Proceedings of the 4th IEEE International Workshop on Satellite-Based Information Services (WOSBIS '99), pp. 48-55, Rio de Janeiro, Brazil, 8 December 1999.

Maximum Delay Bounds for Voice Transport over Satellite Internet Access Networks

Jan Janssen •, Rudi Windey °, Danny De Vleeschauwer •, Guido H. Petit •, Jean-Marie Leroy °

• Alcatel Bell, Corporate Research Center
Francis Wellesplein 1
B-2018 Antwerp, Belgium
E-mail: {jan.janssen, danny.de_vleeschauwer, guido.h.petit}@alcatel.be

° Alcatel Bell Space
Berkenrodelei 33
B-2660 Hoboken, Belgium
E-mail: {rudi.windey, jean-marie.leroy}@alcatel.be

ABSTRACT

Whether or not voice over IP calls of traditional quality can be supported between two users connected to the Internet via satellite access systems depends mainly on the mouth-to-ear delay, an important part of which is consumed by the satellite networks themselves. In this paper, upper bounds for the latter delay component are calculated for several codecs and different voice packet sizes.

1. INTRODUCTION

The quality of Voice over IP (VoIP) calls is mainly determined by the Mouth-to-Ear (M2E) delay, i.e., the time that elapses between the moment the talker utters the words and the moment the listener hears them. Other factors that may impair the quality of a VoIP call are the use of compression techniques, the level of echo and the occurrence of packet loss.

In [1], we calculated the largest M2E delays that can be tolerated for VoIP calls of at least "traditional quality". The resulting delay bounds depend on the type of codec used, the echo control mechanism and the packet loss ratio.

Both the access networks (connecting the two involved parties to the IP backbone) and the backbone network will consume a part of the tolerable M2E delay. Here, we assume that the IP backbone is accessed via satellites at both sides of the VoIP connection. The main objective of this paper is to calculate the part of the tolerable M2E delay budget allowed to be consumed by these satellite Internet access systems, and this for several standardized codecs and different voice packet sizes. The resulting effective bit rates, i.e., the rates at which bits are sent onto these access networks, are also worth investigating in this context.

We start by recalling the main results of [1] in section 2. In section 3, we recall some general facts on satellite Internet access systems and derive the formulae needed to determine the delay budget they may consume in a satellite-PC-to-satellite-PC VoIP scenario. Also a definition of the effective bit rate is given. The obtained results are presented and discussed in section 4. Finally, section 5 contains the main conclusions.

2. TOLERABLE M2E DELAY BOUNDS FOR VOIP CALLS

In this section, we recall the main results obtained in [1], where we used the E-model [4, 12] to predict the subjective quality of a phone call based on its characterizing transmission parameters. More precisely, the E-model combines all possible impairments (caused e.g. by compression of the voice signal, echo, packet loss, delay and loss of interactivity) into a rating factor denoted as R.

ITU-T Recommendation G.109 [9] defines the range of R-factors for a voice call to belong to a certain quality class: best ($90 \le R < 100$), high ($80 \le R < 90$), medium ($70 \le R < 80$), low ($60 \le R < 70$) or poor ($50 \le R < 60$). Voice communications with R-values below 50 are not acceptable. As the same Recommendation mentions that the term "toll quality" is ill-used, we prefer the use of "traditional quality" instead. We have taken an R-value of 72 as the limit for traditional quality. The reason for this choice is that for this value of R, the delay bounds obtained for the G.711 codec with the E-model correspond more or less to the values mentioned in ITU-T Recommendations G.114 [10] and G.131 [8] for voice in uncompressed format.

In Table 1, we recall the tolerable M2E delay bounds T_{M2E} obtained with the E-model for several standardized codecs. Under the assumption that the echo is perfectly controlled and that there is no packet loss in the network, one obtains at least traditional voice quality for delays below these bounds.

