wi-Fi Alliance P2P specification (Wi-Fi, 2010) defines a variation on the infrastructure BSS called a peer-to-peer (P2P) group. A P2P group is an ad-hoc arrangement of stations where one station, the group owner (GO), assumes a coordinating role similar to the AP. In this book we are primarily concerned with the infrastructure BSS and P2P group; however, much of the discussion applies equally to IBSSs. The term BSS is loosely used to refer to both an infrastructure BSS and a P2P group.

A station becomes aware of the existence of a BSS through scanning, that is passively seeking beacon transmissions or actively probing for the existence of an AP through a Probe Request/Response exchange.

A station's membership of a BSS is dynamic. The station may turn on and off, or the station may be mobile and move in or out of the area covered by the BSS. A station becomes a member of a BSS by becoming "associated" with the BSS. On leaving the BSS, a station becomes "disassociated." In an ESS comprised of multiple infrastructure BSSs, a station may migrate from one BSS to another BSS within the ESS through "reassociation."

8.2.1 Beacons

The AP in an infrastructure BSS and the GO in a P2P group periodically broadcast Beacon frames. The beacon period defines a fixed schedule of target beacon transmission time (TBTTs) and the Beacon frame itself is transmitted on or as close to the TBTT as possible subject to the medium being idle (Figure 8.2).

The Beacon frame carries regulatory information, capability information and information for managing the BSS.

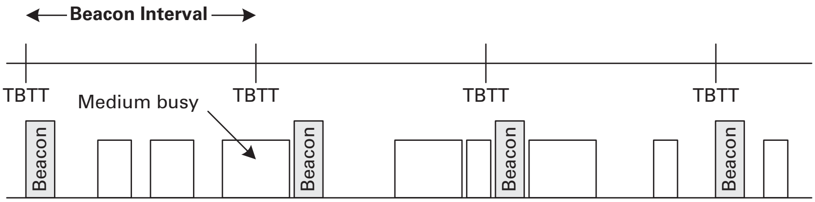

Figure 8.2 Beacon transmission on a busy network.

8.2.2 Scanning

Scanning is the process by which a station discovers a BSS and the attributes associated with that BSS. Two forms of scanning are possible: passive scanning and active scanning.

Passive scanning is a receive only operation that is compatible with all regulatory domains. With passive scanning the station looks for Beacon transmissions and may switch channels to find these transmissions. Beacon frames include, among other things, information on the country code, maximum allowable transmit power, and the channels to be used for the regulatory domain. Once the station has discovered the AP through its Beacon transmission and has this regulatory information it may probe the AP directly for additional information using a Probe Request/Response exchange if that additional information is not present in the Beacon frame itself.

Active scanning may be used when it is permitted by the regulatory domain in which the station operates. With active scanning a station transmits Probe Request frames on each of the channels where it is seeking a BSS. Depending on the generality of the search, the Probe Request frame includes the following addressing information:

- **SSID (service set identifier)**. The SSID in the Probe Request may be the SSID of the specific ESS for which the station is seeking BSSs or it may be the wildcard SSID.
- **BSSID (BSS identifier)**. The BSSID in the Probe Request may be the BSSID of a specific BSS or it may be the wildcard BSSID.
- **DA (destination address)**. The DA of the Probe Request frame is the broadcast address or the specific MAC address of the AP station.

An AP that receives a broadcast Probe Request sends a Probe Response to the station making the request if the following conditions are true: (a) the SSID is the wildcard SSID or matches the SSID of the ESS and (b) the BSSID is the wildcard BSSID or the AP's BSSID. Multiple APs may respond to a Probe Request using normal channel access procedures to avoid collisions.

8.2.3 Authentication

Authentication is the process by which two stations (one of which is usually an AP or GO) that wish to communicate establish their identity to a mutually acceptable level. The original 802.11 specification supported two authentication methods operating at the link level: open system authentication and shared key authentication. With the former any station may be admitted as a member of the BSS. With the latter, stations rely on the wired equivalent privacy (WEP) protocol to demonstrate knowledge of a shared encryption key.

With open system authentication, a station joining the BSS sends an Authentication frame requesting open system authentication and the AP responds with an Authentication frame with status "success."

With shared key authentication using WEP, a station joining the BSS participates in a four-way exchange of Authentication frames. The station initiates the exchange, sending

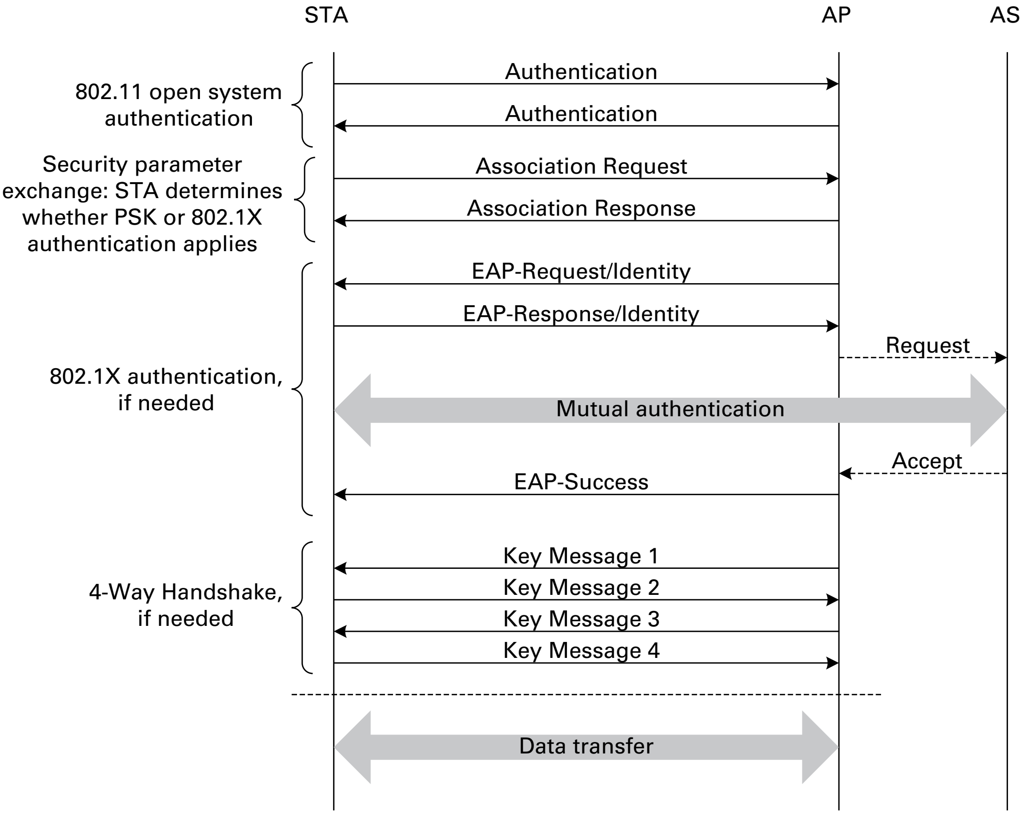

Figure 8.3 Modern 802.11 association exchange.

an Authentication frame that requests shared key authentication. The AP responds with an Authentication frame that carries challenge text. The station responds in turn with a third Authentication frame where the challenge text is encrypted using the shared WEP key. The AP decrypts the challenge text using its WEP key and if it matches the challenge text sent in the second Authentication frame, authentication is successful and a fourth Authentication frame is sent to the station with status “success.” In 2001, WEP was shown to be insecure and the newer security techniques were introduced with the 802.11i amendment. With these newer techniques, the Authentication frame exchange prior to association proceeds as it would for open system authentication, with actual authentication performed later in the association procedure. The association procedure using the new exchange is shown in Figure 8.3.

8.2.4 Association

Before a station is allowed to send data via an AP it must associate with the AP. Association provides a mapping between the station and AP that allows messages within the DS (distribution system) to reach the AP with which the station is associated and ultimately the station itself.

Association begins with the station sending an Association Request to the AP. If the station is admitted, the AP responds with an Association Response. With the Association Request and Response exchange, the station and AP exchange capability information (support for optional features) and the AP informs the station of specific operating parameters within the BSS.

