IEEE 802.11
Wireless Access Method and Physical Specification

Title: **DCF proposal with Active priority signalling.**

Prepared by:

Wim Diepstraten
AT&T-GIS (NCR)
WCND-Utrecht
Nieuwegein The Netherlands
Tel: (31)-3402-97482
Fax: (31)-3402-97555
Email: Wim.Diepstraten@utrecht.ncr.com

Tim Phipps
Symbionics Networks Ltd.
Cambridge, UK
Tel: (44)-223 421025
Fax (44)-223 421031
Email: tgcp@symbionics.co.uk

Abstract: This paper is a proposal for a DCF channel access method that provides priority support needed for Distributed Time Bounded Service (DTBS). Support of DTBS functionality was approved by the committee in the March meeting. Selection of the priority mechanism is scheduled for the July meeting.

Background:

This proposal for a priority based DCF is jointly authored by AT&T and Symbionics and builds forward on the priority mechanism evaluation as was presented in the May meeting in Toronto. A priority based access mechanism is needed to support the Distributed Time Bounded Service (DTBS) that allows optional Time Bounded support without the network overlap restrictions that apply for the PCF based CF-TBS as specified in the Foundation MAC.
This proposal has also been proposed as the so called Channel Access Method (CAM) for Hiperlan in Europe being developed by ETSI RES10. Recognising that a common world-wide medium access scheme offers great potential benefits, our companies are committed to keeping the Channel Access Mechanisms used in Hiperlan and IEEE 802.11 closely aligned. In particular this will be of interest in a potential future 5.2 GHz band in the US, where both Hiperlan and IEEE products should be able to co-exist.

Introduction

This document describes the principal features of a Channel Access Mechanism (CAM) based on CSMA with active priority signalling and random back-off. The contention resolution method proposed has been proven to work well in a radio environment and is simple to implement. Priority signalling is added to provide multiple hierarchically independent levels of channel access priority. Hierarchical independence means that increasing load from lower priority classes does not degrade the performance of higher priority classes.
The described mechanism is independent of the use of RTS/CTS, and need to be followed for every DCF access.
Priority signalling was first proposed by Rom and Tobagi in 1981. It was informally suggested by Apple (Kerry Lynn, Larry Taylor) as possible enhancement of the CSMA scheme and it has since been accepted in ETSI RES10 for Hiperlan.

Distributed Coordination Function (DCF) Description

CSMA relies on the detection of the transmissions of other nodes within an area that is larger than the coverage area (the area in which data transmission is possible) to avoid collisions during medium access. In the variety of CSMA proposed here, if the medium is found free for a specified period then a node can transmit its data immediately. If a node finds the medium is busy, it waits until the medium becomes free and then attempts to transmit its data according to a set of channel access rules. These rules provide multiple levels of channel access priority and allow contention resolution for transmission attempts at the same access priority.

Medium Sensing

A key factor in this CAM is the fast and reliable determination of the state of the medium: idle or busy. The minimum resolution of the state of the medium is determined by the transmitter turn-on time, signal propagation time over the maximum operating distance and signal detection time in the receiver. This resolution period is known as the medium sensing slot.

Note 1: Medium access efficiency is optimised when transmissions start at slot boundaries but this is not required

Note 2: The slot size need not to include the time for antenna diversity switching. Typically the first antenna will give a high chance (e.g. 95 %) of detecting that a signal is present. Where this is not the case the signal will be detected on the second antenna. This doubles the detection time but reduces the collision probability introduced by imperfect medium sensing. Since the impact of the detection time is multiplied by the number of medium sensing slots, the lesser impact of a slightly increased collision probability is preferable.

DCF Channel access algorithm

If the medium has been idle for a period MFC (Medium Free Condition) when a node has a frame ready to transmit, then that frame will be transmitted immediately. This scenario is illustrated in figure 1 below. The length of the period MFC is defined later in this paper.

