Improving VoIP Quality over Satellite Transmission with Dynamic Bandwidth

Franco Brandelero and Evelio M. G. Fernandez
Electrical Engineering Department
Federal University of Parana
Curitiba, Brazil
e-mails: francobrandelero@yahoo.com, evelio@ufpr.br

Abstract— **This work presents a dynamic bandwidth allocation scheme for DVB-RCS satellite standard aiming to avoid the conflict that arise when VoIP transmission with Voice Activity Detection (VAD) is used. A proposal is presented to improve the quality perceived by the user through bandwidth pre-allocation and validated by computer simulations using the Network Simulator (NS2).**

Keywords— DVB-RCS; VoIP; VAD.

I. Introduction

Recently it was adopted by the Brazilian Government, through decrees 2.592/1998 and 4.769/2003, the General Plan of Targets for Universalization (PGMU in Portuguese) that sets, among others, the following target: every region in Brazil with up to three hundred people must have at least one public phone installed on an accessible place 24 hours per day, with the capacity to originate and receive national and international long distance calls.

The faster way to take telephony services to these remote regions is through a satellite network. The satellite standard most used nowadays is the DVB-S (Digital Video Broadcasting - Satellite) that allows the use of VoIP (Voice over IP).

The bandwidth cost of a satellite network is very expensive, but with VoIP there is a great economy in relation to conventional TDM due to voice compression from the codec and silence suppression. Silence suppression, however, is not commonly used on DVB networks because of its incompatibility with the standard scheme used for dynamic bandwidth allocation.

In this paper, a bandwidth allocation scheme that allows the use of silence suppression is presented which is based on finding the minimum number of voice channels that fits a reasonable number of simultaneous calls. The proposed scheme makes a more efficient use of the available channels by saving bandwidth through the use of silence suppression and at the same time keeping the voice quality on acceptable levels.

The rest of this paper is organized as follow. Section II makes a brief explanation of the DVB-RCS standard and its effect over the perceived quality of a VoIP call when the voice activity detection (VAD) is enabled and presents some studies that tried to improve this quality. Sections III and IV present the proposed model to predict the bandwidth of VoIP calls with silence suppression. On sections V and VI the method developed is applied on a simulated system. A discussion about the results and suggestion for future works are presented in section VII.

II. DVB and VAD

The DVB-S is a widely accepted standard in the forward link of broadband satellite communication systems. The second generation (DVB-S2) [1] includes the transmission of multimedia contents and a variety of unicast and multicast services. The aforementioned wide applications of DVB-S2 need interactivity and thus, a return link over satellite is mandatory. The current counterpart of DVB-S2 for the return link is the DVB-return channel satellite (RCS) standard.

Bandwidth reservation messages are sent from the return channel satellite terminals (RCST) to the network control center (NCC) and a centralized scheduler applies the dynamic bandwidth allocation (DBA) algorithms. The bandwidth allocation is sent back to the RCSTs at least one round trip time (RTT) after the request was sent. Requests generated by the RCSTs depend on the queued traffic at the MAC queues of each terminal and are sent using the standard defined satellite access control (SAC) messages.

The DVB-RCS standard defines the following types of capacity request:

- Constant rate assignment (CRA);
- Rate based dynamic capacity (RBDC);
- Volume based dynamic capacity (VBDC);
- Absolute Volume based dynamic capacity (AVBDC);
- Free capacity assignment (FCA).

For VoIP applications the most appropriate scheme is a combination CRA + RBDC, where CRA guaranties a minimal bandwidth for VoIP signaling and RBDC is used to allocate bandwidth on demand.

Although the utilization of VAD offers a solution to the issue of bandwidth efficiency, when combined with a standard RBDC requesting mechanism it can lead to additional delay. This is one of the reasons for the performance degradation together with the fact that satellite systems are asymmetrically sized (typically 8:1 or 4:1), offering less bandwidth on the return path.

If a nominal RBDC scheme is used with VAD, it is not possible to sustain a continuous rate, as capacity is not requested during an ”off” period. This leads to an additional re-occurring delay of 0.4 - 1.5 secs every time capacity is re-requested for an ”on” period.

