Adaptive Bandwidth Allocation Methods in the Satellite Environment

Raffaele Bolla°, Franco Davoli°, Mario Marchese*

°Department of Communications, Computer and Systems Science (DIST)

*CNIT - Italian National Consortium for Telecommunications

University of Genoa, Italy

Via Opera Pia 13, 16145 Genova (Italy)

Ph. +39-010-3532985, Fax. +39-010-3532154

° lelus, franco@ dist.unige.it; * mario.marchese@cnit.it

Abstract - **The paper proposes two alternative bandwidth allocation schemes suited for the Ka-band satellite environment. The aim is to provide a control mechanism to compensate rain fade. The main idea is considering the effect of the attenuation as a reduction of the bandwidth 'seen' by a single land station and using a supplementary portion of bandwidth to guarantee a rain margin. One of the earth stations involved plays the role of master. It monitors the available resources and contains the Centralized Network Control Center, which manages the bandwidth allocation. Both schemes are organized in two hierarchical levels. The upper level, located in the master station, assigns a portion of the overall bandwidth to the earth stations. At each of these, the lower level shares the received portion between guaranteed and non-guaranteed traffic. The allocation mechanisms are aimed at keeping the call blocking probability of the guaranteed traffic below a given threshold and at reducing the packet dropping probability of the non-guaranteed, best-effort traffic. An extensive performance analysis has allowed investigating the behavior of the two strategies and to evidence the differences between them.**

I. INTRODUCTION

Numerous advantages of the satellite environment with respect to terrestrial networks justify the interest to deliver multimedia services via satellite. In particular the less utilized bandwidths as Ka-band (20-30 GHz) have been the object of experimentation since 1991. The first tests were performed over the Italian satellite ITALSAT, which is also taken as a reference in this work. In these last few years many national and international programs and projects in Europe, Japan and USA have concerned satellite networks and applications over the Ka-band. NASA ACTS (summarized in [1] and [2]), and CNIT-ASI [3], which partially supports the present work, deserve a particular attention, among many others.

In order to provide multimedia services over satellite it is important that the used systems manage efficiently the various satellite and bandwidth resources [4]. Differently from cabled and also wireless networks for personal communications, satellite channels vary their characteristics depending on the weather and the effect of fading heavily affects the performance of the whole system [5], in particular for systems operating in the Ka-band [6, 7], which guarantees wider bandwidth and smaller antennas, but is very sensitive to rain fading. The practical effect is on the quality of service offered to the users. Many user applications require a high degree of quality and techniques to provide compensation for rain attenuation are needed. Actually, there are many ways reported in the literature to provide compensation for rain attenuation: using extra transmission power in areas affected by rain and using a portion of the system bandwidth to have a rain margin are two of them. The latter is the method chosen in this work. The basic idea of this work is considering the fading effect as a reduction of the real bandwidth seen by the stations and giving a supplementary portion of bandwidth to the stations affected by fading, if a high level of quality of service is required. The starting point is a bandwidth allocation scheme for cabled networks [8, 9]. Taking care of the fading effect in the allocation process makes the mechanism suited for a satellite environment. A first proposal about it is contained in [10], where the ABASC (Adaptive Bandwidth Allocation in Satellite Channels) philosophy is introduced. This paper proposes a new strategy and analyses its performance.

The satellite network is composed by earth stations connected through a geostationary satellite [11]. An earth station (or the satellite itself, if switching on board is allowed) represents the master, which manages the resources and provides the other stations with a portion of the overall bandwidth; each of the latter shares the assigned portion between the traffic flows of the single station. Two types of traffic are considered: a QoS guaranteed traffic, modeled as synchronous transmission operating at a fixed speed (in kbits/s) and a non-guaranteed best-effort traffic, modeled by a self-similar Pareto distribution [12]. The fading is modeled by assigning a probability of channel degradation to each link, along with a weighting coefficient to 'measure' the degradation itself. The system considered has a star topology, where all the traffic flows are transmitted to a master station.

The paper is structured as follows. Section II contains the description of the general framework: the characteristics of the network topology, of the bandwidth allocation scheme and of the traffic models used. Section III describes the control system: the cost functions and the optimization problem. The results are contained in section IV. Section V reports the conclusions.