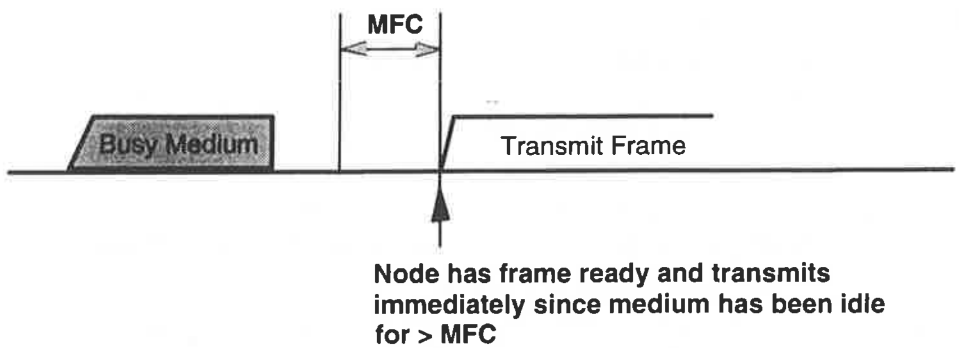


Figure 1: DCF behaviour when medium has been idle for longer than MFC

If the medium is sensed busy when a node has a frame to transmit then the node continues to monitor the state of the medium until it becomes idle. If more than one node has a frame for transmission and has sensed the medium busy, these nodes will contend for access to the medium after it is sensed idle. These nodes wait for a further period, termed the Distributed Inter-Frame Spacing (DIFS) before entering priority resolution (figure 2).

Please note that this definition does not impact the definition and operation of the SIFS and PIFS delays that are used for system responses and Contention free operation using the PCF.

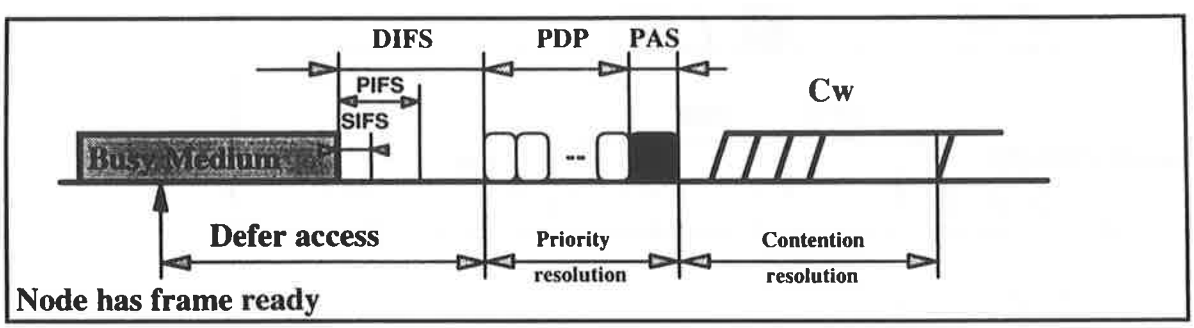


Figure 2 : DCF Priority and Contention Resolution

Priority Resolution

Priority resolution is designed to resolve between multiple DCF accesses with differing transmission priorities. Note that the SIFS and PIFS delays as defined in the Foundation MAC represent absolute priority levels which do not need to involve contention resolution, because a single source is assumed to use it in response to a previous frame.

Contending nodes indicate the transmission priority level by means of a Priority Assertion Signal (PAS). A node indicates the priority level of the frame to be sent by transmitting a PAS at a specific position relative to the end of the DIFS. The highest level of priority is indicated by transmission of the PAS immediately following the DIFS. The lowest level of priority is indicated by the PAS being absent.

During the period between the end of the DIFS and the transmission of the PAS the node listens for PAS transmissions from contending nodes. This period is known as the Priority Detection Period (PDP). If node A detects a PAS from contending node B during the PDP, then node B takes priority and A defers immediately to the next priority resolution/contention resolution cycle. A contending node will transmit its PAS only if it does not see a Priority Assertion Signal during the Priority Detection Period.

The overhead due to this priority resolution mechanism is principally dependent on the PAS duration and number of priority levels, and is not sensitive to Tx/Rx turnaround time.

