Load balancing approach for wireless IEEE 802.11 QoS enhancement

Issam JABRI[1,2], Nicolas KROMMENACKER[1], Adel SOUDANI[2], Thierry DIVOUX[1]

[1]CRAN (CNRS ESA 7039), UHP Nancy I, BP 239, 54506 Vandoeuvre, France
{issam.jabri, Nicolas.krommenacker, Thierry.divoux}@cran.uhp-nancy.fr
[2]Laboratoire EµE, Faculté des Sciences de Monastir, 5019 Monastir, Tunisie
adel_soudani@yahoo.fr

Abstract. In the few last years, the deployment of IEEE 802.11 WLAN in hotspots environment had becoming a useful solution providing practical and attractive communication characteristics. However the problem of user bandwidth availability arises as one of the most limit of this solution. In fact, the IEEE 802.11 standards do not provide any mechanism of loading distribution among different Access Points of the network. Then an AP can be heavily overloaded leading to station throughput degradation.
This paper deals with this problem. It focuses on the presentation of QoS (Quality of Service) management solution for wireless communication system. It, mainly, presents a protocol structure between mobile and AP. This protocol is intended to provide best resources allocation and efficiency on communication metrics. An SDL description and MSC simulation is provided as a first step in the development of this protocol.
Key words. Load Balancing, QoS Protocol, Wi-fi, SDL.

1   Introduction

In the last few years the IEEE 802.11 technology becomes very interesting. One of its popular uses is its cheap hardware infrastructure price promoting to provide practical and efficient Hotspots environment [1].
The research work [1, 2, 3, 5] carried in this context had proved that additional effort is yet required to build up a system with a high service quality. A specification of further interaction in the IEEE 802.11 protocol between AP and the mobiles mainly in call admission will help to ensure some QoS parameters such as loading distribution and packet losses [1, 2, 4, 19, 20]. A new standard has IEEE 802.11e [21] has been defined to ensure quality of service in Wireless LAN.
This paper presents a protocol specification managing the QoS in the context of Hotspots communication environment. The first part presents an overview of the actual quality of service mechanisms for the IEEE 802.11 wireless LAN. The second part focuses on the description of the general hotspots environment architecture. The third part presents the definition of new protocol primitives between the mobile and the access point managing QoS metrics. Then we present the description of this protocol with the SDL (Specification and Description Language) language [8, 9] and some MSC (Message Sequence Charts) simulation results of the behaviour of this protocol. We finish by highlighting future contributions in this field.

2   Overview of QoS mechanisms for IEEE802.11 wireless LAN

1.2   QoS limitations of IEEE802.11 wireless LAN

The most important functions of a wireless MAC layer include controlling channel access, maintaining Quality of Service, and providing security. Wireless links have specific characteristics such as large packet delay and jitter, high loss rate, bursts of frame loss, packet reordering. Furthermore, the wireless link characteristics are not constant and vary over time and place [12]. Mobility of users may cause the end-to-end path to change when users are roaming. Users expect to receive the same QoS once changing their point of attachment. This implies that the new path should also support the existing QoS, and problems may arise when the new path cannot support such requirements [12]. The original IEEE 802.11 networks (DCF) are best effort networks and do not support QoS for time critical applications [22]. All stations in a BSS have the same priority to access the channel. There are no differentiation mechanisms to guarantee bandwidth, packet delay or jitter for high priority stations with times-bounded applications or multimedia flows. In [12], authors have make simulations on an ad hoc topology in which stations transmit three types of traffic (audio, video and background traffic) to each other. These simulations clearly shows that there is no throughput or delay differentiation between different flows since only one queue is shared by all the three flows, thus they all experience the same delay. So, there is no way to guarantee the QoS requirements for high-priority audio and video traffic unless admission control is used.

A PCF mode has been designed to support time-bounded multimedia applications, but it has many problems that lead to poor QoS performances [22, 17, 18]. In this mode wireless resources are wasted since all communications between stations in the same BSS have to go through the Access Point. This mode must be implemented with the DCF mode. Cooperation between Contention Period and Contention Free Period may lead to unpredictable beacon delays [18]: to switch from DCF to PCF, the wireless channel must be idle. The access point is not authorised to stop an established communication to make on the PCF mode and then we have no guarantee on the DCF mode duration. With PCF mode, it is difficult to an access point to define time needed by

each polled station to transmit data frames. The transmission time of polled stations is difficult to control since the physical rate can be changed according to the varying channel status.

All these limitations for both DCF and PCF led to a large number of research activities to enhance the performance of 802.11 MAC.

2.2 QoS mechanisms for IEEE802.11 wireless LAN

Most existing QoS mechanism for 802.11 can be classified into three categories [23]:

Service differentiation. basically, service differentiation is achieved by two main methods: priority and fair scheduling [24]. While the former binds channel access to different traffic classes by prioritized contention parameters, the latter partitions the channel bandwidth fairly by regulating wait times of traffic classes in proportion according to given weights [23]. Used parameters for both approaches are CW size, backoff algorithm and inter frame space. The main service differentiation mechanism is the upcoming 802.11e standard. A new access method called Hybrid Coordination Function (HCF) is introduced. It is a queue-based service differentiation that uses both DCF and PCF enhancements. HCF describes some enhanced QoS-specific functions, called contention-based HCF channel access and polling-based HCF access channel. These two functions are used respectively during both CP and CFP for transfers with QoS. Enhanced DCF (EDCF) is the contention-based HCF channel access. The goal of this scheme is to enhance DCF access mechanism of IEEE 802.11 and to provide a distributed access approach that can support service differentiation. The proposed scheme provides capability for up to eight types of traffic classes. It assigns a short CW to high priority classes in order to ensure that in most cases, high priority classes will be able to transmit before the low-priority ones. For further differentiation, 802.11e proposes the use of different IFS set according to traffic classes. Instead of DIFS, an Arbitration IFS (AIFS) is used. Classes with smallest AIFS will have the highest priority.

