A Multimedia Architecture for 802.11b networks

P. Berthou, T. Gayraud, O. Alphand, C. Prudhommeaux, M. Diaz
LAAS-CNRS
7, Avenue du Colonel Roche
31077 Toulouse, cedex 4, France
email : {berthou, gayraud, oalphand, cprudhom, diaz}@laas.fr

Abstract - **Wireless networks have become increasingly popular for the past few years, especially IEEE 802.11 as the most widespread wireless technology. This article proposes a new architecture for broadband 802.11 networks to support distributed multimedia applications. After discussing the differences between multimedia and traditional applications, we show how the 802.11b mechanism ensuring enhanced robustness connectivity for low signal to noise ratio (SNR) areas, penalizes multimedia stream transmissions. A new multimedia architecture disabling Dynamic Rate Shifting mechanism and using a performance enhancing proxy is then proposed.**

I. INTRODUCTION

Mobile devices are certainly the major revolution of the past few years and the trend is towards smaller devices like PDAs, cellular phone which offer high quality multimedia services. 802.11b belongs to the short range technologies that are cheap and easy to deploy in small and difficult to wire areas.

To cope with high loss rates that characterize wireless networks, 802.11b uses the Dynamic Rate Shifting (DRS) mechanism. The mechanism and multimedia applications characteristics are presented in the first section. The second section studies the consequences of disabling the DRS to support a better bandwidth utilization for multimedia applications. A solution which both promotes multimedia applications and keeps an acceptable level of performance for long distance TCP connections is then worked out.

Finally, this new multimedia architecture is validated in the last section through a set of simulations realized with the NS simulator.

II. MULTIMEDIA COMMUNICATION IN 802.11 WIRELESS NETWORKS

A. Multimedia communication QoS requirements

Multimedia data are distinguishable from classical data (text, binary data, etc.) by their processing unit: Multimedia data are handled as *streams*, whereas texts or binary data are handled as files.

Generally speaking, multimedia streams are characterized by their *Quality of Service (QoS)*. This generic term regroups notions like *temporal constraints, quality restitution* and *transmission constraints* :

(a) The two most significant *temporal constraints* are: *delay* and *jitter.* From the network **delay** depends the multimedia **application interactivity**. (b) The *quality restitution* underlines the fact that multimedia streams can tolerate degradations without affecting a media presentation : for instance, an isolated image loss is not perceptible by a spectator during a movie diffusion. From the network **reliability** depends the application **restitution quality**. (c) *Transmission constraints* represent the transmitted stream parameters. The most important is the data throughput generated by the application. The video one can vary from several kilo bits per seconds to some Mbps, depending on the video coding scheme, the video size and the requested quality. Video streams, even compacted with MPEG or H263 are extremely bandwidth consumer. The network has therefore to provide a sufficient **bandwidth**.

Figure 1 provides typical multimedia data streams quality of service requirements.

	End-to-end delay	*Tolerated jitter*	*Reliability requirement*	*Throughput*
Audio stream	<400 ms	< 20 ms	> 90 %	13 kbps for GSM quality 128 kbps for CD quality
Video stream	<400 ms	< 20 ms	> 80 %	Depends on quality, typical rate are: 320 kbps for QCIF & H263 1.2Mbps for NTSC & MPEG
Picture	x	High	> 80 %	Minimal
Data stream	x	High	100 %	Minimal

Figure 1 - Multimedia data typical QoS

Each QoS parameters has to be provided by the network in order to support multimedia applications. It is, however, a complex task since QoS parameters are sometimes correlated. The following section shows how the throughput and the reliability are linked in 802.11b networks.

B. Quality of service in 802.11b networks

This part points out several 802.11 basic mechanisms which can interfere with new multimedia services and application deployments.

B.1 Dynamic Rate Shifting

Originally, 802.11 standard only provided two data rates (1 and 2 Mbps), considered as being too slow to meet most business and multimedia requirements. The IEEE therefore ratified a new standard known as 802.11b HR (High Rate) in 1999, which added two higher bit rates (5,5 and 11 Mbps)[8][9]. The 802.11b WLANs could then reach Ethernet levels of performance, throughput and availability.