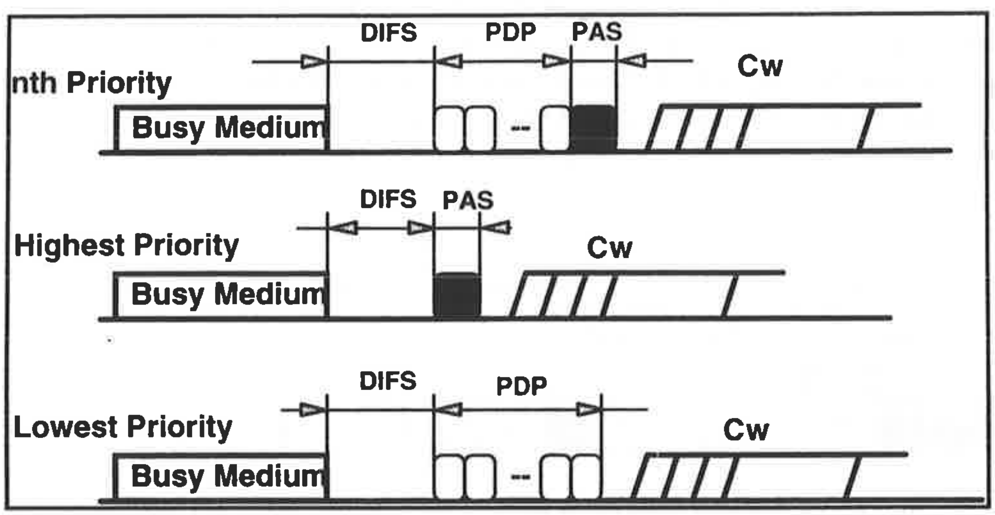


Figure 3: Priority Signalling

PaS Generation and Detection

A PaS signal could be generated by simply turning on the transmitter for the specified time. This will cause the generation of a modulated signal with part of its PHY-preamble. This would not require any additional functionality in the MAC/PHY interface.

As for the PaS detection, we could consider the use of the CCA interface signal, to indicate any activity. However this may have effect on the CCA detection algorithm. It should be noted that the receiver needs to be able to detect multiple simultaneous arriving PaS signals from multiple sources. This will likely require some hybrid CCA detection schemes that react on modulated energy, or any more specific characteristic of received frame indications. This may require an interface with the MAC to report multiple types of CCA indications, to distinguish between for instance PaS and frames.

Contention window and random back-off

Contention resolution is designed to resolve between accesses that have the same transmission priority. Contention resolution takes place after priority assertion, i.e. after the PAS has been transmitted or after the PDP for the lowest priority class.
When a node initially defers due to medium busy, a random delay is selected (expressed as a number of medium sense slots) before its next attempt at transmission. The maximum delay is known as the contention window (CW).
When the node senses that the medium is idle for the duration of at least one medium sense slot, the delay to transmission is decremented by an equal number of slots. This decrementing process is started one medium sensing slot after the PAS has been sent or, in case of the lowest priority level, PDP has elapsed. If the node senses that the medium is busy for longer than PAS (indicating that another node won), then the node defers and participates in the next priority resolution cycle. The delay is not decremented during the defer period, while the medium is busy and during the DIFS that follows. When the delay reaches zero, the node will transmit its frame.
The effect of this back-off algorithm is that nodes already in back-off will have shorter average back-off delays than nodes selecting a back-off delay.

Overload stability provision is needed to assure stable network throughput in the event of an instantaneous high load. This is done by a binary exponential increase of the contention window after each retransmission, as is specified in the Foundation MAC.

Medium Free Condition (MFC)

The Medium Free Condition is defined in order to prevent nodes which are not in contention from pre-empting an active contention window. MFC is defined as continuous medium idle for the concatenation of the DIFS, the maximum PDP and [1/n] x CW.
1/n has an effect on the medium access delay. The larger n the smaller that delay. However, when 1/n becomes small enough, pre-emption of active contention windows may occur. The chance of pre-emption goes down with increasing load levels (because the contention window tends to become fully used). At lower loads the delay of higher priority traffic caused by occasional pre-emption by lower priority traffic may be considered acceptable.