8.2.5 Reassociation

Reassociation supports BSS-transition mobility, allowing a station to move from a current association with one AP to another within the same ESS. This keeps the DS informed of the current mapping between AP and station. Reassociation may also be performed to change attributes of the station association such as station capability information.

Reassociation is initiated by the station with the station sending a Reassociation Request to the AP. The AP responds with a Reassociation Response.

8.2.6 Disassociation

Disassociation terminates an existing association and may be performed by either the station or the AP. Stations should attempt to disassociate when they leave the network. However, because loss of communication may prevent this, a timeout mechanism allows the AP to disassociate the station without a message exchange should the station become unreachable.

To disassociate a station from the BSS, the AP or station sends a Disassociation frame. Disassociation is not a request, thus the other party merely acknowledges reception of the frame.

8.2.7 802.1X Authentication

A full discussion of 802.1X authentication is beyond the scope of this book; however, a brief overview is provided here to show the basic protocol exchange and its relationship to other messages exchanged during association. With 802.1X, a station desiring access to a BSS authenticates with an authentication server (AS) using the extensible authentication protocol (EAP). The AS may be co-located with the AP or, as is more typical, a separate system. 802.1X defines the protocol framework and encapsulation method for authentication but does not dictate the actual authentication method used. The method used is negotiated during the 802.1X exchange with the AS and numerous methods are supported today, some examples being EAP-Transport Layer Security (EAP-TLS), the Lightweight Extensible Authentication Protocol (LEAP) and EAP-MD5.

Following the Association Request/Response exchange (see Figure 8.3), the AP sends an EAP Request, challenging the station to identify itself. The station responds with an EAP Response that is forwarded to the AS. This is followed by an EAP authentication exchange between the station and the AS, with the AP re-encapsulating and forwarding the messages as appropriate. If the station successfully authenticates, the AS sends a message to the AP indicating “success,” together with information for a pairwise master key (PMK). The AP forwards an EAP-Success message to the station, signaling successful completion of the authentication. If authentication fails, the AS informs the AP and the AP sends an EAP-Failure message to the station followed by a Disassociation frame.

8.2.8 Key distribution

With modern security techniques, stations encrypt data frames using keys that are valid for a limited time period. Fresh keys are periodically generated (typically every 24 hours) to reduce the chance that a key can be discovered by analyzing encrypted traffic.

Two types of keys are used: a pairwise transient key (PTK) to protect traffic between the AP and an individual station and a group transient key (GTK) to protect the broadcast and multicast traffic sent by the AP.

IEEE 802.11 uses a four-way handshake to distribute the PTK and GTK and a two-way handshake to distribute the GTK when sent alone. Note that the PTK itself is never sent, only information which, together with pre-shared information, can be used by both sides to derive the PTK. The four-way handshake proceeds as follows:

- **Key Message 1 (AP to station)**. Carries the ANonce, a random number generated by the AP and used only once. On receipt of Key Message 1, the station generates a random number called the SNonce. Using the information in Key Message 1 (including the ANonce), the SNonce and knowledge of the PMK, the station derives the PTK.
- **Key Message 2 (station to AP)**. Carries the SNonce generated by the station. Using the information in Key Message 2 (including the SNonce), the ANonce and knowledge of the PMK, the AP derives the PTK. Key Message 2 also carries a message integrity check (MIC) generated using the PTK and by which the station demonstrates to the AP that it knows the PMK.
- **Key Message 3 (AP to station)**. Carries a group temporal key (GTK), if it is needed, encrypted using the PTK. Key Message 3 also carries a MIC by which the AP demonstrates to the station that it too knows the PMK. By sending Key Message 3, the AP indicates that it is satisfied that the station is who it says it is.
- **Key Message 4 (station to AP)**. The station confirms receipt of the GTK (if included in Key Message 3) and that it is satisfied that the AP is who it says it is.

With the successful delivery of Key Message 4 to the AP, a secure session is established between the AP and the station. Data frames can now be encrypted in both directions using the PTK.

The group key handshake consists of the Group Key Message 1 sent by the AP to an individual station carrying the GTK encrypted using the PTK. The station responds with Group Key Message 2, confirming receipt.

The four-way handshake is performed whenever the PTK is refreshed. The two-way group key handshake is performed when the GTK alone is refreshed.

8.3 Distributed channel access

The specific CSMA/CA mechanism used in the 802.11 MAC is referred to as the distributed coordination function (DCF). A station that wishes to transmit first performs a clear channel assessment (CCA) by sensing the medium for a fixed duration, the DCF inter-frame space (DIFS). If the medium is idle then the station assumes that it may take ownership of the medium and begin a frame exchange sequence. If the medium is busy, the station waits for the medium to go idle, defers for DIFS, and waits for a further random backoff period. If the medium remains idle for the DIFS deferral and the backoff period, the station assumes that it may take ownership of the medium and begin a frame exchange sequence.

The random backoff period provides the collision avoidance aspect. When the network is loaded, multiple stations may be waiting for the medium to go idle having accumulated packets to send while the medium was busy. Since each station probabilistically selects a different backoff interval, collisions where more than one station begins transmission at the same time are unlikely.

Once a station has gained access to the medium, it maintains control of the medium by keeping a minimum gap, the short inter-frame space (SIFS), between frames in a sequence. Another station will not gain access to the medium during that sequence since it must defer for a fixed duration, DIFS, which is longer than SIFS. Rules limit the types of frame exchange sequences that are allowed and the duration of those sequences to prevent one station from monopolizing the medium.

Fundamental to CSMA/CA is the carrier sense. The DCF uses both physical and virtual carrier sense functions to determine the state of the medium. The physical carrier sense resides in the PHY and uses energy detect and preamble detect with frame length deferral to determine when the medium is busy. The virtual carrier sense resides in the MAC and uses reservation information carried in the Duration field of the MAC headers announcing impeding use of the medium. The virtual carrier sense mechanism is called the network allocation vector (NAV). The medium is determined to be idle only when both the physical and virtual carrier sense mechanisms indicate it to be so.

The DCF also makes use of the immediate feedback provided by the basic acknowledgement mechanism that has the responder send an ACK frame in response to the initiator's data or management frame. Not receiving the ACK response frame is a likely indication that the initiator's transmission was not correctly received, either due to collision or poor channel conditions at the time of the data transmission.

To further minimize the chance of collisions, and as a more robust collision detect mechanism, the initiating station may begin a sequence with a short control frame exchange using robustly modulated RTS and CTS frames. This sets the NAV in the stations surrounding both the initiator and responder, some of which may be hidden nodes unable to detect the more remote station's transmissions and thus only able to defer for frame transmissions from nearby nodes.

The DCF provides a distributed contention-based channel access function. Stations compete for channel access without the need for a central coordinator or arbiter. This mechanism is remarkable efficient and fairly apportions bandwidth among the active stations.

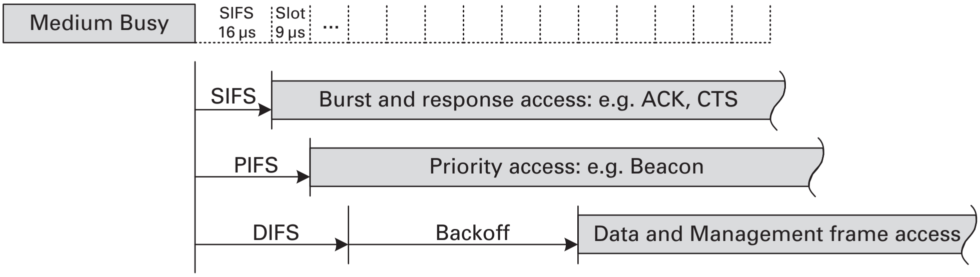
Medium Busy
SIFS 16 µs
Slot 9 µs
...
SIFS
Burst and response access: e.g. ACK, CTS
PIFS
Priority access: e.g. Beacon
DIFS
Backoff
Data and Management frame access

Figure 8.4 Basic channel access priorities with associated timing.