Origin	Standard	Type	R_{cod} (kb/s)	T_F (ms)	T_{LA} (ms)	B_F (bits)	T_{M2E} (ms)
ITU-T	G.711	PCM	64	0.125	0	8	379
	G.726, G.727	ADPCM	16	0.125	0	2	NA
			24			3	NA
			32			4	305
			40			5	356
	G.728	LD-CELP	12.8	0.625	0	8	192
			16			10	305
	G.729(A)	CS-ACELP	8	10	5	80	278
	G.723.1	ACELP	5.3	30	7.5	158	203
		MP-MLQ	6.3			189	237
ETSI	GSM-FR	RPE-LTP	13	20	0	260	192
	GSM-HR	VSELP	5.6	20	0	112	NA
	GSM-EFR	ACELP	12.2	20	0	244	324

NA : corresponding quality is Not Attainable

Table 1: Tolerable M2E delay values T_{M2E} below which traditional quality is obtained and characterizing parameters (see section 3.2) for several standardized codecs.

When the echo control is not perfectly controlled or when packet loss occurs in the network, similar delay bounds (below which at least traditional quality is attained) can be calculated with the E-model. Intuitively, we expect the resulting delay bounds to be smaller in both cases. For the case of non-perfect echo control, we refer to [1] for specific values. The quantitative study of the influence of packet loss is not yet completed as the impairments associated with it are still in the phase of standardization [11].

3. SATELLITE INTERNET ACCESS SYSTEMS

3.1 Introduction

For the satellite-PC-to-satellite-PC VoIP scenario considered in this paper (and illustrated in Figure 1), the total M2E delay can be divided into several components, one of them being the delay $T_{satellites}$ that includes all delays introduced by the satellite Internet access systems at both sides of the connection. This delay component consists of e.g. interleaving, (de)modulation, propagation, serving and queuing delays. We assume that the voice information is transmitted over a shared medium such that in the upstream direction also a Medium Access Controller (MAC) delay occurs.

The IP packets of a voice flow leave the talker's satellite PC at a constant packet rate. In the satellite access networks and the IP backbone, each individual IP packet is delayed over a stochastic time such that the flow is jittered, that is, the voice packets do not arrive at the listener's side at a constant rate. Since the decoder needs a constant flow of packets, a dejittering buffer is necessary in the listener's satellite equipment to compensate for this difference in delays. For the VoIP scenario considered here, the dejittering delay T_{jit} (the time the first voice packet is held up in this dejittering buffer) compensates for the difference in delays encountered between the talker's and the listener's satellite PCs. In order to limit the packet loss (due to packets arriving after they were supposed to be read out), it must not be chosen too small. On the other hand, in order not to increase the M2E delay too much, the dejittering delay must not be taken too large. In this paper, we approximated the dejittering delay T_{jit} by dividing it in two parts, compensating for the difference in delays in the IP backbone ($T_{jit}^{backbone}$) and the satellite access networks ($T_{jit}^{satellites}$), respectively. Although the latter component is not really consumed by the satellite Internet access networks, it is included in the delay budget $T_{satellites}$ foreseen for them.