State Machine inputs

The following is a state diagram that illustrates the medium access procedure that supports priority.
It shows when the backoff procedure is initiated, and the priority and contention resolution phases that start after the DIFS period. After a PaS is detected during the PDP, then a state is entered in which the station waits for the detection of the frame of the higher priority level, and will subsequently defer for that. If no frame is detected during a timeout period (could be when PaS detection was a false alarm), then the station will continue with its backoff phase.
During the contention phase the station will decrement its backoff while the medium is free. If a busy medium is detected, then the backoff delay is not decremented, but the station will not defer waiting for the next priority and contention phase, unless the busy detection is long enough to conclude that an other station won the contention and has obtained the medium. This is to prevent that a station will leave its contention resolution phase in response to a PaS signal that is not in sync with the timing of this station.
This can for instance occur when other stations did go into the priority and contention resolution phase earlier, because for instance they did not "see" the Ack of the previous frame.
This is one of the problems that were identified in [2], and will result in leakage between traffic of different priority levels, because the priority signalling periods are outof sync.

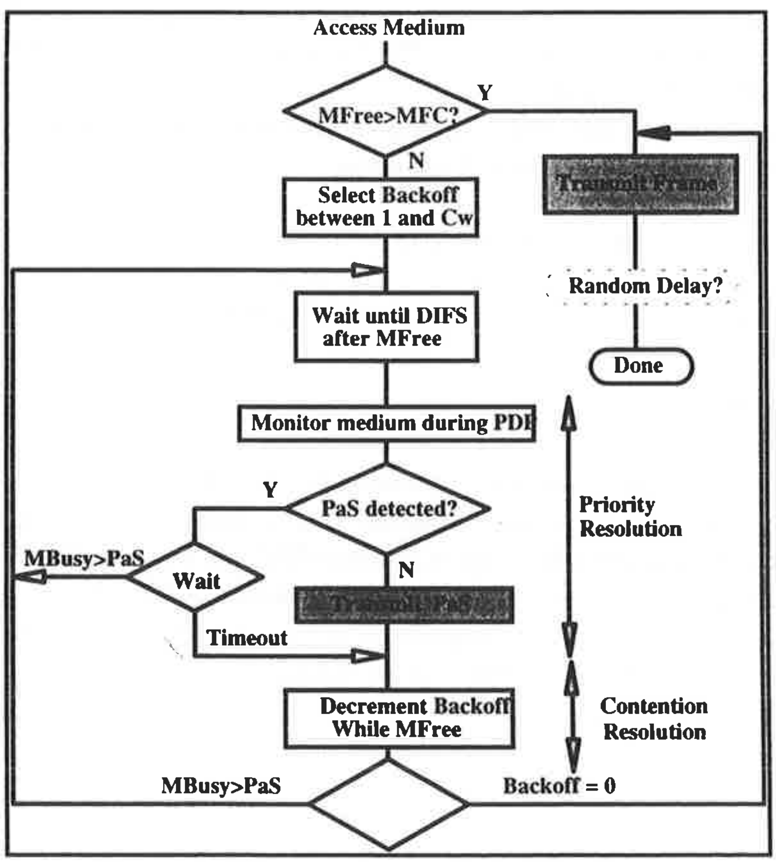


Figure 4; Channel Access State Machine

The figures 5 an 6 shows how the state machine description in the current draft standard [4] is to be modified to support this proposal.
It replaces the C07 branch. It includes a "Wait PDP" state in which the node is sensitive for priority signalling of higher priority, and a "Generate PaS" state in which all stations of equal priority will send out a PaS. Please note that depending of the priority level the PDP and PaS periods can be zero.