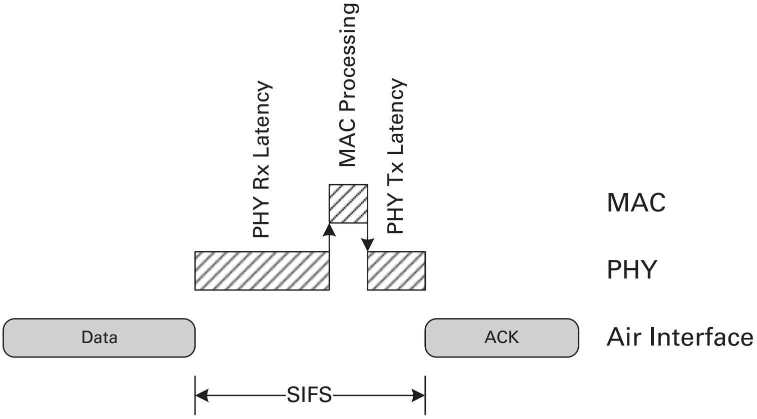

Figure 8.5 PHY and MAC latencies generating a response frame.

8.3.1 Basic channel access timing

Basic channel access timing from the original 802.11 specification is illustrated in Figure 8.4. The different inter-frame space (IFS) durations effectively provide access to the wireless medium at different priority levels.

8.3.1.1 SIFS

The short inter-frame space (SIFS) is used to separate a response frame from the frame that solicited the response, for example between a data frame and the ACK response. SIFS is designed to be as short as possible but still accommodate the latencies incurred in a reasonable implementation. These latencies include the decode latency in the PHY for demodulating the received frame, the MAC processing time for the received frame and building the response, and the transmitter startup time to send the response (Figure 8.5).

SIFS is also used to separate individual frames in a back-to-back data burst. Stations accessing the medium using SIFS timing do not check if the medium is busy, but simply switch to transmit mode (if not already in transmit mode) and begin transmission at the SIFS boundary.

The SIFS duration for a particular PHY is defined by the aSIFSTime parameter. For the 802.11a, 802.11g, 802.11n, and 802.11ac PHYs the value is 16 µs.

8.3.1.2 Slot time

Timing for the other IFS durations is SIFS plus an integral number of slot times with transmission beginning on the slot boundary. In practice, propagation delays and, to a small extent, implementation inaccuracies mean that each station sees a slightly different boundary. The slot duration is designed to accommodate this variability and provide enough time for a transmitting station's preamble to be detected by neighboring stations before the next slot boundary. During each slot time, stations not yet transmitting remain in receive mode and check that the medium remains idle.

The slot time for a particular PHY is defined by the aSlotTime parameter. For the 802.11a, 802.11g, 802.11n, and 802.11ac PHYs the value is 9 µs.

8.3.1.3 PIFS

The PCF inter-frame space (PIFS) defer provides the next highest access priority following SIFS and is used to gain priority access to the medium. PIFS is defined by the following equation:

$$\mathrm{PIFS} = \mathrm{aSIFSTime} + \mathrm{aSlotTime} \tag{8.1}$$

The AP uses the PIFS defer to gain access to the medium to send a Beacon, start a contention free period, or to regain access to the medium if an expected response frame is not received during a contention free period. Despite its name, which reflects its original use for the point coordination function (see Section 9.1), PIFS is now also used for other priority operations, such as by a station that needs to send a Channel Switch Announcement frame (802.11h).

8.3.1.4 DIFS

The DCF inter-frame space (DIFS) is used by stations operating under the DCF to transmit data frames and management frames and is defined by the following equation:

$$\mathrm{DIFS} = \mathrm{aSIFSTime} + 2 \times \mathrm{aSlotTime} \tag{8.2}$$

A station using the DCF is allowed to transmit if it determines that the medium is idle for the duration of the DIFS, or if it determines that the medium is idle for the duration of the DIFS plus the remaining backoff time following the reception of a correctly received frame.

8.3.1.5 Random backoff time

When the medium transitions from busy to idle, multiple stations may be ready to send data. To minimize collisions, stations wishing to transmit select a random backoff count and defer for that number of slot times. The random backoff count is selected as a pseudo-random integer drawn from a uniform distribution over the interval [0, CW], where CW, an integer value, is the contention window.

The contention window (CW) parameter takes the initial value CWmin and effectively doubles on each unsuccessful MPDU transmit, for example each time an ACK response is not received for a data frame. If the CW reaches CWmax it remains at that value until it is reset. The CW is reset to CWmin after every successful MPDU transmit.

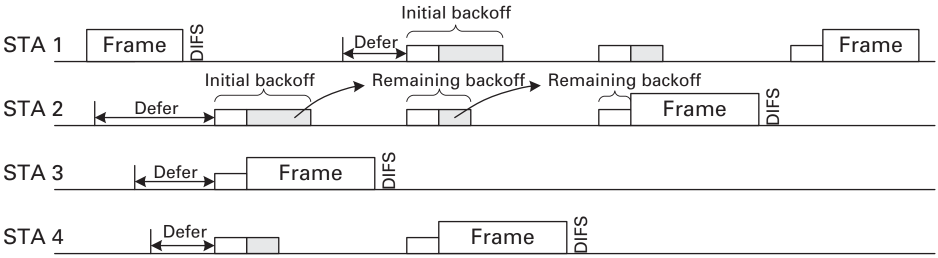

Figure 8.6 Backoff procedure.

CW, CWmin, and CWmax may take values that are a power of 2 less 1. For the DCF, CWmin and CWmax are specified according to the particular PHY used. For 802.11a, 802.11g, 802.11n, and 802.11ac PHYs, CWmin is 15 and CWmax is 1023. CW would thus start with the value 15 and when "doubled" take on the next higher power of 2 less 1 until it reaches 1023, i.e. 15, 31, 63,. . ., 1023. The CW was defined this way for easy implementation: using binary notation, "doubling" is effectively a left shift operation with a lower order one inserted and the backoff value is obtained by using CW to mask a full word random number.

8.3.1.6 Random backoff procedure

To begin the random backoff procedure, the station selects a random backoff count in the range [0, CW]. All backoff slots occur following a DIFS during which the medium is determined to be idle. During each backoff slot the station continues to monitor the medium. If the medium goes busy during a backoff slot then the backoff procedure is suspended. The backoff count is resumed when the medium goes idle again for a DIFS period.

The effect of this procedure is illustrated in Figure 8.6. When multiple stations are deferring and go into random backoff, then the station selecting the smallest backoff count (STA 3) will win the contention and transmit first. The remaining stations suspend their backoff and resume DIFS after the medium goes idle again. The station with the next smallest backoff count will win next (STA 4) and then eventually the station with the longest backoff count (STA 2). A station that begins a new access (STA 1 again) will select a random backoff from the full contention window and will thus tend to select a larger count than the remaining backoff for stations (such as STA 2) that have already suspended their backoff from a previous access attempt.

8.4 Data/ACK frame exchange

Transmission over a wireless medium is error prone. Data transfer benefits from a low latency, link level repeat mechanism that allows for the retransmission of frames that have not been successfully demodulated at the receiver. The basic mechanism by which this is achieved is to have the station that correctly receives a data frame addressed to it send an immediate, positive acknowledgement in the form of an ACK frame. If the

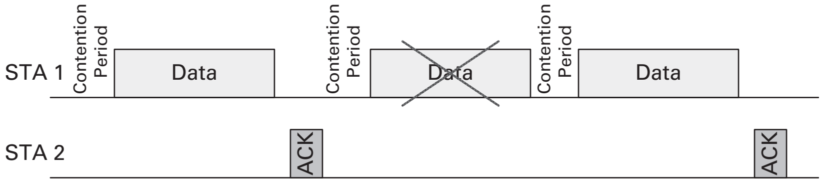

Figure 8.7 Basic data/ACK frame exchange sequence.

station sending the data frame does not receive the ACK frame, it assumes the data frame was not received and may retransmit it.

Not all data frames can be acknowledged in this way. Broadcast and multicast data frames are directed to all or a subset of the stations in a WLAN and cannot be acknowledged this way. In an 802.11 WLAN, broadcast and multicast frames do not benefit from the additional reliability that the acknowledgement mechanism provides.