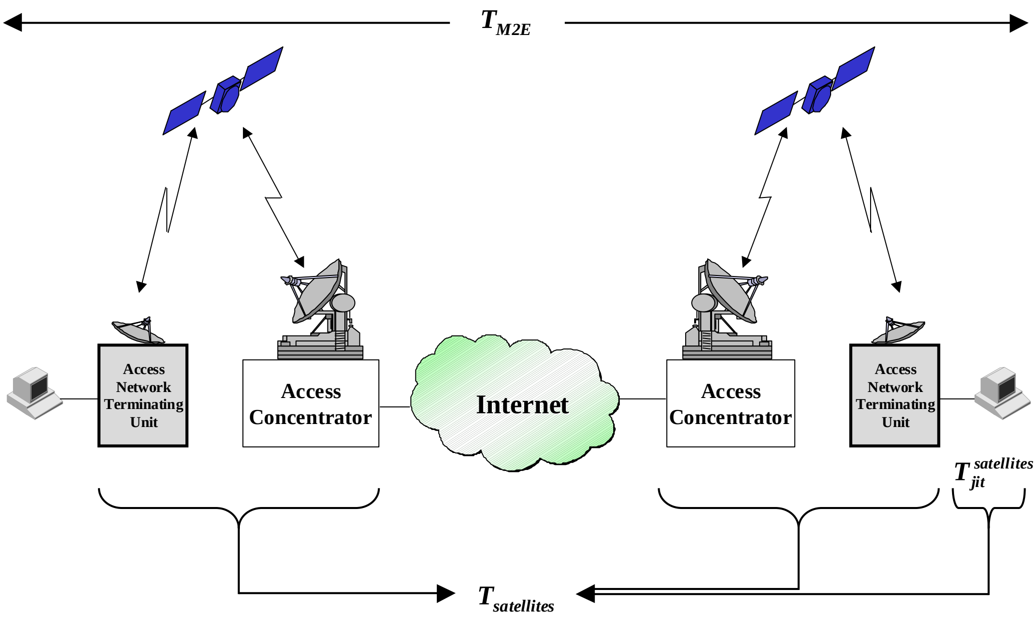

Figure 1: The satellite PC-to-satellite PC VoIP scenario.

The largest delay component in a satellite access system is the propagation delay between the satellites and the ground equipment, i.e., the satellite PC and the IP backbone network. An easy calculation shows that for geostationary satellites with an altitude of approximately 36000 km, the propagation delay over the radio link equals at least (without taking into account possible elevation angles) 2 x 36000 km x 3.33 µs/km = 240 ms at one side of the connection. This is the reason why satellite Internet access system designers will likely choose for low-orbit satellites (with an altitude of 1500 km), whose minimum corresponding propagation delay equals about 10 ms. The delays caused by the necessary handovers in such low-orbit satellite systems are also included in the delay budget $T_{satellites}$.

Finally, we mention that in case of multi-media service offering, the size of the majority of the IP packets will be too large to send them as one single burst over the radio link. Therefore, fragmentation is needed. In the sequel of this paper, we assumed that the voice packets are fragmented into ATM cells, and sent over ATM Adaptation Layer (AAL) 5 in the satellite Internet access networks.

3.2 Delay Budget

In order to calculate the portion of the M2E delay allowed to be consumed by the satellite Internet access systems, we identify several other important components of the M2E delay below.

All voice codecs work according to the same principle. They first collect a few quantized samples of speech, referred to as a voice frame of length T_F. Sometimes they also need some voice samples after the ones being encoded, referred to as the look-ahead of length T_{LA}, in order to better encode the samples of the current voice frame. Then the codecs calculate a code word of length B_F. This calculation takes an encoding time T_{enc}. At the receiver side the decoder uses the code word to produce a close copy of the original voice frame. Again it takes a certain time to perform this operation, referred to as the decoding time T_{dec}. We define the codec delay T_{cod} as $T_{enc} + T_{dec} + T_{LA}$. For the VoIP scenario considered in this paper, the encoding and decoding processes will be implemented in the end user's equipment (satellite PC), whose processing power is assumed to be large enough in order to perform both of them in a negligible amount of time. Hence, we have

$$T_{cod} = T_{LA} \quad . \tag{1}$$

An overview of the characterizing parameters (T_F, T_{LA} and B_F) for several codecs was included in Table 1. Observe that the codec bit rate R_{cod} equals B_F/T_F.

The packetization delay T_{pack} is the time needed to fill an IP packet with N_F voice code words produced by the encoding process. Therefore, one easily derives that

$$T_{pack} = N_F T_F \quad . \tag{2}$$

In order to characterize the service delay $T_{serv}^{backbone}$ in the IP backbone, we assume that this backbone network consists of N_{stag} nodes with link speed R_{link} (in kb/s). We then can write

$$T_{serv}^{backbone} = N_{stag} \frac{S_{link}}{R_{link}} = N_{stag} \frac{N_F B_F + O_{link}^{backbone}}{R_{link}} = N_{stag} \frac{N_F T_F R_{cod} + O_{link}^{backbone}}{R_{link}} \quad , \tag{3}$$

with S_{link} the packet size and $O_{link}^{backbone}$ the amount of overhead (in bits) on the link layer of the IP backbone. The latter equals 376 bits, stemming from 12 bytes RTP-header, 8 bytes UDP-header, 20 bytes IP-header and 7 bytes PPP-header [7]. Here (as well as in the rest of this paper), we neglect the fact that in IP packets are always rounded to an integer number of bytes.