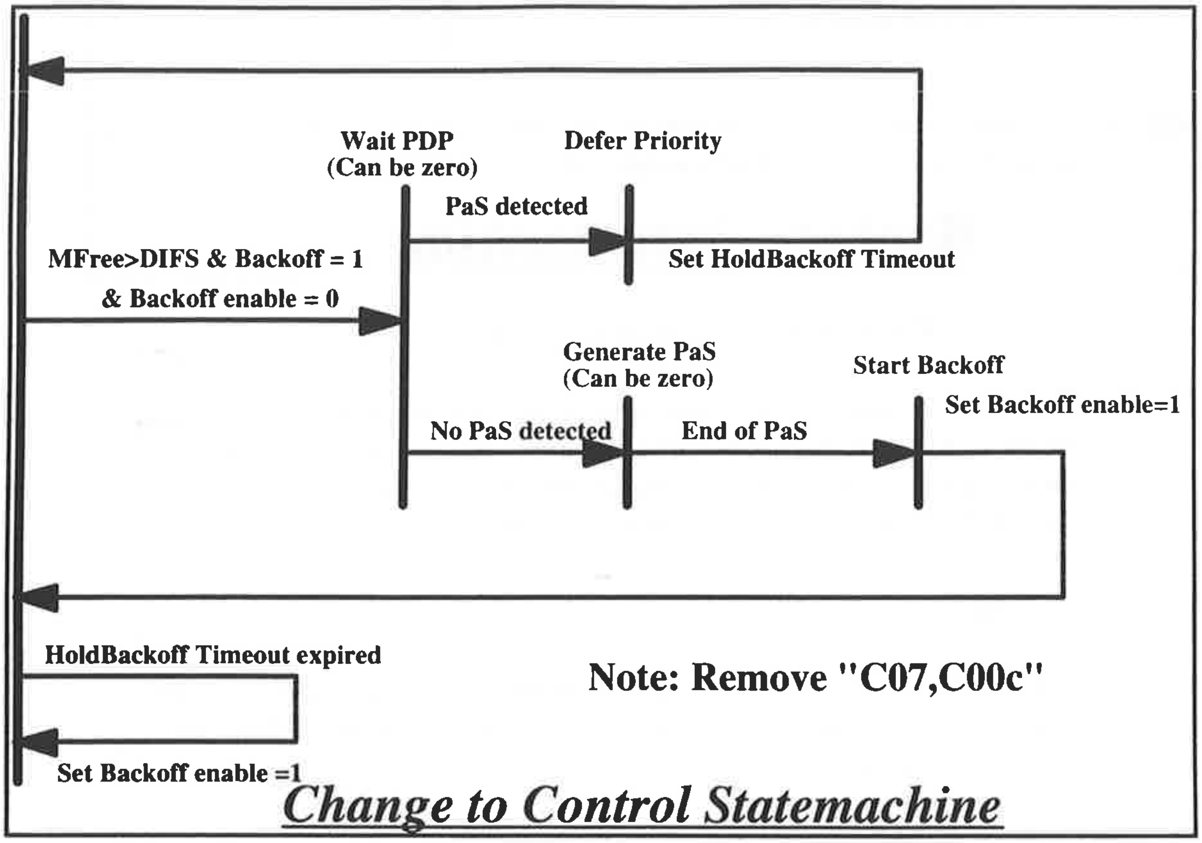


Figure 5: Control State Machine change

Another change needed in the Control State Machine is to change the term "Media Free DIFS" into MFC.

Figure 6 shows a separate state machine description for the backoff counter operation. Its operation is controlled by a "Backoff enable" flag to indicate when the control state machine is in the contention resolution phase.
The Backoff counter is only decremented while the medium is sensed idle. This state is left after the medium is sensed active sufficiently long to conclude that an other station has won the contention. The idea is that a station should be insensitive to PaS signalling in this phase to prevent pre-emption of the contention period due to possible synchronization loss between a number of contending stations.