II. GENERAL FRAMEWORK

A. Control requirements and scope

The network is composed of the following elements:

- A Master Control Station (MCS), which contains the Centralized Network Control Center (CNCC) and has the role of checking and monitoring the status and the available resources of the overall network. MCS plays the role of master also for the traffic, if required; i.e., it receives all the

0-7803-7097-1/01/$10.00 ©2001 IEEE

inbound traffic and forwards the outbound traffic. Even if the MCS is not necessarily also the master station concerning the traffic, it is simpler both from a technical and a conceptual point of view to join the two roles in one device. To allocate the bandwidth to the other stations the CNCC should know the amount of traffic and the weather forecast, to get an estimation of the rain fading level.

- The remote stations, which manage the portion of the bandwidth assigned by the CNCC. Each station conveys the traffic flows coming both from traffic sources (i.e., PCs, workstations, video cameras) directly connected to the station and from LANs.

The traffic considered may be synchronous or asynchronous. Synchronous flows require a certain level of Quality of Service (QoS) and have to be completely guaranteed (as voice or real-time video); asynchronous traffic has no strict performance requirements and the network does its best to provide a minimum level of quality ("best-effort" traffic). The fading effect is interpreted as a reduction in the bandwidth availability of each single station. Two parameters, defined in the next section, are used to describe the channel degradation. Each parameter is defined for each station. The number N represents the number of bandwidth portions to allocate. In the following, it will correspond with the number of land stations to simplify the interpretation of the results. From the point of view of the physical transmission, an ATM-based frame is assumed, also concerning the on-board switching [13].

B. The bandwidth allocation scheme

The designed control scheme manages the inbound available bandwidth and, for each land station, it is aimed at

- Keeping the call blocking probability of the guaranteed traffic below a given threshold and minimizing the packet discarding probability of the best effort portion.

- Guaranteeing, if possible, a minimum grade of service for both traffics involved even in case of degradation of the satellite channel.

The control architecture is organized in two hierarchical levels along similar lines as in [9]. The upper level, called Centralized Bandwidth Allocator (CBA), periodically allocates the available bandwidth C [bits/s] by assigning a portion $C^{(i)}$ [bits/s] of the total bandwidth to each earth station i. The allocation is performed by minimizing a cost function, defined in detail in section III, which considers the channel status seen by each earth station. The CBA acts with slow timing and is located in the CNCC of the MCS. Dedicated portions of the transmission frame or high priority control channels are available to the MCS, through which it receives the necessary information to perform the control mechanism and dictates resource allocation.

The lower level, called Local Controller (LC), acts with a faster timing than the CBA and is located in each remote earth station i. It shares the bandwidth $C^{(i)}$, allocated to station i, between guaranteed ($C_g^{(i)}$ [bits/s]) and non-guaranteed ($C_{ng}^{(i)}$ [bits/s]) traffic, by minimizing a function (see section III) that takes into account the call blocking probability of the synchronous connections and the dropping rate of the best effort traffic. It performs Call Admission Control (CAC) of the incoming guaranteed calls and measures the statistics necessary for successive allocations. Let $N_{max}^{(i)}$ be the maximum number of guaranteed traffic calls acceptable at station i to provide a certain grade of service.

Time t has been dropped to simplify the notation; actually all the quantities used may be thought as time variant ($C^{(i)}(t), C_g^{(i)}(t), C_{ng}^{(i)}(t), N_{max}^{(i)}(t)$).

The multiplexer model used by the LC is shown in Fig. 1, for station i. $Q^{(i)}$ is the buffer dimension (in ATM cells).

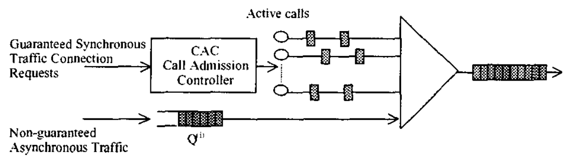

Fig. 1. Generic LC Multiplexer.