Thus, the 802.11b multiple data transfer rate capabilities enable wireless network interface cards to adapt their performance to the radio link quality. This is called *DRS* (Dynamic Rate Shifting). It consists in a physical-layer mechanism transparent to the user and to the protocol stack upper layers. According to the signal interference and fading, 802.11b devices transmit at lower speed in order to improve data transfers reliability and eliminate frame drops.

B.2 A difficult cohabitation between low and high rates

Unfortunately, DRS may lead to drastic lower performances. The following section highlights this problem in calculating the theoretical maximum throughput of a 802.11b network.

There are two ways for accessing the radio channel : DCF (Distributed Coordination Function) and PCF (Point Coordination Function) [6]. We focus on the DCF mode which is the only one supported today. It is based on CSMA/CA : When a host wishes to transmit, it senses the channel and it waits a laps of time equals to a *DIFS* (Distributed Inter Frame Space) then transmits if the medium is still free. If the packet is received properly, the receiving host sends an ACK frame after a *SIFS* (Short Inter Frame Space) time. If a collision occurs or the data packet is corrupted, the packet is retransmitted after a random delay. When a packet is successfully sent, the other stations wait for the channel to become idle during a DIFS time and then start to decrease their back-off time as long as the medium remains free. When the countdown is over, a new transmission begins. This mechanism is summarized in

Figure 2.

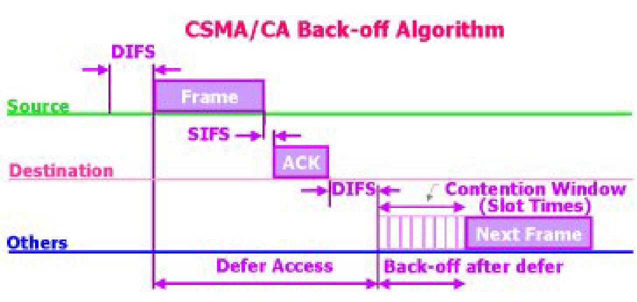

Figure 2 - Basic scheme of transmission in CSMA/CA

The frame consists in 3 fields : the PCLP (Physical Layer Convergence Procedure) Preamble and Header sent at 1Mbps for reliability reasons and the MPDU (MAC Packet Data Unit) part (see Figure 3). This latter is sent, according to the DRS, at 1, 2, 5.5 or 11 Mbps.

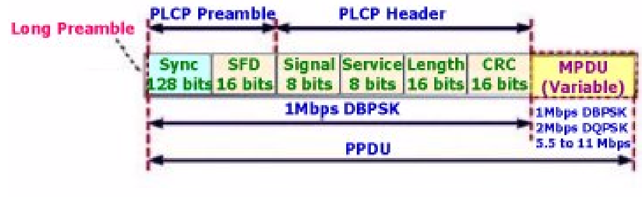

Figure 3 -The PLCP Frame Format

Thus, throughput at the MAC level can be approximated and majored assuming that propagation delays are neglected and that the back-off algorithm is not included in the literal calculation [7]. The estimation is based on transmission delays of a 1534-byte 802.11 MPDU.

The following table presents the calculus details of the transmission time:

IFS or frame transmission time	*Literal expression*	*IFS duration or frame transmission time*
SIFS	SIFS	10 µs
DIFS	SIFS + 2 time slots	30 µs
Tplcpframe	Theader +Tdata	192µs + 1534*8 / Throughput (DRS mechanism)
Tack	Theader + Tackfram	192µs + ACKsize*8 / 1Mbps (ACK sent at the lowest rate)
Ttransmission	SIFS + DIFS + Tplcpframe + Tack	536µs + 1534*8 / Throughput (ACK is 14 bytes long)

Figure 4 - Transmission times of a 1534 bytes frame in 802.11b

The following table presents the transmission time for the available access rates:

	1 Mbps	*2Mbps*	*5,5 Mbps*	*11 Mbps*
Ttransmission (ms)	12.808	6.672	2.767	1.651

Figure 5 - Transmission delay vs access rates

Assuming that a low-rate host and a high-rate host succeed in accessing the medium one after the other (fair access) and that the back-off algorithm is not taken into account, Figure 6 summarizes the medium access process.

