Back to sessions

4:3 VoIP Transmission in Fixed Wireless Access networks based on DECT Packet Radio Service

Luis M. Contreras

Francisco Escrihuela

Alcatel España, Radio Communications Division
Ramirez de Prado, 5 – 28045 Madrid, Spain
{luis.contreras@alcatel.es}

1. Introduction

The possibility of using networks originally conceived for data communications as transport networks for voice and other real time services has evolved from a technological curiosity to a real business in very few years. Originally, one of the main reasons for the rapid growth of voice over IP (VoIP) was the cost-savings that this technology may bring for end-users. However, further motivations are fuelling the interest on VoIP such as the flexibility offered by a packet-switching technology or the potential added value applications to be provided.

The Fixed Wireless Access (FWA) networks were born motivated by the provision of an optimal cost solution for the rapid deployment of access infrastructures to telecommunication networks, mainly PSTN. Nevertheless, technological evolution and efforts in standardisation have resulted in the introduction of efficient packet-mode transport mechanisms in wireless networks.

The situation today is that FWA technologies are able to support efficiently both circuit and packet traffic, and solutions for IP transport over wireless networks are appearing on the market. This means that we can consider the FWA networks as IP Networks (IPN) or as Switched-Circuit Networks (SCN), depending on the nature of the access technology used.

The aim of the paper is to analyse the performance of a FWA network transporting VoIP. On this issue converge many hot topics driving a lot of research and industry effort today, such as wireless data transport, VoIP, local loop unbundling.

2. DECT Packet Radio Services

The Digital Enhanced Cordless Telecommunication (DECT) technology is a general radio access standard for wireless telecommunications that can be used by many different applications and can be connected to different networks as well. The DECT technology has been standardised by the European Telecommunications Standardisation Institute (ETSI) [1]. It can be effectively implemented in a range from simple residential cordless telephones up to large systems providing a wide range of telecommunication services, including FWA, comprising a comprehensive set of protocols that provide the flexibility of inter-working among numerous different applications and networks.

The existing DECT-based data-related standards have been re-composed into a common packet-mode standard, called DECT Packet Radio Service (DPRS). In July 1999, DPRS was approved as new standard by ETSI [2]. The DPRS standard defines the way to implement packet-data services on a DECT access network. DPRS offers certain modularity through different feature

Go to page 1

subsets. These subsets should be combined with Application Specific Access Profiles (ASAPs) devoted to specific applications.

The DPRS adds a set of features to that provided by DECT, as high speed data transfer capabilities, multi-bearer and asymmetrical operation, simultaneous support of voice (circuit-switched) and data (packet-switched) services, co-existence of multiple instances of data services in the same user termination, bandwidth negotiation for an active connection, etc. In addition, DPRS is able to support both Frame Relay (datagram oriented) and character (asynchronous data oriented) services. The former comprises the following types: Ethernet, Token Ring, IP and PPP.

DPRS, as packet-switching technology, allows make an efficient use of the physical resources by releasing the physical channel (when there is no user information to be sent) while holding the call at logic level. In that way, DPRS is able to distinguish, for a connection, two different levels named *virtual call* and *physical connection*.

The virtual call is established at Network layer (level 3), being conceptually equivalent to a circuit-switched call. The connection establishment has to do with the point in time when the virtual call is decided be started up. The network layer is not aware about the packet transport mechanism in the system, which is in fact hidden to it.

The physical connection is the set of physical resources that can be assigned to a virtual call. The access to the air interface is completely controlled at MAC layer (level 2). Connection resumption and suspension are operations related to the allocation and retrieval of radio resources for a specific physical connection. Such operations are triggered by the presence of data to be transmitted.

3. VoIP transport on DPRS

The protocol allows to any of the peers the range of usage from 1 up to 23 full slots (from the available 24). This fact provides a net bit rate ranging from 24 kbps up to 552 kbps with current 2-level modulation. This bit rate will be even higher with expected (and standardised) 4- and 8-level modulation schemes. These data transfer speeds are obtained by the capability in DPRS of operating in a multi-bearer and asymmetrical mode. The figure 1 shows this capability.

In present implementations of DECT-based FWA networks, voice calls are served by using circuit-switched transport mode. One voice call permanently occupies one full duplex slot during the call life. When considering the VoIP transport in packet-switched mode, the usage of the same capacity for VoIP, i.e. one slot, can be viewed as a bound in order to keep the voice packet-based service at least as efficient as the circuit-based one. Further improvements on efficiency can be expected from the capability of DPRS of suspending the physical connection when no data has to be transmitted. For voice calls this could be the case of using Voice Activity Detector (VAD) devices.

