CONTROL-BASED RESOURCE MANAGEMENT PROCEDURES FOR SATELLITE NETWORKS

Francesco Delli Priscoli, Antonio Pietrabissa
Università di Roma "La Sapienza" – Dipartimento di Informatica e Sistemistica (DIS)
via Buonarroti 18, 00184 Roma, Italy; Fax: +39 44585367; e-mail: {dellipriscoli, pietrabissa}@dis.uniroma1.it

Keywords: resource management, connection admission control, satellite networks, Kalman filtering

Abstract*

This paper describes the resource management of a DVB-RCS geostationary satellite network. The functional modules of the access layer aim at efficiently exploiting the link resources while assuring the contracted Quality of Service (QoS) to the traffic entering the satellite network. The main novelty is the integration between the Connection Admission Control and the Congestion Control procedures. Both them exploit the estimation of the traffic load, performed by a Kalman filter. The proposed solution has been analysed via computer simulations, which confirmed their effectiveness.

1 Introduction

In this paper, some *resource management* procedures of multimedia satellite networks are proposed. The assumed protocol stack is compliant to the DVB-RCS (Return Channel via Satellite-Digital Video Broadcasting) standard ([1]). IP (Internet Protocol) traffic entering the UESs is mapped onto connections, which are transported through the following DVB Class of Services ([1]): *Real-Time* (RT) (QoS demanding), *Jitter Tolerant* (JT) (QoS demanding) and *Best effort* (BE) (non QoS demanding).

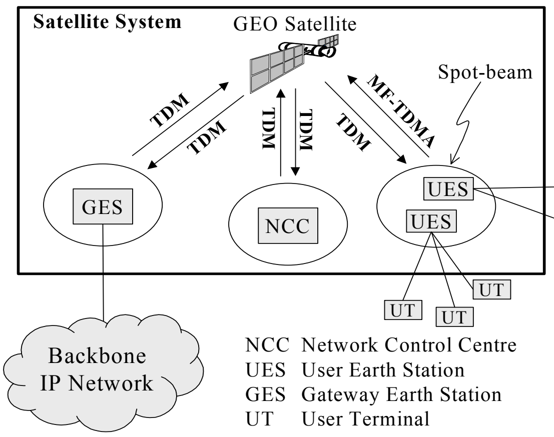

Fig.1 : Reference Satellite System scenario

The considered network, shown in Figure 1, consists of:

- a multi-beam GEO satellite with on-board packet-switch;
- a Network Control Centre (NCC) performing several key control tasks relevant to resource management;
- hundreds of User Earth Stations (UESs), each one providing the access to a few User Terminals (UT);
- a limited number of Gateway Earth Stations (GESs), providing the access to backbone IP networks.

The uplink capacity consists of a set of carriers (*uplink carriers*). GES uplink access is TDM (Time Division Multiplexing), while UES uplink access is MF-TDMA (Multi Frequency / Time Division Multiple Access). Each UES carrier is organized into *frames* having a constant length and organized in an integer number of *time-slots*; each *time-slot* is used to transmit a single Packet. The NCC assigns the *time-slots* to the UESs. The downlink capacity consists of a set of TDM carriers (*downlink carriers*) associated to thye downlink spot-beams. For each downlink carrier, 3 on-board downlink buffers are placed at the output ports of the on-board packet-switch with the following priorities: RT > JT > BE.

The resource management procedures aim at regulating the network connections, which can be classified in *QoS-demanding connections,* characterized by QoS requirements, and *non QoS-demanding connections,* also referred to as *Best Effort (BE) connections*, with no QoS requirement. The main QoS requirements are:

(i) *QoS Bandwidth requirements* specifying the minimum bandwidth which has to be guaranteed;
(ii) *QoS Delay requirements* specifying the maximum tolerated Packet Transfer Delay (CTD);
(iii) *QoS Loss requirements* specifying the maximum tolerated Cell Loss Ratio (CLR).