Data transfer using the Data/ACK exchange is illustrated in Figure 8.7. Here STA 1 is transferring data to STA 2. STA 1 accesses the medium after a contention period during which it defers for DIFS followed by a random backoff period. If the medium remains idle, STA 1 transmits a data frame addressed to STA 2. If STA 2 detects and correctly demodulates the frame then it responds with an ACK. When STA 1 receives the ACK it knows that the frame was correctly received and begins channel access again in order to transmit the next frame. If, as with the second data frame in the figure, STA 2 fails to successfully demodulate the frame then STA 1 will not receive an ACK and will then begin channel access again to retransmit the data frame.

The number of retransmission attempts on a particular MSDU is limited. The transmitting station maintains a count of the number of retransmission attempts on an MSDU and when that count exceeds a configured retry limit the MSDU is discarded.

To enhance the reliability with which acknowledgement feedback is provided, the ACK frame is modulated robustly, i.e. it is sent using a lower PHY data rate than data frames sent to the same station. The additional overhead incurred with robust modulation is relatively small since the ACK frame itself is very short.

8.4.1 Fragmentation

Fragmentation is used to break up large MSDUs to improve the chance that the MSDU will be received correctly and to reduce the overhead of retransmission. At low data rates, an unfragmented MSDU can occupy a large amount of air time. For example, a 1500 byte data frame sent using the 1 Mbps 802.11b rate takes 12 ms to transmit, making it susceptible to changing channel conditions. A bit error in the frame would result in the entire frame being retransmitted. With fragmentation the MSDU would be broken into smaller segments and each segment encapsulated in an MPDU. Each MPDU is sent in a separate PPDU with the preamble of each PPDU providing a new channel estimate. A bit error would result in only the MPDU carrying the errored segment being retransmitted.

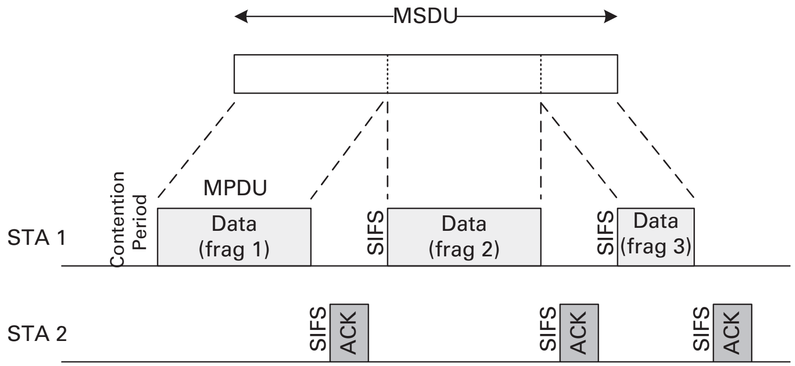

Figure 8.8 Fragment burst.

The fragments making up an MSDU are sent as individual MPDUs. A station may send each fragment in a separate channel access, or the fragmented MSDU may be sent as a burst of data MPDUs following a single channel access, as illustrated in Figure 8.8.

An MSDU is fragmented when its length exceeds a threshold specified by the dot11FragmentationThreshold attribute. Each fragment contains an even number of bytes and all fragments are the same size, except the last fragment, which may be smaller. The fragments are delivered in sequence.

The transmitter of a fragmented MSDU maintains a timer. The attribute dot11MaxTransmitMSDULifetime specifies the maximum amount of time allowed to transmit a MSDU. The source station starts the timer on the first attempt to transmit the first fragment of the MSDU. If the timer exceeds dot11MaxTransmitMSDULifetime then all remaining fragments are discarded.

The receiver of a fragmented MSDU also maintains a timer. The attribute aMaxReceiveLifetime specifies the maximum amount of time allowed to receive an MSDU. The receive MSDU timer starts on reception of the first fragment of the MSDU. If the timer exceeds aMaxReceiveLifetime then all the fragments of the MSDU are discarded. Additional fragments which may be received later are also discarded.

8.4.2 Duplicate detection

With retransmission there is the possibility that a frame that was correctly received may be received again, for example if the transmitter retransmits a frame because the ACK response itself was not correctly demodulated. To detect duplicate frames the Data frame includes a Retry bit and a Sequence Control field consisting of a sequence number and a fragment number. The Retry bit is set on any frame that is retransmitted. The sequence number is generated as an incrementing sequence of integers assigned to MSDUs and management frames. If the MSDU or management frame is fragmented then each fragment receives the same sequence number with an incrementing fragment number.

To detect duplicate frames, the receiving station keeps track of the sequence number and fragment numbers of the last MSDU or management frame that it received from each station communicating with it. In other words, it maintains a cache of <transmit address,

sequence number, fragment number> tuples for each fragment received for the last sequence number seen from each transmit address other than broadcast or multicast addresses. If the station receives an MPDU with the Retry bit set that matches an entry in the cache then it rejects the MPDU as a duplicate.

8.4.3 Data/ACK sequence overhead and fairness

The basic Data/ACK frame exchange has a fixed overhead associated with it. Since most data frames are successfully transmitted, this overhead includes the contention period during which the medium is essentially idle, the overhead associated with transmitting the data frame itself, the radio turnaround time at the receiver (SIFS), and the transmission of the ACK frame. While this overhead is essentially fixed, the duration of the data frame is not fixed since it depends on the modulation and coding scheme (data rate) used by the PHY. The higher the data rate the shorter the data frame duration and the greater the fixed overhead relative to the overall duration of the transfer.

The distributed channel access mechanism promotes fairness in the sense that all stations on the network with data to send will, on average, each send the same number of data frames. If they are all using the same packet size they will see the same throughput irrespective of their individual PHY data rates. So, for example, suppose there are two stations on the network, STA 1 and STA 2, both competing for access to send data to STA 3 (Figure 8.9). Suppose also that STA 1 is using a high data rate while STA 2 is using a low data rate. Each station competes for channel access to send one data frame and each station will on average get the same number of transmit opportunities. However, because STA 2 is using a lower data rate it will use proportionately more air time to send its data frames than STA 1.

8.5 Hidden node problem

The distributed nature of channel access in 802.11 WLANs makes the carrier sense mechanism critical for collision free operation. The physical carrier sense, which logically resides in the PHY, is responsible for detecting the transmissions of other stations. However, in some situations it may not be possible for the physical carrier sense to detect

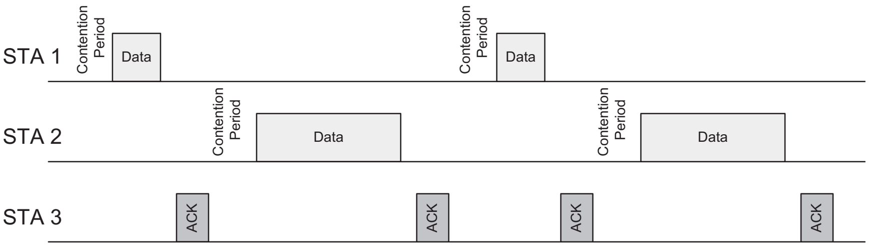

Figure 8.9 Two stations competing for access.

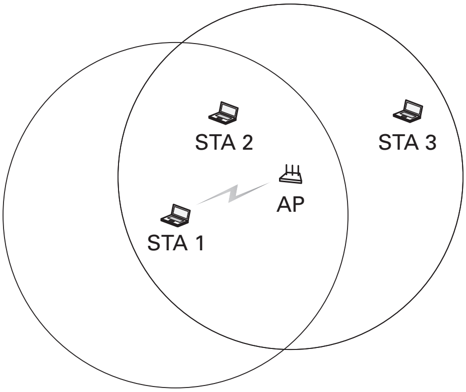

Figure 8.10 Hidden node problem.

the transmissions of all stations. Consider the situation in Figure 8.10 where there is data transfer between STA 1 and the AP. Transmissions from STA 1 can be detected by the AP and STA 2. A distant node, STA 3, can detect transmissions from the AP but not from STA 1. STA 3 is a hidden node with respect to communication between STA 1 and the AP. When STA 1 transmits a frame to the AP there is a chance that STA 3 would still see the medium as idle and also begin a frame transmission.