The stochastic queuing delay $T_{queue}^{backbone}$ incorporates the time spent waiting behind other voice and data packets in the IP backbone network. If we assume that voice packets get non-pre-emptive Head-Of-Line (HOL) priority over data packets (arriving voice packets jump over all queuing data packets but a data packet in service is never interrupted), it is upper bounded by

$$\begin{aligned} T_{queue}^{backbone} &= D(\rho, N_{HL}, P)\frac{S_{link}}{R_{link}} + N_{stag}\frac{8(MTU_{data}+7)}{R_{link}} \\ &= D(\rho, N_{HL}, P)\frac{N_F T_F R_{cod} + O_{link}^{backbone}}{R_{link}} + N_{stag}\frac{8\left(MTU_{data}+7\right)}{R_{link}} \quad . \end{aligned} \tag{4}$$

The first term in this formula refers to the possible queuing behind voice packets in the IP backbone. It equals the (1-P)-quantile (with P very small) of the probability density function of the sum of N_{HL} statistically independent waiting times in M/D/1-queues. N_{HL} stands for the number of heavily loaded nodes and corresponds to the number of nodes at which the voice load reaches a certain maximum admissible voice load ρ. For nodes with lower voice loads, the queuing delay contributions turn out to be negligible. For more details, we refer the interested reader to [2, 5, 6], the first reference of which contains an explicit definition of $D(\rho, N_{HL}, P)$.

The second term of eq. (4) corresponds to the possible queuing behind a data packet in service in every node, with MTU_{data} the largest possible data packet size (in bytes) on the IP backbone, including 20 bytes TCP-header and 20 bytes IP-header. The 7 bytes of the PPP-header need to be added explicitly.

As mentioned already in section 3.1, the total dejittering delay is approximated by the sum of two terms, compensating for the difference in queuing delays encountered in the satellite access networks and the IP backbone, respectively. The dejittering delay component $T_{jit}^{backbone}$ is chosen equal to the difference between the "maximum" (read: (1-*P*)-quantile of eq. (4)) and the minimum queuing delay (i.e. 0) in the IP backbone, that is,

$$T_{jit}^{backbone} = T_{queue}^{backbone} \quad . \tag{5}$$

By doing so, we ensure that almost no (read: $P \times 100$ %) packets arrive too late in the dejittering buffer, while the contribution of the dejittering delay to the M2E delay is as minimal as possible.

Finally, we accumulate all delays that do not occur in the satellite access network and that are not discussed above in the so-called other delay T_{oth}. Its main component is the propagation delay, incurred when a bit traverses a certain distance over a transmission line. Here, it is assumed that a terrestrial propagation delay of 5 μs/km is introduced. Also included in this component are e.g. digitization, switching and echo control delays, which are maximally of the order of a few ms.

Now that all delay components are specified, the tolerable M2E delay bound can be written as

$$T_{cod} + T_{pack} + T_{serv}^{backbone} + T_{queue}^{backbone} + T_{jit}^{backbone} + T_{oth} + T_{satellites} \leq T_{M2E} \quad . \tag{6}$$

By inserting eqs. (1)-(5), we can derive the delay budget that may be consumed by the satellite Internet access networks, that is,

$$\begin{aligned} T_{satellites} \leq\ & T_{M2E} - T_{LA} - N_F\, T_F \left(1 + \frac{N_{stag}\, R_{cod} + 2\, D(\rho, N_{HL}, P)\, R_{cod}}{R_{link}} \right) - \frac{N_{stag}\, O_{link}^{backbone}}{R_{link}} - \\ & 2 \left(\frac{D(\rho, N_{HL}, P)\, O_{link}^{backbone} + N_{stag}\, 8\, (MTU_{data} + 7)}{R_{link}} \right) - T_{oth} \ . \end{aligned} \tag{7}$$