Admission control and bandwidth reservation. service differentiation is helpful in providing better QoS for multimedia data traffic under low to medium traffic load conditions. However, due to the inefficiency of IEEE 802.11 MAC, service differentiation does not perform well under high traffic load conditions [17]. In this case admission control and bandwidth reservation become necessary to in order to guarantee QoS of existing traffic. These two approaches are quite difficult to realise due to the nature of the wireless link and the access method. Admission control schemes can be broadly classified into measurement-based and calculation-based methods. In measurement-based schemes, admission control decisions are made based on the measurement of existing network status, such as throughput and delay. On the other hand, calculation-based schemes construct certain performance metrics or criteria for evaluating the status of the network [23].

Link adaptation. 802.11 specify multiple transmission rates but it intentionally leaves the rate adaptation and signalling mechanisms open. Since transmission rates differ with the channel conditions, an appropriate link adaptation mechanism is desirable to maximize the throughput under dynamically changing channel conditions [23]. Most link adaptation mechanisms focus on algorithms to switch among transmission rates specified in the Physical Layer Convergence Procedure.

These different mechanisms aiming to enhance the quality of service support in the IEEE 802.11 wireless LAN treat the network locally. For example parameters differentiations are made at the node level. The keystone of our approach is to consider the wireless LAN as a hole. We try to make a fair distribution of the load among overlapping cells. So we can fulfil n increasing number of accepted applications with guaranteed quality of service level.

3 General approach presentation

The QoS management on hotspots environment becomes vital as many new emerging applications such as mobile information access, real time multimedia communications, networked games, immersion worlds and cooperative work require a minimum level of QoS [12, 13, 17, 18].

The hotspots environment can be described as a set of access points covering overlapping cells and offering connection to a variable number of mobile stations. User's applications are not similar in terms of QoS requirements so that a fair distribution of the mobile stations among active access points can guarantee a minimum level of quality of service. The bandwidth effectively offered by an access point is given by the following formula (Shanon). BP is the bandwidth defined by the IEEE 802.11 standard.

$$Cmax = BP \times \log_2 (1 + SNR)$$

So, due to the wireless environments (interferences, obstacles…) bandwidth is scare and channel conditions will be time-varying and sometimes highly lossy.

Unfortunately, in the actual IEEE 802.11 protocol, a mobile station is associated to the access point offering the best Signal by Noise Ratio (SNR) independently of the load being applied to the access point by other users.

This can cause, in many cases, unbalanced load between access points. Some access points will be over loaded, others are under loaded. For the first ones applications requirements are not fulfilled. The keystone of our approach is to associate mobile station to access points with a minimum SNR threshold and offering the best QoS level.
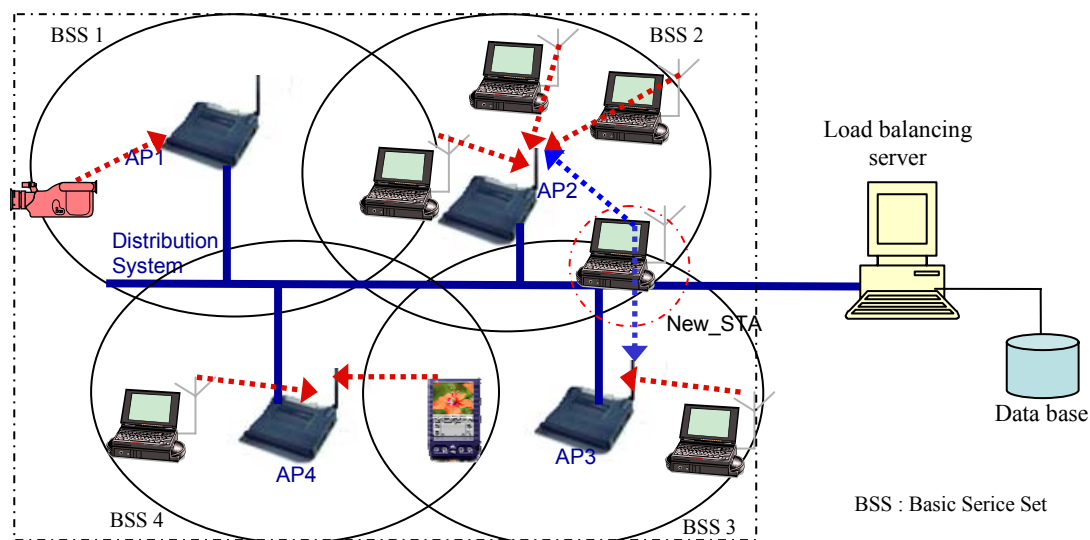

Fig. 1: IEEE 802.11 target architecture

Figure 1 sums up the idea we develop in this paper. A new mobile station (New_STA) reaching the WLAN must be associated with an access point. The association procedure is always initiated by the station (mobile-controlled handover) and the station can be associated with only one access point. The New_STA must discover which access points are present and then requests to establish an association with one of them. Thus, first the station initiates a scanning process that can be either active or passive [6]. Once the scanning process has finished, the station has an updated list of access points in range (AP2 and AP3).