The resource management procedures are the following:

i. Connection Admission Control (CAC)
The CAC procedure runs at the NCC whenever a set-up attempt relevant to a new QoS-demanding connection occurs. The CAC procedure is in charge of deciding whether to accept or to reject the new connection.

ii. Downlink Congestion Control
The *Downlink Congestion Control* procedure, aimed at regulating the transmission of the BE traffic, has to transmit as much as possible BE traffic to maximize the exploitation of the satellite capacity left available from QoS-demanding traffic while avoiding, as far as possible, BE traffic losses due to the overflow of the downlink buffers.

iii. Uplink Bandwidth on Demand (BoD)
At connection set-up, a certain amount of uplink bandwidth (i.e. a certain number of time-slots) is assigned for the connection lifetime (*Fixed Channels*). The remaining amount of bandwidth (*BoD Channels*) is dynamically assigned via the *BoD procedure* [3].

iv. Terminal Scheduling and Buffer Management procedures
These procedures are performed within the GESs/UESs, and decide which packets to transmit among the buffered ones, basing on the connection QoS Delay requirements,

---

* Work partially founded by the European IST Programme GEOCAST (MultiCAST over GEOstationarty satellites), IST-11754.

and which packets to discard, respectively [4].

v. Satellite System Congestion Control procedure

This procedure aims at controlling the overall traffic relevant to the in-progress connections trying to enter the satellite system. This procedure aims at admitting as much as traffic as possible into the satellite system while avoiding congestions of the satellite network [5].

This paper proposes an innovative control based CAC procedure, as well as its efficient interworking with the Downlink Congestion Control procedure presented in [2].

Section 2 focuses on two key resource management procedures, namely the Connection Admission Control (CAC) and the Downlink Congestion Control. Section 3 present the simulation results, while, in Section 4, the conclusions are drawn.

2 Integrated Resource Management (IRM)

The first basic issue of this section will be the description of an original CAC procedure. CAC algorithms can be grouped in 2 categories: parameter-based admission control (PBAC) and measurement-based admission control (MBAC) ([12]):

- PBAC algorithms ([7], [10]) exhibit a reduced computational cost, but need accurate models of the generated traffic. Generally, the available parameters are used to define a *deterministic bound* on the QoS requirements, which are fulfilled even in the worst-case source behaviour. The drawback is that the capacity is not efficiently used.
- MBAC algorithms ([6], [8] [9]) allow better bandwidth exploitations but generally require more complex implementations. MBAC define a *probabilistic bound* on the QoS requirements: they grant that the actually used bandwidth does not exceed the available one with a defined probability ε, leading to a certain degree of statistical gain in the bandwidth utilization.

The *effective bandwidth* is defined accordingly: the actually used bandwidth exceeds the *effective bandwidth* with probability ε. Different approaches in the estimation of the *effective bandwidth* have been investigated ([12]), characterized by different assumption on the statistical description of the traffic: binomial distribution ([13]), large-deviation approach ([14], [15]), poisson distribution ([16]), gaussian distribution ([8], [7], [17]).

In this paper, the gaussian approximation has been selected, since it works well for our objectives: (i) then there are many traffic sources, it provides a good estimation of the actual bandwidth requirement ([12], [7]); (ii) when there are few traffic sources, it over-estimates the actual bandwidth requirement ([12], [16]).

Among the proposals based on the gaussian assumption, the MBAC implementation in [8], which uses a Linear Kalman Filter, provides good resource efficiency; however, it requires that the network is capable of measuring the transmission rates of each connection, which, in a satellite framework, would require unacceptable processing capabilities.

The objective of the proposed CAC algorithm is to obtain the good performance results of [8], while proposing a scalable solution, suitable for a large number of connections).

Section 2.1 will be devoted to the description of the proposed algorithm for the Effective Bandwidth estimation, while Section 2.2 will disclose the actual CAC procedure. Section 2.3 describes the proposed interworking between the CAC and the Downlink Congestion Control procedures.

2.1 Downlink Effective Bandwidth Estimation Algorithm

The NCC receives measurements on the actual transmitted traffic from the satellite switch, and estimates the mean value M_n and the variance V_n of the bit rate of the aggregate traffic entering each downlink buffer *n*; hereafter, since all the variables will refer to the generic downlink buffer *n*, for notation simplicity, the index *n* will be neglected.

The estimation process is re-initialised whenever a new connection is admitted, or whenever an in progress connection is released. Let the *j-th step* be the time interval elapsing between the *j*-th and the (*j*+1)-th connection setup/release. In the following of this section, all considerations are referred to the generic *j*-th step.

The NCC periodically updates the mean and variance estimations with period T_{ITER}. Let us define as *k-th iteration* the *k*-th estimation updating. The stochastic process yielding the *measured* mean and variance of the bit rate is modelled as a linear, time variant system, whose states are the *actual* values of the above-mentioned mean and variance. Then, Kalman filtering theory is used for determining the optimal *estimation* of such system states.

