IEEE 802.11
Wireless Access Method and Physical Layer Specifications

Title: **Transmitter Priority in the MAC Layer**

Author:
Rick White
Motorola, Inc.
Wireless Data Group
50 E. Commerce Dr.
Schaumburg, Illinois 60173
Tel: 1-708-576-7878
Fax: 1-708-576-7907
Email: rick_white@wes.mot.com

Abstract: This submission proposes an approach of allowing a station that has accessed the media successfully and has more data frames to transmit, to have priority after the reception of the ACK. This falls into the "Improvement" category. The author would like to see a priority defined for a transmitting station based on the rules provided in this paper.

Introduction

Windowing was discussed in [1]. It proposed an approach that allowed a station to send multiple packets before receiving an acknowledgment. Windowing is used to deliver a multiple data frames with minimal delay. The windowing proposal contained two key elements. The first was allowing a station to send multiple frames before receiving an acknowledgment. This allowed a reduced amount of overhead in sending multiple frames. The second key element was a station would only contend for the media for the first frame of multiple data frames. This allowed for minimal delay of multiple data frames.

This paper proposes an implementation that allows a station to maintain priority on the medium while transferring multiple data frames. It is based on the windowing discussion in [1] and transmitter priority described in [2] and [3]. When used with fragmentation, described in [4], whole MAC Service Data Units (MSDU) can be fragmented into multiple data frames and sent with minimal delay. For a multiple fragment MSDU, the sending device would only have to contend for the channel for the first data frame of each MSDU. Transmitter priority has equal benefits for both frequency hopping and direct sequence PHYs.

Transmitter Priority

Transmitter priority was originally discussed in [2]. This priority allows a station to send multiple asynchronous frames without being preempted by another station. In order to minimize the amount of time that a station can "hold" the channel, a limited number of frames can be sent. When a station has a MSDU to send, it will contend for the channel in the contention window. Once the station has successfully sent the first fragment of a multiple fragment MSDU, it is allowed to send the remaining fragments of the same MSDU after the acknowledgment of the previous fragment. If an acknowledgment is not received, the station executes the backoff algorithm and contends for the channel in the contention window. It then retransmits the fragment that did not receive an acknowledgment and any remaining fragments for the same MSDU. This allows the station to contend for the channel only once for each MSDU, assuming no transmission errors.

Transmitter priority reduces the delay introduced in a wireless network. When using a request-reply protocol, the data throughput is greatly influence by the delay from the time that a request is sent to the time a reply is received. In a wireless network, a large portion of the round-trip delay is determine by delay in the wireless portion of the network connection. The delay is not only due to the lower bits rates of the wireless LAN but also the contention problem. In an ethernet environment, a station knows if there was a collision during the header of the packet due to the fact that a station can detect collisions. In the 802.11 wireless LAN, a station will only know if there was a collision if it does not receive an acknowledgment from the destination station.

If a station can only transmit one data frame each time it contends for the channel, the end-to-end delay of a multiple date frame transfer will increase. In order to reduce this delay, transmitter priority can be used to allow a station to transmit more than one data frame after it has contended for the channel. In a radio environment it is better to send smaller frames from both a bit error rate and interference perspective. A station that sends multiple small data frames has a better chance of getting the frames to the destination station then sending one large data frame.

Transmitter priority allows data frames of a fragmented MSDU to be transferred "back-to-back" without introducing any significant additional delay. Once the transmitting station has contended for the channel, it can send multiple data frames of the same MSDU without contending for the channel again.

Transmitter priority is beneficial for both Frequency Hopping and Direct Sequence systems since bit error rates and transmission delays are similar.

Implementation of Transmitter Priority

The Foundation MAC provides three different accesses priorities which are distinguished using different values for the minimum silence period. They are described in [5] and restated here. Figure 1 shows the three Inter Frame Spaces (IFS). The three priorities are Short Inter Frame Space (SIFS), Point Coordination Function Inter Frame Space (PIFS) and Distributed Coordination Function Inter Frame Space (DIFS).