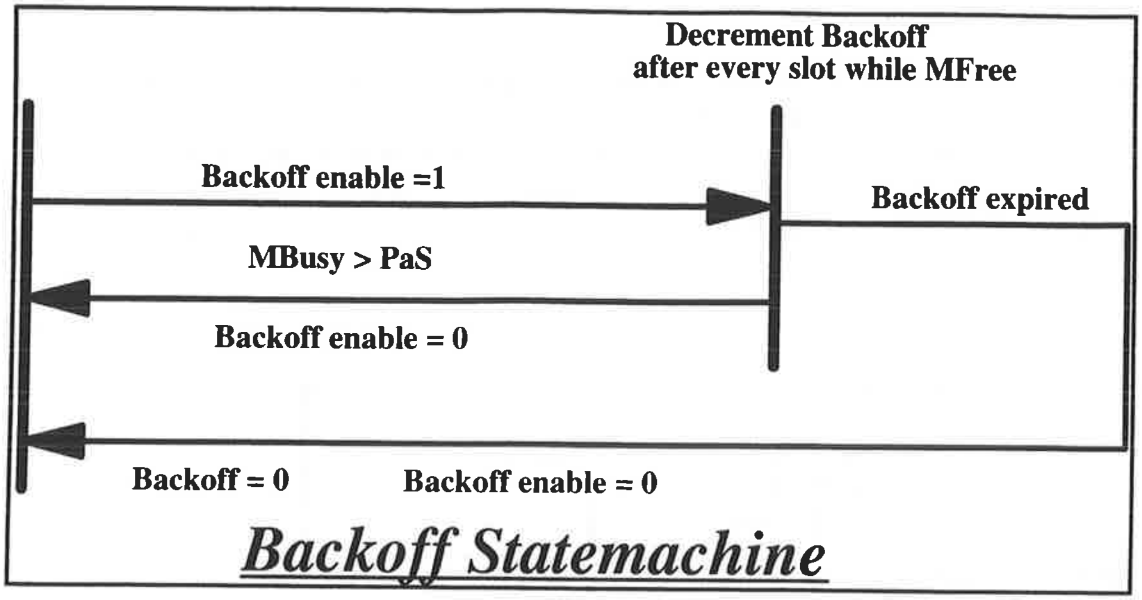


Figure 6: Backoff State Machine

The characteristics of the backoff mechanism are such that stations that are longer in contention (other stations won previous contentions) would have relative higher access probability then stations that just joined the contention by selecting their backoff slot. This is a desirable characteristic from a fairness point of view.

DCF Parameters

The DCF has a number of parameters that can be set to optimise performance under certain conditions. These parameters affect interoperability and therefore they must be part of the standard specification.

- **Slot** This is determined by the properties of the Tx and Rx hardware as well as by the propagation delay over the medium.
  Note: The DCF is tolerant for implementation variations in transceiver design. For example, implementors can choose to use a longer slot time. This does not affect compatibility but it may affect the throughput performance (however the PDP and PAS signalling must be the same in the complete network).
- **SIFS** This period is determined by the need to separate the ACK transmission from subsequent contention based access. The dominant factor is the expected to be the Tx-Rx turnaround time of the transceiver. The SIFS duration is expected to be less than one slot period.
- **DIFS** Determined by the length of the ACK frame; a typical value might be 8 slot periods.
- **PAS** Determined by the need for reasonable reliable communication of the priority signal; a first estimate is two slot periods.
- **PDP** Determined by the number of priority levels; the duration will be (n-1) slots + sync tolerance for an n priority level system.

CW Determines the collision probability at the level of priority. A default CW value of 32 slots is suggested as offering an adequate collision probability.

MFC This condition consists of a continuously idle medium during the concatenation of IFS, the maximum PDP and [1/n] x CW.

DIFS duration:

The DIFS duration will likely need to be longer then SIFS+Ack duration to prevent Ack jamming due to hidden stations generating PaS signals as illustrated in [2]. This situation is shown in figure 7, where station T that wants to transmit, does "hear" the frame from Tx, but not the Ack from Rx, so sends its PaS after the frame. Station B however will see both, and will delay its PaS (if any) until after the Ack.

