IEEE 802.11

Wireless Access Method and Physical Layer Specifications

Title: **Supporting Multiple Outstanding MSDUs**

Presented by: **Johnny Zweig**
Xircom, Inc.
2041 Landings Drive
Mountain View, CA 94043
Tel: (415) 691-2500
Fax: (415) 691-1064
Email: jzweig@xircom.com

Abstract: **This (revised) paper discusses support in the MAC for multiple outstanding MSDUs. It specifies text changes to D3.1 (as updated through Wednesday afternoon of the Waltham meeting) that describe rules restricting which MSDUs may simultaneously be eligible for transmission by the MAC, and clarifies the distinction between the backoff mechanism and the MPDU transmission mechanism.**

Action: **Adopt text changes described here for incorporation into D4.0.**

Introduction

In discussion by the MAC group of how to resolve Matt Fischer's comment number 38 on clause 9 of D3.0, the consensus was reached that, as long as it does not lead to "gratuitous" reordering or discarding of MSDUs, a MAC implementation should be allowed to process multiple outstanding MSDUs. Here, the term "outstanding MSDU" means either an MSDU (or MMPDU) one of whose fragments (which may be the entire MSDU) is eligible for transmission. The MAC already has to be able to receive multiple MSDUs simultaneously, since if the transmission of one MSDU is interrupted, another station may grab the medium and transmit a different MSDU to that station. The text below places limits on how a station may give up on an unsuccessful attempt to send one outstanding MSDU and attempt instead to transmit another. The motivation for doing this is that, rather than trying many, many times to reach a "dead" STA or one that is on the fringe of the reception area -- with each attempt typically taking much longer than the previous one -- a STA should have the opportunity to try sending to a different station. If the problem with the first STA was collisions or interference, the transmitting station is no worse off (it will still be backed off to the same extent). If the problem was that the receiving station could not hear the transmitter, the other MSDUs may be delivered successfully. This can result - in the latter case - in significantly increased performance. In essence, this text provides a mechanism for distinguishing between the two general causes of failure for a transmission attempt, and increase performance by not backing off unnecessarily (backing off in response to failure to be heard by a unit that is simply too far away is unnecessary, since it does not increase the likelihood of getting through on the next try).

The text below prevents a transmitting station (including an AP) from reordering MSDUs transmitted by any station to another station, including group addressed frames (broadcast or multicast). That is, suppose that a station sends an MSDU to another station, then a broadcast, then another directed MSDU to the first station, and finally a different directed MSDU to a second destination. If the transmitting station follows the guidelines below, neither of the intended recipients (they are both intended recipients of the broadcast) will receive the frames it is supposed to receive out of order. They may, however, be delivered to the two intended destinations in a different order (in this case, only the last two directed frames may safely be reordered; the presence of the broadcast will hold up processing of the last frame until after it has been transmitted).

It should be noted that the restrictions imposed here are stricter than those under which 802.11 has been operating, and has the effect of preventing peer LLC entities from observing reordering. The proposed clause 9.8 could easily be modified to loosen this restriction, if the sentiment that reordering group-addressed (broadcast or multicast) frames is not worth preventing.

The very last sentence of clause 9.8 is meant to reflect Rule 7 below, but it does not define what "normal operating conditions" are. Pablo suggested that a mechanism be defined whereby an AP could tell STA in its BSS what aMax-Receive-MSDU-Lifetime value to use. Given that an implementation will always have to make a decision -- based on resource availability and history -- as to what value to use, I don't think adding text to the effect that poor implementations will end up throwing stuff away if they behave stupidly is necessary.

The rules that each station must follow are summarized here:

1) There is only one Current Contention Window Value (CCWV), one associated Backoff Timer, one Station Short Retry Count (SSRC), and one Station Long Retry Count (SLRC) maintained per STA -- so that medium-access contention rules are adhered-to regardless of what MPDU is being transmitted -- and the Backoff Timer cannot be chosen anew after it has been set, until it has been decremented to zero
2) The Short Retry Count (SRC), Long Retry Count (LRC) and the Lifetime Timer must be maintained per-MSDU -- so that MSDUs are only discarded after being attempted the appropriate number of times over the appropriate interval
3) The CCWV is reset to CWmin after successful transmission of an MPDU; the SSRC is reset to 0 whenever a CTS is received in response to an RTS, an ACK is received in response to transmission of any MPDU, or a group-addressed MSDU is transmitted; the SLRC is reset to 0 whenever an ACK is received in response to transmission of a "long" MPDU, or a group-addressed MSDU is transmitted -- the CCWV is ***not*** reset to CWmin upon the initial attempt to transmit each MSDU
4) The CCWV is reset to CWmin whenever the SSRC is equal to aShort-Retry-Max or the SLRC is equal to aLong-Retry-Max (both Station Retry Counts are also reset in this case); the SSRC is incremented whenever any SRC associated with an outstanding MSDU is incremented, and the SLRC is incremented whenever any LRC associated with an outstanding MSDU is incremented
5) There can be no more than one directed MSDU from a particular SA to a particular individual RA outstanding at a time (a simpler way to implement this is to only allow one outstanding MSDU per individual RA, which preserves this invariant but restricts a STA from making some frames eligible for transmission that safely could be) -- this is simply a literal restatement of the requirement that peer LLCs do not suffer gratuitous frame reordering
6) Any group address (broadcast or multicast) is treated as being addressed to all possible individual RAs, so by analogy with Rule 5 there can be no more than one group-addressed MSDU from any particular SA outstanding, and such an MSDU cannot be outstanding if there is a directed MSDU from that SA outstanding
7) The transmitting station should take into account how many outstanding MSDUs it is able to process when it chooses a value for aMax-Transmit-MSDU-Lifetime, since any particular MSDU may take a time proportional to the number of simultaneous MSDUs being processed

Note that clause 9 has been revised by the work that has taken place this week. I have based my revisions on a draft (from Wednesday afternoon) of the new version of clause 9, since it has changed so significantly from D3.1.

The text following this sentence represents the text changes proposed for incorporation into the next revision of the 802.11 specification. These changes will need to be harmonized with any other pending text revisions to the affected sections that take place during this meeting.

9. MAC Sub-layer Functional Description

Here, the MAC functional description is presented. Clause 9.1 introduces the architecture of the MAC sublayer, including the distributed coordination function, the point coordination function and their coexistence in an 802.11 LAN. Clauses 9.2 and 9.3 expand on this introduction and provide a complete functional description of each. Clauses 9.4 and 9.5 cover fragmentation and reassembly. Multirate support is addressed in clause 9.6. Clause 9.7 lists the allowable frame exchange sequences. Clause 9.8 describes a number of additional restrictions to limit the cases in which MSDUs are reordered or discarded.

9.2.4 Random Backoff Time

A STA desiring to initiate transfer of data and management MPDUs shall utilize both the physical and virtual carrier sense functions to determine the state of the medium. If the medium is busy, the STA shall defer until after a DIFS is detected, and then generate a random backoff period for an additional deferral time before transmitting. This process minimizes collisions during contention between multiple STA that have been deferring to the same event.

An exception to this procedure shall be followed for the case of transmission of multicast frames by an AP. In this case, the AP shall utilize both the physical and virtual carrier sense functions to determine the state of the medium. If the medium is busy, the AP shall defer until after a PIFS period elapses with the medium sensed to be free, and then, without generating a random backoff period, shall immediately commence transmission of the multicast frame. This techinique shall be applied by the AP to the transmission of all Data and Control Type frames with a group address in the Address1 field and the To_DS bit clear, including Beacon frames.

Backoff Time = INT(CW x Random()) x Slot time

where:

CW = An integer between the values of MIB attributes aCWmin and aCWmax
Random() = Pseudo random number between 0 and 1
Slot Time = The value of MIB attribute aSlottime

The Contention Window (CW) parameter shall take an initial value of aCWmin~~ for every MPDU queued for transmission~~. Every station shall maintain a Station Short Retry Count (SSRC) as well as a Station Long Retry Count(SLRC), both of which shall take an initial value of zero. The SSRC shall be incremented whenever any Short Retry Count associated with any MSDU is incremented. The SLRC shall be incremented whenever any Long Retry Count associated with any MSDU is incremented. The CW shall take the next value~~, or higher,~~ in the series (or a higher value)~~at~~ every time an unsuccessful attempt~~retry~~ to transmit~~send a particular~~ an MPDU causes either of the Station Retry Counts to increment, until it reaches the value of aCWmax. A retry is defined as the ~~entire~~ sequence of frames sent separated by SIFS intervals~~ to~~ in an attempt to deliver an MPDU, as described in clause 9.7. The CW shall remain at the value of aCWmax until it is reset~~for the remaining retries~~. This improves the stability of the access protocol under high load conditions. See Figure 40.

