SIP-based Resource Allocation for Interactive Multimedia Applications over DVB-S2/RCS Satellite System

M. Gineste, F. Nivor, P. Berthou, and T. Gayraud, *LAAS-CNRS University of Toulouse, Toulouse, France*

***Abstract*— This paper introduces a SIP-based approach to offer QoS guaranties to interactive multimedia applications over an efficient DVB-S2/RCS satellite access assignment scheme (on-demand). It focuses both on the communication opening which represents the weakness of on-demand capacity allocation and optimizes the remaining communication in the satellite context (due to important delays). It finally presents analytical and experimental results of the various proposed enhancements.**

***Index Terms*— multimedia application, quality of service (QoS), satellite, DVB-RCS, DAMA, RBDC, SIP.**

I. Introduction

GEOSTATIONARY satellites offer practical and easy-to-deploy way for Internet communications over the world. Quality of Service (QoS) guarantee on the return link has been often implemented using static resource reservation (e.g. for TV journalists working where terrestrial infrastructure is not available). Nevertheless, due to limited return link resources with high cost, the static resources management scheme is not suitable for interactive multimedia application customers, such as audio-video conferencing. In order to provide a more efficient use of the return link resources and to reduce their cost, a dynamic distributed approach is proposed in [1].

The basis architecture of such DVB-S2/RCS access satellite system is the following: the Satellite Terminals (STs) behave as access routers to the satellite network for the user traffic on the return link channel. A Gateway (GW) centralizes the whole traffic in the satellite network and interconnects with core networks. The resource management (RM) of the satellite network is done by an entity installed in the Gateway (GW-RM). When STs want to emit data on the return link channel, they emit periodical resource requests to the GW-RM. This latter periodically responds by broadcasting the resource allocation plan of the return link channel. However, this allocation scheme introduces large delays (at least 600ms in the best case) due to the request/response cycle (c.f. Fig. 1): 250ms for the transmission of the resource request; 100ms for the computation of the resource allocation plan; 250ms for the broadcast of the resource allocation plan.

Such delays impact interactive multimedia applications. ITU-T organization specifies that such latencies are not supported by interactive applications like audio conference or video conferencing [2].

In this context, several propositions have been expressed to improve performances of the access scheme for multimedia communications: [3] uses a predictive dynamic scheme that allocates free bandwidth based on an estimation of the positive varying trend of Internet traffic at each station. It uses the statistical property of Internet traffic to predict traffic arrivals. [4] is based on a Fractional Gaussian Noise. It uses the distribution of the incoming bytes number in future frames. [5] proposes an efficient solution by adopting control methodologies to model the satellite network entities and to conceive a load adaptive BoD protocol. [6] and [7] use an adaptive predictive control for the input traffic flow. They formulate the DBA problem as an optimal control problem with cost trading off queue occupancy and bandwidth waste.

These works allow reducing this delay during the communication. They anticipate the future traffic and request more resources without waste. However, there is currently no anticipation solution for the beginning of the stream (c.f. Fig. 1). The blue curve represents traffic leaving the application, the red curve traffic leaving the ST. The difference is buffered in the ST transmission queue. At the beginning, the ST requests more resources to the GW-RM in order to drain the first buffered packets of the multimedia stream. So the first resources for the beginning of interactive multimedia are available in the best case at least 600ms.

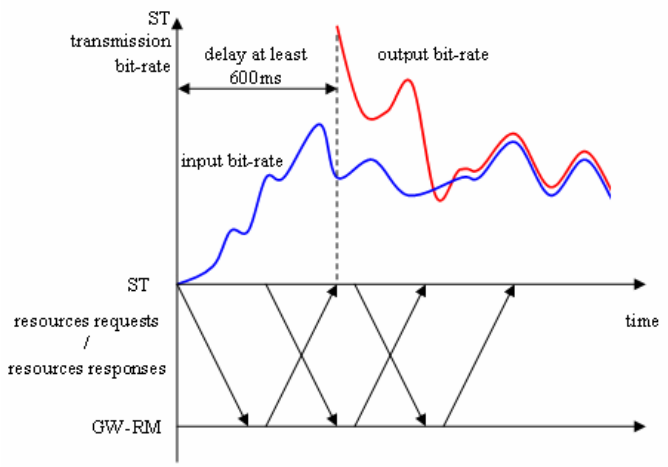


Fig 1. Beginning of a multimedia flow with a dynamic allocation scheme

Therefore this is an incompressible delay because of the real-time property of two-way conversational voice and video flows. It's not possible to accelerate the stream in order to catch up the delay. This initial delay degrades the overall interactivity of interactive multimedia applications and directly increases the end to end delay.