Let $Z(k)=[M_m(k) \quad V_m(k)]^T$ denote the measured mean and variance of the bit rate, computed by the NCC on the grounds of the measurements received by the satellite during the *k*-th iteration. Let $X(k)=[M(k)\ V(k)]^T$ denote the *actual* state of the above-mentioned system at the *k*-th iteration. $M(k)$ is assumed to be independent of $V(k)$. Let $X_e(k)=[M_e(k)\ V_e(k)]^T$ represent the *estimated* state of the system at the *k*-th iteration. Let $N_S(k)$ and $N_O(k)$ denote determinations, occurring at the *k*-th iteration, of two independent Gaussian random processes with zero mean and variance equal to one. Let $F(k)$ and $G(k)$, denote two 2 x 2 matrices; the matrices $F(k)\,F(k)^*$ and $G(k)\,G(k)^*$, represent the *model error covariance matrix* and *measurement error covariance matrix,* respectively.

The linear, time variant system model for the process is expressed by the following equations:

$$X(k+1) = X(k) + F(k)\,N_S(k) \qquad (4.1)$$

$$Z(k) = X(k) + G(k)\,N_O(k) \qquad (4.2)$$

The aim of the proposed model is to achieve a reliable estimation $X_e(k)$ of the actual system state $X(k)$.

Let $P(k)$ denote the *estimation error covariance matrix,* i.e.:

$$P(k) = \mathrm{E}\{(X(k) - X_e(k))\,(X(k) - X_e(k))^*\}$$

For given covariance matrices $F(k)\,F(k)^*$ and $G(k)\,G(k)^*$, the solution of the problem of estimating the state of system (4.1), (4.2), which minimizes the estimation error covariance matrix $P(k)$ under the hypothesis that the stochastic process relevant to the bit rate of the Packets entering into the considered downlink buffer is Gaussian (with mean $M(k)$ and variance $V(k)$), is given by the *Linear Kalman Filter* (LKF) for discrete time systems ([11]):

$$X_e(k) = X_e(k\text{-}1) + K(k)\,(Z(k) - X_e(k\text{-}1)) \qquad (4.3)$$

$$K(k) = P(k)\,(P(k) + G(k)G(k)^*)^{-1} \qquad (4.4)$$

$$P(k) = P(k\text{-}1) + F(k)\cdot F(k)^* \quad (4.5)$$

$$P(k) = (I - K(k))\, P(\mathrm{k}) \quad (4.6)$$

Once the covariance matrices $F(k)F(k)^*$ and $G(k)G(k)^*$ are known, equations (4.3)-(4.6) allow the determination of $X_e(k)$. Fig. 2 shows the system model, the measurement model and the Kalman filter.

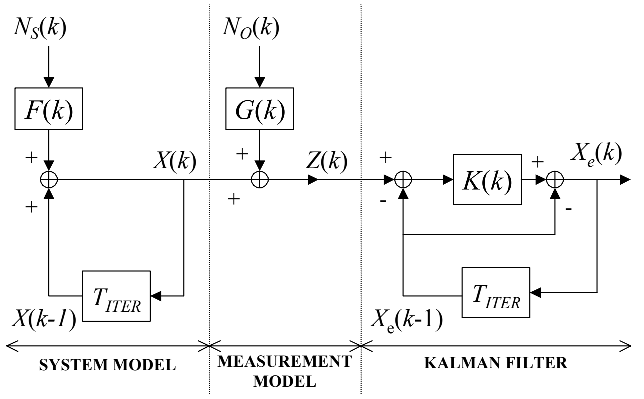

Fig. 2: System model, measurement model and Kalman filter

We underline that the *optimality* of the above estimator is reached under the above-mentioned Gaussian assumption. Since the traffic flow entering a certain downlink buffer consists, in general, of Packets relevant to a plurality of connections having different characteristics, the application of the Central Limit Theorem is partially justified. As demonstrated by extensive simulations performed in various traffic conditions, the Gaussian assumption is not perfectly met, so that the estimation $X_e(k)$ deduced by means of the above-mentioned estimations is just a sub-optimal estimation of the actual state X.

A basic issue for achieving a tight estimation of the actual state $X(k)$ is the determination of appropriate covariance matrices $F(k)\, F(k)^*$ and $G(k)\, G(k)^*$.