This information is used by the station to associate with the access point that provides the higher Signal-to-Noise Ratio (SNR).

Let's suppose that AP2 is chosen by New_STA. The load distribution across access points will be highly uneven [1]. This can cause a performance degradation perceived by the other stations attached to AP2. Quality of service contracts (bandwidth, loss rate…) may be violated. It will be attractive to associate the new host to AP3 which has lower SNR but is under loaded. So that we fulfill the QoS requirements of both old and recent associated stations: the available bandwidth of the WLAN link depends strongly on the number of active stations and their traffic. To achieve this balancing, in terms of quality of service offered to the stations (load, loss rate…) among access points, we have to compute a balancing algorithm each time a new event such as the arrival of new stations or the mobility of existing stations. This algorithm has to find the best state of associations between access points and mobile stations that offers the best quality of service level for user's applications. Thus, we have to get information on associated stations, traffic coursed by access points and users quality of service requirements (Figure 2). This information has to be exchanged between WLAN entities and stored in an updated data base.
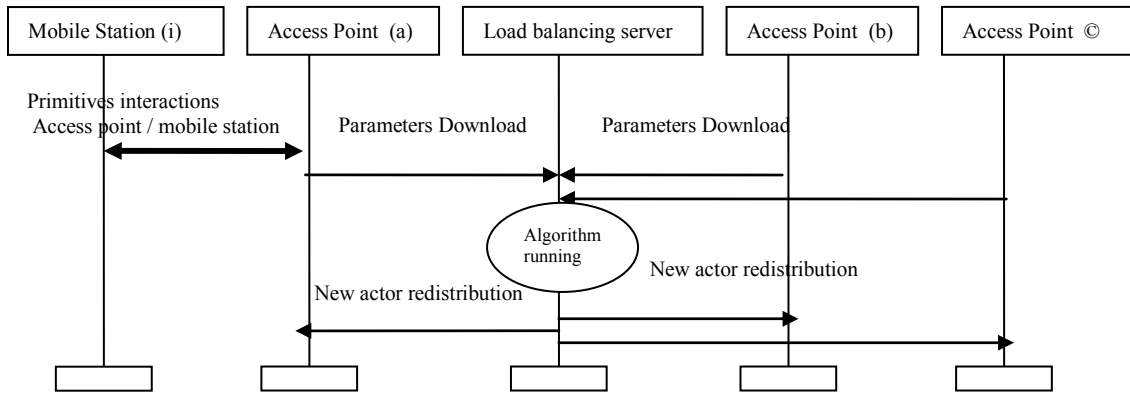
Fig. 2: System protocol interactions

In this architecture the load balancing server should periodically download a set of specific parameters from each access point. It runs up the balancing algorithm finding the best mobile station sharing among the available access points. The result will then be broadcasted in the system. Then, we have defined for this architecture a set of new metrics to quantify the quality of service and primitives exchanging these parameters for association and disassociation between mobile station and access point. These primitives that should be inserted into the MAC layer to improve the IEEE 802.11 standard [14] define a new MAC quality of service policy for wireless LANs.
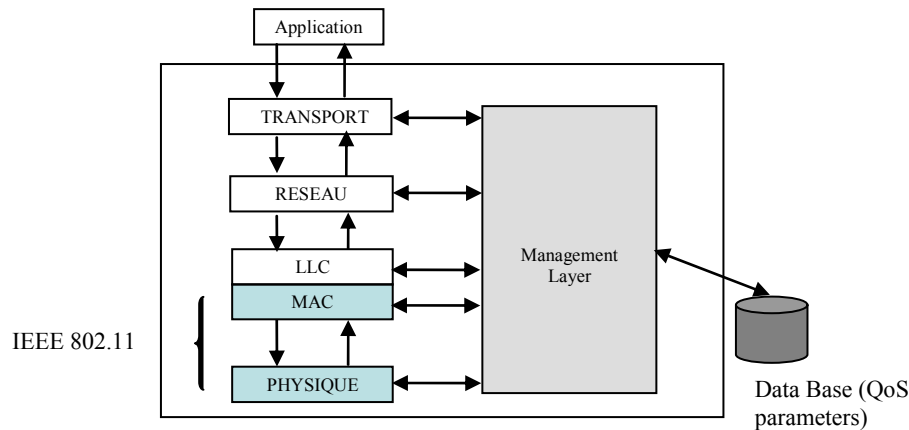

Fig. 3: System exchange design

4   Load balancing algorithm description

The load balancing algorithm [15, 16] is computed by the load balancing server every time a new distribution is needed in the wireless LAN. This will occurred (i) when a new mobile station enters the wireless LAN and aims to associate with an access point, (ii) when an associated station is moving from one to another BSS (iii) and when the applications requirements in a mobile station are changing. The downloaded parameters from the access points and mobile station applications will be useful to find the best distribution of mobile stations among wireless LAN access points.