To fill in this gap, this paper presents an approach that improves the quality of interactive applications over satellite systems: (I) by provisioning on-demand resources for the beginning of the flow transmission; (II) by improving the resource request algorithm during the flow transmission. This approach takes advantage of information already included within the session establishment signaling.

The paper is structured as follow: Section 2 introduces the deployment of a session establishment signaling protocol (SIP) [8] over an access satellite system. Section 3 based on SIP details the enhancements for interactive multimedia applications using the on-demand access scheme for the beginning of the multimedia flow. Section 4 also based on SIP details the enhancements concerning the remainder of the multimedia flow. Finally section 5 presents simulation and experimental results: impacts on QoS are measured through the end-to-end (e2e) delay of applicative data.

II. SIP OVER A SATELLITE SYSTEM

SIP (Session Initiation Protocol) is a session establishment signaling used to initiate, manage and terminate multimedia session. The following scenario (Fig. 2) presents how SIP can be used to establish a multimedia session over a satellite system.

Knowing that the GW-RM represents the entry point of the satellite network, an Outbound SIP Proxy is deployed in the GW-RM side. It is set up with "loose route" option and is in charge of the overall SIP sessions in the Satellite System. Another SIP Proxy is deployed in each ST side. It is in charge of SIP sessions from its user network.

The following scenario describes a session initiation between a caller SIP client in the satellite network side, and a callee SIP client in the other terrestrial side. The caller SIP client transmits a SIP INVITE message to the SIP Proxies (1). They transmit the invitation to the callee SIP Client (2, 3). The callee SIP Client responds with a SIP OK message (4). The SIP Proxies forward the response to the caller SIP Client (5,6). The caller SIP Client acknowledges by transmitting a SIP ACK message to the callee SIP Client (7,11). The caller SIP Client starts to transmit the multimedia flow (8). The ST requests resources to the GW-RM in order to transmit the data (9). The GW-RM responses with the resources allocation plan (10). The ST transmits the multimedia data to the GW-RM (12). The GW-RM transmits the multimedia data to the remote SIP Client. By using the classical SIP architecture, the beginning of the multimedia flow is delayed at least 600ms. In order to limit this initial delay, we propose to provision resources by using information integrated in SIP session establishment.

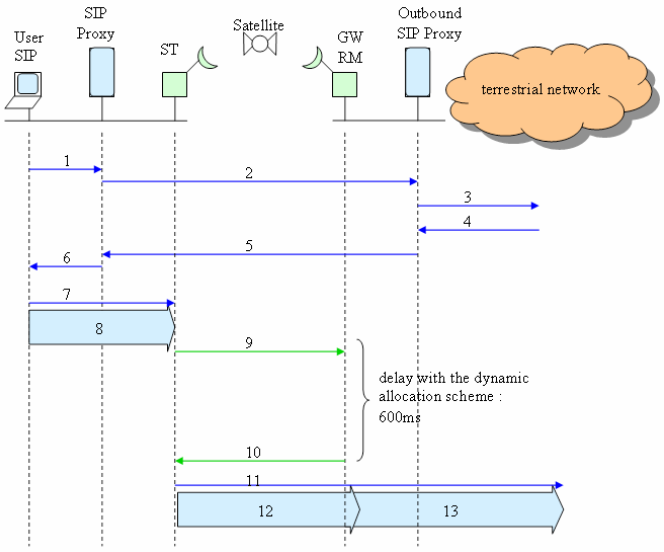


Fig 2. SIP session with a dynamic allocation scheme in satellite system

III. SIP-BASED PROVISIONING

Our first proposition of SIP-based resources provisioning is realized in two steps (c.f. Fig. 3):

First, the Outbound SIP Proxy makes a resource request to the GW-RM: To do so, the Outbound SIP Proxy located on the signaling path receives the SIP OK message (4). It translates the qualitative information (e.g. application types, media types, codec names) into precise quantitative QoS parameters (e.g. bit-rate, packet size, session id) (5). To achieve this translation, the Outbound SIP Proxy stores a local list of association between <application, media type, codec name> and <transmission bit-rate>. So the amount of data required for this session before the multimedia flow starts is known. The Outbound SIP Proxy is able to make a precise resource reservation request to the GW-RM (6a) on behalf of the appropriate ST.