C. Traffic models

Synchronous Guaranteed Traffic

Each synchronous connection is defined as Continuous Bit Rate (CBR) flows at B kbits/s, which should be reserved to each connection to guarantee the proper level of quality of service. The traffic of each station is considered independent of the others. Let $\lambda^{(i)}[s^{-1}]$ be the arrival rate of the connection requests and $\frac{1}{\mu^{(i)}}$ [s] the average duration of each connection, for station i. Due to the fact that only one type of synchronous traffic is considered, no index is introduced to distinguish the possible flows. Exponential distributions are used both from the inter-arrival time and the service time, with the time variable assumed continuous. If, as defined in the previous sub-section, $N_{max}^{(i)}$ is the maximum number of calls acceptable at station i, the call blocking probability experienced by it is

$$P_B^{(i)}(N_{max}^{(i)}) = \frac{\dfrac{\left(\lambda^{(i)}/\mu^{(i)}\right)^{N_{max}^{(i)}}}{N_{max}^{(i)}!}}{\displaystyle\sum_{n=0}^{N_{max}^{(i)}} \frac{\left(\lambda^{(i)}/\mu^{(i)}\right)^{n}}{n!}} \quad (1)$$

Asynchronous Non-guaranteed Traffic

We consider a self-similar traffic model, which represents the superposition of an infinite number of on-off sources, with Pareto-distributed 'on' time and exponentially-distributed 'off' time, respectively[12, 14]. The model may represent a flow of IP packets, coming from various sources connected to the station, which are segmented into fixed-size transmission units (e.g., ATM cells) upon entering the buffer. Let us define:

$$T = L/B_{ng} \tag{2}$$

where L is the length of the non-guaranteed traffic transmission unit, and B_{ng} is the peak generation rate of each asynchronous source. Then, the duration, in multiples of T, of the 'on' period τ (assumed with equal characteristics for each station i, so as to avoid further indexing) is such that:

$$\Pr\{\tau = t\} = c \cdot t^{-\alpha-1} \tag{3}$$

where $c = 1 \Big/ \sum_{i=1}^{\infty} i^{-\alpha-1}$ is the normalization constant. The average value is

$$\tau^m = E\{\tau\} = \sum_{t=1}^{\infty} c \cdot t^{-\alpha} \tag{4}$$

The detailed description of the model can be found in [12, 14]. In this context, the quantity of interest is the cell loss probability of the non-guaranteed traffic in the queue of station i (Fig. 1). If $\lambda_{burst}^{(i)}$ is the rate in [bursts/s] of the asynchronous traffic at station i, the birth rate of the process $\lambda_{ng}^{(i)}$, i.e., the average number of sources, which become active, entering an 'on' period, during the interval T is

$$\lambda_{ng}^{(i)} = \lambda_{burst}^{(i)} \cdot T \tag{5}$$

Let us define the variable $X^{(i)} = \frac{C_{ng}^{(i)}}{L} \cdot T = \frac{C_{ng}^{(i)}}{L} \cdot \frac{L}{B_{ng}} = \frac{C_{ng}^{(i)}}{B_{ng}}$, $C_{ng}^{(i)}$ being the capacity variable for the asynchronous traffic of station i, in [bits/s]; if $\lambda_{ng}^{(i)} \cdot \tau^m < X^{(i)}$, it is true that [14]:

$$P_{over}^{(i)} \le \frac{c \cdot \lambda_{ng}^{(i)}}{\alpha \cdot (\alpha - 1) \cdot \left(X^{(i)} - \lambda_{ng}^{(i)} \cdot \tau^m\right)} \cdot (Q^{(i)})^{-\alpha+1} \quad \text{asymptotically} \tag{6}$$

where $P_{over}^{(i)}$ represents the overflow probability of the buffer dedicated to the asynchronous traffic and Q is the length of the buffer itself (Fig. 1). Formula (6) shows a bound in case of very large buffer and describes an asymptotic behavior. Then, we define the following quantity, to be used in the cost functions in the following:

$$P_{loss}^{(i)}(X_v^{(i)}) = \begin{cases} \min\left\{\frac{c \cdot \lambda_{ng}^{(i)}}{\alpha \cdot (\alpha - 1) \cdot \left(X^{(i)} - \lambda_{ng}^{(i)} \cdot \tau^m\right)} \cdot (Q^{(i)})^{-\alpha+1}, 1\right\} & \text{if } X^{(i)} > \lambda_{ng}^{(i)} \cdot \tau^m \\ 1 & \text{otherwise} \end{cases} \tag{7}$$