As the link speeds of future nodes in IP backbone networks are likely to be of very high speed (of the order of Gb/s and higher), it makes sense to replace R_{link} by ∞ in eq. (7). This results in

$$T_{satellites}^{\infty} \leq T_{M2E} - T_{LA} - N_F\, T_F - T_{oth} \quad . \tag{8}$$

For further reference, we finally define the effective bit rate of a codec in the satellite access networks as

$$R_{eff} = R_{cod} \frac{\left\lceil \dfrac{N_F\, B_F + O_{link}^{access}}{48 \times 8} \right\rceil \times 53 \times 8}{N_F\, B_F} \ , \tag{9}$$

with $\lceil x \rceil$ the smallest integer larger than *x* and O_{link}^{access} the amount of overhead (in bits) on the link layer of the satellite Internet access networks. It equals 400 bits (12 bytes RTP-header, 8 bytes UDP-header, 20 bytes IP-header, 2 bytes PPP-header [3] and 8 bytes AAL5-header). Observe that the effective bit rate equals the rate at which bits are effectively put onto the satellite Internet access network.

4. RESULTS AND DISCUSSION

For the parameter values R_{link} = 34 Mb/s, N_{stag} = 15, N_{HL} = 8, ρ = 0.8, MTU_{data} = 1500 bytes, $P = 10^{-5}$ and T_{oth} = 40 ms, Table 2 shows the maximum values of $T_{satellites}$ for different codecs and two different values of N_F. For each specific codec, the first value of N_F is the smallest value for which the effective bit rate does not exceed 3 times the original bit rate of the used codec. The second reported value of N_F is the largest value for which the budget for $T_{satellites}$ stays above 25 ms. Obviously, these assumptions on maximum effective bit rates and minimum satellite delay should be adjusted for every specific satellite system. The choices made here only serve illustration purposes. Table 2 also reports the maximum value of $T^{\infty}_{satellites}$ and the corresponding IP packet size (in bytes)

$$S_{IP} = \left\lceil \frac{N_F B_F}{8} \right\rceil + O_{IP} \ , \qquad (10)$$

with O_{IP} = 40 bytes RTP, UDP and IP-headers.

For the cases considered in Table 2, the complete decomposition of the M2E delay is illustrated in Figure 2. In this figure, we depict two bars per codec, a left one and a right one which correspond respectively to the smallest and largest value of N_F reported in Table 2.

First, we observe that $T^{\infty}_{satellites}$ is a rather good approximation of $T_{satellites}$, especially for the small values of N_F. Obviously, this approximation will gradually get better if the link speed R_{link} in our IP backbone is taken larger than 34 Mb/s. Therefore, all the following observations are valid both for $T_{satellites}$ and $T^{\infty}_{satellites}$.

	N_F	S_{IP} (bytes)	R_{eff} (kb/s)	$T_{satellites}$ (ms)	$T^{\infty}_{satellites}$ (ms)
G.723.1@5.3kb/s	2	80	14.22	82	95
	3	100	14.22	51	65
G.723.1@6.3kb/s	3	111	14.13	85	99
	4	135	14.13	54	69
G.729@8kb/s	4	80	21.20	179	193
	18	220	11.78	35	53
GSM-EFR@12.2kb/s	2	101	31.80	229	244
	11	376	17.35	41	64
G.728@12.8kb/s	53	93	38.40	105	118
	175	215	19.38	25	42
GSM-FR@13kb/s	2	105	31.80	97	112
	5	203	21.20	34	52
G.728@16kb/s	29	77	46.79	233	246
	343	469	19.78	25	50
G.726@32kb/s	71	76	95.55	242	256
	1622	851	37.64	25	62
G.726@40kb/s	57	76	119.02	295	308
	1932	1248	47.40	25	74
G.711@64kb/s	53	93	192.00	318	332
	1936	1976	73.59	25	97

Table 2: S_{IP}, R_{eff} and maximum values of $T_{satellites}$ and $T^{\infty}_{satellites}$ for two values of N_F. The first value of N_F is the smallest value for which $R_{eff} \leq 3\ R_{cod}$; the second value of N_F is the largest value for which $T_{satellites} \geq 25$ ms.