Due to the protocol stack used for transmitting data when the packet mode is considered, a number of overhead bytes must be added to each of the transmitted data units. These bytes can be accounted as follows: 12 bytes for RTP, 8 bytes for UDP, 20 bytes for IPv4, 7 bytes for PPP. In order to reduce the overhead on the voice traffic over IP, header compression mechanisms can be introduced. By using header compression mechanisms defined in [3], the overhead originated by RTP+UDP+IPv4 can be reduced from 40 bytes to 2 bytes.

For current 2-level modulation, this bound means that the minimum granularity of the “pipe”, i.e. 24 kbps, will allow transmit 30 bytes every 10 ms.

Go to page 1

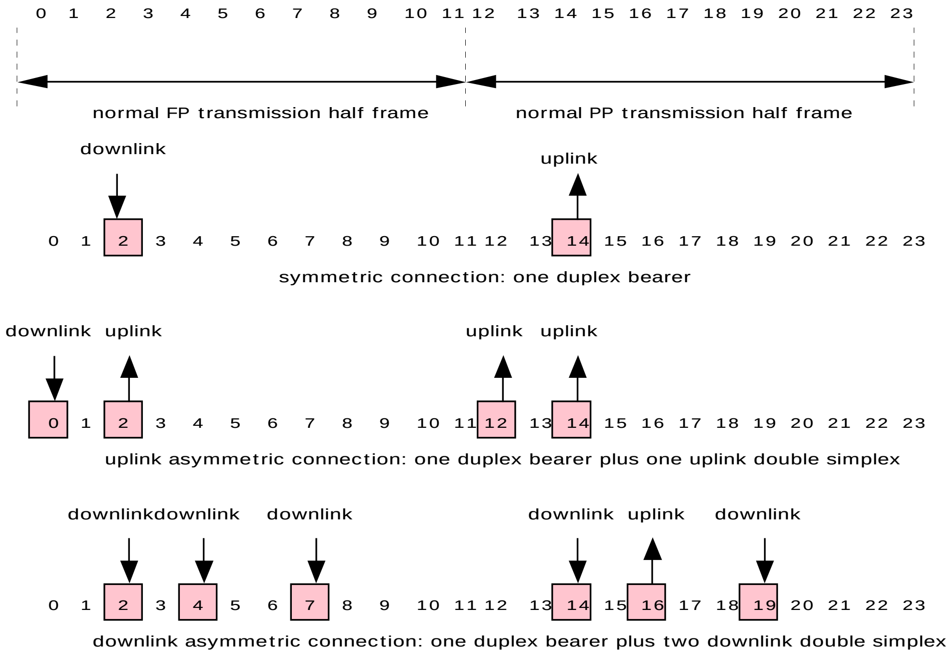

Figure 1. DPRS operating capabilities.

The suitable codec values will be that which VoIP kbps in table 1 are less than 24 kbps for either header compressed or non header compressed packets. Furthermore, the voice packet interval should be controlled because a long packet causes big framing delay. The estimation used here for transmitting voice packets in a sequence will be the margin 10~40 ms.

4. VoIP scenarios

This paper is mainly concerned with the provision of Telephony service over IP based networks. The user of the IP Telephony will normally expect the same QoS that is currently provided by other telecommunication networks. For the provision of IP Telephony, inter-working between IP networks and other telecommunications networks will normally be expected. In this way, several architectural configurations can take place in a real deployment of a network. We propose here to evaluate the performance achieved by two different network configurations accessed by a FWA network based on DPRS and providing VoIP for speech services. The configurations under consideration will be:

- Communication between two FWA networks (DPRS based) through an IPN, forming an all IP network.
- Communication between two FWA networks (DPRS based) through a SCN, inter-working the packet- and circuit-switched worlds by mean of gateways.

4.1. All IP scenario

The first voice communication scenario to be evaluated presents the configuration shown in figure 2. This scenario allows two FWA networks based on DPRS to communicate with each other through an IPN, thus forming an all IP network. The voice communication along the path will be packet-switched based.

The voice signal on the originating side is packetised and transmitted via a radio link to the IP network. In the IP network, the packet stream transporting the voice signal will be routed towards the terminating side of the communication. There, the incoming packet flow will be de-jittered, and the voice signal will be extracted from it.