The bandwidth allocated to the non-guaranteed traffic at station i ($C_{ng}^{(i)}(t)$) at time t is actually a random variable, as it depends (see (8) below) on the number $n^{(i)}(t)$ of guaranteed traffic connections in progress at station i at the same instant.

$$C_{ng}^{(i)}(t) = C^{(i)}(t) - B \cdot n^{(i)}(t) \tag{8}$$

The quantities in (8) have already been defined , except for the time dependence. Thus, also the quantity $X^{(i)}$ should be assumed time variant. Time has been inserted to underline the fact that the non-guaranteed traffic always takes the residual bandwidth not used, at any given instant, by the guaranteed traffic. The time index will be dropped again where not explicitly necessary.

The process $n^{(i)}(t)$ can assume only discrete values from 0 to $N_{max}^{(i)}$; as a consequence $C_{ng}^{(i)}(t)$ will assume only discrete values with a certain probability, depending on the probability of having $n^{(i)}(t)$ connections in progress at time t at station i. If we indicate by $X_j^{(i)}(t)$ the realization of the variable $X^{(i)}(t)$, corresponding to $n^{(i)}(t) = j$, we have:

$$X_j^{(i)}(t) = (C^{(i)} - B \cdot j)/B_{ng}, \quad j = 0,1,\ldots N_{max}^{(i)} \tag{9}$$

and

$$\Pr\left\{X^{(i)}(t) = X_j^{(i)}(t)\right\} = \Pr\left\{n^{(i)}(t) = j\right\} \tag{10}$$

where $\Pr\left\{n^{(i)}(t) = j\right\}$ is given by the stationary distribution of a $M/M/N_{max}^{(i)}/N_{max}^{(i)}$ queueing system.

We assume as an indication of the packet loss rate at station i the quantity defined in (7), averaged over the number of guaranteed connections.

$$\begin{aligned} \overline{P}_{loss}^{(i)}(C_{ng}^{(i)}(t), N_{max}^{(i)}) &= \sum_{j=0}^{N_{max}^{(i)}} P_{loss}^{(i)}(X_j^{(i)}(t)) \cdot \Pr\left\{X^{(i)}(t) = X_j^{(i)}(t)\right\} = \\ &= \sum_{j=0}^{N_{max}^{(i)}} P_{loss}^{(i)}(C^{(i)} - B \cdot j) \cdot \Pr\left\{n^{(i)}(t) = j\right\} \end{aligned} \tag{11}$$

It is important to note the dependence of the loss probability on the threshold $N_{max}^{(i)}$ and on the overall bandwidth $C^{(i)}$ allocated to station i. As already noted, even the latter quantities may be (slowly) time-variant, should the bandwidth be reallocated on-line, on the basis of changes in the channel characteristics; however, the holding time of given values of $C^{(i)}$ (and ,consequently, of $N_{max}^{(i)}$) may well be considered infinite with respect to the dynamics of the process $n^{(i)}(t)$. Moreover, it is worth noting that, in writing (11), we have exploited the fact that the time scales of the guaranteed and non-guaranteed traffic are widely different, in order to use independent stationary distributions for both traffics (see [8] and [15] for more detailed discussions on this point).