A solution to improve voice quality presented in [2] consists of a combination of RBDC + VBDC. In such a case the role of VBDC is to support best effort traffic if it exists as well as maintain RBDC by continuously sending capacity requests. The drawback of this approach is that to alleviate the effect of transition between silence and talk-spurt it is necessary a considerable quantity of simultaneous calls (more than 20), which makes this method impractical in regions with low traffic density.

In [3] a modification on the bandwidth request algorithm is proposed to estimate the mean rate used by the calls. The problem here is that the required bandwidth not always corresponds to the mean rate, mainly in conditions of low traffic where the mean is much lower than the needed bandwidth.

Another solution appears in [4] where it is proposed a SIP Proxy with extended functionalities to automatically make a resource reservation for a multimedia session. The disadvantage of this method is that the bandwidth allocated corresponds to a session without VAD, which really improves the perceived quality, but makes an inefficient use of the bandwidth.

To solve this problem keeping the call quality within acceptable levels and saving bandwidth (when possible), a pre-allocation must be made at the beginning of the call. The amount of bandwidth that must be allocated will be determined in the next sections.

III. Bandwidth allocation

When there are several simultaneous calls the probability of all of them being on talk-spurt state is small, so a statistical allocation is more appropriated.

VoIP traffic with VAD activated presents an ON/OFF profile; ON and OFF times having exponential distribution with random duration [5].

When the number of calls in the talk-spurt state is q, from a total of n calls, the system is said to be on state E_q. Let P_q be the proportion of time that the system remains on state E_q. Being L the rate of change to the talk state for one source, so the rate on which the transition $E_q \rightarrow E_{q+1}$ occurs is nLP_q.

Considering now the downward transition $E_{q+1} \rightarrow E_q$. Being α^{-1} the mean duration of the talk-spurt, the rate at the transition $E_{q+1} \rightarrow E_q$ occurs is $(q+1)\alpha P_{q+1}$.

Applying the principle of conservation of flow, the equation of static equilibrium [6] is obtained:

$$nLP_q = (q+1)\alpha P_{q+1}, \qquad q = 0, 1, \ldots, N-1. \tag{1}$$

The result that expresses each P_q in terms of P_0 is:

$$P_q = \frac{(nL/\alpha)^q}{q!} P_0. \tag{2}$$

Since P_q are proportions, they must sum to unity:

$$\sum_{q=0}^{N} P_q = 1. \tag{3}$$

Hence, P_0 can be found as follows:

$$P_0 = (\sum_{k=0}^{n} \frac{(nL/\alpha)^k}{k!})^{-1} \tag{4}$$

and the proportion of time P_q that q calls are in talk state is

$$P_q = \frac{(nVaf)^q/q!}{\sum_{k=0}^{N}(nVaf)^k/k!}, \qquad q = 0, 1, \ldots, N \tag{5}$$

where $Vaf = L/\alpha$ is the voice activity factor.

If there are q channels allocated, equation (5) can be rewritten to represent the blocking probability:

$$P\left(n, Vaf, q\right) = \begin{cases} 1, & q = 0 \\ \frac{(nVaf)^q/q!}{\sum_{k=0}^{q}(nVaf)^k/k!}, & 0 < q < n \\ 0, & q = n. \end{cases} \tag{6}$$

By setting the maximum value that $P(n, Vaf, q)$ can assume, the lowest value of q that fits n simultaneous calls can be obtained. Once found q, the bandwidth that must be allocated is computed as

$$B = \frac{q \cdot packet\_length}{T} \tag{7}$$

where T is the packetization period of the codec.

How to determine the higher acceptable value of $P(n, Vaf, q)$ is explained in next section.

IV. Packet level model

For VoIP calls with silence suppression activated, the probability distribution function (PDF) of the interarrival time for the superposition of n simultaneous calls is [5]:

$$A\left(t\right) = \begin{cases} 1 - (1-\lambda t)^{n-1}, & 0 < t < T \\ 1 - \frac{(1-p)^n \exp\left[-\beta n(t-T)\right]}{(\beta T + 1 - p)^{n-1}}, & t > T \end{cases} \tag{8}$$

and its probability density function (pdf) can be found by means of its derivative:

$$a(t) = \frac{dA(t)}{dt} \tag{9}$$

where

$p = \frac{N-1}{N}$ is the probability of packet arrival;

$N = \frac{\alpha^{-1}}{T}$ is the mean number of packets in a talk-spurt;

$\lambda = \frac{Vaf}{T}$ is the mean interarrival rate for one call in packets per second and,

β^{-1} is the mean duration of the silence period.