8.5.1 Network allocation vector (NAV)

One mechanism defined to overcome the hidden node problem is the network allocation vector (NAV). The NAV is a function that logically resides in the MAC and provides a virtual carrier sense to augment the physical carrier sense. Each MAC frame carries a Duration field that is used to update the NAV in any station other than the addressed station that successfully demodulates the frame. The Duration field holds a time value that indicates the duration for which the sending station expects the medium to be busy referenced from the end of the last symbol of the PPDU carrying the MAC frame.

All frames[2] include the Duration field and may set the NAV in neighboring stations. However, to do so the frame must be successfully demodulated by the neighboring stations. The NAV is most effectively set in neighboring stations using robustly modulated control frames, such as the RTS/CTS exchange, rather than data frames.

8.5.1.1 RTS/CTS frame exchange

To protect a station's transmissions from hidden nodes, a station may begin a sequence with an RTS/CTS exchange as illustrated in Figure 8.11. The RTS (request to send) is sent by the initiator (STA 1) and the station addressed by the RTS (STA 2) responds with a CTS (clear to send). The RTS frame occupies less air time than the data frame and is thus less susceptible to collision than the longer data frame transmitted alone. Also, loss

[2] Except the PS-Poll frame, which uses this field for other purposes.

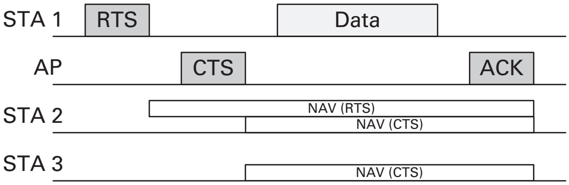

Figure 8.11 RTS/CTS exchange for hidden node protection.

of the RTS to collision would be quickly detected. The RTS and CTS are robustly modulating so that they are broadly received.

The Duration field of the RTS frame carries a NAV setting to cover the CTS response plus the time needed for the subsequent frame exchange. The CTS response has its Duration field set to the Duration field value seen in the RTS less SIFS and the duration of the CTS response itself. In the diagram, the hidden node (STA 3) would receive the CTS frame and set its NAV to defer for the subsequent frame exchange. STA 2 sees both the RTS and CTS.

The RTS/CTS exchange is required when the length of a data or management frame exceeds the threshold set by the dot11RTSThreshold attribute. The dot11RTSThreshold is a local management attribute and may be set to 0 so that all MPDUs are delivered with an RTS/CTS exchange, to the maximum allowed MPDU length so that the RTS/CTS need not be used at all, or any value in between.

8.5.2 EIFS

Another mechanism used to protect against hidden nodes is the extended inter-frame space (EIFS). A station uses EIFS instead of DIFS to defer if a frame is detected but not correctly received, i.e. the MAC determines that the frame check sequence (FCS) is invalid. EIFS is defined as:

$$\text{EIFS} = \text{aSIFSTime} + \text{ACKTxTime} + \text{DIFS} \tag{8.3}$$

where ACKTxTime is the time required to transmit an ACK frame at the lowest mandatory PHY data rate. EIFS is intended to prevent a station from transmitting over the ACK of a hidden node when a station is unable to demodulate the data frame and thus correctly set its NAV. If during the EIFS defer a valid frame is received (for example, the ACK) then a DIFS defer is used following the actual frame instead of continuing with EIFS. EIFS usage is illustrated in Figure 8.12.

8.6 Enhanced distributed channel access

Enhanced distributed channel access (EDCA) is an extension of the basic DCF introduced in the 802.11e amendment to support prioritized quality of service (QoS). The EDCA mechanism defines four access categories (ACs). Each AC is characterized

Table 8.1 AC relative priorities and mapping from 802.1D user priorities

Priority	802.ID user priority	802.ID Designation	AC	Designation
Lowest	1	BK	AC_BK	Background
	2	–		
	0	BE	AC_BE	Best effort
	3	EE		
	4	CL	AC_VI	Video
	5	VI		
Highest	6	VO	AC_VO	Voice
	7	NC		

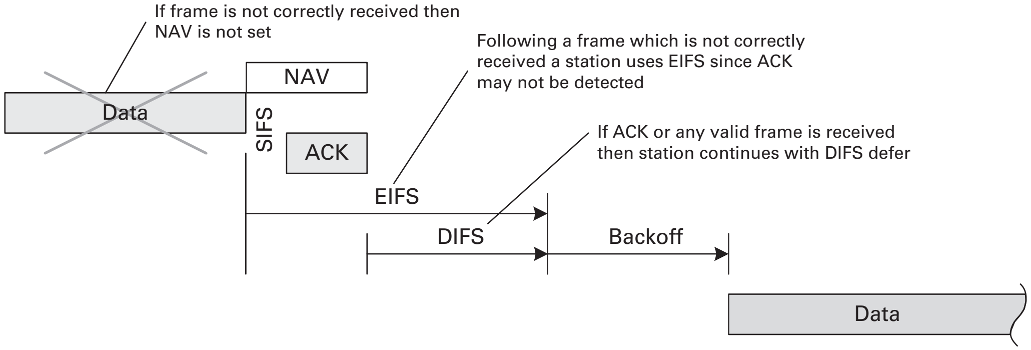
If frame is not correctly received then NAV is not set
Data
NAV
SIFS
ACK
Following a frame which is not correctly received a station uses EIFS since ACK may not be detected
If ACK or any valid frame is received then station continues with DIFS defer
EIFS
DIFS
Backoff
Data

Figure 8.12 EIFS usage.

by specific values for a set of access parameters that statistically prioritize channel access for one AC over another. The relative access priorities of the four ACs and the mapping of 802.1D (MAC bridging) user priorities to ACs are give in Table 8.1. An MSDU with a particular user priority is said to belong to a traffic category (TC) with that user priority.

Under EDCA, egress traffic (traffic leaving the system) is sorted logically into four queues, one for each AC (Figure 8.13). An instance of the EDCA access function operates for each queue contending for access with that AC's access parameters when the queue is non-empty. The EDCA access functions, like DCF, compete for access to the medium by deferring for a fixed period, the arbitration inter-frame space (AIFS), when the medium goes idle and then for a random backoff period. The parameters for EDCA access are similar to the parameters that are used for the DCF, but defined per AC. The AIFS value for each AC is referenced as AIFS [AC]. The contention window from which the random backoff count is selected is referenced as CW[AC].

The contention window for a particular AC, CW[AC], starts with the value CWmin[AC]. If a frame transmission for a particular AC is not successful CW[AC] is effectively doubled as described in Section 8.3.1.5. If CW[AC] reaches CWmax[AC] it remains at that value until it is reset. CW[AC] is reset to CWmin[AC] after a successful MPDU transmit in that AC.

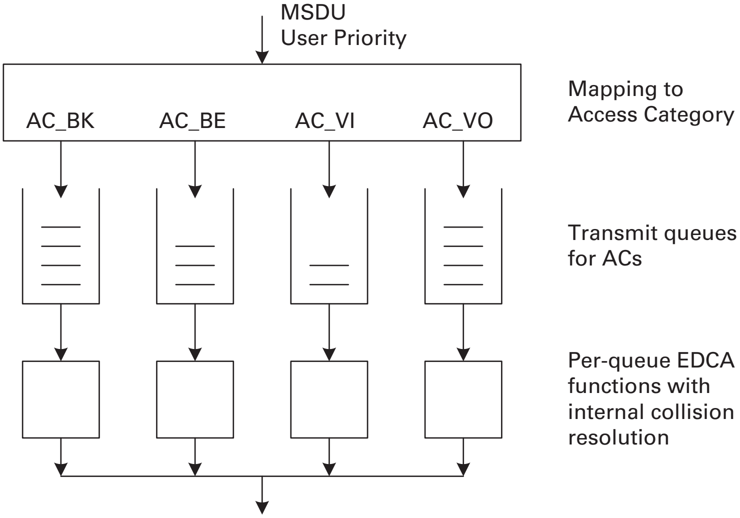
MSDU
User Priority
AC_BK
AC_BE
AC_VI
AC_VO
Mapping to Access Category
Transmit queues for ACs
Per-queue EDCA functions with internal collision resolution

Figure 8.13 EDCA reference implementation. Reproduced with permission from IEEE (2012). IEEE Std 802.11–2012, © 2012 IEEE. All rights reserved.

