A Robust Service for Delay Sensitive Applications on a WLAN

Fanilo Harivelo and Pascal Anelli

IREMIA, Université de La Réunion BP 7151, 15 Avenue R. Cassin, 97715 Saint
Denis Messag 9, France

Abstract. Technological advances in mobile terminals and the large spreading of Internet have led to the growing need of a certain level of a quality of service for the applications. Wireless networks characteristics make this task difficult. Thus, the classical protocols and models of QoS became inaccurate in this type of networks. This article presents a mechanism that guarantees a service corresponding to the Expedited Forwarding PHB (Per Hop Behavior) in a wireless network. Simulations under NS-2 are carried out to evaluate the performances of the solution.

1 Introduction

With the proliferation of mobile terminals and the popularity of Internet access, the IEEE 802.11 Working Group has proposed a standard [1] for wireless local area networks. It proposes two access methods: DCF (Distributed Function Coordination) and PCF (Polling Function Coordination). DCF is available in infrastructure mode as well as in ad hoc mode and is based on CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) method. Before initiating a transmission, a station senses the medium and executes an exponential backoff algorithm to avoid collisions. With DCF mode, no priority exists among the stations. Besides, a station with a low transmission rate, while capturing the channel, can penalize the other stations on the long run [2]. PCF method tackles with delay sensitive data transmissions and is limited to the infrastructure mode. In PCF, time is divided into superframes. A superframe consists of a period called CFP (Contention Free Period) during which the coordinator, generally the access point, polled each station if it has packets to send, and a CP (Contention Period) period during which DCF mode is used as access method. PCF is complex and some ambiguities remain in its specification. This article proposes a service for delay sensitive application aiming to support flows marked as EF (Expedited Forwarding) according to DiffServ Architecture [3]. This service is provided in a wireless network without access point. In the following, the considered network consists of stations having the same diffusion domain and the hidden station problem is supposed solved by a mechanism such as RTS/CTS. Section 2 gives a state of the art of QoS in wireless networks. Section 3 details the proposed architecture, which will be validated in section 4. The results are summarized in section 5.

2 Related Works

Many studies have been drawn to introduce QoS in the wireless networks. The IEEE 802.11e working group has defined improvements [4] to IEEE 802.11 standard which introduce two new access methods, namely, EDCF (Enhanced DCF) and HCF (Hybrid Coordination Function). In EDCF which derives from DCF, QoS is obtained by the use of eight levels of TCs (Traffic Categories). At the MAC level, the packets are transmitted via separate instances of the backoff algorithm, each instance having parameters set according to the priority level. Although EDCF ensures a better service for higher priority traffic, it does not offer any quantitative guarantee. Moreover, under high load, many collisions may occur even for the priority traffic. HCF function adopts the same principle as PCF and allows a hybrid coordinator, localized generally at the access point, to poll the stations having priority traffic for CFP period. Some of the drawbacks of PCF remain with HCF. To mitigate these shortcomings, [5] proposes a mechanism derived from EDCF, called AEDCF (Adaptive EDCF), which takes into account the contention level of the channel. AEDCF adjusts the size of the contention window and the persistence factor according to the number of collisions. [6] introduces a solution to support the real-time traffics. CFP Period will be used for the transmission of real-time traffic while the CP period is exclusively reserved for the Best-Effort traffic. To ensure a better bandwidth usage and to avoid starving the Best Effort traffic, [7] introduces the concept of free space which defines the unused bandwidth by the privileged traffic, that can be recovered by the Best Effort traffic. To a privileged packet can be piggybacked lower priority packets sharing the same next hop. [8] presents an architecture supporting EF and AF PHB (Per Hop Behaviors). EF PHB is ensured by allocating a low IFS to the corresponding stations. To alleviate the contention among EF flows, two jamming sequences are transmitted by each EF station. That which has the longest sequences will access the medium. [9] proposes to support EF PHB in a wireless network This approach consists in setting a map of the bandwidth usage in the network using an exchange of messages and in deducing the local BE traffic rate. For a better use of the resources, the unused bandwidth by the EF traffic is recovered by the BE traffic. The delay constraints are ensured by anticipating possible load increases by the introduction of a thresholds system that allocates a bandwidth margin to EF traffic.