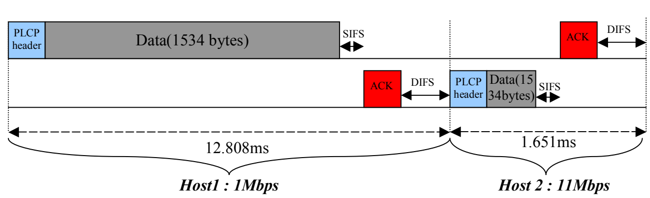

Figure 6 - Transmission times on a 802.11 WLAN shared by two users

With one 11Mbps user (H1) and one 1Mbps user (H2), the MAC throughput available for each of them is then :

$$Throughput = \frac{MPDUsize}{H1\_Ttransmission + H2\_Ttransmission}$$

i.e. **0.848 Mbps/user** for delivering the same amount of data (1534 bytes).

Consequently, when a low-rate host succeeds in transmitting over the medium, it penalizes the other stations awaiting its transmission end. Its low transfer monopolizes the channel whereas the others could transfer the same amount of data nearly ten times faster. Moreover, the overhead access increases with the number of host which also degrades the performance.

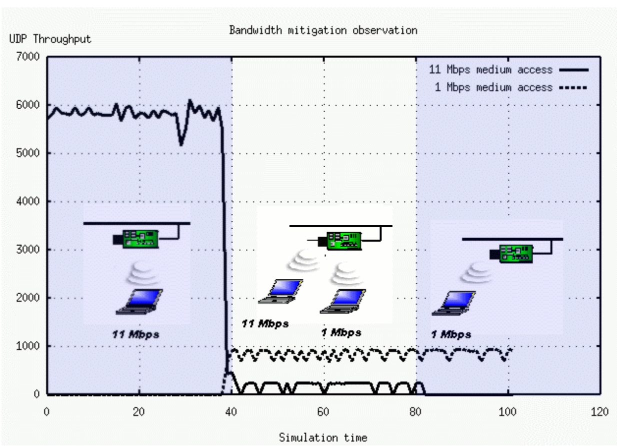

Figure 7 - Consequences of Dynamic Rate Shifting in a BSS with two users sharing the same 802.11b link

The Figure 7 shows the throughput measured at the application layer when two hosts, with different access rates, are communicating over a 802.11b network. On the left part, the first host accesses the network at a 11 Mbps rate and succeeds a 6Mbps data transfer. When the 1 Mbps host arrives, the global bandwidth decreases to 1.3 Mbps. The throughput remains constant when the first one leaves. Similar effects have also been analyzed with TCP.

Consequently, 802.11b DRS mechanism increases the network reliability at the expense of its rate. However, suppressing it has major drawbacks.

TCP is massively used in data transmissions. Its fully reliability and ordered service are a basis for a lot of applications as files transfers, web services or secured transactions. It is estimated to be responsible for over than 80 % of the Internet traffic.

TCP congestion control algorithm is based on the assumption that packet losses are only arising when underlying networks are congested. Applying it in a high bit error rate (BER) wireless link context seriously limits the connection throughput. Furthermore the higher the network delay is the more significant the throughput reduction is. This argument is the main justification for the 802.11b DRS.

B.3 802.11 drawbacks for multimedia communication

New compression algorithms enable multimedia communication over not so large bandwidth networks. The 11 Mbps service of 802.11b networks could theoretically run about twenty good quality videoconferences.

However, a network user roaming or starting its computer in an area affected by a poor SNR has a bad consequence. Because of the bandwidth mitigation, the global network throughput is drastically reduced and multimedia applications are no longer supported.

Multimedia communications and data exchanges have opposite requirements, requiring either a large bandwidth or a reliable network. By implementing the DRS mechanism, the 802.11b specification community has clearly privileged data exchanges.

The next part presents a multimedia architecture based on 802.11b networks that can fulfill multimedia applications requirements without disturbing TCP connections.