The determination of $G(k)\, G(k)^*$ relies on the measurements $Z(k)$. Every T_{MONIT}, the samples of the Packet bit rate are sent from the satellite to the NCC. Let d_i refer the generic sample and let us assume that $T_{ITER} = N\cdot T_{MONIT}$ where N is a positive integer number. Then, we propose to compute the parameters $Z(k)$ and $G(k)\, G(k)^*$ as follows:

$$Z(k)=[M_m(k)\ V_m(k)]=\left[\frac{1}{N}\sum_{i=1}^{N} d_i \quad \frac{1}{N}\sum_{i}^{N}(d_i - M_m(k))^2\right] \quad (4.7)$$

$$G(k)\, G(k)^* = \begin{bmatrix} \frac{(V_m(k))^2}{N} & 0 \\ 0 & \frac{\sum_i^N (d_i - M_m(k))^4}{N^2} - \frac{(V_m(k))^2}{N} \end{bmatrix} \quad (4.8)$$

where the two non-null elements (1,1) and (2,2) in the matrix of (4.8) are obtained by computing the two variances $E\{(M(k) - M_m(k))\ (M(k) - M_m(k))^*\}$ and $E\{(V(k) - V_m(k))\ (V(k) - V_m(k))^*\}$, respectively.

As for the model error covariance matrix FF^*, we have assumed that the elements (1,1) and (2,2) of FF^* are linearly proportional to the sum, P_{sum}, of the peak bit rates of the connections constituting the estimated aggregate, i.e.:

$$F(k)F(k)^* = \begin{bmatrix} \alpha_{11}\cdot P_{sum} & 0 \\ 0 & \alpha_{22}\cdot P_{sum} \end{bmatrix} \quad (4.9)$$

where α_{11} and α_{22} are two constants which has to be properly determined. Note that the peak bit rate of a certain connection can be straightforwardly deduced from the Peak Packet Rate, namely a traffic descriptor declared at connection set-up.

By so doing, each connection set-up (release) brings a positive (negative) contribution to the elements of the model error covariance matrix; therefore, such a matrix is constant during the time intervals between a connection set-up/release and the next one. This empiric procedure has been validated by extensive simulation runs (see Section 3).

At the k-th iteration, once $X_e(k) = [M_e(k)\ V_e(k)]^T$ has been computed, the so-called *Downlink Effective Bandwidth*, denoted as $B_e(k)$, of the aggregate traffic relevant to the considered downlink buffer, can be calculated. The computation of $B_e(k)$ is based on the parameter $B'_e(k)$ such that the probability that the actually used downlink bandwidth B is greater than $B'_e(k)$ is not greater than a given value ε; under the Gaussian assumption, this probability is equal to:

$$\Pr\{B > B'_e(k)\} = \frac{1}{2} erfc\left(\frac{B'_e(k) - M_e(k)}{\sqrt{2\cdot V_e(k)}}\right) \le \varepsilon \quad (4.10)$$

Then:

$$B_e(k) = \min\{B'_e(k), P_{sum}\} \quad (4.11)$$

where (4.11) takes into account that the gaussain approximation can not catch the fact that the bit rate of the traffic entering the considered downlink buffer is anyhow included in the range $[0, P_{sum}]$.

The value of ε has been selected equal to the maximum tolerated Bit Loss Ratio relevant to the Service Class the downlink buffer in question refers to. Note that such a Bit Loss Ratio can be directly deduced from the maximum Packet Loss Ratio which is a QoS parameter declared at connection set-up. This choice assures that if, during the k-th iteration, the actually used downlink band B is not greater than the Downlink Effective Bandwidth $B_e(k)$, then, during such iteration, the satellite system is experiencing a tolerable packet loss ratio (i.e. lower then the maximum tolerated one).

Whenever the transition from the j-th to the (j+1)-th step occurs, the estimation process has to be re-initialized. Let us indicate with a superscript the step of the estimation process and let us denote as $X_e^{j-last} = [M_e^{j-last}\ V_e^{j-last}]$ the last estimate of the j-th step. Then, at the (j+1)-th step, the new estimation process is re-initialised as follows:

$$K^{j+1}(0) = 0 \quad (4.12a)$$

$$P^{j+1}(0) = 0 \quad (4.12b)$$

$$X_e^{j+1}(0) = \begin{cases} X_e^{j-last} & \text{if connection release} \\ X_e^{j-last} + [PR_{new}\ \ 0]^T & \text{if connection admission} \end{cases} \quad (4.12c)$$

where PR_{new} denote the peak bit rate of the new admitted connection. Initial conditions (4.12) apply even for the first step (j=0) occurring when the first connection is admitted, provided that we assume X_e^0 = [0 0].