This algorithm checks if the new distribution is balanced mainly by computing the balance index (β). The balance index appeared in the first time in [11] and it is used in [2, 15] as a performance measure. The balance index reflects the used capacity in each access point. Let Ti be the total traffic of the APi. Then, the balance index is:

$$\beta j = (\Sigma_i Ti)^2 / (n * \Sigma_i Ti^2)$$

βj is the balance index of an overlapping zone j.
   Ti is the total traffic of an APi overlapping with other access points in the zone j.
The proposed distribution of mobile stations is balanced if the balance indexes of all the overlapping cells converge to 1. At this step, the algorithm has to send the new distribution to the access points which will be authorised to dissociate, associate and reassociate mobile stations.

5 Protocol specification

5.1 QoS protocol parameters

In this approach, the QoS management is based on the idea that some added primitives must be ensured at the connection level between the mobile station and the access point. Then, each mobile in the wireless LAN may be able to propose a level of QoS and to modify it when needed.

In this architecture, the mobile station defines four variables managing its QoS state. The communication process will then base its negotiation with the access point on these parameters to build up clause for service quality. Table 1 sums up these parameters and their functions.

Parameter	Function
QoS_{max}	the maximum quality of service that the mobile station can offer to the user
$QoS_{negociated}$	the quality of service used by the mobile station at time t
$QoS_{expected}$	the quality of service wanted by the user or the application
$Old\_QoS_{negociated}$	It is necessary to conserve the old quality of service to make comparisons in case of voluntary changes or new offers of QoS.

Table. 1: Quality of service parameters

The following inequality describes the logical relation between these parameters

$$QoS_{max} \geq QoS_{expected} \geq QoS_{negociated}$$

From the part of the access point, some other parameters must be provided to enable QoS management (Table 2).

Parameter	Function
D_{max}	the higher throughput that can be provided by the access point according to his hardware capabilities
D_{min}	the lower throughput agreed for each user (the Best Effort service)
Da	the reserved throughput, that means the required throughput for a mobile station in an attachment attempt added to the current throughput
D_r	the reserved throughput, that means the required throughput for a mobile station in an attachment attempt added to the current throughput.

Table.2: Access point parameters

So, we can propose rules that enable the management of the stations access according to the requirements and the availability of QoS: $D_r = D_u + QoS_{expected}$

$$D_r < D_{max} - D_{min}$$

5.2 Device identification

In this approach, to ensure QoS management in the WLAN some identifiers should be joined to the parameters describing present and old quality of service states in each mobile. These identifiers are maintained in a specific database both in the access point and the mobile station.

Each mobile station will then discuss the attachment attempt responses of the access point according to its own QoS parameters.

We describe in table 3 these parameters from both the access point and the mobile station point of view.

Parameters	Access point	Mobile Station	Function
My_Id_{AP}	*		The access point identifier
$Id_M(X)$	*		The identifier of mobile station number X
My_Id_M		*	Defines the mobile station identifier
$QoS_{negociated}(X)$	*	*	The Quality of service negotiated with the mobile station X
$Old\_QoS_{negociated}(X)$	*	*	the old level of QoS being agreed for a mobile station number X
St_Moving(X)	*		Describes the state of moving state of the X mobile station
St_Reserved(X)	*		Describes presence state of the mobile station X
Timer(X)	*	*	For actions limited in time

<div align="center">Table 3: New wireless entities parameters</div>

These parameters have to be saved in a specific data base managing the whole environment of the wireless device. This database communicates with the other layers defined in the IEEE 802.11 model to ensure coordination in call admission processes.

6 QoS protocol primitives

The IEEE 802.11 suffers from lack of specific QoS primitives. The only parameter on which is based the connection negotiation between the AP and the mobile station is the SNR ratio. The satisfaction of only this parameter in the connection phase don't meet necessary the QoS requirements of the application. So it appears indispensable to specify new protocol primitives to enable the integration of other communication parameters in the connection decision and then in loading redistribution.

We have then, defined a set of new primitives expressing general requirements.

Primitives	Access Point	Mobile Station	Parameters
ASK_ATTACH.conf	*		$(Id_M, Id_{AP}, QoS_{negociated})$:
ATTACH.conf:	*		$(Id_M, Id_{AP}, QoS_{negociated})$
WAIT	*		(Id_M, Id_{AP})
ASK_RATTACH.req:	*		$(Id_M, Id_{AP},$ liste_d'AP_possibles)
ASK_ATTACH.req		*	$(Id_M, Id_{AP}, QoS_{expected})$
ATTACH.req		*	$(Id_M, Id_{AP}, QoS_{negociated})$
ATTACH.req		*	$(Id_M, Id_{AP}, QoS_{negociated})$
ASK_RATTACH.conf		*	$(Id_M, New\_Id_{AP}, Old\_Id_{AP})$
LEAVE		*	(Id_M, Id_{AP})
OK	*	*	$(Id_M, Id_{APa}, Id_{APb})$
MOD_QoS.req	*	*	$(Id_M, Id_{AP}, QoS_{proposed}, Temps)$
MOVE.req	*	*	(Id_M, Id_{AP})
MOVE.conf	*	*	$(Id_M, New\_Id_{AP}, Old\_Id_{AP})$