III. The Control Scheme

A. The lower level optimization problem

The aim of this sub-section is defining the optimization problem for the LC layer at the generic station (i). LC, as said in section II, will share the bandwidth $C^{(i)}$ allocated to station i between guaranteed and non-guaranteed traffic. The lower level optimization problem will evaluate the threshold $N_{max}^{(i)}$ for a specific value of $C^{(i)}$. The problem has been formalized as follows: given the bandwidth $C^{(i)}$, allocated to station i, evaluate the maximum number of acceptable connections $N_{max}^{(i)}$, such that the call blocking probability (1) be, if possible, lower than $\gamma^{(i)}$. $N_{max}^{(i)}$ may assume values in the range $\left[0,\ldots,\left\lfloor C^{(i)}/B \right\rfloor\right]$. Let $C_{min}^{(i)}$ be the minimum bandwidth needed to obtain a call blocking probability lower than $\gamma^{(i)}$.

$$C_{min}^{(i)} = \arg\min_{X^{(i)}} \left\{ X^{(i)} \in \Re : P_B^{(i)}\left(\left\lfloor X^{(i)}/B \right\rfloor\right) \le \gamma^{(i)} \right\} \quad (12)$$

The value of $C_{min}^{(i)}$ may be computed off-line.

The minimum threshold will be

$$N_{max}^{(i)} = \begin{cases} \left\lfloor C_{min}^{(i)}/B \right\rfloor & \text{if } C_{min}^{(i)} < C^{(i)} \\ \left\lfloor C^{(i)}/B \right\rfloor & \text{if } C_{min}^{(i)} \ge C^{(i)} \end{cases} \quad (13)$$

If the bandwidth allocated by the upper layer is higher than the minimum threshold, the available bandwidth is shared between the guaranteed and non-guaranteed flow, assuring a portion sufficient to guarantee the required QoS (represented by $\gamma^{(i)}$) for synchronous traffic. If the value of $C^{(i)}$ is lower than the minimum bandwidth required, all the bandwidth is given to synchronous traffic.

Both from the computation and the technological point of view, it is simpler to assume $C^{(i)} \in \Re : C^{(i)} = k \cdot mpb,\ \forall k \in N,\ C^{(i)} \le C$, where mpb is the minimum portion of bandwidth that can be allocated, and represents the granularity of the algorithm. If mpb is very small, the algorithm is very flexible but the computational load increases.

B.Channel modeling

The fading effect on the channel 'seen' by station i is modeled as a reduction of the capacity $C^{(i)}$ allocated. The real capacity utilized by station i may be written as $C_{real}^{(i)} = \beta_{level}^{(i)}(t) \cdot C^{(i)}$, where $\beta_{level}^{(i)}(t)$ is a coefficient to weight the channel degradation. The index 'level' identifies the level of the degradation. The time identification has been re-introduced just to focus on the time dependence of the quantity. It will be dropped again in the following. A certain probability $p_{level}^{(i)}$ is associated to each level of degradation.

C. The upper level optimization problem

The upper level optimization problem is aimed at defining the values $C^{(i)}$ for each station i. Four different approaches are proposed to allocate the bandwidth. They differ in the definition of the cost function used in the optimization and for the performance they offer. The first two methods have to be considered only as a reference.

Let $Z^{(i)}$ be a generic variable and N be number of stations that compete for the bandwidth. The function $N_{max}^{(i)}(Z^{(i)})$ is defined in (14) below.

$$N_{max}^{(i)}(Z^{(i)}) = \begin{cases} \left\lfloor C_{min}^{(i)}/B \right\rfloor & \text{if } C_{min}^{(i)} < Z^{(i)} \\ \left\lfloor C^{(i)}/B \right\rfloor & \text{if } C_{min}^{(i)} \ge Z^{(i)} \end{cases} \quad (14)$$

It can be noted that the formula is the same as (13), but, whereas in (13) a specific number has to be computed after receiving the value $C^{(i)}$ from the upper level controller, (14) is just a function and $C^{(i)}$ is the value to be computed.

The different approaches proposed are listed in the following:

RPA (Request Proportional Allocation)

The bandwidth is allocated proportionally to the guaranteed traffic intensity of each station ($\frac{\lambda^{(i)}}{\mu^{(i)}}, i = 1,\ldots,N$).

$$C^{(i)} = \left\lfloor C \cdot \left(\lambda^{(i)}/\mu^{(i)}\right) \Big/ \sum_{j=1}^{N} \lambda^{(j)}/\mu^{(j)} \right\rfloor \quad (15)$$

The residual bandwidth ($C - \sum_{j=1}^{N} C^{(j)}$), if present due to the granularity approximation, is allocated to the station $k : C^{(k)} > C^{(i)}, \forall i \ne k, i = 1,\ldots,N$.