As most VoIP codecs use packets of constant size during the call, the pdf of the service time is given by

$$b(t) = \delta(t - \frac{1}{\mu}) \tag{10}$$

where μ is the capacity of the output link in packets per second and $\delta(t)$ the unit impulse function.

The waiting time distribution $w(t)$ of the $(j+1)$th packet that arrive on a RCST can be achieved by solving the integral Lindley's equation [7]:

$$w_{j+1}(t) = \begin{cases} w_j(t) * c(t), & t > 0 \\ \delta(t) \int_{-\infty}^{0^+} [w_j(t) * c(t)] dt, & t = 0 \\ 0, & t < 0. \end{cases} \quad (11)$$

where

$$c(t) = a(-t) * b(t), \quad (12)$$

$$w_1(t) = \delta(t), \quad (13)$$

and $*$ denotes convolution.

For a queue in statistical equilibrium,

$$w_{j+1}(t) = w_j(t) = w(t). \quad (14)$$

The estimation of MOS (Mean Opinion Score) quality level can be done through the E-Model [8] where the R factor can be obtained, considering only the network impairments such as delay and packet loss, as [9]:

$$R = 93.2 - Id - Ieef + A \quad (15)$$

where Id is the impairment caused by the absolute delay assuming perfect echo cancelation, A is the advantage factor and $Ieef$ is the effective equipment impairment factor which is calculated as:

$$Ieef = Ie + (95 - Ie) \frac{P_{pl}}{P_{pl} + Bpl} \quad (16)$$

where Ie and Bpl are tabulated values depending on the codec and P_{pl} is the packet loss probability and can be estimated as:

$$P_{pl} = \int_{J_B}^{\infty} w(t) dt = 1 - \int_{0}^{J_B} w(t) dt \quad (17)$$

where J_B is the size of the jitter buffer in time units.

To achieve the higher possible MOS level, P_{pl} must be zero. For a fixed number of calls, P_{pl} must be computed changing the value of μ in the range:

$$n\lambda < \mu \leq \frac{n}{T}. \quad (18)$$

The lower limit comes from the convergence limit imposed by queuing theory, i.e., below this limit both the queue size and waiting time go to infinity [6]. It is also important to note that this value corresponds to the mean rate for n calls.

The upper limit is the rate used if VAD were deactivated; above this value the waiting time is always zero.

Once found the lower value of μ that makes $P_{pl} = 0$, apply on (6) to find the higher acceptable blocking probability.

V. Allocated and Needed Bandwidth

To verify the reliability of the presented method 100 simulations over a terrestrial link using NS-2 (Network Simulator Version 2.33) were performed. Each simulation was configured with a different number of simultaneous calls. The simulation parameters were selected in a way to simulate a G.729annexB codec with VAD activated [10] with packetization period of 40 ms, means of 360 ms and 650 ms for ON and OFF periods respectively, resulting on a voice activity factor of 35.6%, according to studies of [3] and [5]. For each simulation the maximum rate used was measured and compared with the allocated bandwidth calculated through equation (7).

A blocking probability of 3% was found suitable for this configuration, i.e., for a jitter buffer of 40 ms the packet loss calculated using (17) is zero. The results are shown in Figure 1.

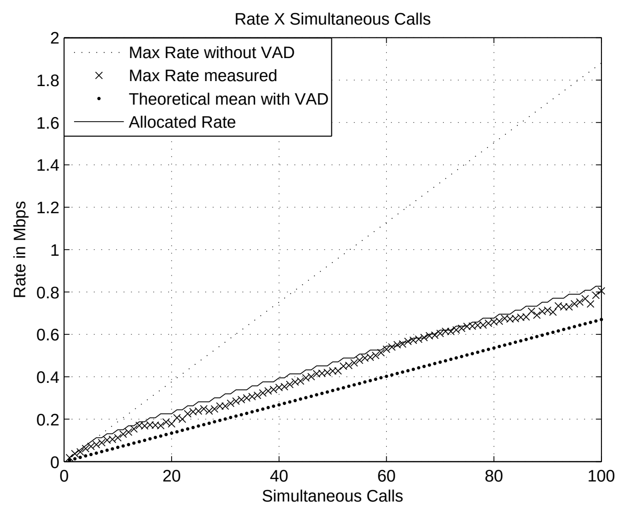

Fig. 1. Comparison between measured and calculated rates for bandwidth allocation.