III. A MULTIMEDIA ARCHITECTURE FOR 802.11B

A. Disabling DRS in multimedia context

Disabling DRS has advantages and drawbacks. On the one hand, the possibility to operate at a unique data rate (11 Mbps) leads to a more efficient bandwidth utilization since there would not be low or high-rate users anymore. Multimedia applications which tolerated packet losses would not be penalized as they have an acceptable quality of service. Indeed, thanks to disabling DRS, the radio channel could not be monopolized by 1 Mbps rate hosts anymore and would be really fairly shared.

On the other hand, whereas multimedia applications based on UDP tolerate such a degradation, TCP connections suffer from packet losses. Errors in wireless context are not due to congestion but still result in a reduction of the end-to-end throughput by reducing the congestion window size, invoking slow start and congestion avoidance, which is totally irrelevant.

B. Performance Enhancing Proxy to restore QoS

Using a proxy could solve the problems of disabling DRS.

TCP performance degradation over high error rate links have already been deeply studied. It leads, to several recommendations to improve TCP stack implementation for wireless network [1].

B.1 Proxy principle

The same IETF working group has defined the notion of Performance Enhancing Proxies (PEPs) [2] to cope with the network heterogeneity induced by wired and wireless network interconnections, this approach goal is to increase end to end performances by breaking the end to end (TCP) connections into several segments with optimized parameters for each segment. Each segment is a complete TCP connection. Data streams are forwarded from one segment to another and possibly buffered.

Several PEPs types have been described as well as the mechanisms used to improve their performance in [2].

The two main advantages of this architecture are the following :

Easy TCP stack adaptation

The Internet TCP stacks, especially those located in computer belonging to traditional wired networks, will probably not implement the TCP options recommended for wireless network (TCP Sack [3], Fast Retransmit/Fast Recovery [4], etc.).

The PEP principle offers a good alternative to this problem. Only PEPs and TCP stack for wireless networks have to implement these options. As illustrated in Figure 8, when a wired computer uses classical TCP to communicate with a wireless laptop, the TCP connection is intercepted by the PEP and a new TCP connection (TCP*) is initiated between the PEP and the laptop. This connection should implement TCP option for wireless network. Thus the global connection is optimized and the wired computer avoid any network stack configuration.

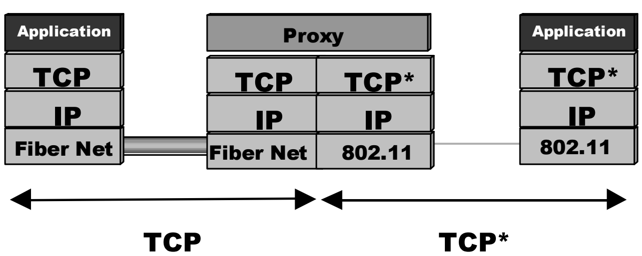

Figure 8 - TCP stack adaptation mechanism

Local retransmissions on the first/last hop

Local TCP retransmissions are done by the PEP when data are lost between the PEP and the receiving end systems. This mechanism limits the retransmission delay by retransmitting as soon as possible. Moreover, in most cases, wireless networks are local networks and are located on the first (or on the last) hop of a connection. Retransmission delays are then even shorter.

C. A strategic location

Theoretically, the best PEP efficiency is obtained when it is located at the edge of the wireless network in order to isolate the congestion errors on the wired side and the high-loss link errors on the other side (see Figure 9).

Thus, on the wireless link, the corrupted packets are locally retransmitted by the proxy. On a LAN the retransmission delay of lost packets is short enough to keep the efficiency of the TCP error control mechanism. Moreover, on the wired part, the long delay characterizing MAN and WAN is well handled by TCP when errors are only due to congestion.

Hence, the global TCP behaviour is greatly improved as shown by the simulation results in the following section.