If two (or more) instances of the EDCA access function gain access simultaneously, the internal collision is resolved by the highest priority AC gaining access and the other AC behaving as if an external collision occurred by doubling its contention window and re-arming for another access attempt.

8.6.1 Transmit opportunity

A key concept introduced in the 802.11e amendment is the transmit opportunity (TXOP). A TXOP is a bounded period during which a station may transfer data of a particular traffic class. Under EDCA, a TXOP is obtained by the station through the channel access procedure using access parameters for the particular class of traffic for which the TXOP will be used. Once the TXOP has been obtained, the station may continue to transmit data, control, and management frames and receive response frames, provided the frame sequence duration does not exceed the TXOP limit set for that AC. A TXOP limit of zero means that only one MSDU or management frame can be transmitted before competing again for access.

The TXOP concept promotes resource fairness rather than throughput fairness in that all stations accessing the network with traffic of the same class will on average receive the same amount of air time. Suppose two stations are competing for access, one with a high PHY data rate and the other with a low PHY data rate as shown in Figure 8.14. Both stations will on average receive the same amount of air time, but the station with the higher PHY data rate will see higher throughput than the station with the lower PHY data rate. Contrast this with the situation in Figure 8.9 where both stations see the same effective throughput but with unequal resource utilization.

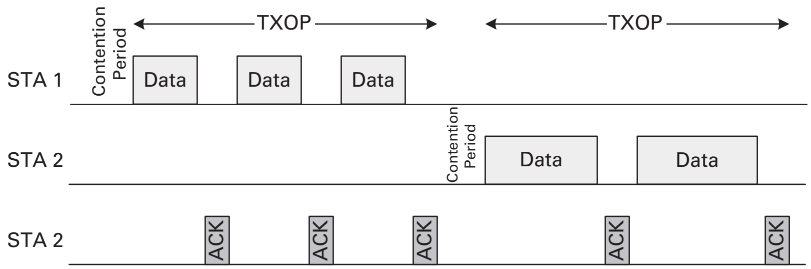

Figure 8.14 TXOP usage with different PHY data rates.

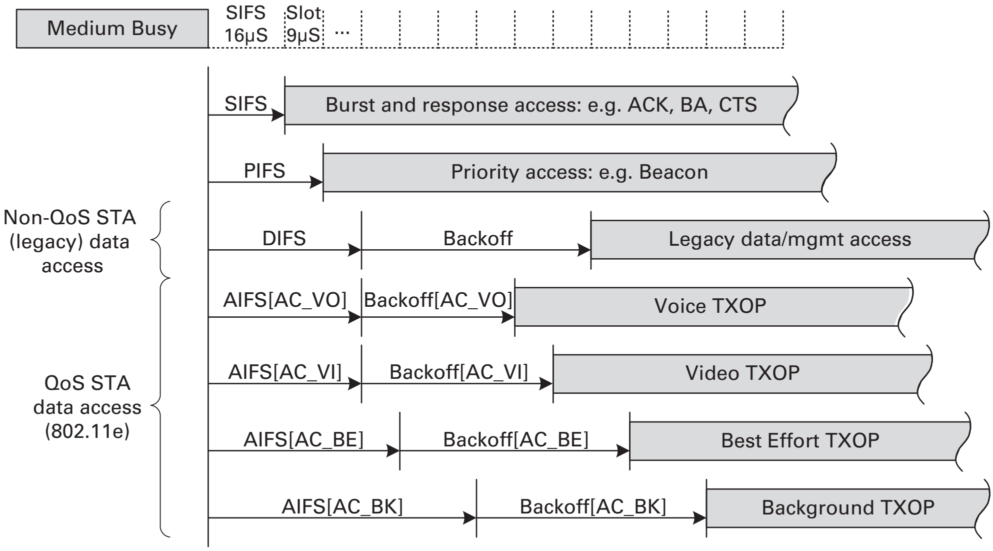

Figure 8.15 Channel access priorities including EDCA with associated timing.

8.6.2 Channel access timing with EDCA

Extending Figure 8.4 to include EDCA access timing we arrive at Figure 8.15. This figure shows the access priorities for the four ACs in relation to each other and to the DCF.

The arbitration inter-frame space (AIFS) for a particular AC is defined by the equation

$$\text{AIFS[AC]} = \text{aSIFSTime} + \text{AIFSN[AC]} \times \text{aSlotTime} \tag{8.4}$$

where AIFSN[AC] is the slot count.

8.6.3 EDCA access parameters

The EDCA access parameters are provided in the EDCA Parameter Set information element that is present in Beacon and Probe Response frames. Stations on the BSS use the last seen version of the parameters and the AP may adjust the parameters over time, for example based on network load or the number of associated stations.

Table 8.2 Default EDCA access parameters for 802.11a, 802.11g, and 802.11n PHYs

AC	CWmin	CWmax	AIFSN	TXOP limit
AC_BK	31	1023	7	0
AC_BE	31	1023	3	0
AC_VI	15	31	2	3.008 ms
AC_VO	7	15	2	1.504 ms
legacy	15	1023	2	0

The default EDCA access parameters for the 802.11a, 802.11b, and 802.11n PHYs are given in Table 8.2. The default EDCA parameters are used if the AP does not broadcast a different set of parameters. While not an access category, the table also shows for comparison the equivalent parameters for the DCF (labeled legacy).

The EDCA access parameters determine the degree by which one AC is prioritized over another. The AIFSN parameter provides weak differentiation. ACs with lower values gain access more frequently than ACs with higher values, other parameters being equal. The CWmin parameter provides much stronger differentiation. The random backoff count is selected from the range [0, CW], where CW is typically equal to CWmin. Increasing the range from which the backoff count is selected has a bigger effect on the overall defer period than differences in AIFSN. The TXOP limit is also a strong differentiator. An AC with a large TXOP limit will receive more air time than an AC with a small TXOP limit assuming equal allocation of TXOPs.

8.6.4 EIFS revisited

Under DCF, a station must defer for EIFS instead of DIFS following a frame that is detected but not successfully demodulated. The intent is to protect a possible ACK response to the unsuccessfully demodulated frame from a more distant node that may not be detected by the station. To effect equivalent behavior under EDCA, a station must defer for EIFS − DIFS + AIFS[AC] following a frame that is detected but not successfully demodulated.

The EIFS and EIFS − DIFS + AIFS[AC] defer are a convoluted way of saying that following an unsuccessfully demodulated frame a station must defer for SIFS plus the duration of an ACK frame before performing the usual DIFS or AIFS defer.

8.6.5 Collision detect

When a station obtains a TXOP it may transmit for the duration of the TXOP, either as a single transmission or as a burst of back-to-back data frames. There is, however, a chance that two stations will obtain channel access simultaneously and the resulting colliding transmissions are likely to be unintelligible to the receiving peers.

To minimize the loss of air time due to these collisions, stations must perform a short frame exchange at the beginning of the TXOP to detect a collision. The short frame exchange may be either an RTS/CTS exchange or a short single Data/ACK exchange.

Collision detect is also necessary for the correct operation of the backoff algorithm, which must have its contention window size doubled should a collision occur. Doubling the contention window reduces the chance that a collision will occur the next time the two stations attempt channel access.

The short frame exchange performed at the beginning of the TXOP also allows the two stations involved in the frame exchange to set the NAV of their neighboring stations.

8.6.6 QoS Data frame

To support the QoS features and block acknowledgement (discussed below), the 802.11e amendment introduced a new data frame, the QoS Data frame. The QoS Data frame has the same fields as the regular Data frame, but includes an additional QoS Control field (see Section 12.1.5). The QoS Control field carries various subfields for managing QoS and other features introduced in the amendment. The TID or traffic identifier identifies the TC to which the frame belongs. Under EDCA, the TID field carries the user priority, which is mapped to the AC through Table 8.1.