<div align="center">Table 4: quality of service primitives</div>

7 Scenario description

The primitives that we defined in the last section are used to manage the access of the mobile stations to the wireless LAN via access points. Then we check theses primitives with many communications scenarios. Communications scenarios vary from simple to much complex situations. In this paragraph we describe one of the scenarios. This example is presented in Figure 5. M4 enter the wireless LAN and aim to connect to AP2 (Figure 4). AP2 is not able to offer connection to M4 with the requested quality of service level. It asks the load distribution server to find a new distribution to make possible the connection of M4 to AP2. After the computation of the load balancing algorithm, the load balancing server broadcasts the new distribution of mobile stations onto the access points. AP2 have to dissociate a mobile station M2. This one will be associated to AP2 which is able to give it the required quality of service level. Finally M4 and AP2 complete the connection procedure (Figure 6).
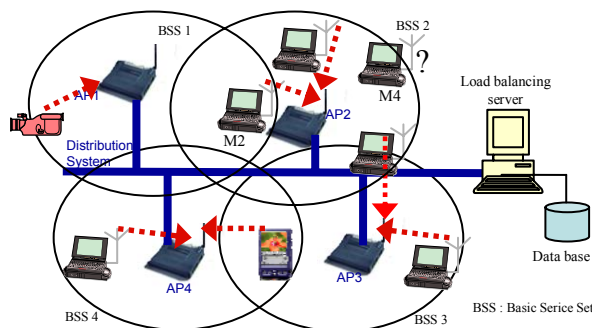
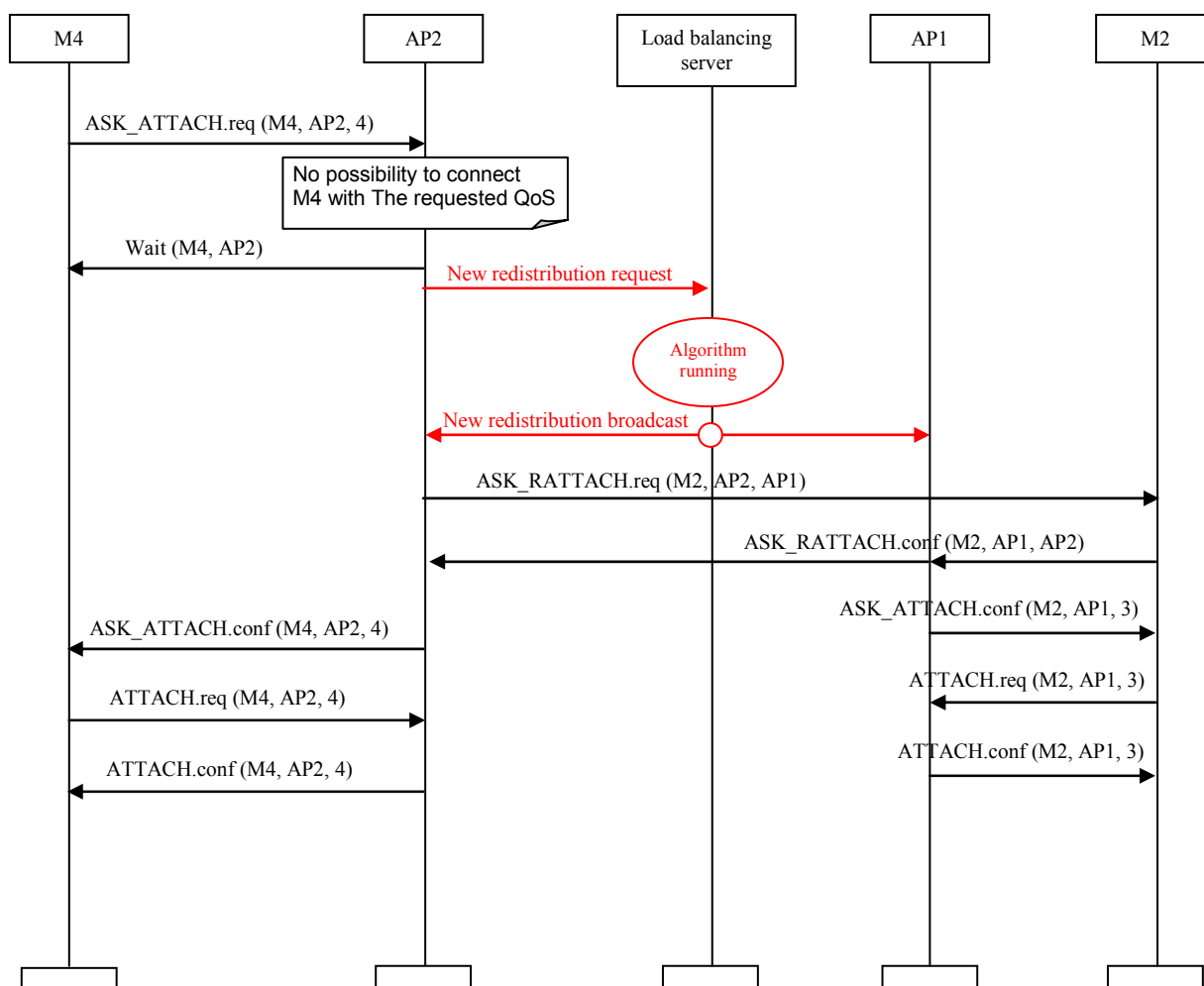<div align="center">Fig. 4: M4 asking for connection</div>
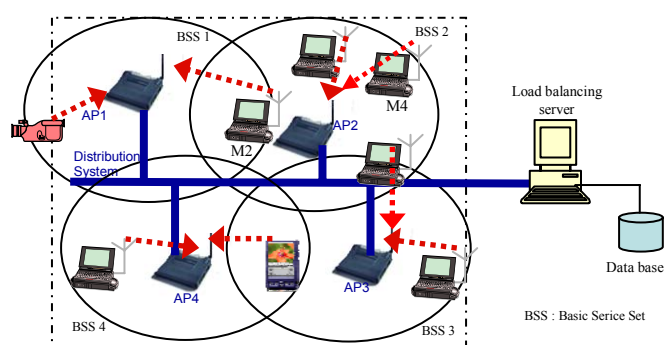
Fig. 5: Example of communication scenario