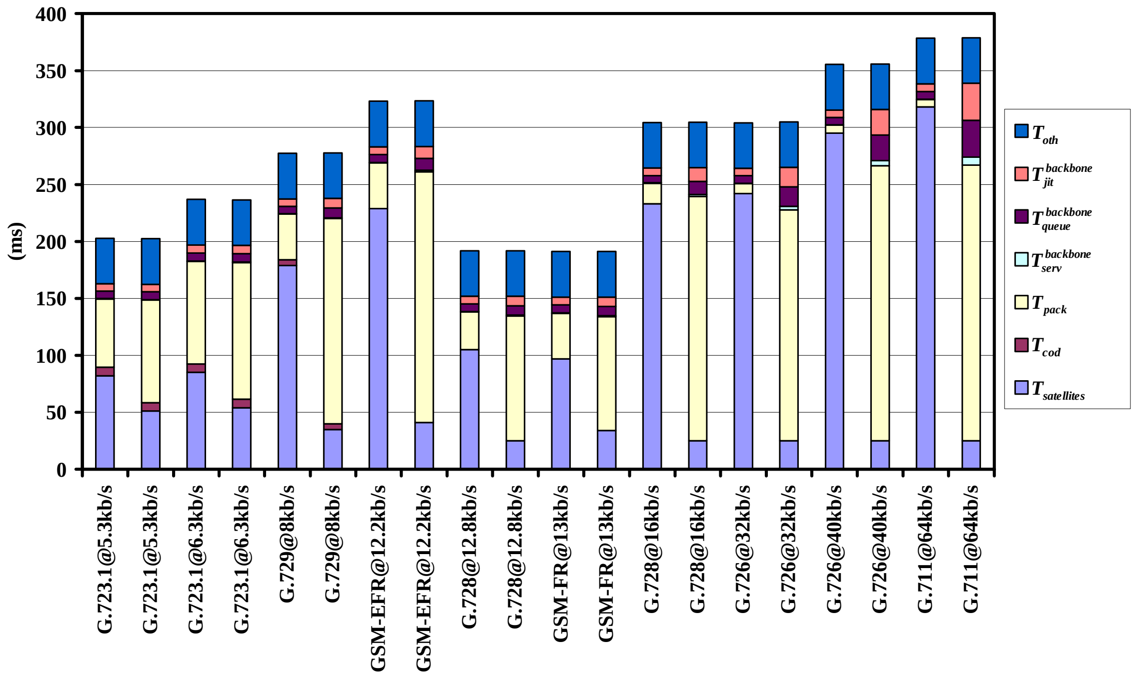

Figure 2: Decomposition of the tolerable M2E delay budget for the cases considered in Table 2.

Next, we see that the more voice code words gathered into one packet, i.e., the larger the voice packet size, the lower the effective bit rate and the less the delay budget left for the satellite access networks. On the other hand, putting fewer voice code words in one packet allows the satellite access networks to consume a larger portion of the M2E delay. Yet, by doing so, the influence of the header overhead increases such that the effective bit rate increases as well. This phenomenon can be explained by examining formulae (7) and (8), which immediately imply that the maximum satellite delay is linearly decreasing with the number of voice code words N_F that are put into one IP packet. It can also be recognized in Figure 2. Indeed, as the sum of T_{cod}, $T_{serv}^{backbone}$, $T_{queue}^{backbone}$, $T_{jit}^{backbone}$ and T_{oth} is almost independent of the packet size, the only way to increase the satellite delay budget is by decreasing the packetization delay.

The maximum satellite Internet access delays results derived above are likely to be a great help during the design process of such an access system. Moreover, they can be of use for system engineers of any type of delay-consuming Internet access network, as the results obtained above for satellite Internet access delays are basically (with only some minor adjustments) also valid for VoIP scenarios with other access networks. We conclude this section with some engineering tips for satellite Internet access systems.