3 Wireless bandwidth access control

Our proposition stands for a limitation of the BE traffic of the network, that aims to guarantee initially fixed bandwidth and delay for the EF class of service. This restriction is done on the basis of the network state and require neither any exchange of messages nor the knowledge of the traffic of the other nodes. Indeed, this is used to prevent the network from congestion and to avoid overload due to signaling mechanism. The traffic control is conveyed to the policy layer in such a way that no queueing delay will be induced to the currently transmitted frame at the MAC layer. A congestion avoidance scheme is, then, applied. A control function computes the sending rate of the local BE traffic. The EF traffic profiles are distributed to the appropriate stations. The method to distribute these profiles is out of the scope of this paper. However the adopted policy must take care not to exceed a certain ratio of the bandwidth [10]. The conformance to the profile is done using a token bucket and the excess traffic will be dropped. The suggested solution is localized between the MAC layer and the IP layer as shown in the Fig. 1. Each station implements this architecture.

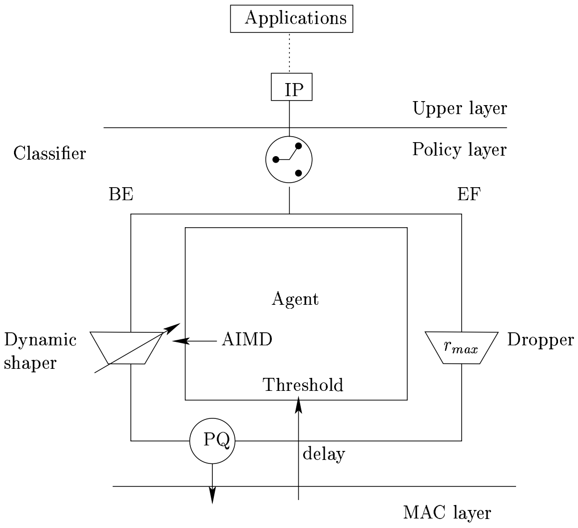


Fig. 1. Global architecture

The information obtained from the MAC layer will be used to determine the network state. This disposal is taken to make it possible for the architecture to deal with wireless network characteristics, while allowing the possibility of combination with a MAC level mechanism to accentuate the service differentiation. Indeed, the QoS support provides at the MAC level tackles with the choice of the node which will acquire the medium while an IP level solution defines the packet which will be transmitted within a node [11]. EF and BE packets are handed over to the MAC layer according to a PQ (Priority Queue) scheduling. The BE traffic limitation is done using a dynamic shaper whose parameters result from a congestion avoidance mechanism. This mechanism is highly interrelated with the control function used by the station to increase or to decrease its BE traffic and it is comprised in the agent localized in each station in such a way that each one reacts in the same manner depending on the network state. The agent estimates the network state and allocates the maximum BE sending rate of the station to ensure a high bandwidth usage while guaranteeing low delays. The packets are classified thanks to the DS field of the IP header.

The congestion avoidance control consists of a thresholds system similar as that of [12] and a binary feedback. The network state is provided, periodically every Δt, by the thresholds system. This information is deduced from the response time of the MAC level, i.e. the delay d taken by a packet to be transmitted, and the initially guaranteed MAC level delay d_{max}. The network load is estimates using the percentage $\hat{\delta} = 100\frac{d}{d_{max}}$. A binary feedback (0 or NC for not congested, and 1 or C for going to be congested) is determined by the stations, so that they can adjust (increase or decrease) their rate r_{BE}, via a control function. If this feedback estimates that the network is not congested, then, the BE traffic rate can increase, otherwise, the BE traffic rate is decreased. The choice of binary feedback is motivated by its simplicity and its efficiency for the resource controller. The thresholds system is used to bring up the measured load so that the network can act before the maximum delay is reached. So the network never enters in congestion. The delay is calculated on the packets successfully received and corresponds to the duration from the handling of the packet by the MAC level and the receipt of the acknowledgment. By considering the model of Fig. 2, let $\hat{\delta}$ equals 75 and the previous state, NC, then the current state will be C, corresponding to a congested network.