For comparison, the dotted line corresponds to the bandwidth that would be needed if VAD were deactivated. The theoretical mean with VAD on the graphic is the multiplication of the bandwidth without VAD by the voice activity factor, which is the calculation commonly done by designers of VoIP networks to estimate the needed bandwidth. But it can be clearly seen that this value is always below the bandwidth measured, which results on voice quality degradation.

In the case of few calls it is worth noting that, in this example, for six or less calls there is no bandwidth saving (see Figure 2), but it is the only solution that keeps voice quality on an acceptable level.

VI. Performance Simulation using NS-2

The satellite network simulated using NS-2 with TDMA-DAMA patch [11] is composed by 20 RCSTs, also known as VSATs (Very Small Aperture Terminal) connected to an Access Gateway (AGW) and to a Local Area Network (LAN).

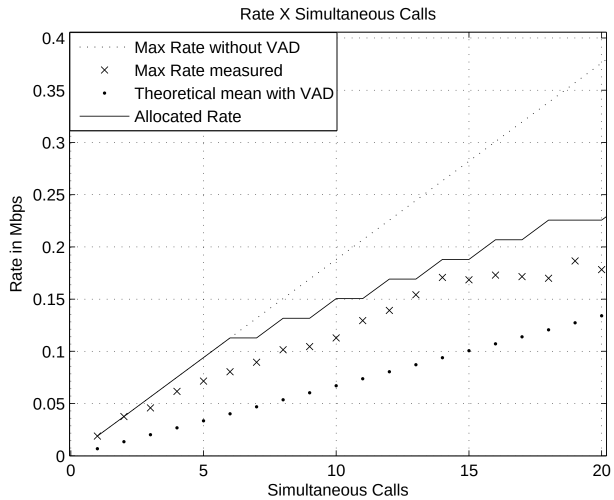

Fig. 2. Comparison between measured and calculated rates for allocation for few simultaneous calls.

All originated data on VSATs are forwarded (through geostationary satellite) to the satellite Hub that is connected to a terrestrial backbone with Internet access and a Trunking Gateway (TGW) for interconnection to the Public Switched Telephone Network (PSTN) as showed in Figure 3.

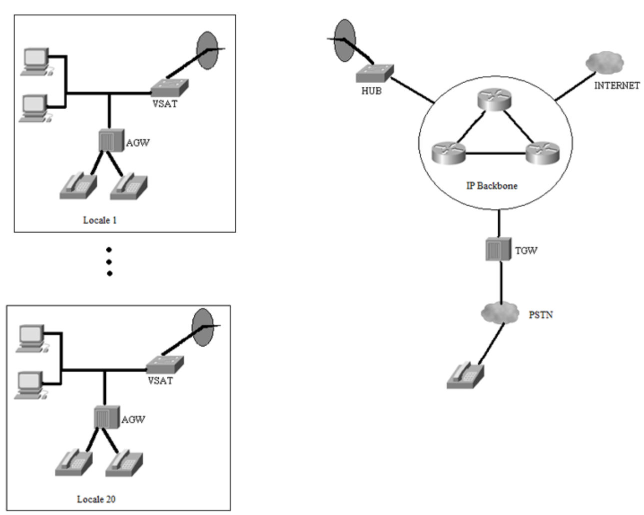

Fig. 3. Topology for Simulation.

On each locale there are four elements (beyond VSAT) belonging to the same LAN, i.e., sharing the same access medium. One will be doing a FTP transfer with the Internet provider, simulating a download. Another element will be doing HTTP requisitions at random intervals and receiving data of random size, simulating Web navigation. The other two elements will be on a bi-directional communication of RTP [12] traffic, as two VoIP calls from a common Access Gateway. This configuration was chosen because it represents a typical use of VoIP together with Web traffic and files transfers, very common on Internet today.