Fig. 6: M4 connected to AP2

The scenarios that we have defined will be described and verified with the SDL and MSC languages in the following sections.

8 SDL protocol description

The SDL pattern is an efficient design language for the development of a communication system. It enables a formal description system by defining a static modular architecture and interactions between different blocks [7]. Systems in SDL language are structured into interconnected entities (system, block, process, and channel) where process system description provides dynamic behaviour for internal task execution. It is based on the model of Extended Finite State Machines (EFSMs) [8]. In its dynamic behaviour, each state is reached after asynchronous signal exchange between blocks [9].

New primitives and exchanges defined in our approach have been described and validated with SDL (Figs. 7 and 8). Figure 7 shows the SDL model of a mobile station. It represents exchanges between the management layer and the data base of the mobile station.
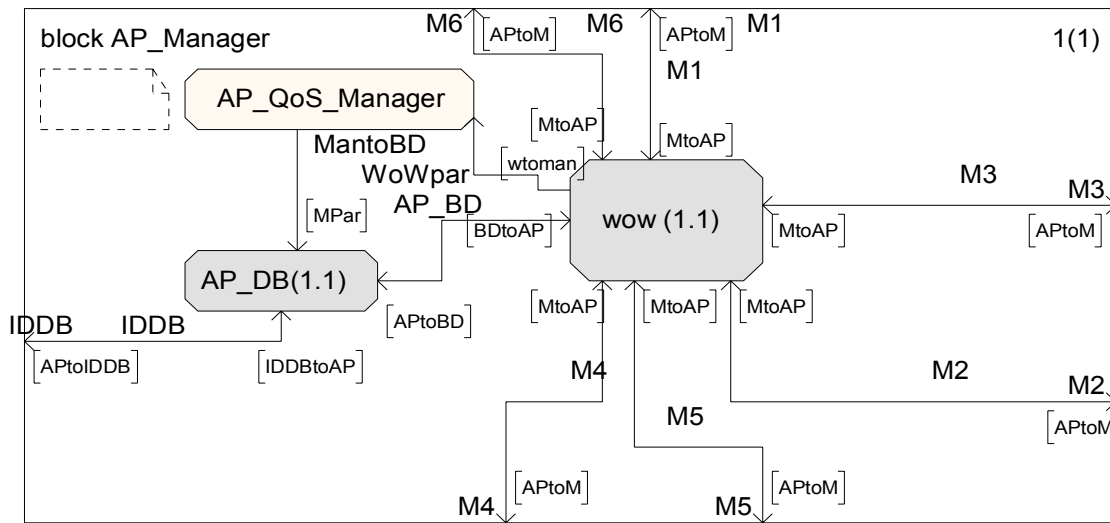

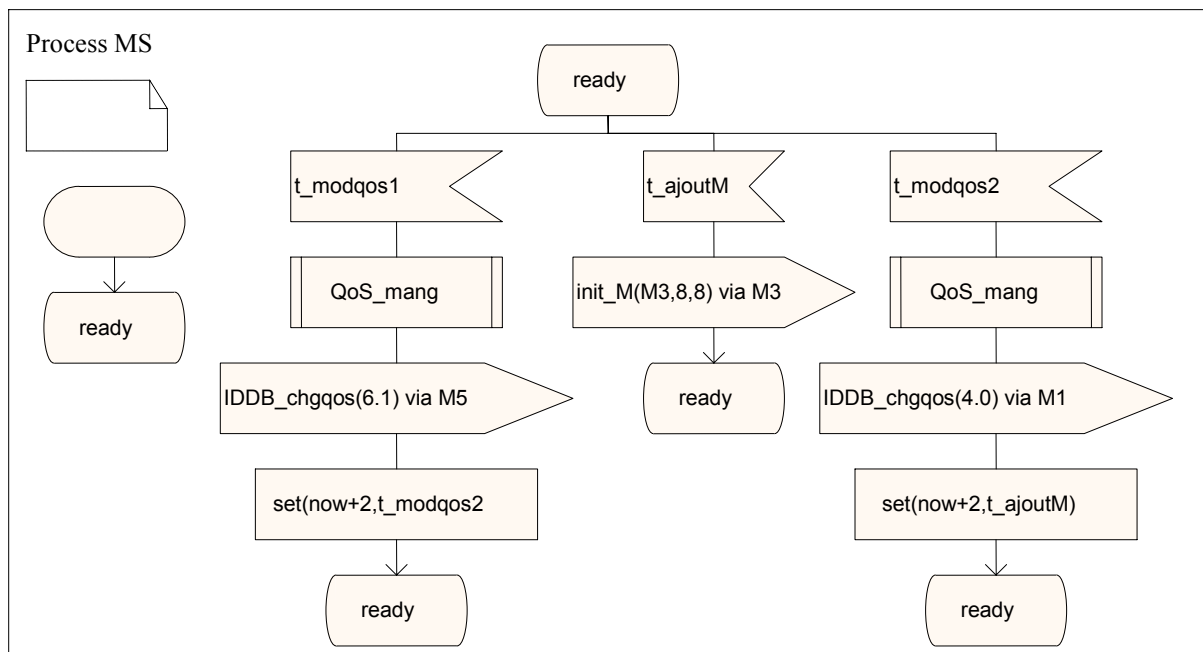**Fig. 7:** SDL model of a mobile station