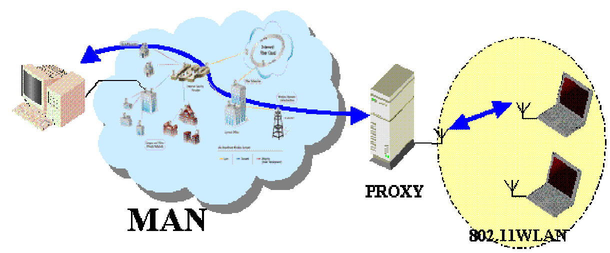

Figure 9 - PEP best location in a MAN/WLAN environment

D. Limitation

Motivations for PEP development and use are described in the draft along with consequences of using them, especially in the Internet context.
On a split connection, that includes transport layer hops, the most negative effect is the end to end semantic break, which is one of the main architectural Internet principle [5].
Nevertheless, this approach is very useful for 802.11 networks. Considering the Diffserv notation, PEPs are located on “border networks”, the WLAN remaining at the first (or last) hop of a connection path. PEPs can therefore be easily managed and finely tuned, their deployment has no more consequences on the Internet than an HTTP proxy deployment.

IV. PERFORMANCE EVALUATION

Because of the unavailability of highly configurable 802.11b network card, it was impossible to disable the DRS mechanism. As consequence, the proposed architecture has been validated thanks to NS network simulator.

The testbed consists in a 11 Mbps 802.11 wireless LAN connected to the Internet through a wired WAN. The network delay is 10 ms for the WLAN and 100 ms for the WAN. To study the behavior of a TCP connection when facing high losses rates, the reliability of the underlying network was decreased every 10 seconds. The simulation starts with a reliable network then decreases the BER to 10^{-4}. The TCP throughput (number of delivered packets per second), i.e. the throughput available for a FTP connection, is measured.

Firstly, the network reliability degradation impact on a TCP connection behavior is analyzed, then our multimedia architecture is evaluated.

A. 802.11 link quality impact over end-to-end TCP connections

The throughput instability of a long delay TCP connection (WAN) when the BER is upper than 10^{-7} appears clearly on Figure 10 (curves at the bottom of the figure). When the BER is over 10^{-5} the TCP throughput is decreasing drastically to a very low rate. Conversely the instability phenomenon is unperceivable for short delay connections (WLAN).

It is due to the good TCP error control efficiency when the retransmission delay is short enough. The throughput is even acceptable for a 10^{-5} BER, but is becoming naturally close to zero after. A lossy network is then not a real problem for local networks.

The throughput difference between LAN and WAN configurations is not significant since the TCP sliding window mechanism links the throughput to the connection delay.

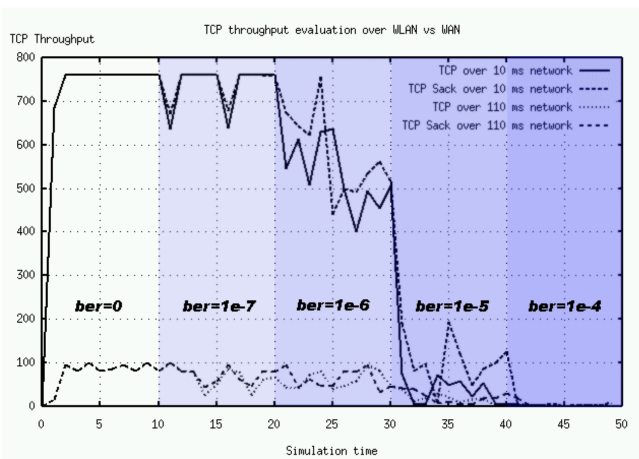

Figure 10 - Quality degradation impact over WLAN and WAN TCP connections

TCP mechanisms have already been extensively studied to cope with high error rates, however Figure 11 shows that, when the connection delay is high (110 ms), TCP Reno or TCP Sack are very reactive to losses as for TCP. End-to-end TCP connection is never a good solution for long delay connection with high error losses.