The voice quality evaluation was made by measuring MOS according to E-Model through the analysis of network parameters (delay, jitter and packet loss) observed on the Trunking Gateway (measures must be done on this element because it is on this direction that the satellite return channel is used).

The return channel have a bandwidth of 2 Mbps shared among all terminals through time division multiple access (TDMA) with a frame being composed by 95 timeslots (94 usefull) of 106 bytes and where the superframe is composed by three frames. The bandwidth available for each terminal is dynamically allocated through assignment of timeslots as requested.

The RTP flow analyzed has the same characteristics of that of session V with 60 seconds of duration. The receiver was using a jitter buffer of 40 milliseconds to soften the delay variation inherent on a VoIP net.

A. *Standard Scheme Allocation without VAD*

For comparison effects the first simulation corresponds to the scenario traditionally used with VAD deactivated and the standard VoIP allocation scheme (RBDC + CRA). Figure 4 shows the packet end-to-end delay.

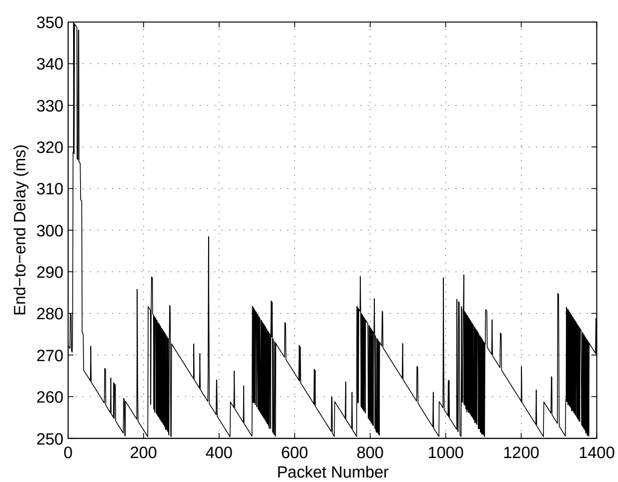

Fig. 4. End-to-end delay with standard scheme allocation.

As the delay of the first packet was 272 milliseconds, all packets with delay higher than 312 milliseconds were dropped at the receiver. At the beginning of the transmission there is a peak of 350 milliseconds caused by the lack of bandwidth at this moment; soon after one more timeslot is allocated for this VSAT, as can be perceived by the stabilization of the delays times. The result is 1.42% of packet loss, jitter of 6.7 milliseconds and the perceived quality on MOS scale is 3.91 which correspond to some users dissatisfied.

B. *Standard Scheme Allocation with VAD*

When the traffic is at ON state there is an allocation of the satellite bandwidth and at OFF state a bandwidth deallocation. Every time an allocation is needed there is an additional delay on the packets, which is the time of the packet arrival till the time that there is enough bandwidth to be delivered. Figure 5 shows the packet delay.

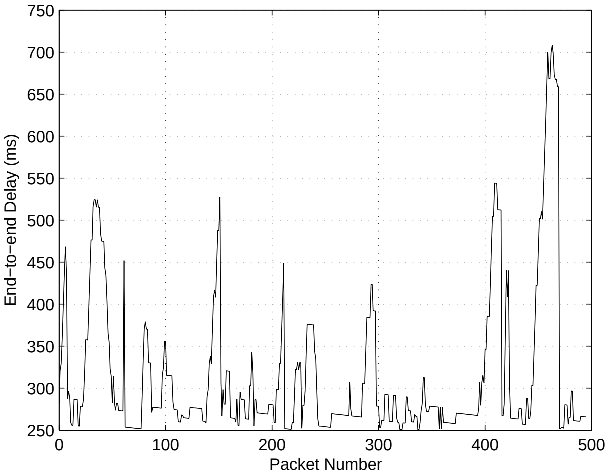

Fig. 5. End-to-end delay with standard allocation and VAD activated.

First packet delay was 281 milliseconds, so every delay above 321 milliseconds are dropped, achieving 26% of packet loss, jitter of 11.73 milliseconds and the perceived quality on MOS scale is 1.74 which corresponds to all users dissatisfied.