Moreover, taking into account that (4.10) entails an over-estimation of the actually used bandwidth, in both cases considered by equation (4.12c) the proposed approach over-estimates the mean of the bit rate of the Packets feeding the considered downlink buffer: in the admission case, it assumes that the new connection always transmits at its peak bit rate; in the release case, it assumes that no mean decrease occurs.. In conclusion, (4.12c) represents a conservative approach in the acceptance of a new connection.

As outlined at the beginning of this section, the proposed approach derives from [8], but introduces the following fundamental innovations:

- The iterations of the algorithm described in [8] are not periodic, as it happens in the proposed solution (the period being T_{ITER}), but they are driven by the admittance and the release of the connections, i.e. a new iteration starts whenever a connection is either admitted/released. [8] considers a single estimation process whose iterations coincide with the steps of the proposed procedure. Conversely, in the proposed solution the estimation process restarts at every connection set-up/release. The advantage of a periodic estimation is that we have a much more frequent update of the estimates (a single step can include many iterations) yielding tighter estimations of the actual state. The availability of tighter estimations during the various steps is very important for the Downlink Congestion Control as detailed in the next Section.
- In [8], the model error covariance matrix FF^* is computed by using the difference between the measured and the declared mean and variance of the last admitted connection bit rate. Such an approach would not be possible in a satellite system, where the NCC cannot avail of the measurements relevant to the single connections, since this would entail an unacceptable signalling overhead for transmitting such measurements from the satellite to the NCC. Thus, we have empirically selected the model error covariance matrix (4.9), which only relies of the connection Peak Packet Rates declared at connection set-up. By so doing, we have not experienced any sensible impairment in the tightness of the estimation with respect to the results in [8].
- Even more, we have the further advantage of making such a matrix independent of the iteration (i.e. independent of k). Thanks to this last issue, we can remove the hypothesis of *slow varying traffic* which was present in [8].

2.2 Connection Admission Control (CAC)

Whenever a new QoS-demanding connection set-up attempt occurs, the NCC checks whether or not the uplink and downlink capacities relevant to involved spot-beams are sufficient to support the connection with the requested QoS and without infringing the QoS requirements of the already in-progress connections. The connection is accepted if and only if the ckecks on both the uplink and the downlink capacities yield a positive result.

Assume that, at a time t_{new}, a new QoS-demanding connection set-up attempt occurs, which is relevant to the downlink buffer associated to then downlink carrier h and to the Class of Service q; hereinafter, proper subscripts indicate the downlink carrier and the Class of Service associated to the various parameters. Then, the NCC computes, according to (4.12c), the estimated state $X_{e,h,q}^{j+1}(0)$ relevant to such a downlink buffer. Afterwards, the NCC determines, according to (4.10) and (4.11), the corresponding Downlink Effective Bandwidth $B_{e,h,q}^{j+1}(0)$, namely; note that in (4.10) ε is set equal to CLR_q, namely the Packet Loss Ratio characterizing the Class of Service q.

Let Q denote the total number of QoS-demanding Class of Services (in the proposed satellite system, we have Q=2, since we have two QoS-demanding Class of Services, namely RT and JT). Let $B_{e,h,i}(t_{new})$ denote the Downlink Effective Bandwidth, at time t_{new}, relevant to the downlink carrier h and to a generic Class of Service i (i=1,...,Q); as explained above, for the Class of Service q involved in the connection set-up attempt we have $B_{e,h,q}(t_{new}) = B_{e,h,q}^{j+1}(0)$. Let $C_{down,h}$ denote the satellite capacity of the downlink carrier h; obviously, the value assumed by this parameter depends on the considered satellite system and is deduced according to link budget considerations.

So, as for the downlink, the new connection set-up is accepted if and only if:

$$\sum_{i=1}^{Q} B_{e,h,i}(t_{new}) \leq C_{down,h}. \qquad (4.13)$$

The rationale behind (4.13) is that the sum of the Downlink Effective Bandwidths has not to exceed the downlink carrier capacity: the left-hand side of (4.13) can be thought as the overall capacity on the downlink carrier h which the CAC procedure estimates to be necessary for supporting both the new connection and the other QoS-demanding connections relevant to the downlink carrier h.

Note that, at time t_{new}, it is only necessary to compute $B_{e,h,q}(t_{new})$, since for the other Downlink Effective Bandwidths $B_{e,h,i}(t_{new})$ (i=1,...,Q, i≠q) the values computed at the last iteration previous to t_{new} can be adopted.

