IEEE 802.11
Wireless Access Method and Physical Specification

Title: Evaluation of priority mechanism and Wintech Etiquette compatibility.

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

Abstract: This paper is evaluating two different implementation idea's for the priority mechanism for supporting Distributed Time Bounded Services (DTBS). It is also evaluated against the Wintech Etiquette rules that apply in the 1.9 GHz unlicensed PCS band.

Introduction:

In the March 94 meeting in Vancouver a proposal was accepted to support Distributed Time Bounded Services (DTBS) based on a priority mechanism in the DCF. A implementation proposal was done in doc. IEEE P802.11-94/58 [1], but acceptance of the approach was postponed until the May meeting to allow investigation of an alternative method. This alternative was briefly addressed in [3] presented by Kerry Lynn (Apple).

This paper is evaluating both methods, by especially evaluating them for compatibility criteria with the Wintech etiquette, because that is one of the concerns that I have with the alternative proposal.

So thereby this paper does implicitly also address Etiquette compatibility issue's of the current proposal.

Current Wintech Etiquette requirements summary.

The goal of the Wintech etiquette that is intended to become the FCC ruling in the 1.9 GHz PCS unlicensed band, is to allow dissimilar systems that are not interoperable to share the band. The intend is that systems with different bandwidth requirements can overlap with each other. The rules address transmit power levels, bandwidth allocation rules and an access mechanism based on a "Listen Before Talk" (LBT) protocol.

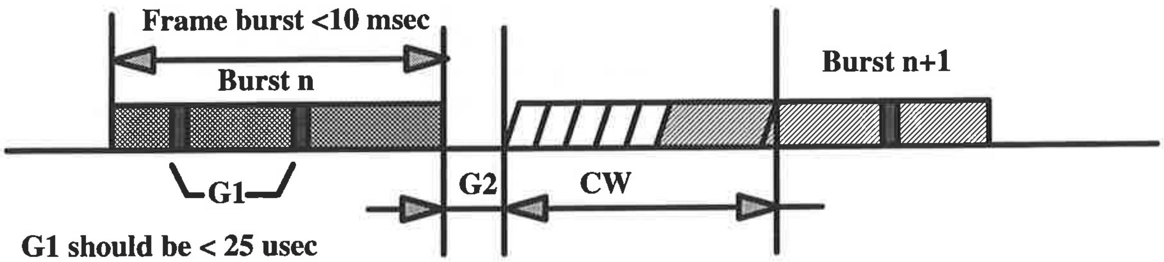
In a nutshell the etiquette access rule requirements for the Async data part of the band are as follows:

The basic access scheme is based on an LBT protocol that uses an energy sensing function, with a threshold that is dependent on the transmit level used for the frame to be transmitted. The energy sensing threshold can be made 1 dB insensitive for every dB that the transmit level is decreased.

The Medium Access Rules are as follows:

- the maximum medium occupancy time of a burst of frames is limited to 10 msec.
- The gaps within the frame burst should be less then 25 usec.
- The gaps between medium occupancy bursts should at initial access attempt be a uniform delay distribution between a minimum of 50 usec and at least 750 usec. After every subsequent attempt the random window period should be doubled up to a maximum of 12 msec.
- The maximum slot time (Tx-on + Energy detect) is not more then 50 usec.

This is represented as follows:



G1 should be < 25 usec

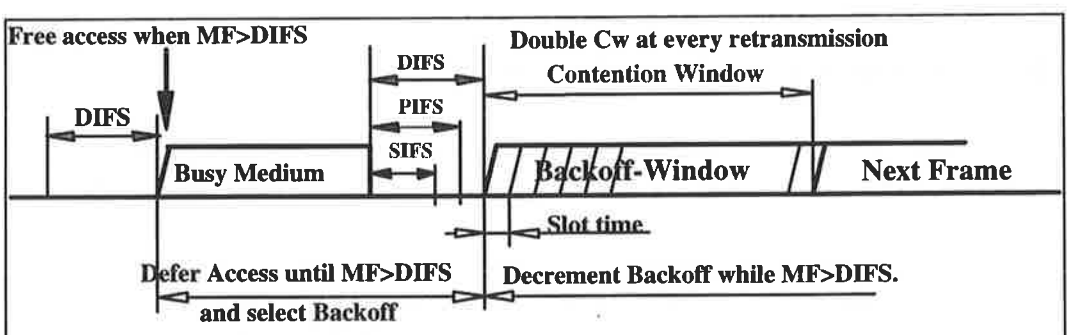
G2 should be >50 usec

CW should initially be 700 usec which doubles on every attempt

Please note that it is allowed that cooperating stations take part in the frame-burst as long as the maximum occupancy duration is not violated, and the minimum gap within the burst elements is smaller then 25 usec. This is done to allow for instance Data-Ack and RTS-CTS-Data-Ack exchanges without interruption.