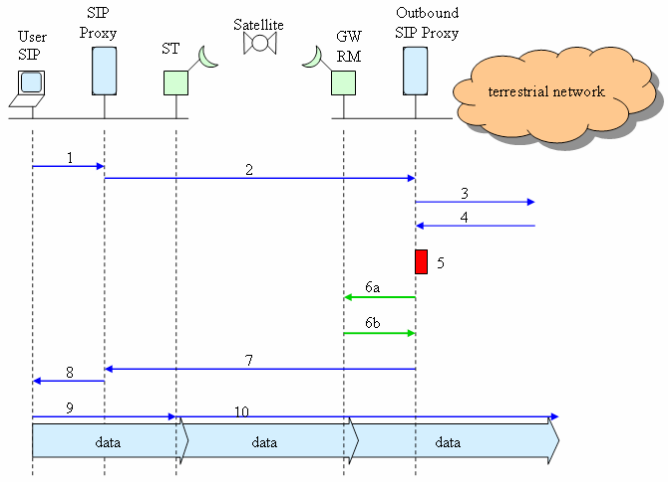


Fig 3. SIP-based provision in satellite system (local initiation)

Then, at the second step, the Outbound SIP Proxy delays the session establishment: the Outbound SIP Proxy keeps the SIP OK message until the resource allocation plan is computed. It's equivalent at user level to a longer call ringing.

The GW-RM warns the Outbound SIP Proxy that the resource allocation plan is about to be sent using a notification (6b). Then, the Outbound SIP Proxy forwards the SIP OK message to the caller SIP Client (7, 8).

The SIP OK message and the resource allocation plan are transmitted at the same time on the forward link. So the ST is able to send immediately the beginning of the multimedia flow on the return link. **This is important because as the opposite of a classical architecture, the multimedia stream is not buffered, but the session initiation is delayed. Then the call setup duration is increased but not the end-to-end delay.**

This proposition is adopted when a SIP Client inside the satellite network initiates the multimedia session. But when the SIP Client inside the satellite system receives a multimedia session invitation, the enhancements are different (c.f. Fig. 4):

The Outbound SIP Proxy receives the SIP OK message (5). It translates the qualitative information into precise quantitative QoS parameters (6). Then it transmits the SIP OK message to the remote SIP Client (7). It performs a resource reservation but cannot make a resource allocation because the session establishment is not yet completed.

Once Outbound SIP Proxy receives the SIP ACK message (8), thanks to the "loose route" option, it performs the resource allocation:

- It makes the request to the GW-RM (9a);
- It waits for the GW-RM notification (9b).

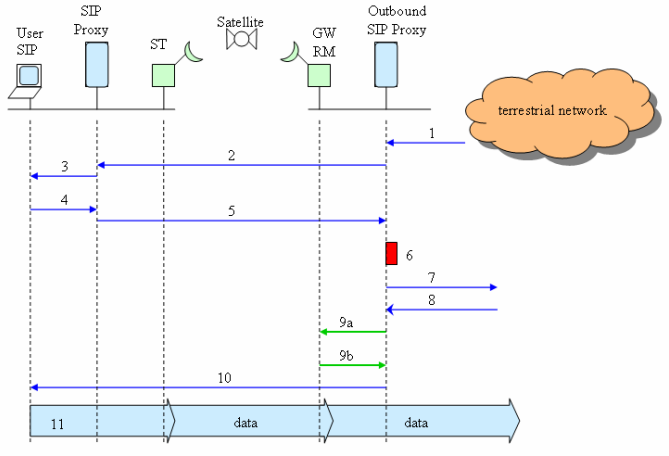


Fig 4. SIP-based provision in satellite system (remote initiation)

Then, the Outbound SIP Proxy forwards the SIP ACK message to the caller SIP Client (10). Thus the ST is able to send immediately the beginning of the multimedia flow on the return link (11).

IV. SIP-BASED ANTICIPATION ALGORITHM

The second proposition optimizes the next part of the communication. Once the STs have the required resources to transmit the beginning of multimedia flows, a resource request algorithm in the ST at MAC level handles the remaining of the multimedia flow. The requests must be periodical to optimize the bandwidth usage and follow as much as possible the bandwidth variation.

The most evolved algorithms use anticipation mechanisms to predict and cope with traffic variations. Our proposition is based on one of them [10] which was retained by the SatSix [11] QoS architecture.

This algorithm uses an α anticipation parameter that increases more or less the resource anticipation. Fixed in the [0;1] interval, this α factor allows balancing the amount of traffic previously received in the MAC queue for the future resource requests.