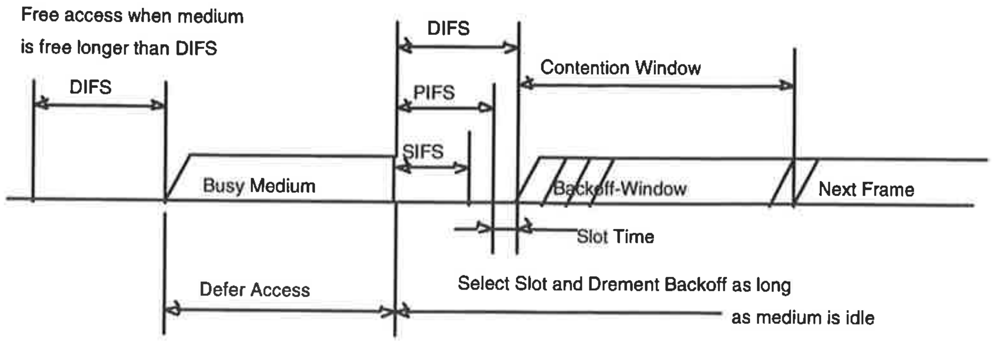

Figure 1: Foundation CSMA/CA Access Method

Short Priority

This priority level is used for all immediate response actions. This priority levels used for an ACK frame, immediately following a received frame, a CTS frame immediately following a received RTS frame, and by a station responding to any polling as is used by the Point Coordination Function (PCF). The corresponding interframe space is called the Short IFS, or SIFS.

PCF Priority

This priority level is used by the PCF in the AP to send any of the Contention Free Period (CFP) frames. The AP will send the next-in-line queued CFP frame after it finds the medium free for the period PIFS (PCF Interframe Space), during a CFP-Burst .

DCF Priority

The DCF priority is used by the Distributed Coordination Function to transmit Asynchronous frames in the Contention Period. Asynchronous stations that want to start transmission of an RTS frame, or a DATA frame (without the RTS/CTS option), will monitor the medium for at least a time DIFS (DCF Inter Frame Space) of silence, after the NAV and the CS function indicates a free medium. When the medium was found busy, then the DFWMAC will defer until an DIFS is detected, and then a random access backoff procedure is started.

Implementation of Transmitter Priority in the Foundation MAC

Transmitter priority is used to provide a efficient MSDU delivery mechanism. Once a station has contended for the channel, it can maintain control of the channel until it has sent all of the data frames of a MSDU. After all frames have been transmitted, the station will relinquish control of the channel.

Once the station has contended for the channel, it will continue to send frames until either all frames of a fragmented MSDU have been sent or an acknowledgment is not received.

Figure 2 illustrates the transmission of a multiple fragment MSDU using transmitter priority.

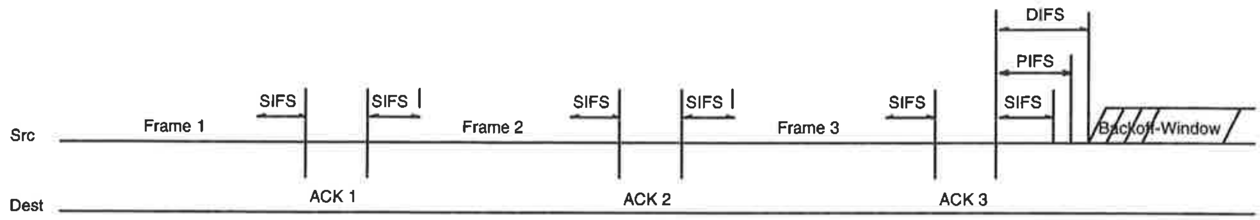

Figure 2 : Transmission of a Multiple Fragment Packet using Transmitter Priority

The source station transmits a data frame then releases the channel and waits for an acknowledgment. When the source station releases the channel following its data frame, it will immediately turn its radio around and monitor the channel for an acknowledgment frame from the destination station.

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

If the source station does not receive an acknowledgment frame, it will attempt to retransmit the data frame at a later time (according to the backoff algorithm). When the time arrives to retransmit the data frame, the source station will contend for access in the contention window.

After a station contends for the channel to retransmit a data frame of a multiple fragment MSDU, it will start with the last data frame that was not acknowledged. The destination station will receive 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 packets. This will happen if the destination station sends an acknowledgment and the source does not receive it. The source will resend the same packet after executing the backoff algorithm and contending for the channel.

A station will transmit after the SIFS only under the following conditions (for transmitter priority):

- The station has just received a frame that requires acknowledging.
- The source station has received an acknowledgment to a previous frame and has more data frame(s) for the same MSDU to transmit.

The following guidelines also apply.

- When a station has transmitted a frame other than a data frame, it does not have priority to transmit on the channel following the acknowledgment for that frame.
- When a MSDU has been successfully delivered, the device does not have priority to transmit on the channel following the last acknowledgment of the last data frame.
- Only unacknowledged data frames need to be retransmitted.

If a multiple frame MSDU does not require an acknowledgment (for example, a broadcast/multicast packet transmitted by the Access Point), the source station will transmit all frames of the MSDU without releasing the channel.

RTS/CTS usage with Transmitter Priority

RTS/CTS can be implemented in two ways when using transmitter priority. In the first approach, the RTS/CTS packets contain the duration of the entire transfer of the MSDU including acknowledgments. This is illustrated in Figure 3.