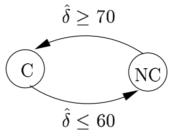


Fig. 2. Thresholds System with 2 states

However, a question arises on the way by which each station lower its rate in the case of congestion. Indeed, the flows having a high sending rate must decrease more their rate compared to the small flows, in other words, the reduction of the rate must be proportional to the rate. This is done by choosing a multiplicative function for the decrease. A similar consideration has to be made regarding the increase. The fairness constitutes the only condition required for BE traffic. The sharing of the bandwidth must be fair among the stations generating BE traffic and independent of the rate currently generated by each source. An additive function is appropriate for the increase in the rate. The choice of AIMD (Additive Increase Multiplicative Decrease) is judicious insofar as [13] shows that this algorithm ensures fairness and convergence. The AIMD control function is summarized in the following expression:

$$r_{be}(t) = \begin{cases} r_{be}(t - \Delta t) + r_{AI} & \text{if} \quad state = NC \\ r_{be}(t - \Delta t)/r_{MD} & \text{if} \quad state = C \end{cases} \tag{1}$$

in which r_{AI} et r_{MD} correspond respectively to the increment value and the decrease factor of BE traffic rate.

4 Performance evaluation

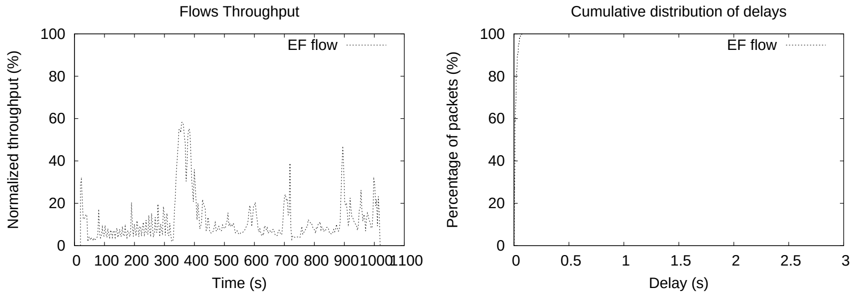
The evaluation of the proposed mechanism was carried out with the NS-2 simulator in a IEEE 802.11 network comprising 6 stations, one of which is used as the destination for overall traffic. The capacity of the medium is set to 1 Mbits/s. 4 nodes generate BE UDP traffic with a packet size equals 512 bytes and a rate of 400 kbits/s. One of the nodes moves during simulation and becomes out of reach of the other nodes. The EF traffic consists of an MPEG encoded movie.

Three cases are considered:

- The first case evaluates the performances of the EF flow when it is the only one being in activity.
- The second case defines a common scenario where BE sources transmit until making network congested.
- The last case corresponds to the use of the proposed mechanism in the previous case to ensure a service to EF flow. The maximum delay imposed for packets at the MAC level is set to of 0.056 s with $\Delta t = 40$ ms. This choice is based on previous works [9]. The increase is done by an increment of $r_{AI} = 400$ bits/s and the decrease by a ratio of $r_{MD} = 1.5$, derived from empirical considerations.

The results are summarized in the table 1 and the curves Fig. 3, 4, 5. In the first case, the bandwidth required by MPEG flow is granted (Fig. 3a) and the delay is low (Fig. 3b) with a maximum value of 80 ms. No packet dropped because of the absence of contention on the medium. In the presence of BE flows, the constraints in term of bandwidth (Fig. 4a) and delay (Fig. 5a) are not satisfied anymore. The EF packets delay are high with a maximum value of 1700 ms, while the flow experiences significant delay variation. However, the bandwidth usage is high, but this causes a large number of collisions. The proposed mechanism respects the constraints in term of bandwidth (Fig. 4b) and delay (Fig. 5b). Indeed, the maximum delay perceived by EF flow equals 130 ms while delay variation (20 ms against 110 ms in the first case) remains low. The rate control handles correctly the abrupt increases in the load of the EF flow, as that occurring at $t = 320$ s. The EF traffic is completely isolated from BE flows. The bandwidth usage has been reduced with an average value of 46.63%. A higher EF load would lead to a better utilization ratio because the maximum value reaches 77.87% vs 76.23% in the case without QoS. Besides, the number of collisions decreases significantly (1716 vs 33505).