Fig. 5 sums up results from simulations in order to evaluate the impact of this α parameter both on the latency and the efficiency of this algorithm. In this experimentation, the input traffic is a video-conferencing flow using H.264 codec [13], with transmission bit-rates varying between 700Kbit/s in average and 2Mbit/s in peak.

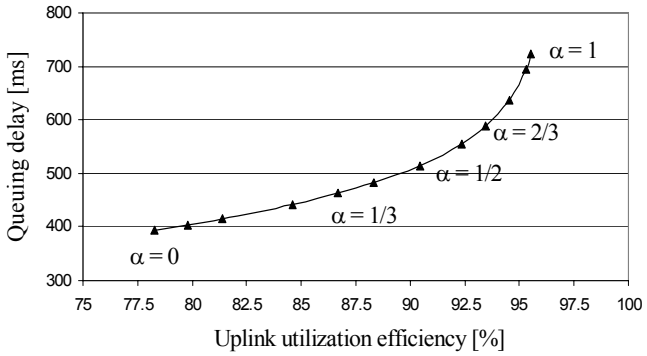


Fig 5. Queuing function of utilization

The queuing delay measured is the latency experienced by a packet crossing the satellite network. This codec produces high variables transmission bit-rates consequently fast transmission bit-rate variations produce breaks in the allocation mechanism which should quickly reacts to the traffic burst arriving in the transmission queue. An α anticipation factor near to 0 increases resources requests for future needs. So the e2e delay is less than 400ms. It is compatible with the strict time constraints associated to interactive multimedia applications [2]. But when the required resources are not finally used, they are lost. It creates an under use of the return link channel in the satellite network: the return link efficiency is less than 77%. In the other extremity, an α anticipation factor near to 1 decreases the future resource needs. So the received resources can not handle the overall incoming traffic. Therefore, e2e delay is more than 700ms which is not compatible with time constraints of interactive flows [2]. However the return link efficiency is optimized to 95%.

Today, the α anticipation parameter is fixed by the network operator and stays constant. This puts the operator in front of dilemma of the allocation optimization of the satellite network and the quality of service for the customers.

We propose to use information integrated in SIP session establishment in order to compute a dynamic α anticipation parameter. The following scenario extends the previous one (c.f. Fig 6).

The SIP Proxy in the ST side receives the SIP OK message (5). It translates the qualitative description of the media used into their transmission bit-rate (6). This latter is passed to the ST MAC layer, in order to compute a dynamic α anticipation factor during the multimedia flow (7).

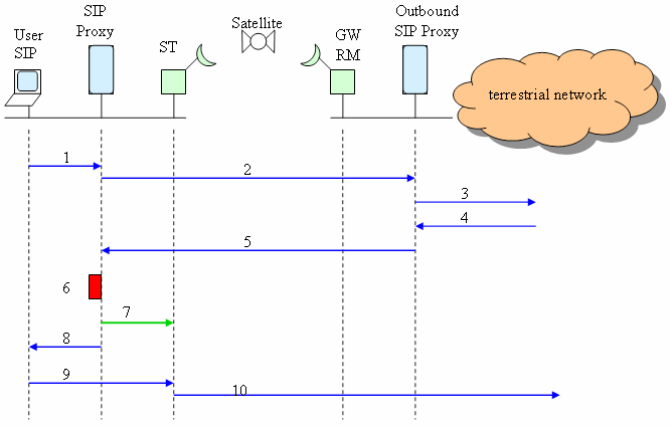


Fig 6. SIP-based anticipation in the ST side

The objective of this method is to obtain a high anticipation when the traffic is variable (e.g. video-conferencing) and a weak anticipation (thus a high utilization) when the traffic is stable (e.g. VoIP).

Thus, we propose to link the arrival transmission bit-rate of packets (d_I), measured on the transmission queue, to the characteristics (transmission bit-rate) announced by the traffic through SIP (d_C). This computation is done before every resource requests thus with the same period of a super-frame (ie. 500ms). The arrival rate in the ST (d_I) is measured from a sliding window with the same period. From the arrival transmission bit-rate of packets (d_I) in the ST and knowing the transmission peak bit-rate (d_C) of each multimedia application, we calculate the α anticipation parameter through the formula:

$$\alpha = \frac{d_I}{\sum d_C}$$

This formula, with the use of dual token bucket at the entry of the MAC queue, guarantees that the anticipation factor is included between 0 and 1. Besides, a MAC transmission queue is reserved to the SIP signaled application in the ST.