Go to page 1

G.723.1 at 5.3 kbps							
no Header Compresion				Header Compression			
samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]	samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]
1	30	67	17.86	1	30	29	7.73
G.723.1 at 6.3 kbps							
no Header Compresion				Header Compression			
samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]	samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]
1	30	71	18.93	1	30	33	8.8
G.729.A + VAD at 8 kbps							
no Header Compresion				Header Compression			
samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]	samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]
1	10	57	45.6	1	10	19	15.2
2	20	67	26.8	2	20	29	11.6
3	30	77	20.53	3	30	39	10.4
4	40	87	17.4	4	40	49	9.8
GSM-FR at 13 kbps							
no Header Compresion				Header Compression			
samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]	samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]
1	20	80	32	1	20	42	16.8
2	40	113	22.6	2	40	75	15
GSM-HR at 5.6 kbps							
no Header Compresion				Header Compression			
samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]	samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]
1	20	61	24.4	1	20	23	9.2
2	40	75	15	2	40	37	7.4
GSM-EFR at 12.2 kbps							
no Header Compresion				Header Compression			
samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]	samples per packet	time[ms]	VoIP size [byte]	VoIP BW [kbps]
1	20	78	31.2	1	20	40	16
2	40	109	21.8	2	40	71	14.2

Table 1. Codec's bandwidth per framing time.

4.2. Mixed IP-SCN scenario

The second voice communication scenario to be evaluated presents the configuration shown in figure 3. This scenario allows two FWA networks based on DPRS (and thus IP-based) to communicate with each other through a SCN, interworking the packet- and circuit-switched worlds by mean of gateways.

The voice signal on the originating side is packetised and transmitted via a radio link to an ingress gateway to the SCN. There, the jitter introduced by the radio link on the packet flow will

Go to page 1

be compensated, and the voice signal transcoded in order to prepare it for adequate transport on SCN towards the terminating side of the communication.

The voice signal will arrive to an egress gateway, where it will be transcoded again for optimising its transport on radio link. Finally, the packet flow will be de-jittered and the voice signal presented to the terminating end.

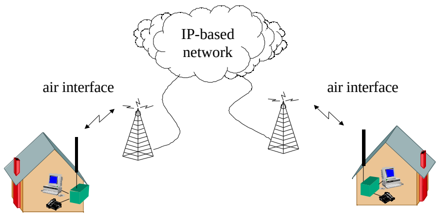

Figure 2. All IP scenario

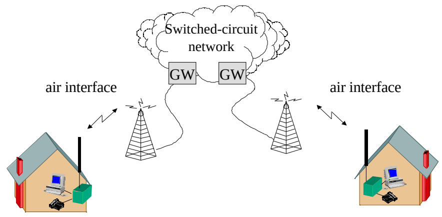

Figure 3. Mixed IP-SCN scenario

5. Performance calculations

The technical issues regarding the implementation of wireless VoIP strongly depends on the Quality of Service (QoS) figures provided by the system. The QoS figures have the property of transitivity in the sense that the costs in quality are propagated along the network. In this way, it is important to distinguish the different contributions coming from the system allowing to operate on them in order to optimise the whole system. The QoS parameters will have influence in the end-to-end characteristics of the system. There are three main components in the QoS characterisation of the systems under study:

- QoS parameters derived from the application (VoIP), such as codec, jitter, etc.
- QoS parameters derived from the packet-switched nature of the communication, such as buffering, routing, etc.
- QoS parameters derived from the wireless nature of the network access, such as channel errors, framing, etc.

All these sets of parameters will influence the speech quality along the system as perceived by the end-user. Nevertheless, for the scope of this paper, which intends a primary incursion to the problem of transporting VoIP by using DPRS, some assumptions will be taken into account in order to ease the tractability of characterising that issue. The performance calculations will be based on the computational E-model [4, 5].

5.1. QoS parameters derived from the VoIP application

In VoIP application, the speech signal is coded in the originating side and passed through the network towards the terminating side where it is decoded. The device able to code and decode the speech signal is the called *codec*. The codec delay is given by [6]:

$$T_{cod} = T_{enc} + T_{dec} + T_{la} \tag{1}$$

where T_{enc} is the encoding delay, T_{dec} is the decoding delay and T_{la} is the look ahead voice frame. Both T_{enc} and T_{dec} are upper bounded by the voice frame, T_f. To form an IP packet, further delay is added to the speech signal. Such delay is called packetization delay, and is defined as [6]:

$$T_{pack} = N\ ?T_f \tag{2}$$

being N the number of voice code words in a packet.