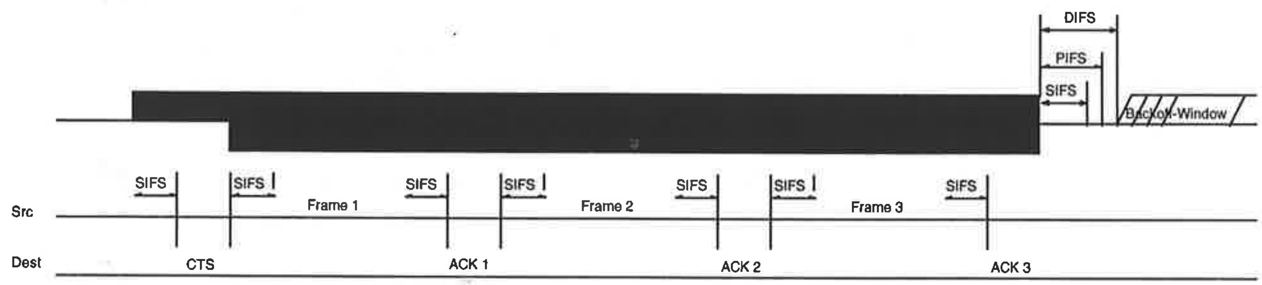

Figure 3 : RTS / CTS with Transmitter Priority - Approach 1

The problem with this approach is if the source station does not receive an acknowledgment for a data frame, it will stop sending and execute the backoff algorithm. If the data frame was not the last fragment of the MSDU, the channel will be marked busy for a considerable amount of time when it is not. For example, if the first acknowledgment was not received by the source station, the channel would be marked busy when it really was not. This is shown in Figure 4.

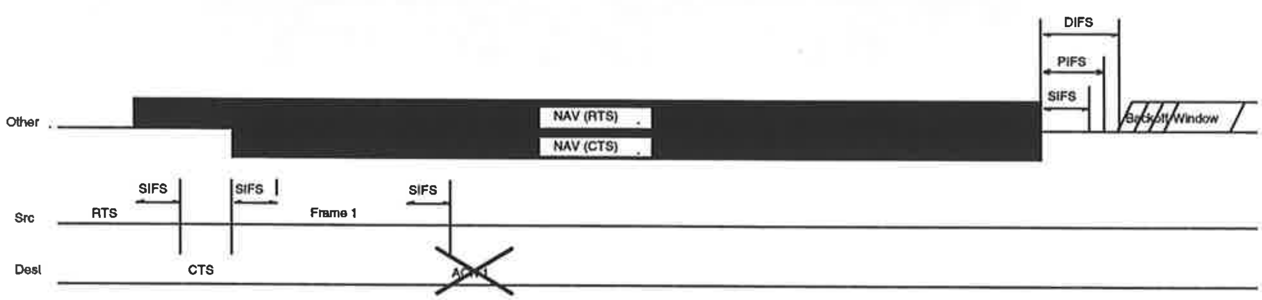

Figure 4 : RTS / CTS with Transmitter Priority with missed Acknowledgment - Approach 1

An alternate approach is to use RTS/CTS for the transmission of the first frame and define an element for the data frame and acknowledgment that specifies the duration of the next frame. This is illustrated in Figure 5.

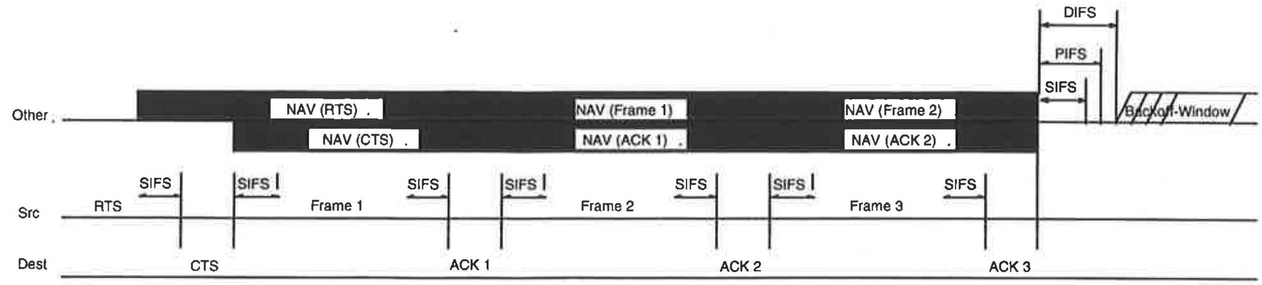

Figure 5 : RTS / CTS with Transmitter Priority - Approach 2

Each frame contains information that defines the duration of the next transmission. The RTS will update the NAV to indicate busy until the end of ACK 1. The CTS will also update the NAV to indicate busy until the end of ACK 1. Both Frame 1 and ACK 1 will update the NAV to indicate busy until the end of ACK 2. This is done by using a duration element in the DATA and ACK frames. This will continue until the last Frame and ACK which will not include the duration element. Each Frame and ACK acts as a virtual RTS and CTS. The Frames and ACKs do not need to be short like RTS and CTS because the channel has already been marked busy.