The CW shall be reset to CWmin after every successful attempt to transmit an MSDU or MMPDU. The SSRC shall be reset to 0 whenever a CTS is received in response to an RTS, whenever an ACK is received in response to an

MPDU transmission, or a frame with a group address in the Address1 field is transmitted. The SLRC shall be reset to 0 whenever an ACK is received in response to an MPDU of length greater than aRTS-Threshold transmission, or a frame with a group address in the Address1 field is transmitted.

The set of CW values are 7, 15, 31, 63, 127, 255, 511, 1023.

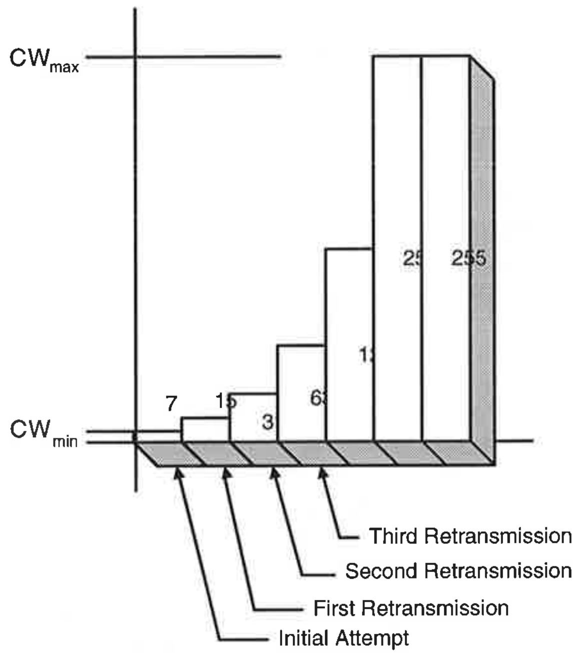

Figure 1, Exponential Increase of CW

aCWmin and aCWmax are settable MAC parameters. Each STA shall update its aCWmin and aCWmax parameters from the CW field contained in each Beacon and Probe Response frame received from its AP.

In an IBSS aCWmin and aCWmax shall be fixed to their default values.

9.2.5.1 Basic Access

Basic access refers to the core mechanism a STA uses to determine whether it may transmit.

~~A~~ In general, a STA may transmit a pending MPDU when it is operating under the DCF access method, either in the absence of a Point Coordinator or in the Contention Period of the PCF, when it detects the free medium for greater than or equal to a DIFS. If, under these conditions, the medium is busy when a STA desires to initiate the initial frame of one of the frame exchanges described in clause **Error! Reference source not found.**, exclusive of the CF period, the Random Backoff Time algorithm described in clause 9.2.4 shall be followed. There are conditions, specified elsewhere, where the Random Backoff algorith shall be followed even for the first attempt to transmit an MPDU.

The basic access mechanism is illustrated in the following diagram.

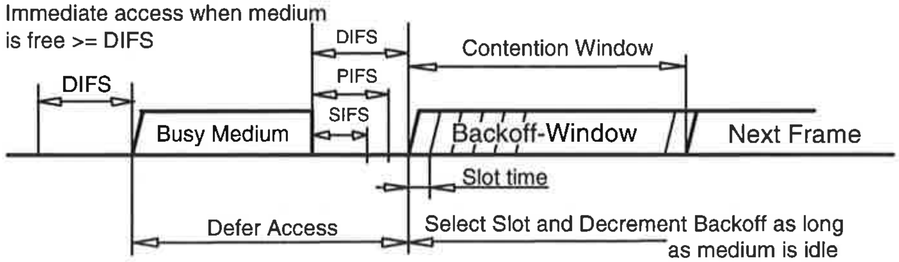

Figure 41, Basic Access Method

9.2.5.1 Backoff Procedure

The backoff procedure shall be invoked whenever a STA desires to transfer a frame and finds the medium busy as indicated by either the physical or virtual carrier sense mechanism (Figure 42), except when the next MSDU to be transmitted has a group address in the Address1 field, the To_DS bit is clear, and the transmitting station is an AP. The backoff procedure shall also be invoked when a transmitting STA detects a failed transmission as defined in subclauses 9.2.5.7 or 9.2.8.