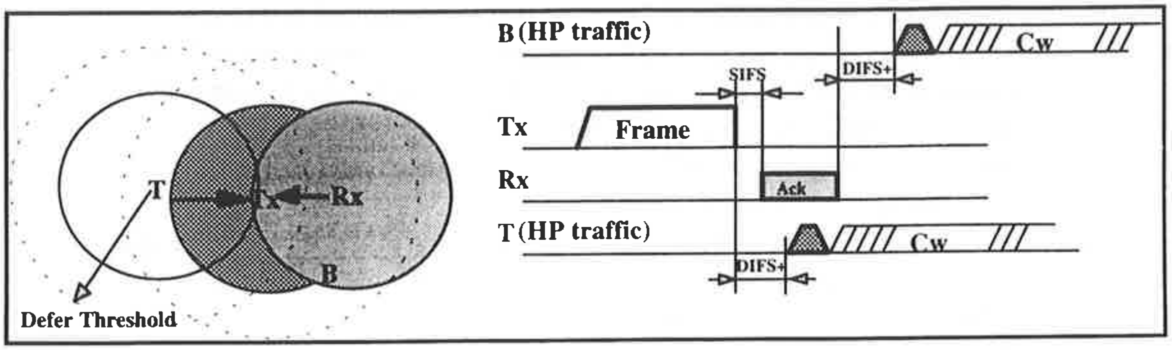


Figure 7: Unsynchronized signalling due to Hidden nodes

Whether the Ack jamming prevention by increasing the DIFS length is needed, will also depend on the Defer Threshold that is specified in the PHY. It could therefore be PHY dependent. Further simulations will be needed with the relevant PHY parameters to determine the tradeoff between threshold and DIFS length.

PDP and PaS duration

The duration of the PDP and PaS periods need to be determined. To a large extend this duration will depend on:

- The "Busy medium"-off detection tolerance (or sync tolerance).
- The medium propagation delay.
- The energy/signal detect time.

For the energy/signal detect time, only detection on one antenna needs to be accounted for. The effect of this would be that there is a probability that a PaS signal in a fade may not be detected. The effect of this would be that there is a small probability that a station with low priority traffic may contend with high priority traffic. This unintentional leakage between priority levels is considered acceptable.

Performance aspects

The proposed priority mechanism "CMSA with Active Priority Signalling - CSMA/APS" has better priority separation characteristics as the method originally proposed in [4].

That method mainly relied on difference in IFS timing between the different priority levels. This generates extra overhead for the lowest priority traffic, so for the mostly used Asynchronous service. This will especially become more of an issue when more then 2 priority levels are required. Although simulations showed relative low transfer delay impact for "low priority only" traffic, it was pointed out that this effect would become more visible when a buffered load model was assumed.

Although no specific simulations have been performed to date with a fully implemented active priority mechanism, it can safely be assumed that the proposed active priority signalling scheme has less overhead, and will consequently have better throughput and delay characteristics as the method proposed in [4]. This is true for the low priority level, because the priority signalling overhead is less then the extra priority signalling overhead. The higher priority level would comparably have a little more overhead, but its effect on delay and throughput will be very insignificant.

It is however recommended to upgrade the RFMACSIM simulator to include the active priority functionality, to verify the access algorithm, and evaluate the relevant DCF parameters, and the impact of hidden nodes.

Number of Priority levels needed in 802.11

The above description is generic for multiple hierarchical independent priority levels. For an n-level system there will be n-1 PaS positions. Within the standard the number of priority levels will need to be defined such that the different services can be supported.

Two different services are distinguished within the 802.11 MAC, the default Asynchronous service, and the optional Time Bounded Service. In order to support this, at least two priority levels are needed. More priority levels could be required, depending on the different Quality of Service (QoS) classes that are distinguished in the DTBS definition.

Apart from the hierarchical independent priority levels that are obtained by the active priority signalling mechanism, other priority mechanisms can be available.