Thus, the proposed mechanism provides a QoS for an EF real VBR flow while offering fairness for the BE flows.



(a) Case 1: EF flow throughput/Time

(b) Case 1: Distribution of delays/Delay

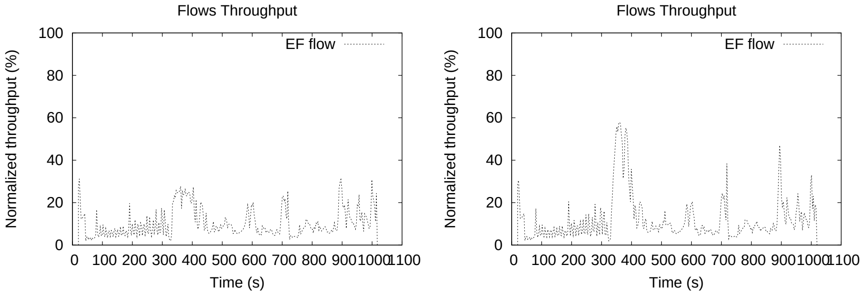
Fig. 3. Throughput and delays curves in the first case

Case	EF alone	WLAN	WLAN + QoS
Bandwidth usage (%)	12.23	69.8	46.63
Max bandwidth usage (%)	58.30	76.23	77.87
# Collisions	0	33505	1716
# Transmitted packets	21267	191650	124493
Exchanged bytes (MB)	14.52	100.23	66.90
# Dropped packets	0	229490	45600
Max EF delay (ms)	80	1700	130
Mean EF delay (ms)	15	90	20
EF standard deviation delay (ms)	13	110	20

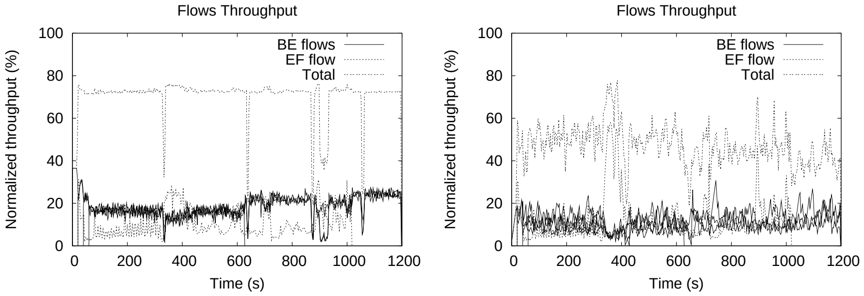
Table 1. Statistics of the networks in the three cases

5 Conclusion

This article presents a robust mechanism for the support of EF PHB in a wireless network and the bandwidth sharing among the BE traffic. The principle consists in avoiding the network to be in a congested state. That is done by the restriction of the BE traffic on the basis of an estimation of the network state thanks to local information, namely, the MAC level delay. The BE traffic rate is decreased or increased according to whether the network is in a congested state or not. A maximum delay, below which a certain level of service can be assured, is initially fixed. To prevent abrupt increase in the load of EF traffic, a thresholds system is set up in order to put a margin on the MAC delay increase. Simulations show that the EF traffic is completely isolated from the BE traffic. The principal advantage of the proposal lies in its ease of implementation and the absence of overload: no signaling is needed. The mechanism works in a totally distributed mode, thus, the motion of a node does not affect the way the other nodes perform