V. EXPERIMENTS AND RESULTS

A. Evaluation of the efficiency of SIP-based resource provisioning approach

First, we propose to evaluate analytically and experimentally our first proposition. Its impact on QoS is measured through the end-to-end (e2e) delay of applicative data. Experiments are conducted on a satellite emulation platform proposed in SatSix European project [11], aiming at functionally validate the DVB-RCS access and network layers definitions. The satellite network platform emulates components and implements communication protocols (IP/MAC/DAMA) of a regenerative satellite system. The scenarios presented in Fig. 3 and 4 are implemented.

To measure the performance of the propositions, an audio-conferencing flow with G711 audio codec at 64kbit/s is used with the dynamic allocation scheme.

Parameters of the emulated satellite system are setup to:

- Satellite link capacity: 2048kbps
- Super-frame periodicity: 500ms
- Resource request period: 500ms
- Resource allocation plan emission period: 500ms

We propose to evaluate the solution through two scenarios:

- Scenario 0 (Fig. 2) is the reference test: it evaluates resources allocation with no anticipation (basic access behavior);
- Scenario 1 (Fig.3) evaluates our solution.

The e2e delay that should be observed for the first packet of scenario 0 is:

$$DelayFirstPackets_{VoIP} = EmissionDelay_{request} + PropagationDelay_{request} + ComputationTime_{request} + propagationDelay_{AllocationPlan} + EmissionDelay_{VoIPPacket} + propagationDelay_{VoIPPacket}$$

$$DelayFirstPackets_{VoIP} = 500\text{ms} + 250\text{ms} + 100\text{ms} + 250\text{ms} + \frac{212 * 8}{64}\text{ ms} + 250\text{ ms} = 1376.5\text{ ms}$$

This first sum represents the e2e delay in the case of the reference scenario (scenario 0). It corresponds for the transmission of the first packets to a delay around 1376.5 ms which is too high for interactive applications [2]. This high delay is the consequence of the on-demand access scheme introducing delay for requests computation and propagation over the satellite system.

The e2e delay that should be observed for the first packet of scenario 1 is:

$$DelayFirstPackets_{VoIP} = EmissionDelay_{VoIPPacket} + propagationDelay_{VoIPPacket}$$

$$DelayFirstPackets_{VoIP} = \frac{212 * 8}{64}\text{ ms} + 250\text{ ms} = 276.5\text{ ms}$$

This second sum corresponds to the e2e delay for the scenario 1, anticipating the resource request on the gateway (GW-RM) side before the interactive flow starts. Thus it avoids all the propagation and computation delays introduced by the first requests of the on-demand access-scheme. This theoretical delay is around 276.5 ms which is compatible with interactive flow constraints [2].

On the experiments, for the reference test without resource anticipation (scenario 0), the average e2e delay is around 1400ms which confirms the analytical results. The e2e delay experimented for this scenario is not compatible with the strict time constraints associated to interactive multimedia flow [2].

For the solution with acknowledgment (scenario 1) the observed e2e delay is around 300ms on average. It confirms as well analytical results and enables to guaranty a delay compatible with time constraints of interactive flows.

B. Evaluations of the dynamic anticipation factor

In this section, the relevance of the dynamic α anticipation factor is estimated. The anticipation factor is applied on two multimedia applications during simulations:

- A video conferencing flow using H.264 codec [13], with an average bit-rate 700Kbit/s and a peak bit-rate 2Mbit/s. This video codec has the particularity to produce high bit-rate variations;
- An audio conferencing flow using GSM codec [14], with a average bit-rate 13Kbit/s; the bit-rate is evolves around the average bit-rate with weak variations;

The α anticipation factor is computed for each of these types of multimedia applications under the period 500ms. Their evolution is observed during time (c.f. Fig 7).

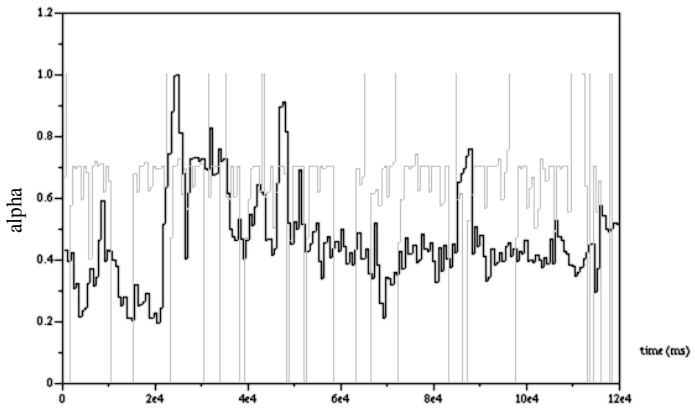


Fig 7. Evolution of α parameter for H.264 codec and GSM codec

Applied to the H.264 video stream (black curve), we notice in the Fig. 7 that the average anticipation α factor calibrates really low (0.35). This type of multimedia application requires some reactivity in its resource requests. Therefore the factor anticipates more resources in order to assure low delays.