2.3 Integration of CAC and Downlink Congestion Control

The proposed Downlink Congestion Control procedure exploits the Downlink Effective Bandwidth computation already used for the CAC procedure.

Let $C_{left,h}(t)$ denote the capacity which, at a time t, is left available, on the downlink carrier h, for BE traffic. Then, at a time t, the proposed Downlink Congestion Control is regulated by the following relation:

$$C_{left,h}(t) = C_{down,h} - \sum_{i=1}^{Q} B_{e,h,i}(t) \qquad (4.14)$$

where $B_{e,h,i}(t)$ denote the Downlink Effective Bandwidth, at time t, relevant to the downlink buffer associated to the downlink carrier h and the QoS-demanding Class of Service i.

By comparing (4.14) with (4.13), it should be clear that, at any time $t = t_{new}$, at which a new connection set-up is accepted, the capacity $C_{left,h}(t)$ coincides with the one which the CAC procedure does not engage for supporting QoS-demanding connections. Nevertheless, (4.14) is not updated just at connection set-ups, but whenever a new iteration occurs for any of the downlink buffers associated to a QoS-demanding Class of Service. By so doing, we have a tight tracking of the capacity actually left available for non QoS-demanding traffic, thus allowing an efficient exploitation of such capacity; in this respect, note that most of downlink Congestion Control procedures proposed so far, updates the available capacity estimate only at the connection set-ups.

For an efficient and flexible Downlink Congestion Control, it is fundamental the ability of $C_{left,h}(t)$ to tightly track the bandwidth actually left available by the QoS-demanding connections, i.e. the ability of the estimated Downlink Effective Bandwidth to tightly track the bandwidth actually used by the QoS-demanding connections. As a matter of fact, since, at time t, the capacity $C_{left,h}(t)$ is assigned to the non QoS-demanding connections, a tight tracking limits, on the one hand, over-assignment to non QoS-demanding connections which could lead to on-board buffer overflows (with consequent Packet losses) and, on the other hand, under-assignment to such connections which could cause bad capacity exploitation. As also shown by simulations (see Section 5) the proposed procedures exhibit such a tight tracking ability.

3 Simulation Results

CAC algorithm described in Section 4.2 has been simulated using OPNET, a discrete-event software simulator specific for telecommunication networks.

The scenario of the simulated satellite network includes 8 uplink spot-beams with 8 terminals in each of them. These terminals can be involved in the following QoS-demanding connection kinds: (i) Audio connections, characterized by constant IP packet size (equal to 29 kB) and constant packet interarrival time (equal to 20 ms); (ii) Video connections, modelled as a 3-state Markov process: each state represents one kind of frame (Intra, Bidirectional and Predictive) of the MPEG coding and has a different average transmission.

For both connection kinds, connection duration is an exponentially distributed random variable with mean equal to 10 s; for each terminal the time interval elapsing from a connection termination (or from a connection set-up rejection) and the next connection set-up attempt is an exponentially distributed random variable with mean equal to 30 s. A connection set-up attempt refers to an audio or to a video connection with probability equal to 0.5 and 0.5, respectively.

Both connection kinds are assumed to belong to the same Class of Service characterized by a maximum tolerated Bit Loss Ratio equal to 10^{-5}. Simulations are set with all the connections loading a single downlink carrier, say the downlink carrier h, whose capacity is equal to 15 Mbps (i.e. $C_{down,h}$ = 15 Mbps) . The monitoring interval T_{MONIT} and the iteration interval T_{ITER} have been set equal to 0.05 s and 0.5 s, respectively.

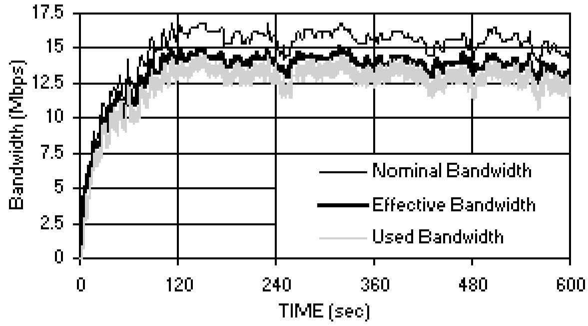

Fig. 3: Used, Nominal and Effective Bandwidth in the solution proposed in this paper