First of all, it follows immediately that in a satellite-PC-to-satellite-PC scenario, VoIP calls of traditional quality cannot be supported over an access system based on geostationary satellites. Indeed, only the sum of the propagation delays over the radio links on both sides of the connection (480 ms) already exceeds the maximum satellite delays reported in Table 2. The former sum even exceeds the total tolerable M2E delay bounds reported in Table 1.

For low-orbit-based satellite Internet access systems in a shared medium configuration, a lot of care has to be taken in designing the MAC, which controls the time window and resources each of the user terminals can use from this shared medium. More precisely, the delay introduced by this MAC protocol should be kept as small as possible.

Finally, it should be recalled that all results presented in this paper do not consider the influence of packet loss. As indicated in section 2, the tolerable M2E delays that do take this effect into account will be smaller than the ones reported in Table 1, resulting in even tighter bounds for the satellite access systems. Therefore, the packet loss ratio (of the satellite access networks) should be kept as small as possible, which can be accomplished by the use of efficient error recovery techniques. Error concealment (implemented for several codecs) can help to diminish the remaining effects of lost voice packets.

5. CONCLUSIONS

In this paper, we proposed a model to calculate the delay budget that may be consumed by satellite Internet access systems for satellite-PC-to-satellite-PC VoIP calls of traditional quality. This budget, which obviously depends on the type of codec, turns out to decrease as the efficiency (amount of overhead versus payload information) obtained on the network increases. That is, the larger the delay budget for the satellite Internet access networks, the less efficient they are used, and, vice versa. Moreover, this simple analysis allows us to conclude that satellite Internet access systems based on geostationary satellites are not feasible for VoIP with traditional quality, while for low-orbit-based satellite access networks a lot of attention has to be paid to error recovery and concealment techniques and the development of the MAC.

ACKNOWLEDGEMENTS

This work was carried out within the framework of the project LIMSON sponsored by the Flemish Institute for the Promotion of Scientific and Technological Research in the Industry (IWT).

REFERENCES

[1] D. De Vleeschauwer, J. Janssen, G.H. Petit, “Delay Bounds for Low Bit Rate Voice Transport over IP Networks”, Proceedings of the SPIE Conference on Performance and Control of Network Systems III, Vol. 3841, Boston, USA, 20-21 Sep. 1999, pp. 40-48.

[2] D. De Vleeschauwer, G.H. Petit, S. Wittevrongel, B. Steyaert, H. Bruneel, “An Accurate Closed-Form Formula to Calculate the Dejittering Delay in Packetised Voice Transport”, Submitted to Networking 2000, 14-19 May 2000, Paris, France.

[3] G. Gross, M. Kaycee, A. Li, A. Malis, J. Stephen, “PPP Over AAL5”, IETF RFC 2364, July 1998.

[4] N.O. Johannesson, “The ETSI Computation Model: A Tool for Transmission Planning of Telephone Networks”, IEEE Communications Magazine, Jan. 1997, pp. 70-79.

[5] L. Kleinrock, “Queueing Systems, Volume 1: Theory”, John Wiley and Sons, New York, 1975.

[6] L. Kleinrock, “Queueing Systems, Volume 2: Computer Applications”, John Wiley and Sons, New York, 1975.

[7] W. Simpson, Editor, "PPP in HDLC-like Framing", IETF RFC 1662, July 1994.

[8] “Control of Talker Echo”, ITU-T Recommendation G.131, Aug. 1996.

[9] “Definition of Categories of Speech Transmission Quality”, ITU-T Recommendation G.109, Sep. 1999.

[10] “One-Way Transmission Time”, ITU-T Recommendation G.114, Feb. 1996.

[11] “Provisional Planning Values for the Equipment Impairment Factor I_e”, Appendix to ITU-T Recommendation G.113 (Draft), Dec. 1998.

[12] “The E-model, a Computational Model for Use in Transmission Planning”, ITU-T Recommendation G.107, Dec. 1998.