To begin the backoff procedure, the STA shall ~~select~~ set its Backoff Timer to a random backoff time ~~from~~ using the equation in clause 9.2.4 Random Backoff Time, or an implementation-dependent algorithm that, on average, selects longer random time values. All backoff slots occur following a DIFS period during which the medium is free for the duration of the DIFS period.

A STA in backoff shall monitor the medium for carrier activity during backoff slots. If no carrier activity is seen for the duration of a particular slot, then the random backoff process shall decrement its backoff time by aSlottime.

If there is carrier activity sensed at any time during a slot, then the backoff procedure is suspended, that is, the backoff timer shall not decrement for that slot; The medium shall be sensed as idle for the duration of a DIFS period before the backoff procedure is allowed to resume. Transmission shall commence whenever the Backoff Timer reaches zero.

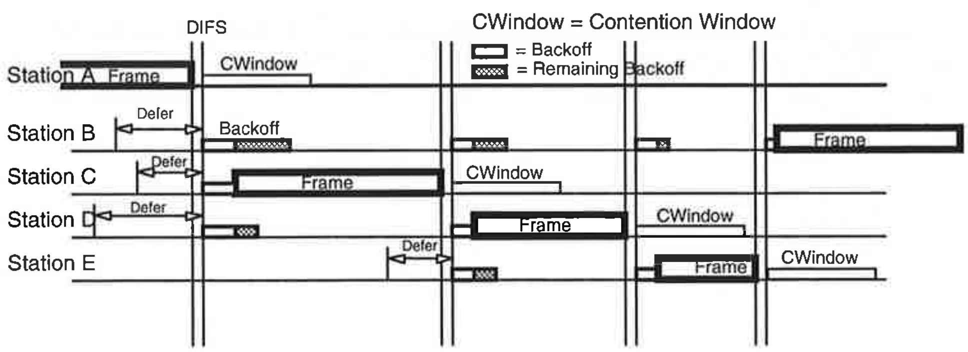

Figure 42, Backoff Procedure

A station that has just transmitted an MSDU and has another MSDU ready to transmit (queued), shall perform the backoff procedure, except when the next MSDU to be transmitted has a group address in the Address1 field, the To_DS bit is clear, and the transmitting station is an AP. This requirement is intended to produce a level of fairness of access to the medium amongst STA. A backoff procedure shall be performed immediately after the end of every transmission of an MPDU with the More Fragments bit set to 0, even if no additional transmissions are currently queued. In the case of successful acknowledged transmissions, this backoff procedure shall begin at the end of the received ACK. In the case of unsuccessful transmissions requiring acknowledgement, this backoff procedure shall begin at the end of the ACK-Timeout interval. If the transmission was successful, the CW value reverts to CWmin before the random backoff iterval is chosen, and the Station Short Retry Count and/or Station Long Retry Counts are updated as described in clause 9.2.4. This assures that ~~TX~~ transmitted frames are always separated by a backoff interval.

The effect of this procedure is that when multiple stations are deferring and go into random backoff, then the station selecting the smallest backoff time~~lowest delay~~ through the random function will win the contention.

In an IBSS, the backoff time shall not decrement in the period from TBTT until the expiration of the ATIM window. Beacon and ATIM frames may be transmitted during this same period.

9.2.5.3 Recovery Procedures and Retransmit Limits

Error recovery is always the responsibility of the STA which initiates a frame exchange sequence, as defined in subclause 9.7. Many circumstances may cause an error to occur which requires recovery. For example, the CTS frame may not be returned after an RTS frame is transmitted. This may happen due to a collision with another transmission, due to interference on the channel during the RTS or CTS frame, or because the station receiving the RTS frame has an active virtual carrier sense condition (indicating a busy medium time period).