The Ack Policy subfield determines how the data frame is acknowledged by the receiving peer and carries one of the following values:

- **Normal Ack** – if correctly received, the recipient responds to the QoS Data frame with an ACK response.
- **No Ack** – the recipient does not respond to the QoS Data frame. This may be useful for traffic that has a low tolerance for jitter or delay and does not benefit from retransmission.
- **No Explicit Ack** – there may be a response frame, but it is not an ACK. This policy is used when polling under centrally coordinated channel access (see Section 10.2).
- **Block Ack** – the recipient takes no action on the received frame except to record its reception. This policy is used under the block acknowledgement protocol.

8.7 Block acknowledgement

The block acknowledgement protocol, introduced with the 802.11e amendment, improves efficiency by allowing for the transfer of a block of data frames that are acknowledged with a single Block Acknowledgement (BA) frame instead of an ACK frame for each of the individual data frames. Unlike the normal acknowledgement mechanism, however, the block acknowledgement mechanism is session oriented and a station must establish a block acknowledgement session with its peer station for each traffic identifier (TID) for which block data transfer is to take place. A particular block acknowledgement session is thus identified by the <transmit address, receive address, TID> tuple.

The 802.11e amendment introduced two flavors of the block acknowledgement protocol: immediate block ack and delayed block ack. The two flavors differ in the manner in which the block acknowledgement control frames are exchanged. Under immediate block ack the Block Ack Request (BAR) frame solicits an immediate Block Ack (BA) frame response, i.e. the BA is returned within SIFS of receiving the BAR and thus within the same TXOP. With delayed block ack the BAR is sent in one TXOP and the BA response is returned in a separate, subsequent TXOP. Immediate block ack provides lower latency and improved performance over delayed block ack, which was primarily defined for ease of implementation.

The block acknowledgement protocol is described in more detail in Section 9.3, but a brief overview is provided here to illustrate the basic concept. Block acknowledgement is enabled in one direction for a particular TID with the exchange of an ADDBA Request and ADDBA Response. The station with data to send – the originator – sends an ADDBA Request to the station that will receive the data – the recipient. The recipient acknowledges the correct receipt of the ADDBA Request with an ACK and then responds some time later with an ADDBA Response to which the originator responds with an ACK. The ADDBA exchange allows the originator and recipient to exchange parameters such as the recipient reorder buffer size. To tear down a block ack session, the originator or the recipient sends a DELBA request which, if correctly received, is acknowledged with an ACK.

Block data transfer occurs as follows. The originator transmits one or more QoS Data frames addressed to the recipient and containing the TID of the block acknowledgement session. The Ack Policy field is set to Block Ack. The block of data frames need not be transmitted in sequence and may include retransmitted frames. The recipient is responsible for reordering the frames and delivering them in sequence to the higher layer and performs this function using a reorder buffer. The recipient will hold frames in the reorder buffer until gaps in the sequence number space are filled. The originator limits the sequence number range for which acknowledgements are outstanding so as not to overrun the recipient reorder buffer.

After sending a block of data frames, the originator sends a Block Ack Request (BAR) frame. The BAR frame performs two functions; it flushes the recipient's reorder buffer and it solicits a Block Ack (BA) frame. The recipient's reorder buffer may need to be flushed to advance passed holes in the sequence number space resulting from MSDUs that did not make it through after exhausting their retransmission count or lifetime limit. Flushing the reorder buffer releases MSDUs that may be held up behind these holes.

The BAR frame includes a Starting Sequence Control field that contains the sequence number of the oldest MSDU in the block for which an acknowledgement is expected. MSDUs in the recipient's buffer with sequence numbers that precede the starting sequence number are either forwarded to the LLC layer (if complete) or discarded (if one or more fragments are missing). The solicited BA frame contains a bitmap that represents the acknowledgement state of the data frames received beginning with the starting sequence number from the BAR frame.

On receiving the BA frame the originator discards acknowledged data frames and requeues data frames not acknowledged for retransmission. The originator may also discard data frames that have reached their retransmission count or lifetime limit.

With a block ack session in place, the originator may still solicit a regular ACK for QoS Data frames by setting the Ack Policy field to Normal Ack.

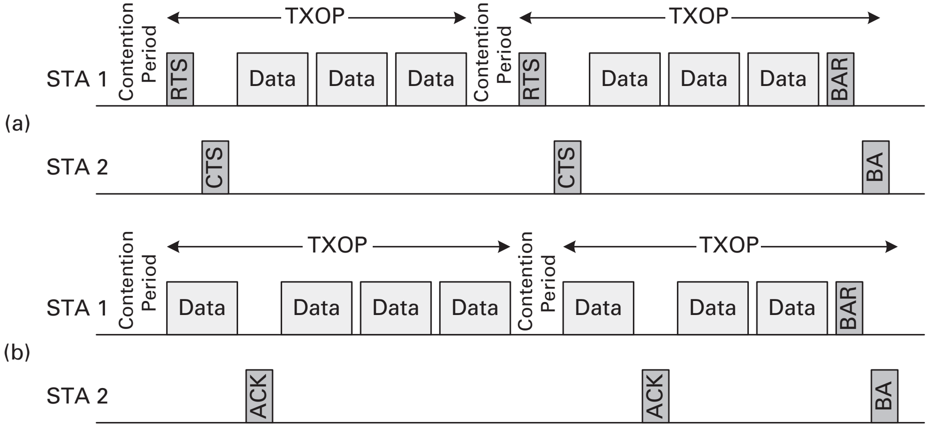

Figure 8.16 Block data frame exchange sequence.

8.7.1 Block data frame exchange

The block data frame exchange using the immediate block ack protocol is illustrated in Figure 8.16(a) with STA 1 transferring data to STA 2. After a contention period, STA 1 gains a TXOP. As a collision detect mechanism and to set the NAV in neighboring stations, STA 1 performs a short frame exchange, in this case RTS/CTS. STA 1 then sends a back-to-back burst of data frames with SIFS separating the individual transmissions until the TXOP limit is reached. Since it has more data to send, STA 1 again accesses the wireless medium and gains a TXOP. An RTS/CTS exchange is again performed followed by the remaining frames in the block sent as a back-to-back burst. STA 1 then sends a BAR frame, which solicits a BA response from STA 2. The BA response indicates which of the data frames in the block were correctly received.

As an alternative to performing the RTS/CTS exchange, STA 1 may use a Data/ACK exchange with STA 2 to effect collision detect, as illustrated in Figure 8.16(b). Collision detect is necessary through one of these mechanisms as it reduces the degradation in network throughput that would otherwise occur if two stations collided for the full duration of the TXOP. The Data/ACK exchange provides more limited protection near the transmitter (due to the higher order modulation used for data frames), but is more efficient than the RTS/CTS exchange with which no information transfer takes place.

It should be noted that the block transfer is independent of the TXOP. The block transfer may occur over multiple TXOPs or it may be contained within a single TXOP.

8.8 Power management

The increased use of 802.11 radios in low power devices (such as mobile handsets) means that power management has become a critical feature for 802.11 operation. The original 802.11 specification included support for power management in client devices, as described below. Further power management features have been added with various amendments, the most important of these are described later in this section.

The 802.11 radio is typically used sporadically. An effective power saving technique is thus to turn the radio off when not actively communicating, which is most of the time, and turn it on periodically to check for buffered traffic at the AP or to transmit traffic to the AP. A station operating in this mode is said be in Power Save (PS) mode. A station that is always available to receive traffic (i.e., does not turn its radio off) is said to be in Active mode. A station can transition between these two modes of operation and declares its current mode using the Power Management field in the Frame Control field of the MPDUs it sends. A station indicates a change in its Power Management mode (PS mode to Active mode or vice-versa) by sending an MPDU addressed to the AP that solicits an acknowledgement. The acknowledgement ensures that the AP correctly received the mode change notification.

A station operating in PS mode can be in one of two states. The station is in the Awake state when its radio is active and it is able to transmit and receive frames. The station is in the Doze state when its radio is not active and it is not able to transmit or receive frames. A station operating in Active mode is always in the Awake state.

The original 802.11 specification assumed that the AP was mains powered and thus always operates in Active mode. As such, it is required to set the Power Management field to 0 in all the frames it transmits.