C. Pre-allocation Scheme with VAD

In the next experiment, the request algorithm used by the VSAT was changed according to (7) to pre-allocate bandwidth at the beginning of each RTP session. The result is showed in Figure 6.

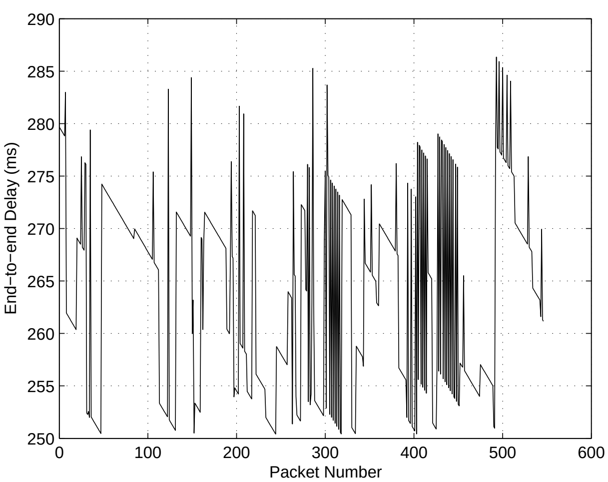

Fig. 6. End-to-end delay with pre-allocation.

Now all delays are below 290 milliseconds which means that there is almost no queue formation on the VSAT. There was no packet loss, jitter of 2.06 milliseconds and MOS equal to 4.11, which means users satisfied, a better result than previous scenarios.

VII. Conclusion

To deliver telephone services to remote regions a good option is to use VoIP over satellite network with DVB-RCS, but traditional networks actually does not benefit of VAD due to deterioration of voice quality. The proposed allocation scheme solves this problem by allocating enough bandwidth to accommodate all existing calls. The allocation is done by the maximum required bandwidth which guarantees that enough bandwidth could be allocated even on conditions of low traffic as observed on remote places with up to 300 inhabitants served by PGMU. The presented method is also useful for designing of non-satellite VoIP networks by accurately estimating bandwidth consumption with VAD activated.

As future work it is proposed a study about the most efficient way to transport information between MAC layer and application layer and how to implement it over a real system.

References

[1] ETSI EN 301 790 v1.3.1, *Digital Video Broadcasting (DVB); Interaction channel for satellite distribution systems* March, 2003.

[2] A. Sathiaseelan, G. Fairnhurst, *Performance of VoIP using DCCP over a DVB-RCS satellite network* In ICC'2007, IEEE International Conference on Communications, pages 13-18.

[3] M. Kalama, G. Acar, B. G. Evans, S. Mourgues, *VoIP over DVB-RCS Satellite Systems: A novel Capacity Request Mechanism for improved Voice Activity Detection* In IEEE VTC Spring 2008, pages 2957-2961.

[4] O. Alphand, P. Berthou, T. Gayraud, S. Combes, *QoS arquitecture over DVB-RCS satellite networks in a NGN framework* In IEEE Globecom 2005.

[5] K. Sriram, *Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data* IEEE Journal on Selected Areas in Communications, Vol. 4, No. 6, pp. 833-846, September 1986.

[6] R. B. Cooper, *Introduction to Queueing Theory* Elsevier North Holland Inc., 1981.

[7] M. H. Ackroyd, *Computing the Waiting Time Distribution for the G/G/1 Queue by Signal Processing Methods* IEEE Transactions on Communications, Vol. Com-28, No. 1, January 1980.

[8] ITU-T Rec. G.107, *The E-model, a computational model for use in transmission planning* May, 2002.

[9] Vitalio Alfonso Reguera, Felix F. Alvarez Paliza, Walter Godoy Jr. and Evelio M. Garcia Fernandez, *On the impact of active queue management on VoIP quality of service* Computer Communications, Vol. 31, Issue 1, pp. 73-87, January 2008.

[10] ITU-T Rec. G.729 Annex B, *A silence compression scheme for G.729 optimized for terminals conforming to Recommendation V.70* November, 1996.

[11] R. Secchi, *Tdma-dama: satellite on-demand bandwidth allocation for ns-2* [online] In http://wnet.isti.cnr.it/software/tdma-dama.

[12] RFC 1889, *RTP: A Transport Protocol for Real-Time Applications* January, 1996.