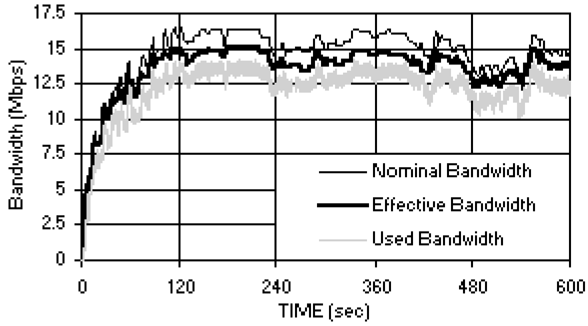

Fig. 4: Used, Nominal and Effective Bandwidth in the solution proposed in [8]

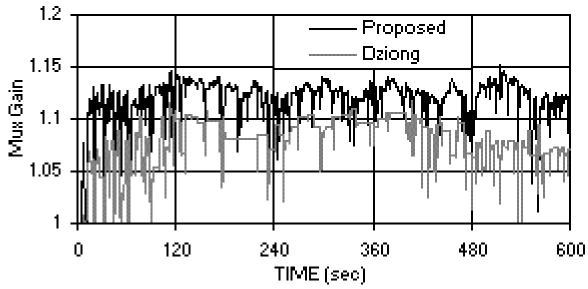

Fig. 5: Multiplexing (Mux) Gain for IRM and Dziong

Fig. 3 shows, as time evolves (at time zero no connection is set-up), (i) the bandwidth *actually used* by the in progress connections, (ii) the sum of the peak bit rates of the in progress connections, hereinafter referred to as *nominal bandwidth* and (iii) the *estimated Downlink Effective Bandwidth* computed according to the approach proposed in this paper, hereinafter referred to as *IRM* (Integrated Resource Management) solution. Fig. 4 shows the same parameters for the approach proposed by Dziong at alii in [8], hereinafter referred to as *Dziong* solution. For the simulations presented in the Figs. 4 and 5, the connection traffic dynamics and the inter-arrival/termination dynamics are identical.

In both solutions, the nominal rate can exceed the downlink capacity of the considered downlink carrier (i.e. 15 Mbps), thanks to the fact that connection acceptance/rejection is based on the Downlink Effective Bandwidth rather than on the peak bit rate.

From these figures, it is evident that the IRM solution provides a tighter tracking of the used bandwidth than the Dziong solution. Such a tighter tracking also occurs at connection set-up times, thus allowing the IRM solution to be less conservative in the connection set-up acceptance (i.e. to avail of a greater connection admission probability) with respect to Dziong solution. Such greater connection acceptance probability entails that the IRM solution assures an average bandwidth exploitation for QoS-demanding traffic of about 3.3 % greater than the Dziong one.

Furthermore, the above-mentioned tighter tracking also allows the assignment of the same amount of traffic to BE connections. In the considered simulation, the average capacity left available for BE connections is equal to 1.7 Mbps in both the IRM and Dziong solutions, in spite of the fact that, as stressed above, in the IRM solution more QoS-demanding traffic has been accepted.

Fig. 5 shows, as time evolves, the multiplexing gain, i.e., the ratio between the nominal and the Effective Bandwidth, for the two solutions; the IRM and the Dziong approaches guarantee an average gain of 11.3% and 7,6%, respectively.

Finally, the Bit Loss Ratio perceived by the QoS-demanding connections has been evaluated by taking into account that Packet loss is caused by on-board traffic overflows. The Bit Loss Ratio has been computed as the integral, extended at the time intervals occurring during the simulation time in which the actually used bandwidth is greater than the Downlink Effective Bandwidth, of the ratio between the actually used bandwidth minus the Downlink Effective Bandwidth and the actually used bandwidth. The Bit Loss Ratio achieved in the simulations so far presented is equal to 10^{-6} , i.e. it is smaller than the maximum tolerated one (10^{-5}). In this respect, it should be noted that the Bit Loss Ratio is kept lower than the maximum tolerated one mainly because the Downlink Effective Bandwidth is computed according to (4.10) with ε equal to the maximum tolerated Bit Loss Ratio and because, due to (4.12c), the Downlink Effective Bandwidth over-estimates the actually used one. In other simulation runs, setting the maximum tolerated Bit Loss Ratio to 10^{-3}, has produced a Bit Loss Ratio equal to about 10^{-4}.

From the above discussions, it should be clear that the efficiency of the IRM solution derives from the fact that both the CAC and the Downlink Congestion Control procedures base on a Downlink Effective Bandwidth guaranteeing a tight tracking of the actually used bandwidth and, at the same time, the respect of the QoS Loss requirement.