In the case of the GSM audio stream (grey curve), we notice on the Fig. 7 that the average α factor is high (0.75) with some peaks when the deletion of silence ends.

VI. CONCLUSION AND FUTURE WORK

This paper has proposed and evaluated enhancements in resources allocation for interactive applications over satellite. It uses a highly efficient (thus less costly) assignment type of dynamic access scheme. The solution concerns (I) the anticipation for provisioning on-demand resources on the communication opening (transitional state) where accumulation of delay introduced by this on-demand assignment type badly degrades applicative quality ; (II) the improvement of the request resource algorithm (permanent state) in order to turn it more dynamic. It satisfies at best both criteria indicators of performance of the on-demand protocol which are: an optimal use of the resources on the return link

channel, and a significant decrease of the latency in the satellite terminals. Both are based on mechanisms of session signaling such as SIP. The analyses and experiments conducted, proved that the solution improves communication quality from user point of view for interactive applications over satellite systems.

Some work is undergoing concerning the renegotiation of a session during the communication (e.g. codec switching) in order to shorten the delay of this renegotiation in a satellite context, using as well SIP capabilities.

This work has been conducted with satellite systems, but the propositions could be generalized over wireless and heterogeneous networks where the resources are scarce.

REFERENCES

[1] ETSI EN 301 790 v1.3.1: Digital Video Broadcast (DVB); Interaction channel for satellite distribution systems, ETSI Norm, January 2003

[2] ITU-T G 1010 series G : transmission systems and media, digital systems and networks ; end-user multimedia QoS categories

[3] Z. Jiang, Y. Li, V. C. M. Leung, "A Predictive Demand Assignment Multiple Access Protocol for Broadband Satellite Networks Supporting Internet Applications," IEEE International Conference on Communications, 28 April-2 May 2002 Page(s):2973 - 2977 vol.5

[4] J. Aracil, D. Morato, E. Magaña, M. Izal, "IP Traffic Prediction and Equivalent Bandwidth for DAMA TDMA Protocols," IEEE Proceedings on Personal, Indoor and Mobile Radio Communications, 7-10 Sept. 2003 Page(s):2514 - 2518 vol.3

[5] F. Delli Priscoli, A. Pietrabissa, "Load-adaptive Bandwidth-on-Demand Protocol for Satellite Networks," IEEE Conference on Decision and Control, 10-13 Dec. 2002 Page(s):4066 - 4071 vol.4

[6] L. Chisci, R. Fantacci, F. Francioli, T. Pecorella, "Predictive Bandwidth Control for GEO Satellite Networks," IEEE International Conference on Communications, 20-24 June 2004 Page(s):3958 - 3962 Vol.7

[7] L. Chisci, R. Fantacci, F. Francioli, T. Pecorella, "Multi-terminal Dynamic Bandwidth Allocation in GEO Satellite Networks," Vehicular Technology Conference, 17-19 May 2004 Page(s):2797 - 2801 Vol.5

[8] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, "SIP: Session Initiation Protocol," June 2002, IETF RFC 3261

[9] Internet Assigned Numbers Authority, unit of Internet Corporation for Assigned Names and Numbers (ICANN) Available: http://www.iana.org

[10] F. Delli Priscoli, A. Pietrabissa, "Design of a bandwidth-on-demand (BoD) protocol for satellite networks modeled as time-delay systems," Automatica, May 2004 Pages: 729-741 Vol 40 Issue 5

[11] Satellite-based communications systems within IPv6, IST FP6 project, 2005-2007 Available: http://www.ist-satsix.org

[12] A. Pietrabissa, T. Inzerilli, O. Alphand, P. Berthou, T. Gayraud, "Validation of a QoS architecture for DVB-RCS satellite networks via the SATIP6 demonstration platform," Computer Networks, Nov 2005, Vol. 49/6, pp 797-815

[13] ITU-T H series H.264 Advanced video coding for generic audiovisual services, September 2005

[14] Global System for Mobile communications codec Available: http://www.voip-info.org/wiki-GSM+Codec