This method considers neither the asynchronous traffic, nor the possible degradation of the channels.

OC-ABASC (Optimal Channel - Adaptive Bandwidth Allocation in Satellite Channels)

This approach adapts the bandwidth allocation to the needs of the non-guaranteed traffic, too. No channel degradation is taken into account. From (11) and (14), let

$$J^{(i)}(Z^{(i)}) = \begin{cases} \overline{P}_{loss}^{(i)}\left[C_{ng}^{(i)}(Z^{(i)}), N_{max}^{(i)}(Z^{(i)})\right] & \text{if } Z^{(i)} \ge C_{min}^{(i)} \\ H & \text{if } Z^{(i)} < C_{min}^{(i)} \end{cases} \quad (16)$$

be the cost function for station i. H is a very large number. N being the total number of stations, the overall cost function is:

$$J_{oc}(Z^{(1)}, Z^{(2)},\ldots,Z^{(N)}) = \sum_{i=1}^{N} J^{(i)}(Z^{(i)}) \quad (17)$$

Even in this case $Z^{(i)} \in \Re : Z^{(i)} = k \cdot mpb,\ \forall k \in N,\ Z^{(i)} \le C$. The aim is to find the particular values of $Z^{(i)} = C^{(i)}$ that minimize the function (17).

The minimization problem is described in (18),

[illegible] (18)

with constraints,

0-7803-7097-1/01/$10.00 ©2001 IEEE

$$\begin{cases} \sum_{i=1}^{N} Z^{(i)} = C \\ Z^{(i)} \geq 0, \forall i \in [1,...,N] \end{cases} \tag{19}$$

The problem deriving from (18) and (19) admits solution if $\sum_{i=1}^{N} C_{min}^{(i)} < C$; otherwise the allocation is performed as follows:

$$\begin{aligned} &\sum_{i=1}^{N} C_{min}^{(i)} = C \Rightarrow C^{(i)} = C_{min}^{(i)} \\ &\sum_{i=1}^{N} C_{min}^{(i)} > C \Rightarrow C^{(i)} = C \cdot C_{min}^{(i)} \Big/ \sum_{i=1}^{N} C_{min}^{(i)} \end{aligned} \tag{20}$$

If the bandwidth is not sufficient to guarantee the QoS requirements, it is allocated proportionally to the minimum capacity required. The optimization problem is common also for the other methods, described in the following.

CSFL-ABASC (Constrained Single Fading Level - Adaptive Bandwidth Allocation in Satellite Channels)

The approach, which has been already introduced in [10] with the simple name of ABASC, being the only algorithm presented there, takes into account the channel status. The extensions allow differentiating the various alternatives. In this case the cost function is

$$J_{csfl}(Z^{(1)}, Z^{(2)}, ..., Z^{(N)}) = \sum_{i=1}^{N} \sum_{level=1}^{L} p_{level}^{(i)} \cdot J^{(i)}(\beta_{level}^{(i)} \cdot Z^{(i)}) \tag{21}$$

The function $J^{(i)}(\cdot)$ is the one defined in (16). L is the number of degradation levels. The minimization problem may be defined as:

$$C^{(i)}, i = 1,...,N : J_{csfl}(C^{(1)},...,C^{(N)}) \leq J_{csfl}(Z^{(1)},...,Z^{(N)}), \forall (Z^{(1)},...,Z^{(N)}) \neq (C^{(1)},...,C^{(N)}) \tag{22}$$

The constraints and the allocation strategy are the same as in (19) and (20).

It is important to note that CSFL-ABASC constrains each single fading level. The function $J^{(i)}(\cdot)$, computed in the variable $\beta_{level}^{(i)} \cdot Z^{(i)}$ (the real bandwidth seen by station i) means that the event $\beta_{level}^{(i)} \cdot Z^{(i)} < C_{min}^{(i)}$ is strongly penalized (the value H is high), for each level of degradation. The effect is that the allocation algorithm tries to avoid the event in any case, even if it has a very low probability to happen. It is sufficient that the generic station i is affected by a strong degradation $\beta_{level}^{(i)}$ even with very low probability ($p_{level}^{(i)}$) to reserve a large amount of bandwidth to station i so that the constraint is respected. The advantage is the performance increase for station i. The drawback is the strong penalization of the other stations, as should be clear from the performance analysis in the following section.