(a) Case 2: EF flow throughput/Time (b) Case 3: EF flow throughput/Time



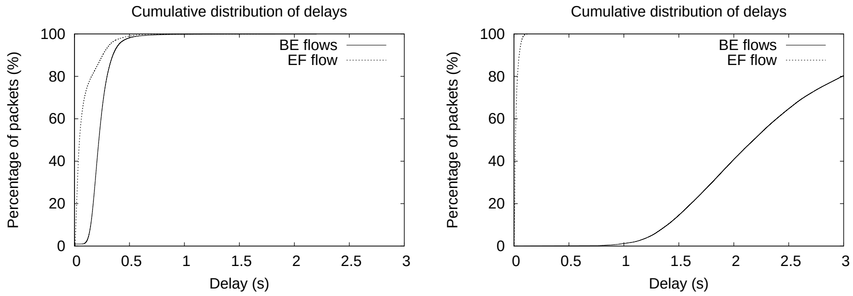
(c) Case 2: Flows throughput/Time (d) Case 3: Flows throughput/Time

Fig. 4. Throughput curves in the second and the third cases

their computation. However, the performance of the mechanism can be improved by combining it with a MAC level solution such as IEEE 802.11e. A better bandwidth usage can also be obtained by making the increase and the decrease factor of the BE load variable with the MAC delay. This can be done by using much richer feedback for the congestion avoidance mechanism. A study on the contribution of the solution in the bandwidth allocation in the hidden station case will also be undertaken. Finally, the support of AF PHB would constitute an additional extension of this architecture.

References

1. IEEE 802.11 WG: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. IEEE standard 802.11, 1999 Edition (1999)
2. Heusse, M., Rousseau, F., Berger-Sabbatel, G., Duda, A.: Performance anomaly of 802.11b. INFOCOM'03, San Francisco, USA (2003)



(a) Case 2: Distribution of delays/Delay

(b) Case 3: Distribution of delays/Delay

Fig. 5. Delays curves in the second and the third cases

3. Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., Weiss, W.: An Architecture for Differentiated Services. RFC 2475, IETF (1998)
4. IEEE 802.11 WG: Draft Supplement to Standard for Telecommunications and Information Exchange between Systems-LAN/MAN Specific RequirementsPart 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC), Enhancements for Quality of Service (QoS). 802.11e Draft 4.1 (2004)
5. Romdhani, L., Ni, Q., Turletti, T.: Adaptive EDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad-Hoc Networks. IEEE WCNC'03 (Wireless Communications and Networking Conference), New Orleans, Louisiana, USA (2003)
6. Velayutham, A., Chang, J.M.: An Enhanced Alternative to the IEEE 802.11e MAC Scheme. www.cs.iastate.edu/ vel/research/E-802.11.pdf (2004)
7. Liu, W., Fang, Y.: Courtesy Piggybacking: Supporting Differentiated Services in Multihop Mobile Ad Hoc Networks. INFOCOM'04 (2004)
8. Banchs, A., Radimirsch, M., Pereza, X.: Assured and Expedited Forwarding Extensions for IEEE 802.11 Wireless LAN . In IWQOS'02 International Workshop on Quality of Service, Monterey, California, USA (2002)
9. Harivelo, F., Grand, G.L., Anelli, P., Wolf, J., Wolfinger, B.: Expedited Forwarding for WiFi. International Symposium on Wireless Communication Systems (2004)
10. Heusse, M., Starzetz, P., Rousseau, F., Berger-Sabbatel, G., Duda, A.: Bandwidth allocation for DiffServ based quality of service over 802.11. Globecom 2003, San Francisco (2003)
11. Le Grand, G., Meraihi, R., Tohmé, S., Riguidel, M.: Intelligent Ad Hoc Networking to Support Real Time Services. VTC, Vehicular Technology Conference (2003)
12. Wolfinger, B., Wolf, J., Le Grand, G.: Improving Node Behavior in a QoS Control Environment for Local Broadcast Networks. Proceedings of SPECTS 2003, Montreal **35** (2003) 361–371
13. Chiu, D.M., Jain, R.: Analysis of the Increase and Decrease Algorithm for Congestion Avoidance in Computer Networks. Computer Networks and ISDN Systems **17** (1989) 1–14