Fig. 8: Example of SDL EFSM development

9 MSC verification and simulation

To check the QoS protocol behaviour based on the defined communication scenarios such as the one defined in Figure 5, we have used the ObjectGeode tool based on SDL and MSC. With SDL, we have validated the new primitives' exchanges between access points and mobile stations. The Figure 9 brings out a part of simulation results of a QoS negotiation between a mobile station (mobile 4) and an access point (AP2). First of all, entities must be set. Here we have a new instance of the mobile station M4 and the access point AP2. Mobile station gets his identifier, QoS_{max} and $QoS_{expected}$ values and begins a search of the access point with the requested QoS. Some other communication scenarios are also verified with SDL and MSC.
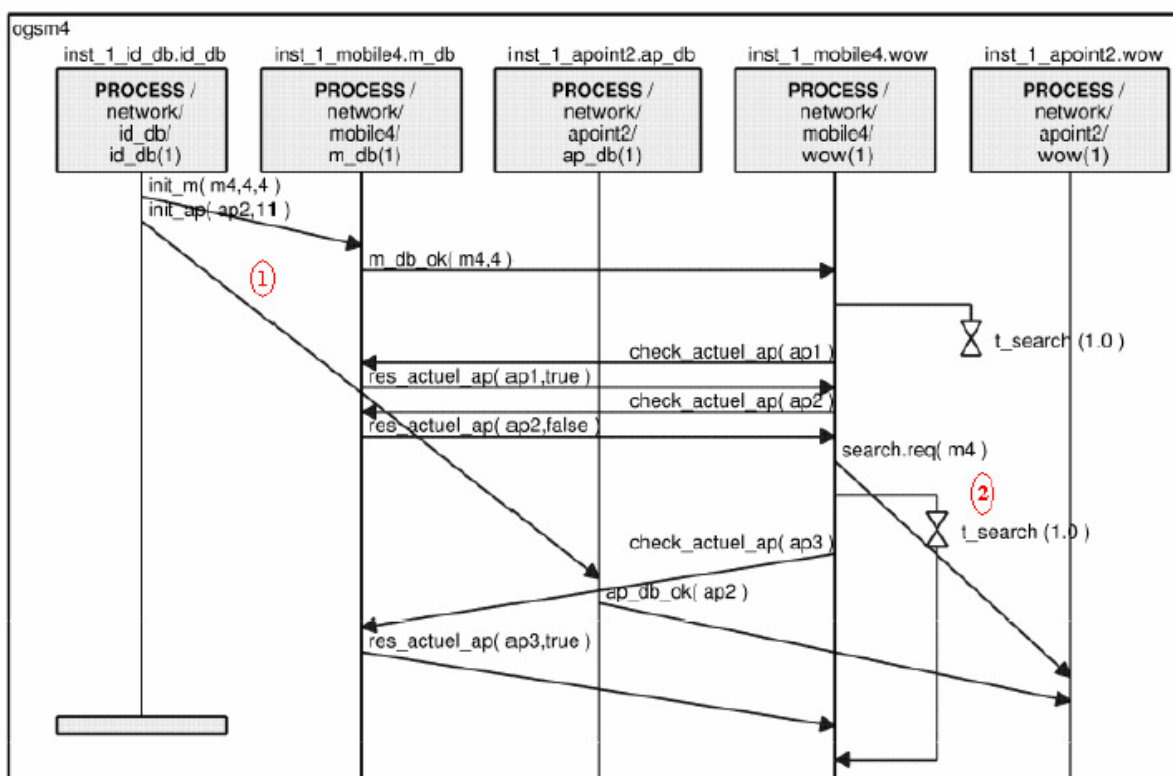
Fig. 9: MSC simulation of an example of exchange between wireless entities

10   Conclusion

This paper addresses the problem of QoS management in the WLAN. It presents a protocol specification between mobile stations and access points to negotiate QoS requirements during the mobile station attachment. This protocol defines some new primitives related to the QoS management that must operate with the IEEE 802.11. The specification of these protocol primitives has been carried out.

The second part of this paper presents an SDL description of this protocol and it shows the behavior verification with MSC simulation.

This work has to be completed with an implementation of this approach in a simulation architecture using an appropriate tool such as Opnet or Network Simulator (NS). This helps to analyze the performances and helps to adjust the parameters of this protocol before the experimentation. Other parameters can be also used to characterize quality of service requirements of the mobile stations such as loss ratio or jitter.