Error recovery shall be attempted by retrying transmissions for frame exchange sequences which the initiating station believes have failed. Retries shall continue, for each failing frame exchange sequence, until the transmission is successful, or until the relevant retry limit is reached, whichever occurs first. Stations shall maintain a Short Retry Count and a Long Retry Count for each MSDU or MMPDU awaiting transmission. These counts are incremen~~e~~ted and reset independ~~e~~antly of each other.

After an RTS frame is transmitted, the STA shall perform the CTS procedure, as defined in subclause 9.2.5.7. If the RTS transmission fails, the Short Retry Count for the MSDU or MMPDU ~~is~~ and the Statio Short Retry Count are incremented.This process shall continue until the number of attempts to transmit that MSDU/MMPDU reaches aShortRetryMax.

After transmitting a frame which requires acknowledgement, the STA shall perform the ACK procedure, as defined in subclause 9.2.8. The Short Retry Count for an MSDU or MMPDU and the Station Short Retry Count shall be incremeneted every time transmission of a MAC frame of length less than or equal to aRTS-Threshold fails for that MSDU or MMPDU. This Short Retry Count and the Station Short Retry Count shall be reset when a MAC frame of length less than or equal to aRTS-Threshold succeeds for that MSDU or MMPDU. The Long Retry Count for an MSDU or MMPDU and the Station Long Retry Count shall be incremeneted every time transmission of a MAC frame of length greater than aRTS-Threshold fails for that MSDU or MMPDU. This Long Retry Count and the Station Long Retry Count shall be reset when a MAC frame of length greater than aRTS-Threshold succeeds for that MSDU or MMPDU. All retransmission attempts for an MSDU or MMPDU that has failed the ACK procedure one or more times shall be made with the Retry field set to 1 in the Data or Management Type frame.

Retries for failed transmission attempts shall continue until the Short Retry Count for the MSDU or MMPDU is equal to aShort-Retry-Limit or until the Long Retry Count for the MSDU or MMPDU is equal to aLong-Retry-Limit. When either of these limits is reached, retry attempts shall cease, and the MSDU or MMPDU shall be discarded.

A station in power save mode initiates a frame exchange sequence by transmitting a PS-Poll frame to request data from an AP. In the event that neither an ACK frame or a data frame is received from the AP in response to a PS-Poll frame, then the station shall retry the sequence by transmitting another PS-Poll frame, at its convenience. If the AP sends a data frame in response to a PS-Poll frame, but fails to receive the ACK frame acknowledging this data frame, the next PS-Poll frame from the same station may cause a retransmission of the last MSDU. This duplicate MSDU shall be filtered at the station using the normal duplicate filtering mechanism. If the AP responds to a PS-Poll by transmitting an ACK frame, then responsibility for the data frame delivery error recovery shifts to the AP because the data is transferred in a subsequent frame exchange sequence, initiated by the AP. The AP shall attempt to deliver one MSDU to the station which transmitted the PS-Poll, using the any frame exchange sequence valid for a directed MSDU. If the power save station which transmitted the PS-Poll returns to Doze state after transmitting the ACK frame in response to successful receipt of this MSDU, but the AP fails to receive this ACK frame, the AP will retry transmission of this MSDU until the relevant retry limit is reached. See Clause 11 for details on filtering of extra PS-Poll frames.

9.2.5.5 Control of the Channel

The Short Interframe Space (SIFS) is used to provide an efficient MSDU delivery mechanism. Once the station has contended for the channel, that station shall continue to send fragments until either all fragments of a singleMSDU or MMPDU have been sent, an acknowledgment is not received, or the station is restricted from sending any additional fragments due to a dwell time boundary. Should the sending of the fragments be interrupted due to one of these reasons, when the next opportunity for transmission occurs the station shall resume transmission~~sending the fragments~~. The algorithm by which the STA decides which of the outstanding MSDUs (as defined in clause 9.8) shall next be attempted after an unsuccessful transmission attempt is beyond the scope of this standard.

Figure 44 illustrates the transmission of a multiple fragment MSDU using the SIFS.