Go to page 1

The jitter is the delay variation due to the changing traffic conditions in a packet network. The jitter is removed by buffering in the system, and so, a buffering delay must be taken into account. The buffering is needed to offer a constant bit stream to the decoder.

Some other aspects will not be taken into account, such as signalling or security mechanisms induced delays, audio/phone card buffering in case of PC-to-PC communication, etc.

5.2. QoS parameters derived from the packet-switched nature of the communication

The packet-switched nature of the communication becomes apparent in two parts of the proposed scenarios. Clearly, the access networks will be packet-based transmission means, but also the transport network is packet-based for the case of the all IP scenario.

When talking of the access networks, the unreliable properties of radio links due to transmission phenomena such as fading can originate errors in the transmitted packets. Different channel conditions over time will affect differently to the sequence of packets sent on the air. Such sequence could result altered, or differently delayed, if some recovering mechanisms are provided. The wireless nature of the communication will be discussed in the following subsection.

When talking of the IPN, the packet-switched nature of the communication implies certain handling or processing of the datagrams passing through the network. The voice communication is an interactive, real-time service, so important aspects such as routing, scheduling or queuing management can impact on the perceived QoS. Those mechanisms handling the voice stream transmission should be optimised in order to guaranteeing some QoS degree as certain delay or packet loss-rate. They will not be taken into account here. On the contrary, the packet-based transport network will be considered to contribute as a whole end-to-end delay.

It should be noted that once a packet with compressed header overcome the radio gap, it is necessary to de-compress it for transmission on the IP backbone network. This data processing will originate an extra delay. Nevertheless, it is considered to be also included in former contribution to end-to-end delay.

For this paper, we will consider the Internet as packet-based transport network. It is clear that more controlled IP-based transport networks (e.g. an Intranet) will perform better. As delay contribution of Internet we will take the value for Internet delay considered in [7], that is 100 ms. Furthermore, for the jitter buffer we will consider the value assumed in [7] as well, i.e. 60 ms.

5.3. QoS parameters derived from wireless nature of the network access

The non-reliable characteristics of the radio channel impose a certain degree of packet loss intrinsic to the wireless systems. The packetised voice stream can be seriously damaged by the packet loss rate of the radio channel in two dimensions. On one hand, since it represents a real-time service, the retransmission of packets with errors will not be longer valid (or if valid, maybe for one retransmission only). On the other hand, voice codecs will offer distinct behaviour to the presence of gaps in the voice stream, affecting the same packet loss rate differently depending on the codec used.

Transmissions errors in radio channel will not be taken into account. This has been also the approach of other papers dealing with the issue of VoIP over wireless systems [7]. As consequence, the results presented here can be viewed as an upper bound of system performance. Furthermore, an extra implication is derived from this assumption. Due to there are no errors in the radio path, all the packets entering it will exit ordered and equally delayed. Thus, there will

Go to page 1

not exist the need of placing a jitter buffer (and consequent delay) at the output of each radio gap. This applies for the rest of the paper.

5.4. E-model

The E-model [4, 5] is a computational tool able to provide us a predicted Mean Opinion Square (MOS) rate respect to the speech quality obtainable from a specified communication scenario as perceived by the listener.

The E-model has been widely used along the literature for characterising distinct VoIP scenarios [6, 8, 9], since it is a versatile tool well adapted to take into account the impairments that appear in VoIP. Also here, it will be used for estimating the relative user satisfaction when transporting VoIP by means of DPRS on radio links.

Basically, the E-model allows to combine multiple sources of impairments caused by the transmission parameters that characterise a specified communication scenario, and quantify the overall impact by providing a rating factor R, which can be related with the subjective user reactions in form of MOS. The combination of impairments is done in an additive manner.

The rating factor is expressed by:

$$R = R_0 - I_s - I_d - I_e + A \tag{3}$$

where R_0 takes into account noise effects suffered by the signal along the communication path (e.g., circuit noise); I_s takes into account impairments that ocurr simultaneously with the voice signal (e.g., too loud side tone); I_d takes into account delayed impairments (e.g., absolute delay); I_e takes into account impairments originated by the use of special equipment (e.g., low bit-rate codecs); and A takes into account how the user is willing to tolerate certain defects on transmission as consequence of perceiving some extra advantages in the system respect to traditional telephony (e.g., mobile component of mobile telephony despite of a poorer speech quality).