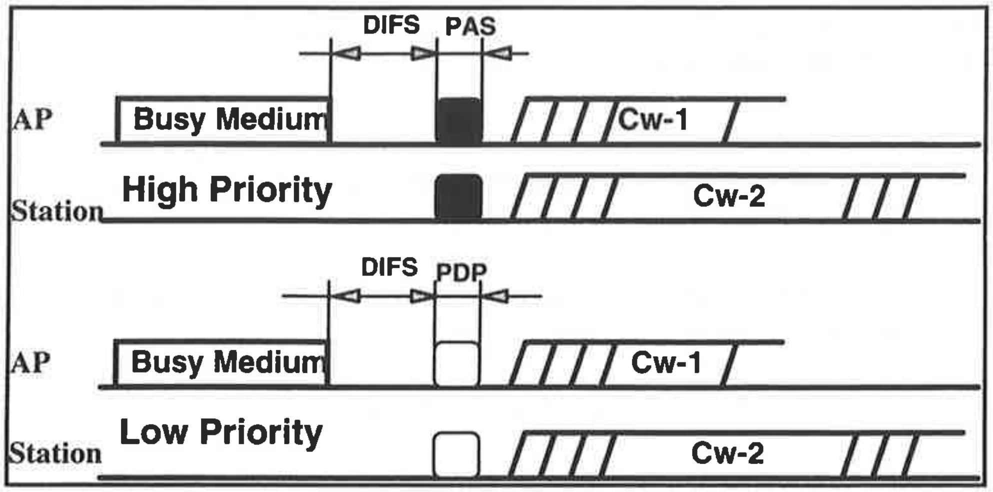
There can be good reasons to assign additional priority mechanisms within the same service level. For instance the Contention Window size (CW) that can be varied, to obtain relative priority differences within one hierarchical priority level.

If hierarchical priority levels are used to separate the Asynchronous and DTBS service, then the CW-size can be used to vary relative priority between different devices. For instance an AP can be given a relative higher priority level for the same kind of traffic then a station. This would make sense because the AP is likely to generate more traffic then individual stations.

The priority scheme would then look as follows:

- DTBS(optional service)	Highest priority	(PDP = 0; PaS = x)
. AP	CW-1	
. Station	CW-2	
- Asynchronous Service	Lowest Priority	(PDP = x; PaS = 0)
. AP	CW-1	
. Station	CW-2	

Please note that the CW values are the initial CW values, which need to be doubled for every re-transmission for a given frame according to the current Foundation MAC definition.



Compatibility:

The proposed priority signalling method has been accepted for the Hiperlan Channel Access Method. The contention resolution methodology is still under discussion. Compatibility between these two standards will allow coexistence between devices of those different standards in the same band. This is considered relevant in a possible future 5.2GHz band in the US.
It is concluded in [2] that the current 1.9 GHz etiquette does not support priority access. In order to support the services defined in 802.11 it is therefore desirable to have an etiquette that can support priority. The authors believe that compatibility between the basic access method of IEEE 802.11 and Hiperlan supporting the required priority access mechanism would be a good basis for upgrading the etiquette for the 5.2 GHz band.

Conclusion:

This document describes a proposal to adopt active priority signalling followed by a binary exponential backoff phase as the priority mechanism to support DTBS.
A system with two hierarchical independent priority levels is expected to be adequate to provide the Asynchronous and DTBS services for 802.11. Further additional relative priority levels can be created that can be used within a service level to give AP's a relative higher priority then stations. This can be benificial because it can be expected that most

traffic will be going through an AP, which will in practise mean that an AP will generate close to 50% of the frames in a BSS.
It is recommended that the RFMACSIM simulator is modified such that this mechanism can be verified by simulations, and to verify the proper parameterization of the access procedure.

References:

[1] DFWMAC Distributed Foundation Wireless MAC Protocol", W. Diepstraten NCR-WCND-Utrecht, G. Ennis Symbol Technologies, P. Belanger Xircom; November 93, IEEE P802.11-93/190. See also P802.11-93/191, P802.11-93/192, P802.11-93/193.

[2] "Evaluation of priority mechanism and Wintech Etiquette compatibility", Wim Diepstraten AT&T-GIS; May 1994 IEEE P802.11-94/116.

[3] "A Channel Access Mechanism for HIPERLAN, CMSA with Active Priority Signalling - CSMA/APS", Simon Black, Tim Phipps Symbionics, Wim Diepstraten, Jan Kruys AT&T-GIS. June 1994 HIPERLAN TTG 94/..

[4] "Priority in CSMA/CA to support distributed Time-Bounded Services", Wim Diepstraten AT&T-GIS; March 1994, IEEE P802.11-94/58.

[5] "Draft Standard IEEE 802.11", P802.11 Editors P802.11-93/20b0.