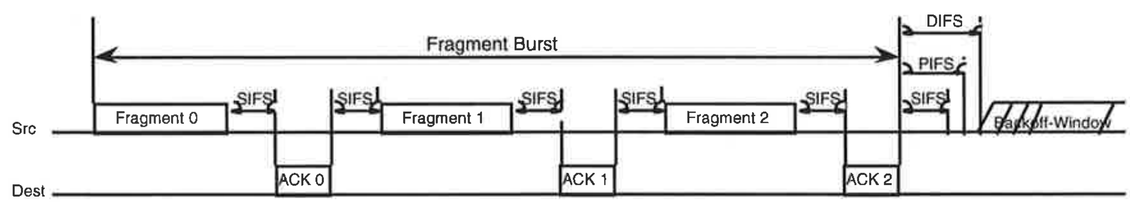

Figure 44, Transmission of a Multiple Fragment MSDU using SIFS

When the source station transmits a fragment, it shall release the channel, then immediately monitor the channel for an acknowledgment as described in subclause 9.2.8.

When the destination station has finished sending the acknowledgment, the SIFS following the acknowledgment shall be reserved for the source station to continue (if necessary) with another fragment. The station sending the acknowledgment~~does not have permission to~~ shall not transmit on the channel immediately following the acknowledgment.

The process of sending multiple fragments after contending for the channel is defined as a fragment burst.

If the source station receives an acknowledgment but there is not enough time to transmit the next fragment and receive an acknowledgment due to an impending dwell boundary, it shall contend for the channel at the beginning of the next dwell time.

If the source station does not receive an acknowledgment frame, it shall attempt to retransmit the failed MPDU or some other eligible (as defined in 9.8) MPDU after performing the backoff procedure and the contention process.

After a station contends for the channel to retransmit a fragment of a MSDU, it shall start with the last fragment that was not acknowledged. The destination station shall receive~~s~~ the fragments in order (since the source sends them ~~one at a time,~~ in order). It is possible however, that the destination station may receive duplicate fragments. It shall be the responsibility of the receiving station to detect and discard duplicate fragments.

A station shall transmit after the SIFS only under the following conditions during a fragment burst:

- The station has just received a fragment that requires acknowledging.
- The source station has received an acknowledgment to a previous fragment, has more fragment(s) for the same MSDU to transmit, and there is enough time before the next dwell boundaryto send the next fragment and receive its acknowledgment.

The following rules shall also apply.

- When a station has transmitted a frame other than an initial or intermediate fragment, that station shall not transmit on the channel following the acknowledgment for that frame, without performing the backoff procedure.
- When an MSDU has been successfully delivered or all retransmission attempts have been exhausted~~used~~, and the station has a subsequent MSDU to transmit, then the station shall perform a backoff procedure, except when the next MSDU to be transmitted has a group address in the Address1 field, the To_DS bit is clear, and the transmitting station is an AP.
- Only unacknowledged fragments shall be retransmitted.

9.8 MSDU Transmission Restrictions

In order to avoid reordering MSDUs between pairs of LLC entities and/or unnecessarily discarding MSDUs, the following restrictions shall be observed by any STA that is able to simultaneously process multiple outstanding MSDUs for transmission. Note that here the term "outstanding" refers to an MSDU or MMPDU that is eligible to be transmitted at a particular time. A STA may have any number (greater than or equal to one) of eligible MSDUs outstanding simultaneously, subject to the restrictions below.

The STA shall ensure that no more than one MSDU or MMPDU from a particular SA to a particular individual RA is outstanding at a time. Note that a simpler, more restrictive invariant to maintain is that no more than one MSDU with a particular individual RA may be outstanding at a time. This latter restriction precludes certain situations where MSDUs or MMPDUs from different LLC sources to a particular individual address, that may safely be reordered by the MAC, are simultaneously outstanding.

The STA shall ensure that there is no group-addressed (multidestination) MSDU or MMPDU outstanding from the SA of any other outstanding MSDU or MMPDU (either directed or group-addressed). Note that this is because a group-addressed MSDU or MMPDU is implicitly addressed to a collection of peer stations that could include any individual RA.

The STA shall select a value of aMax-MSDU-Transmit-Lifetime that is sufficiently large that the STA does not discard MSDUs due to excessive Transmit MSDU Timer values under normal operating conditions.