4 Conclusions

In this paper, an Integrated Resource Management approach (IRM) for geostationary satellite networks has been presented. The key issue of the IRM is the estimation of the effective traffic load: an algorithm based on the Kalman Filter has been proposed which takes into account the declared traffic parameters of the connections and the aggregate traffic measures, without requiring per-connection measures. The main advantages are that, thanks to the cooperation of CAC and Congestion Control modules, a large multiplexing gain is achieved while preserving the QoS perceived by the users; the congestion probability of the best effort traffic is also reduced, due to the following characteristics:

i) The measurement-based CAC approach allows to estimate the capacity effectively used by the high priority traffic.

ii) The *Effective Bandwidth* estimation algorithm is updated even between two consecutive connection set up/releases; the output of the algorithm is continuously used by the Congestion Control to compute the amount of bandwidth which has to be shared among the best effort connection. In this way, the Congestion Control is capable of adapting the bandwidth allocation to the traffic variations.

The IRM algorithms and procedures have been validated by software simulations and compared with an already existing measurement-based CAC protocol.

References

[1] European Broadcasting Union: "Digital Video Broadcasting (DVB): Interaction channel for satellite distribution systems", *ETSI EN 301 790 V1.2.2*, http://www.etsi.org, December 2000.

[2] F. Delli Priscoli, A. Pietrabissa, "Resource management for -based geostationary satellite netwoks with on-board processing", *Computer Networks*, 39(1), pp. 43-60, 2002.

[3] F. Delli Priscoli, A. Pietrabissa, "Control-theoretic Bandwidth-on-Demand Protocol for Satellite Networks", *Proc. of IEEE CCA*, Glasgow, September 18-20, 2002.

[4] C. Santececca, A. Pietrabissa, "Gateway Architecture for DVB-RCS Satellite Networks", in *Proc of Networking 2002*, Pisa, May 20-24, 2002.

[5] A. Isidori, A. Delli Priscoli, "A control-engineering approach to traffic control in wireless networks" in *Proc. Of CDC 2002*, Las Vegas (NE), December 2002.

[6] R. Gibbens, F. Kelly, "Measurement-based connection admission control", *in Proc of 15th Intern. Teletraffic Congress*, (Washington DC), pp. 879-888, June 1997.

[7] R. Guerin, H. Ahmadi, M. Naghshineh, "Equivalent capacity and its application to bandwidyth allocation in high speed networks", *IEEE Journal on Selected Areas in Communications*, vol 9, pp. 968-981, 1991.

[8] Z. Dziong, M. Juda, L. G. Mason, "A framework for bandwidth management in networks - Aggregate equivalent bandwidth estimation approach", *IEEE/ACM transactions on networking*, vol 5(1), pp.134-147, February 1997.

[9] W. Jia, W. Zhao, "Efficient Connection Admission Control Algorithms for Adaptive Qos Real-Time Connections over networks", in *Proc of IEEE ICMCS '99*, vol 10(2), 1999.

[10] G. Kesidis, J. Walrand, C. Chang, "Effective Bandwidths for multiclass Markov fluids and other sources", *IEEE/ACM Transaction on Network*, vol.1, pp424-428, August 1993.

[11] A. De Santis "*Identification Methods*" (in Italian), Roma, Jenuary 1999.

[12] H. J. Chao, "*Quality of Service Control in High-Speed Networks*", J. Wiley & Sons inc., USA, 2002

[13] H. Zhang, E. W. Knightly, "*Providing end-to-end statistical performance guarantee with bounding interval dependent stochastic models"*, in *Proc. of ACM SIGMETRICS*, pp. 211-220, May 1994

[14] F. P. Kelly, "*Effective bandwidths at multi-class queues*", Queuing Systems, vol. 9, pp. 5-16, 1991

[15] A. Elwalid, D. Mitra, R. H. Wentworth, "*A new approach for allocating buffers and bandwidth to heterogeneous regulated traffic in an node*", IEEE J. Select. Areas Commun., vol. 13, no. 6, pp. 1115-1127, Aug. 1995

[16] C. Rasmussen, J. Sorensen, K. S. Kvols, S. B. Jacobsen,"", IEEE J. Select. Areas Commun., vol. 9(3), pp. 351-358, 1991

[17] S. Abe, T. Soumiya, "*A traffic control method for service quality assurance in an network*", IEEE J. Select. Areas Commun., vol. 12, no. 2, pp. 322-331, Feb 1994