This approach also helps when hidden stations move into the area of the source or destination stations. Since the NAV is updated by each frame exchange, not just the RTS and CTS, the hidden station will have it's NAV updated after the first frame exchange that it hears.

In the case where an acknowledgment is not received by the source station, the NAV will be marked busy for next frame exchange. This is the worst case situation. This is shown in Figure 5. If the acknowledgment is not sent by the destination station, stations that can only hear the destination station will not update their NAV and be free to access the channel. All stations that hear the source will be free to access the channel after the NAV from Frame 1 has expired.

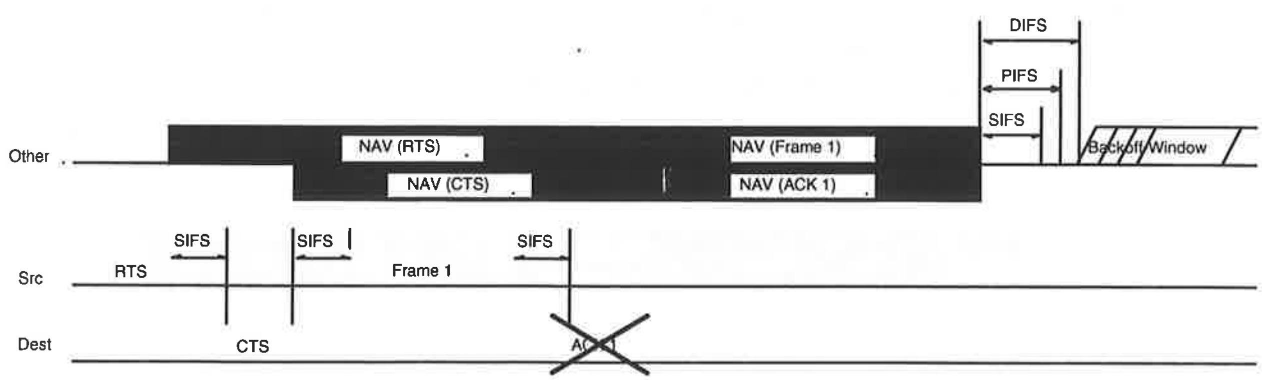

Figure 6 : RTS / CTS with Transmitter Priority with missed Acknowledgment - Approach 2

The second approach will not reserve the channel any longer than the first. In many cases, it will reserve the channel for a shorter time if there is an error in transmission. This is due to the fact that the NAV is only update to mark the channel busy for the duration of the next data frame and acknowledgment. There is a small amount of additional overhead in each data frame and acknowledgment required to define the duration of the next frame exchange when RTS / CTS is used.

Conclusion

Transmitter priority provides an efficient method for transmitting a multiple frame MSDU with minimal delay. There are no new interframe spaces defined, only another use of the short interframe space. The implementation of transmitter priority is straight forward. It does not require a device to be concerned with receiving fragments out of order since each is acknowledged. The source station will send the next frame only after the current frame has been acknowledged.

There are two approaches for using RTS/CTS with transmitter priority. The first approach defined in this paper is simpler to implement but the second provides a better representation of the activity on the channel.

The author would like to have transmitter priority with the second RTS/CTS approach added to the Foundation MAC.

References

[1] Rick White - Motorola, "Frame Windowing at the MAC Layer", Doc IEEE P802.11-94/38 March. 1994.

[2] Rick White - Motorola, "Frame Prioritization in a CSMA/CA Media Access Control Protocol", Doc IEEE P802.11-93/159 Sept 1993.

[3] Rick White - Motorola, et al, "A Complete Description of Frame Prioritization in a CSMA/CA MAC Protocol", Doc IEEE P802.11-93/208 Nov. 1993.

[4] Mark Demange - Motorola, et al, "Packet Fragmentation", Doc IEEE P802.11-94/37 March. 1994.

[5] Wim Diepstraten - NCR, Greg Ennis - Symbol Technologies, & Phil Belanger - Xircom, "DFWMAC - Distributed Foundation Wireless Medium Access Control", Doc IEEE P802.11-93/190 Nov. 1993.