Current DFWMAC etiquette compatibility:

The DFWMAC access protocol is currently as follows:



CSMA/CA Access Methodology with PCF priority

Please note that direct access is possible when MF (medium free) is longer then the DIFS. So a station does not wait DIFS before it actually does access the medium, but will immediately access the medium when the above condition was satisfied.

The SIFS is being used to generate the Ack, and would be required to be smaller then 25 usec to comply with the etiquette.

The PIFS is specified, but could not be used unless it is also less then the 25 usec specified in the etiquette. The PIFS would typically be used by a PCF that assures that a burst of traffic is generated between different stations, but such that the gap between frames will never be more then 25 usec. The other criteria would be that the burst should be limited to 10 msec duration.

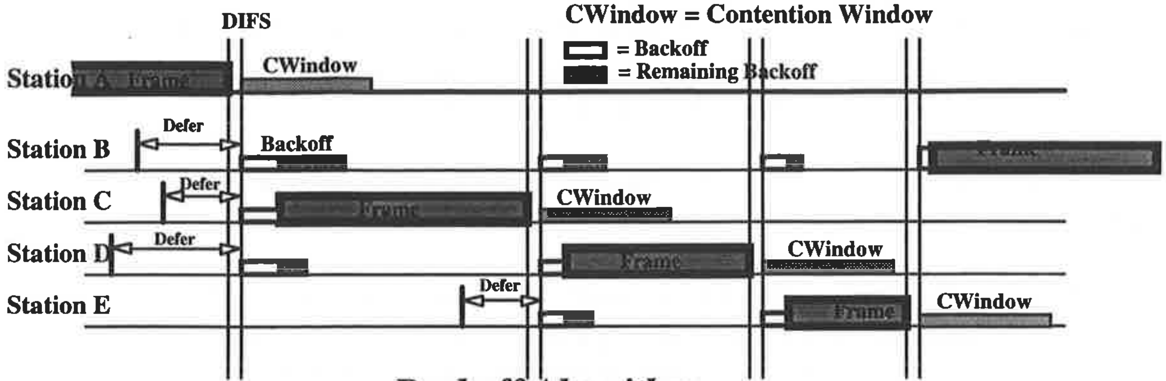
The DIFS is required to be 50 usec minimum according to the etiquette.

The slot time should at least be 50 usec, but can be smaller to improve contention resolution performance.

The contention window should be such that at least a random delay between 50 and 750 usec is generated. This means that the actual CW may be longer. Also the slot time can be smaller then the 50 usec, but the 50 usec is an upper limit that will determine the worst case collision window.
So we could have for instance a parameterization where:

- SIFS <= 25 usec
- PIFS <= 25 usec
- DIFS >= 50 usec
- 1 slot time = 25 usec
- Cw = 32 slots or between 50 and 32*25+50=850 usec.

Please note that there is a subtle difference between the above specified access algorithm and the etiquette required behaviour. The difference is that the backoff delay is only decremented under the condition that MF>DIFS. The effect is that stations that have already deferred for a busy medium (have contended before) have a relative higher access probability. This backoff mechanism is illustrated in the next picture:



Backoff Algorithm

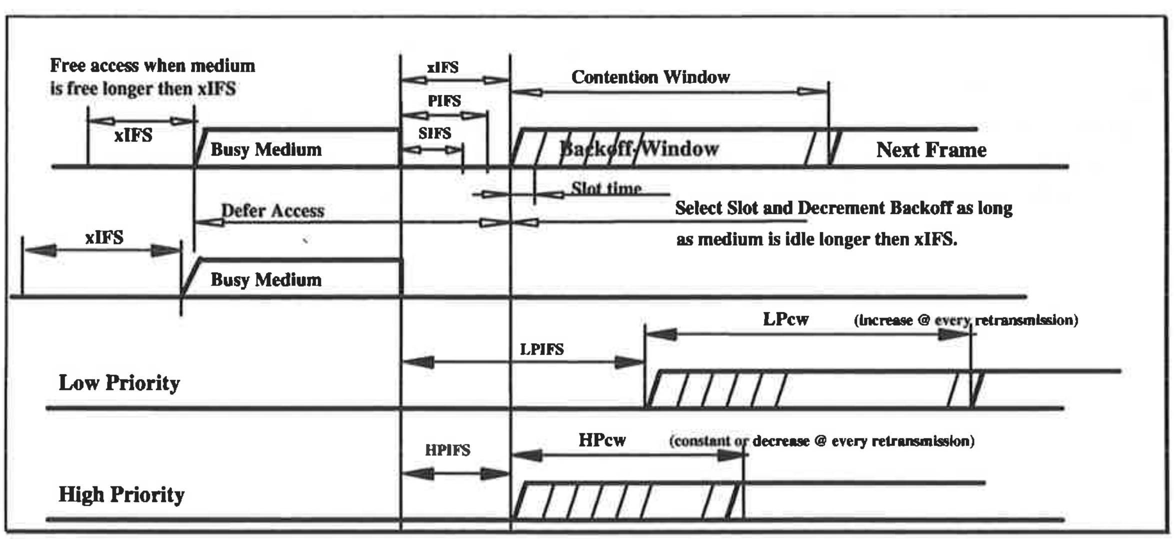
So the practical procedure is that a station will select a backoff delay when it finds MF<DIFS at initial access time, and will decrement its backoff delay whenever MF>DIFS at every slot, and will access the medium when the backoff counter becomes zero. When that transfer is not successful, and the frame needs to be retransmitted, then the Contention Window CW will be doubled.
Please note that no new delay will be generated when the contention is lost and the station is again deferring. This seems to be the main difference with the etiquette.

Question is whether this would be acceptable external behaviour, or whether we need to change the algorithm for 1.9 etiquette compliance?

Please note that the original WMAC proposal did have an identical backoff behaviour apart from parameterisation.

Priority mechanism effect:

The Asynchronous access protocol with priority would look as follows, according to our latest proposal documented in IEEE P802.11-94/58 [2].



Typically the relative priority would be controlled by the overlap ratio of the LPcw and HPcw, which is controlled by the value of LPIFS, and the CW value's. Suggested value's are:

- HPIFS = 3 slots	HPcw = 32 slots	High priority DTBS
- LPIFS = 3+16=19	HPcw = 32 slots	Medium priority Async (from AP)
- LPIFS = 3+16=19	HPcw = 64 slots	Low priority Async (station)

The main issue/comment with the above algorithm is that the majority of traffic will be Async traffic, which would suffer from extra delay due to a larger LPIFS in the LP or MP only traffic case. My simulations do however show only minor impact on delay and throughput, as can be seen in doc. 58.

Priority mechanism compatibility with Etiquette.

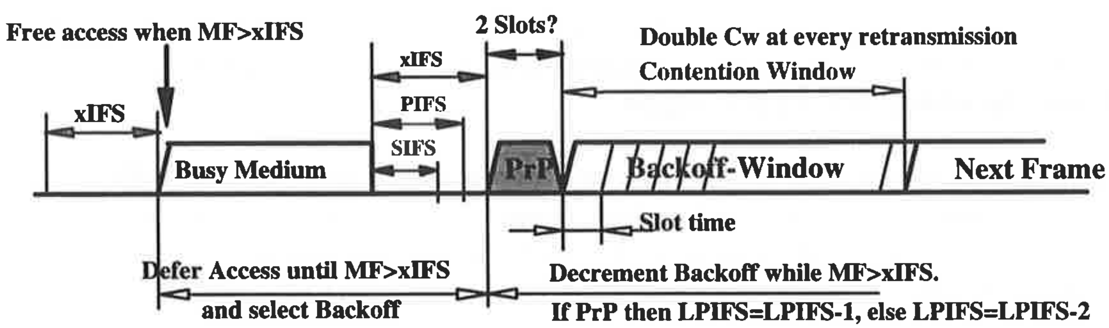
In an Etiquette configuration, the high priority traffic will get fair access with the parameterization as discussed above, but the Medium and Low priority traffic will get lower priority also compared to the rest of the dissimilar traffic following the etiquette.

It should be understood however that this can not be avoided, because the etiquette has no priority mechanism. An 802.11 system can be made compliant with the Etiquette, to allow operation in the 1.9 PCS band, but a choice may be that priority traffic is not supported.

Evaluation of an alternative priority mechanism.