REFERENCES

1. A. Balachandran, G.M. Voelker, P. Bahl, and P.V. Rangan, "Characterizing user behavior and network performance in a public wireless LAN", June 15-19 2002, ACM SIGMETRICS international conference on measurement and modelling of computer systems (Marina Del Rey, California).
2. A. Balachandran, G.M. Voelker, P. Bahl, "Hot spot congestion relief in public-area wireless networks", June 20-21 2002, 4th workshop on mobile computing systems and applications (Callicoon, NY, USA) pp 70-80.
3. D. Kotz and K. Essien, "Characterizing usage of a campus-wide wireless network", September 23-28 2002, 8th annual ACM international conference on mobile computing and networking (Westin Peachtree Plaza, Atlanta, Georgia, USA), pp 107-118.
4. D. Tang and M. baker, "Analysis of a local area wireless network", August 6-11 2000, 6th annual ACM international conference on mobile computing and networking (Boston, Massachusetts), pp 1-10.
5. D. Tang and M. baker, "Analysis of a metropolitain area wireless network", March-May 2002, Wireless Networks (Kluwer Academic Publishers), Vol. 8 (2-3), pp 107-120.
6. Matthew S. Gast, "802.11 Wireless Networks : the definitive guide", First Edition April 2002 (O'Reilly and associates Inc.), ISBN: 0-596-00183-5.
7. R. Gotzhein, P. Schaible, "Pattern-based development of communication systems", 1999, Annales Télecommunication N° 54.
8. R.L. Probert, H. Ural, A.W. Williams, "Rapid generation of functional tests using MSCs, SDL and TTCN", February 15 2001, Computer Communications (Elsevier) Vol. 24 No. 3-4, pp 374–393.

9.   C. Bourhfir, E. Aboulhamid, F. Khendek, R. Dssouli, May 2001, "Test cases selection from SDL specifications", Journal of Computer Networks (Elsevier), Vol. 35, N°. 6, pp 693– 708.

10. E. Persiali, "Insertion d'une politique de qualité de service au niveau MAC d'un protocole de réseau sans fils", 2003, Rapport de DEA, Centre de Recherches en Automatiques de Nancy, Université Henri Poincaré Nancy I.

11. Chiu, Dah-Ming, Jain, "Analysis of the Increase and Decrease Algorithms for Congestion Avoidance in Computer Networks", June 1989, *Journal of Computer Networks and ISDN*, vol. 17, N° 1, pp. 1-14.

12. Qiang Ni, L. Romdhani, T. Tureletti, "A survey of QoS enhancements for IEEE 802.11 wireless LAN", 2004, Journal of Wireless Communications and Mobile Computing (John Wiley and Sons Ltd), Vol. 4, Issue 5, pp 547-566.

13. C. R. Lin, M. Gerla, "Real time support in multihop wireless networks", March 1989, ACM/ Wireless Networks, Vol. 5, N°. 2, pp 125-135.

14. IEEE 802.11 WG. ANSI/IEEE Std 802.11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications : medium access control (MAC) Enhancements for quality of service (QoS) IEEE 802.11/D2.0. 2001.

15. H. Velayos, V. Aleo, G. Karlsson, "Load balancing in overlapping wireless cells", June 20-24 2004, IEEE International Conference on Communications (Paris, France).

16. G. Bianchi, I. Tinnirello, "Improving Load Balancing mechanisms in Wireless Packet Networks", 28 April- 2 May 2002, IEEE International Conference on Communications (New York, USA), Vol. 2, pp 891-895.

17. A. Lindgren, A. Almquist, O. Schelen, "Evaluation of quality of service schemes for IEEE 802.11 wireless LANs", November 14-16 2001, 26th Annual IEEE Conference on Local Computer Networks (Tampa, Florida, USA), pp 348-351.

18. S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz, L. Stibor, "IEEE 802.11e wireless LAN for quality of service", February 2002, European Wireless (Florence, Italy), Vol. 1, pp 32-39.

19. D. Blinn, T. Henderson and D. Kotz, "Analysis of a Wi-Fi Hotspot Netoork", June 5 2005, the International Workshop on Wireless Traffic Measurements and Modelling (Seattle, WA, USA), pp1-6.

20. F. Guo, et Tzi-Cker Chiueh, "Scalable and Robust WLAN Connectivity Using Acess Point Array", 28 June- 1 July 2005, International Conference on Dependable Systems and Networks (Yokohama, Japan), pp 288- 297.

21. IEEE 802.11 WG, 2003. Draft Supplement to STANDARD FOR Telecommunications and Information Exchange Between Systems-LAN/MAN Specific Requirements - Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), IEEE 802.11e/Draft 4.2.

22. IEEE Std. 802.11-1999, part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, Reference number ISO/IEC 8802-11:1999(E), IEEE Std. 802.11, 1999 edition, 1999.

23. H. Zhu, M. Li, I. Chlamtak and B. Prabhakaran, "A survey of quality of service in IEEE 802.11 networks", August 2004, IEEE Wireless communications, Vol. 11, No. 4, pp. 6-14.

24. W. Pattara-Atikom, P. Krishnamurthy, and S. Banerjee, "Distributed mechanisms for quality of service in wireless LAN", June 2003, IEEE wireless communications, Vol. 10, pp. 26-34.