6. Performance evaluation

This chapter provides the estimated MOS values for each of the scenarios. Standardised codec parameters and impairment factors associated to them are presented in table 2.

The table 3 presents the results obtained from the E-model computational tool. For this calculation, the only parameters distinct from default values in [5] have been the one way absolute delay and the impairment factor.

6.1. All IP scenario

In the all IP scenario, the packet stream containing the voice signal is transported to the far end without any modification, except that derived from compressing/de-compressing the RTP/UDP/IP header, if applied.

The voice signal will only suffer one codification/de-codification process along the communication path. In this way, the contribution of such process to the total delay in the end-to-end communication path is given by the sum of eqs. (1) and (2).

Go to page 1

Codec	T_f (ms)	T_{la} (ms)	I_e
G.723.1 @ 5.3	30	7.5	19
G.723.1 @ 6.3	30	7.5	15
G.729.A + VAD	10	5	11
GSM-FR	20	0	20
GSM-HR	20	4.4	23
GSM-EFR	20	0	5

Table 2. Codec's parameters and associated impairment factors.

As stated in chapter 3, the usage of only one slot (that is, 24 kbps) to keep this service at least as efficient as the circuit-based one, will force in some cases to use more than one frame to transport completely the voice packet among the radio channel ends. Each of the frames used will add 10 ms to the end-to-end delay (the radio channel is considered error free, so retransmissions and consequent delay increments are not taken into account). This contribution to the system latency will happen twice due to the presence of two radio-access networks, each of them in one end of the communication scenario.

Once the voice packets pass the first radio gap, they enter the IPN. There, the packets can be routed by different paths and, thus, can experience distinct delay. A jitter buffer before sending the packets over the radio channel towards the terminating side will smooth that variation on delay allowing also for re-ordering the packets. The jitter buffer is placed there because assuming no errors in the radio channel, the packets transported on it will not suffer any jitter.

6.2. Mixed IP-SCN scenario

Some considerations have to be done regarding to the SCN. Here, the SCN is used as core transport allowing the communication among other two access networks. Thus, it assumes that SCN is composed of digital segments only. In the interconnection with the access networks, the 4-wire circuits will provide a close approximation to echo-free connections, assuming adequate acoustic coupling across the handset.

The delay estimation for such SCN should take into account both processing and propagation times. For a typical national connection, the associated processing time should be kept below 50 ms [10]. On the other hand, the transmission time in a purely digital network is given by [10]

$$3 + (0.005 \cdot distance\ [km])\ ms \qquad (4)$$

for optical fibre systems. In the following calculations, we set the distance value to the existing geographic distance between e.g. Madrid and Barcelona, i.e. 498 km.

Now, in the mixed IP-SCN scenario, the voice packets have to be transcoded when passing from the packet-based world to the circuit-based one, and vice versa. This fact has two consequences. On one hand, the delay due to the voice codification/de-codification process is twice the delay in the previous scenario. On the other hand, the total impairment factor is also twice the value than that considered in the scenario before., due to the impairments are additive in eq. (3). Note that we are using one particular codec with G.711 format in between.

Another difference with the all IP scenario will be the absence of dejittering buffer. This is due to the fact of considering the radio path as error-free channel. All the packets entering the radio

Go to page 1

path will be present at the output ordered and equally delayed. Of course, the SCN does not introduce any kind of jitter.