An AP buffers traffic addressed to stations in the PS mode. In addition, if any station on the BSS is in PS mode, the AP will deliver broadcast traffic at predictable intervals so that stations in PS mode can wake during those periods to receive the traffic.

8.8.1 AP TIM transmissions

An AP indicates that it has buffered traffic for a station in PS mode using the traffic indication map (TIM). The TIM is a partial virtual bitmap where each bit represents a station on the BSS. As STAs join the BSS they are assigned an association ID (AID), an integer between 1 and 2007. The bit position in the TIM indexed by the AID indicates whether or not traffic is buffered for the corresponding station. The first bit in the TIM (indexed by AID 0) represents the presence of buffered group addressed (broadcast and multicast) traffic. AID 0 is thus never assigned to a station joining the BSS.

The AP transmits a TIM in every Beacon frame. In every *n*th Beacon frame, the TIM is transmitted as a delivery traffic indication map (DTIM). In format, a DTIM is simply a TIM with 0 in the DTIM Count field. A TIM that is not a DTIM has a DTIM Count that indicates the number of beacons, including the current beacon, until the next DTIM.

The presence of a DTIM in a Beacon frame signifies that group addressed traffic will be delivered immediately following the Beacon frame. The interval between beacons carrying DTIM is the DTIM interval.

8.8.2 PS mode operation

A station operating in PS mode wakes periodically to receive Beacon frames. The standard does not define which Beacon frames the station should receive as the heuristics

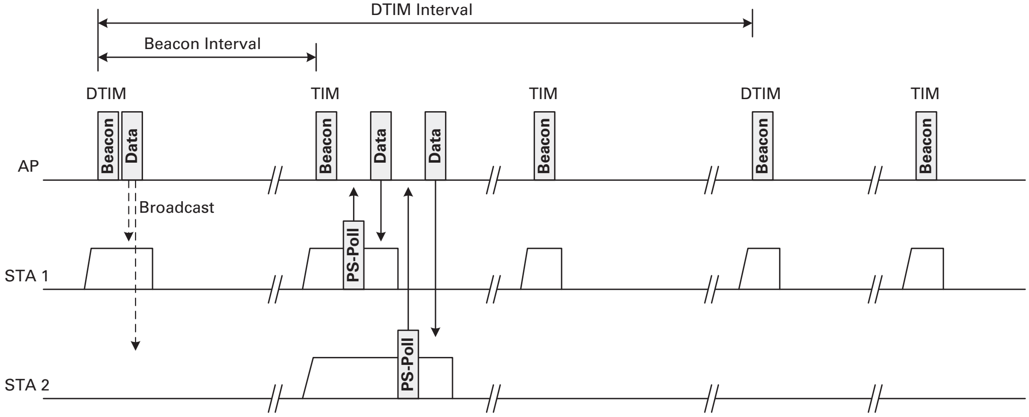

Figure 8.17 An example of AP and STA power management activity.

used are dependent on tradeoffs between latency, whether or not the station needs to receive group addressed traffic and the level of power save the station needs to achieve. The station must, however, periodically interact with the AP to maintain its association.

To receive group addressed traffic, a station needs to wake for each DTIM. In practice, devices can tolerate some loss of group addressed traffic and thus improve power saving by not waking for all DTIMs. Bit 0 of the TIMs leading up to a DTIM and the DTIM itself indicate whether or not any group addressed traffic is available. Buffered group addressed traffic is delivered following the Beacon frame carrying the DTIM. Each buffered group addressed Data frame is transmitted with the More Data field set to 1 except the last, which has the More Data field set to 0.

To receive individually addressed traffic, a station periodically wakes to receive a Beacon frame. The bit in the TIM indexed by the station's AID indicates whether or not traffic addressed to that station is buffered at the AP. To retrieve buffered traffic, the station sends a PS-Poll frame to the AP. The AP responds immediately to the PS-Poll either with a buffered data frame or an ACK followed in a separate TXOP by the buffered data frame at the head of the transmit queue. If there are additional buffered data frames, the AP indicates this by setting the More Data field in the delivered data frame to 1. The station acknowledges successful receipt of the data frame. The station continues to poll the AP until it receives a data frame with the More Data field set to 0.

An example of AP and station power management activity is illustrated in Figure 8.17. The first Beacon frame carries a DTIM, identified as such by the DTIM Count field in the TIM information element being 0. Bit 0 of the partial virtual bitmap is 1, indicating that group addressed traffic is available. STA 1 is awake to receive this Beacon frame and stays awake following the beacon to receive the group addressed traffic. The second beacon carries a TIM with a DTIM Count of 2 indicating that there are two beacons (including this one) before the next DTIM. The TIM also indicates that individually addressed traffic is buffered for both STA 1 and STA 2. Since both STA 1 and STA 2 are awake to receive this beacon, they see the indication and poll the AP for their traffic. Note that each station acknowledges the data frame they receive, although this is not shown in the diagram.

In this example, STA 1 wakes for each beacon while STA 2 does not, conserving more power. STA 2 misses the group addressed traffic and may experience higher latency on its individually addressed traffic.

8.8.3 WNM-Sleep

WNM-Sleep was introduced with the 802.11u amendment, which dealt with wireless network management. This feature has little to do with wireless network management, nevertheless it is saddled with that name. WNM-Sleep allows a STA to miss DTIMs without missing the associated group addressed traffic.

To support a STA operating in WNM-Sleep mode, an AP converts certain group addressed frames to individually addressed frames and delivers them to the STA using the normal PS mode operation. The STA indicates which group addressed frames are to be delivered this way using traffic classification (TCLAS) filters in a TFS Request frame. A TCLAS filter specifies the fields in the data frame that must match to convert the frame for unicast delivery.

Because group addressed traffic can be converted to individually addressed traffic, a STA in WNM-Sleep mode need not receive any protected group addressed traffic and thus need not participate in the group key updates.

8.8.4 SM power save

Another way to conserve power, besides completely turning off the radio, is to turn off all but one receive chain. A protocol supporting this was added with the introduction of spatial multiplexing in the 802.11n amendment.

A STA operating in dynamic SM power save mode keeps one receive chain on while listening to the medium. It turns on the remaining receive chains following a CTS response to an RTS frame addressed to it. The RTS/CTS exchange prefixes the start of a TXOP and with all the receive chains on during the TXOP, the STA is able to receive high rate spatially multiplexed data. The STA switches back to a single receive chain after a gap greater than SIFS in the frame exchange sequence.

A STA operating in static SM power save mode operates with a single active receive chain.

A STA indicates a transition from active mode to SM power save mode (dynamic or static) or vice-versa using the SM Power Save Mode frame.

8.8.5 Operating Mode Notification

Two limitations of SM power save are (1) that it is all or nothing; either all receive chains are active or only one receive chain is active and (2) that it does not support changes in operating channel width. Changing operating channel width is another way to reduce power consumption; a listening STA sampling 80 MHz of bandwidth consumes more power than if it were sampling 20 MHz of bandwidth. The Operating Mode Notification frame and Operating Mode Notification element were added with the 802.11ac amendment to address these limitations.

The Operating Mode Notification frame indicates the STAs current operating mode: the number of space-time streams the STA is capable of receiving and the operating bandwidth.

The Operating Mode Notification frame, being a management frame, is not acted upon immediately, but typically with some delay by management software in the recipient. Because of this possible delay, a STA intending to decrease the number of active receive chains or operating bandwidth, should not do so immediately on sending the frame, but instead wait for a timeout or until it sees traffic conforming to the new operating mode (narrower bandwidth and/or reduced space-time streams).

The Operating Mode Notification element has the same function as the Operating Mode Notification frame, but can be included in various other management frames, such as the Association Response frame.

References

IEEE (2012). *IEEE Standard for Information Technology – Telecommunications and Information Exchange Between Systems – Local and Metropolitan Networks – Specific Requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE Std 802.11^{TM}-2012 (Revision of IEEE Std 802.11-2007).

Wi-Fi (2010). Wi-Fi Peer-to-Peer (P2P) Technical Specification Version 1.1, October 2010.