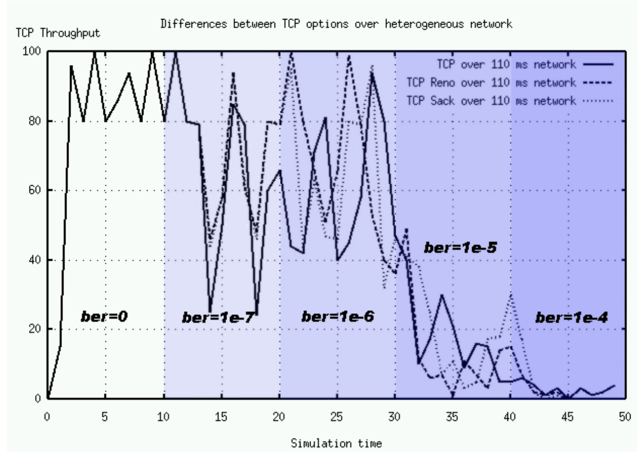

Figure 11 - TCP implementations comparison when facing high losses rate

B. The multimedia architecture evaluation

Improvements obtained when using our architecture are shown on Figure 12. The bandwidth utilization and the resistance to network losses are better. The connection throughput is constant (between 80 to 100 pps) even if the error rate is high. With a 10^{-5} BER the bandwidth utilization is still good, as it was observed (Figure 10) for a LAN. A proxy at the edge of the wireless LAN provides enough QoS to ensure a good behavior of long delay TCP connections.

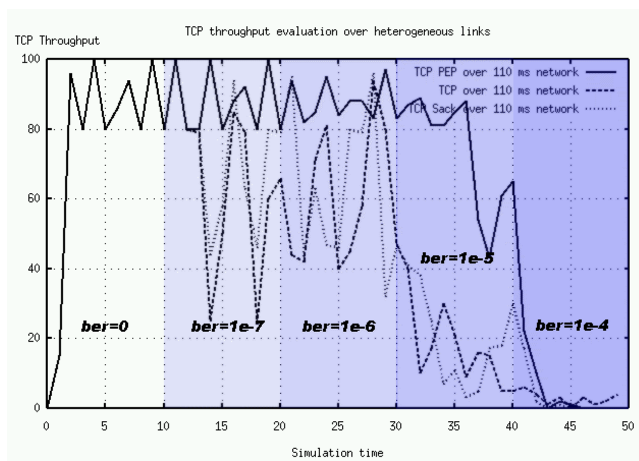

Figure 12 - The multimedia architecture performances evaluation

V. CONCLUSION

This proposition intends to enable multimedia communications over 802.11b wireless networks. The 802.11b dynamic rate shifting mechanism is proposed to be disabled and replaced by a *Performance Enhancing Proxy* at the wireless network edge. The architecture has been simulated and validated through network simulations.

This multimedia architecture could be a cheap solution for the already deployed 802.11b networks. It is easily deployable because its configuration is under the unique responsibility of the WLAN administrator. If configured with a transparent proxy, this architecture implies no changes outside the wireless network. Moreover, the proxy could be enhanced with new sessions services, as connection persistence for instance, or coupled with already existent proxies, as a SIP proxy. This architecture designed for 802.11b networks could also be applied for 802.11a networks which use a similar rate shifting mechanism.

VI. REFERENCES

[1] S. DAWKINS, G. MONTENEGRO, M. KOJO, V. MAGRET, N. VAIDYA, "End-to-end Performance Implications of Links with Errors", RFC3155, August 2001.

[2] J. BORDER, M. KOJO, J. GRINER, G. MONTENEGRO, Z. SHELBY, "Performance Enhancing Proxies", RFC 3135, June 2001.

[3] M. MATHIS, J. MAHDAVI, S. FLOYD, A. ROMANOW, "TCP Selective Acknowledgement Options", RFC2018, April 1996.

[4] W. STEVENS, "TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms", RFC2001, January 1997.

[5] J.H. SALTZER, D.P. REED, D.D. CLARK, "End-To-End Arguments in System Design", ACM TOCS, Vol. 2, No. 4, pp. 277-288, November 1984.

[6] B.P. CROW, I. WIDJALA, J.G. KIM, P.T. SAKAI, "IEEE 802.11 Wireless Local Area Networks", IEEE Communications Magazine, pp. 116-126, September 1997.

[7] J.A. GARCIA-MACIAS, F. ROUSSEAU, G. BERGER-SABBATEL, L. TOUMI, A. DUDA, "Quality of Service and Mobility for the Wireless Internet", Proc. First ACM Wireless Mobile Internet Workshop, Rome, Italy, 2001.

[8] ANSI/IEEE Std 802.11, 1999 Edition

[9] IEEE 802.11b-1999™/Cor 1-2001, IEEE Standard for Information technology, corrigendum 1, 2001.