G.723.1 at 5.3 kbps							
All IP scenario				Mixed IP-SCN scenario			
no Header Compression		Header Compression		no Header Compression		Header Compression	
samples per packet	estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS
1	3.134	1	3.25	1	2.364	1	2.491
G.723.1 at 6.3 kbps							
All IP scenario				Mixed IP-SCN scenario			
no Header Compression		Header Compression		no Header Compression		Header Compression	
samples per packet	estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS
1	3.338	1	3.338	1	2.783	1	2.783
G.729.A + VAD at 8 kbps							
All IP scenario				Mixed IP-SCN scenario			
no Header Compression		Header Compression		no Header Compression		Header Compression	
samples per packet	estimated MOS	samples per packet	Estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS
1	N.A.	1	3.965	1	N.A.	1	3.694
2	N.A.	2	3.917	2	N.A.	2	3.691
3	3.654	3	3.762	3	3.575	3	3.647
4	3.6	4	3.708	4	3.527	4	3.616
GSM-FR at 13 kbps							
All IP scenario				Mixed IP-SCN scenario			
no Header Compression		Header Compression		no Header Compression		Header Compression	
samples per packet	estimated MOS	samples per packet	Estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS
1	N.A.	1	3.305	1	N.A.	1	2.656
2	2.957	2	3.068	2	2.292	2	2.419
GSM-HR at 5.6 kbps							
All IP scenario				Mixed IP-SCN scenario			
no Header Compression		Header Compression		no Header Compression		Header Compression	
samples per packet	estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS
1	N.A.	1	3.25	1	N.A.	1	2.392
2	2.886	2	3.003	2	2.06	2	2.182
GSM-EFR at 12.2 kbps							
All IP scenario				Mixed IP-SCN scenario			
no Header Compression		Header Compression		no Header Compression		Header Compression	
Samples per packet	estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS	samples per packet	estimated MOS
1	N.A.	1	3.986	1	N.A.	1	4.081
2	3.701	2	3.795	2	3.803	2	3.906

Table 3. Estimated MOS from E-model.

Go to page 1

7. Conclusions and further work

The aim of the paper has been to analyse the suitability of DPRS as technology for transporting VoIP in FWA networks. Some assumptions have been taken into account in order to ease the tractability of characterising that issue. Thus, the results presented in here can be viewed as an upper bound in the system performance.

To evaluate the DPRS technology, The E-model has been considered and different MOS estimations, depending on network configuration, codec, voice packet framing and overhead penalisation, have been calculated. The E-model establishes in R=70 (MOS=3.6) the limit for an adequate transmission quality, that is, the range "some users dissatisfied" in the user satisfaction scale.

From all the reported codecs, only G.729.A (mainly in all IP case using header compression) and GSM-EFR seem to perform adequately in the framework of the analysis. The advantage of GSM-EFR is its low impairment factor. This is fundamental for the good figures obtained in the mixed IP-SCN scenario, where the value of I_e should be doubled regarding the all IP case. The advantage of G.729.A resides in its lower codec frame size, which translates to a lower latency in the end-to-end communication path. Also, a not too much high I_e value (if compared with the other codecs distinct to GSM-EFR) allows it to be above MOS 3.6 in the mixed IP-SCN scenario.

It should be noted that in the two network configurations, the parts acting as backbone introduce an excessive delay. For the all IP scenario, a more controlled backbone than Internet will perform better in terms of delay and jitter. For the mixed IP-SCN scenario, the processing time of the SCN will be usually much less than the considered 50 ms.

The results showed in table 3 do not take into account possible errors in the radio channel. The effect of such errors will be twofold. On one hand, recovery mechanisms should be enabled, having repercussion on the end-to-end latency (by the mechanisms themselves and by dejittering buffering). On the other hand, the codecs will be affected in distinct manner by the absence of voice codec words in the system, that having repercussion on the achieved voice quality. The study of the impact of errors in the radio channel is left for further study.

8. References

[1] EN 300 175-1. "Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 1: Overview". V1.4.2. June, 1999.
[2] EN 301 649. "Digital Enhanced Cordless Telecommunications (DECT); DECT Packet Radio Service (DPRS)". July, 1999.
[3] S.Casner, V. Jacobson. "Compressing IP/UDP/RTP headers for low-speed serial links". RFC 2508.
[4] ETR 250. "Transmission and Multiplexing; Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks". July, 1996.
[5] ITU-T Rec. G.107. "The E-model, a computational model for use in transmission planning". Dec., 1998.
[6] D. De Vleeschauwer, J. Janssen, G.H. Petit. "Delay bounds for low bit rate transport over IP networks". *Proc. SPIE Conf. On Performance and Control of Networks Systems III*, vol. 3841, pp. 40-48, Boston (MA), 20-21 Sept., 1999.
[7] J. Parantainen, S. Hamiti. "Delay analysis for IP speech over GPRS". *Proc. Fall Veh. Tech. Conf.*, pp. 829-833, Amsterdam, Sept. 1999.
[8] M.E. Perkins et al. "Speech transmission performance planning in hybrid IP/SCN networks". IEEE Comm. Mag., pp. 126-131, July, 1999.
[9] A. Kankkunen et al. "Voice over MPLS framework". <draft-kankkunen-vompls-fw-00.txt>. Work in progress. March, 2000.
[10] ITU-T Rec. G.114. "One-way transmission time". Feb., 1996