The alternative modification of the priority mechanism that is under evaluation is something like the following:

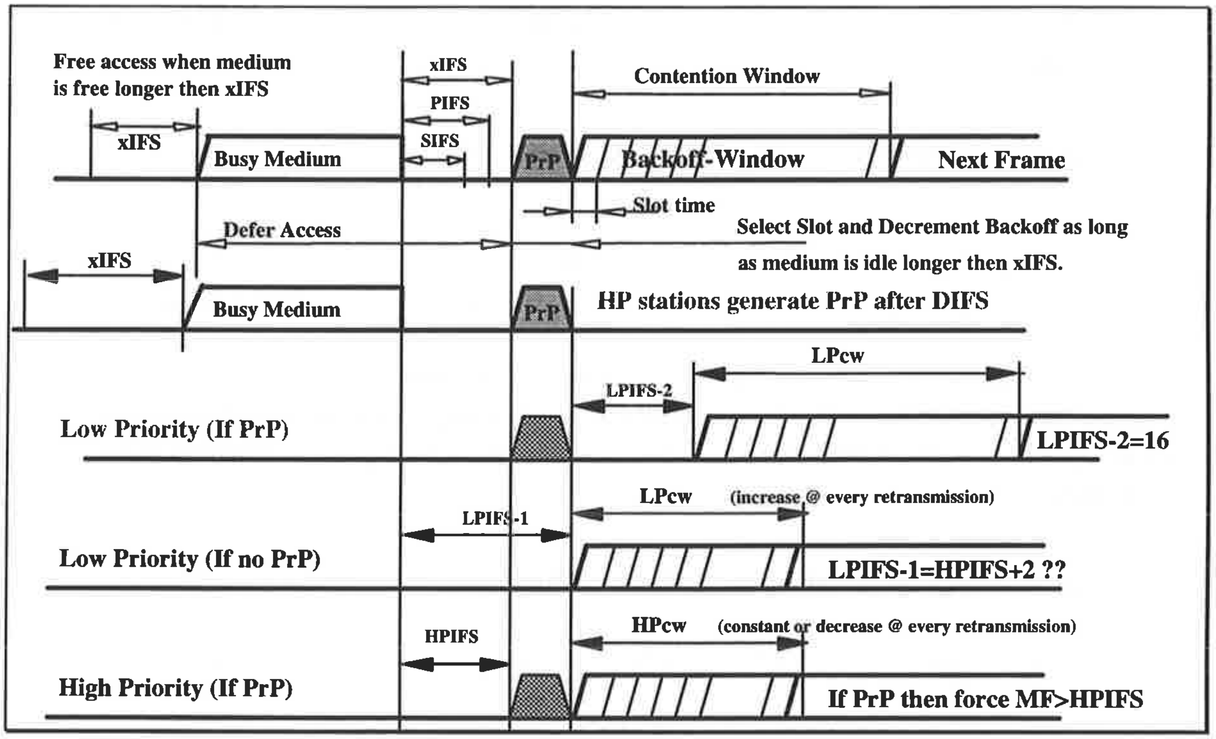
PrP=Priority request Pulse



CSMA/CA Access Methodology with "Priority request Pulse"

The actual algorithm still has to ripe, but the idea is that all Priority traffic will access the medium immediately when MF>xIFS. When MF<xIFS then stations will defer until after the xIFS. Stations with High Priority traffic will generate the PrP (Priority request Pulse of modulated carrier) immediately after the HPIFS.

This will be detected by the low priority stations using the CCA signal from the PHY indicate to all Low priority (and MP) stations that they should use LPIFS-2 (which introduces an offset compared to the HPcw, like 16 or 32). High Priority stations would start their contention window immediately, by forcing the condition MF>xIFS.
The main advantage of this would be that there is less overhead for the LP(and MP)-only traffic, which results in a lower transfer delay.
The behaviour of this scheme with and without the presence of the PrP bit is illustrated below.



If a station that wants to transmit LP traffic detects the PrP, then LPIFS-2 should be used, such that the backoff timer decrement is delayed with LPIFS-2. After the CW period, the LPIFS could be initialised to LPIFS-1 again.

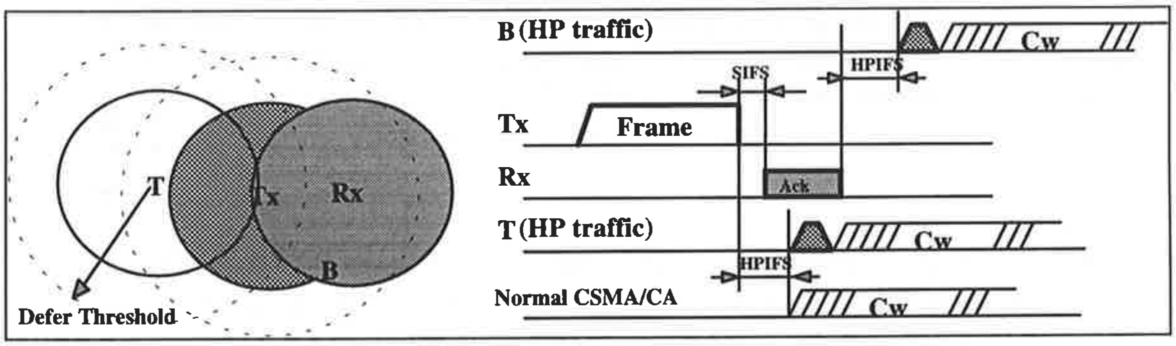
The above scheme would yield priority with only minor added overhead. It still allows instant access when the medium was free long enough at initial access attempt. The PrP only needs to be inserted by the stations having high priority traffic to transmit, and only then the Low priority traffic has to defer an extra amount of time (LPIFS-2).

Issue's related to the PrP bit priority mechanism.

The following are issue's/concerns associated with this scheme:
Issue's:
- The access algorithm is more complicated/less straight forward.

- It is required that all stations are well synchronized to generate and detect the PrP bit. This is an additional requirement for the PHY. The PrP bit can be more then one slot to decrease the synchronization requirement criticality.
- An other type of synchronization is the case where not all stations do see the transmitted frame and the Ack, so that the PrP pulse is not generated at the same time.
- In a hidden node situations, a number of stations may not see the Ack, and start generating the PrP within HPIFS after the transmitted frame, thereby increasing the probability that the Ack is jammed on.

The main issue here is the integrity of the Ack. This will largely depend on the sensitivity level of the defer threshold. The following situation will cause problems:



The above diagram illustrates the situation where a transmitter T that has high priority traffic to transmit does only hear the frame from Tx, and not the Ack from Rx, so it will generate the PrP signal at HPIFS following the previous frame. This will likely jam over the Ack that can therefore not properly be received by Tx. In normal CSMA/CA these situations can also occur, but they will not persist. The probability that this occurs will depend on the defer threshold, and the network load. The probability that T would select a backoff slot that potentially overlaps with the Ack would for instance be 2/32 when the overlap is two slots and the contention window is 32.

An other aspect that is shown in the timing diagram above is that the PrP pulse synchronization is not really possible in wireless. The stations T and B would see a different medium busy status (B does see the Ack, while T does not), and so they will generate the pulse at different times. They will not be detected as such by part of the stations, so the purpose of the PrP is diminished. Again this does not play a major role for the normal CSMA/CA stations that work with different IFS value's to control priority.

Etiquette Compatibility

In the Etiquette, it is allowed that stations that are organized such that they follow the same protocol, are allowed to generate a burst as long as the gap between medium occupation remains smaller then 25 usec. So if the PrP is generated within 25 usec following the previous(802.11) frame then this is allowed. This is apparently only allowed when that frame was positively identified as one that belongs to the same interoperable system.

This can be an issue, because this "own system traffic" recognition can probably only be done up to a "Data sensitivity" range or a little further, but possibly not until the defer threshold, as would be desirable for normal system operation to prevent hidden nodes. The "own system traffic" recognition could be based on modulation detection, or frame detection.
An other issue is that the separation in SIFS, PIFS and DIFS should be all within 25 usec. This requires a fast energy/carrier detector in the PHY, to distinguish between them.
So for etiquette compatibility, the PrP pulse can only be generated when the previous frame was identified as an 802.11 frame, and must then be generated within 25 usec. If the previous frame was not positively identified as such, then the PrP but may not be used, or should be randomized between 50 and 750 usec delay according to the etiquette access rules. Note that this will also occur when the contending system is an 802.11 station that cam in too weak to be identified/recognised as such.

Conclusion:

The conclusion of this evaluation is that the alternative PrP based DTBS priority implementation method has some serious problems associated with it. In addition this method does require additional complexity in the PHY and the MAC.
It also has more problems to be compatible with the Wintech etiquette, that will also increase the MAC complexity. Therefore it is recommended to adopt the priority mechanism as proposed in [2] as the basis for the standard.
The evaluation of the etiquette compatibility shows that there is a mismatch in the backoff mechanism between the adopted DFWMAC method and the etiquette. It needs to be further analysed what the extend of this problem is.
Also it should be understood that the etiquette does not resolve priority. The effect would be that our low priority stations would also have a low priority compared to the other dissimilar system it is competing with. One possible implementation that is valid could be a system without DTBS, which is modelled with only one DIFS value.

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] Priority in CSMA/CA to support distributed Time-Bounded Services, W. Diepstraten AT&T-GIS-WCND-Utrecht; March 94, IEEE P802.11-94/58.

[3] A Distributed Time Bounded Service", Philip Rakity, Larry Taylor, Kerry Lynn Apple Computer; January 1994, IEEE P802.11-94/21