CAP-ABASC (Constrained Average Probability - Adaptive Bandwidth Allocation in Satellite Channels)

The idea on which this method is based is that a strong penalization is needed only if the average call blocking probability is above the fixed threshold. The constraint is explicitly indicated in the new penalty function, which is introduced in (23).

$$F_{cap}^{(i)}(Z^{(i)}) = \begin{cases} 0 & \text{if } \overline{P}_B^{(i)}(N_{max}^{(i)}(Z^{(i)})) \leq \gamma^{(i)} \\ H & \text{if } \overline{P}_B^{(i)}(N_{max}^{(i)}(Z^{(i)})) > \gamma^{(i)} \end{cases} \tag{23}$$

The function $\overline{P}_B^{(i)}(\cdot)$ is defined as the call blocking probability of station i, averaged over the fading levels.

$$\overline{P}_B^{(i)}(N_{max}^{(i)}(Z^{(i)})) = \sum_{level=1}^{L} p_{level}^{(i)} \cdot P_B^{(i)}(N_{max}^{(i)}(Z^{(i)})) \tag{24}$$

$P_B^{(i)}(\cdot)$ has been defined in (1). L is the number of degradation levels.

The cost function of station i is:

$$J_{cap}^{(i)}(Z^{(i)}) = \overline{P}_{loss}^{(i)}\left[C_{ng}^{(i)}(Z^{(i)}), N_{max}^{(i)}(Z^{(i)})\right] + F_{cap}^{(i)}(Z^{(i)}) \tag{25}$$

The global cost is

$$J_{cap}(Z^{(1)}, Z^{(2)}, ..., Z^{(N)}) = \sum_{i=1}^{N} \sum_{level} p_{level}^{(i)} \cdot J_{cap}{}^{(i)}(\beta_{level}^{(i)} \cdot Z^{(i)}) \tag{26}$$

and the minimization problem is expressed as

$$C^{(i)}, i = 1,...,N : J_{cap}(C^{(1)},...,C^{(N)}) \leq J_{cap}(Z^{(1)},...,Z^{(N)}), \forall (Z^{(1)},...,Z^{(N)}) \neq (C^{(1)},...,C^{(N)}) \tag{27}$$

with the constraints and the allocation strategy of the previous cases.

CAP-ABASC constrains the average call blocking probability and allows controlling the performance of the system, for each specific station, without imposing restrictive constraints as in CSFL-ABASC.

IV. RESULTS

The numerical results are divided into two parts. The first part shows the behavior of the CAP-ABASC and CSFL-ABASC bandwidth allocation strategy if the degradation of the channel is varied. It is aimed at highlighting the advantages with respect to the strategies that do not take into account the channel degradation (RPA and OC-ABASC). Four stations are involved. The minimum portion of bandwidth (mpb) that can be allocated has been fixed to 128 kbits/s. C=8 Mbits/s; the following parameters have been used:

$$\gamma^{(i)} = 0.05, \frac{1}{\mu^{(i)}} = 1200\text{s}, B = 128 \text{ kbits/s}, Q^{(i)} = 8000 \text{ cells}, \quad \forall i \in [1, N]$$

$$L^{(i)} = 424 \text{ bits}, \alpha^{(i)} = 1.5, B_{ng}^{(i)} = 324 \text{ kbits/s}.$$

The threshold $\gamma^{(i)} = 0.05$ means a limit of 5% on the call blocking probability of the guaranteed traffic. The channels of three stations are not subject to degradation at all, while the degradation is increasing with the number of the test, concerning the fourth station. The traffic flows imposed are reported in the following: λ=0.006 [conn/s] and λ_{burst}=1200, for station 1 and 2; λ=0.003 [conn/s] and λ_{burst}=600, for station 3 and 4. The channel behavior of the first three stations is fixed and reported in Table 1. Table 2 contains the degradation levels 'seen' by station 4. Eleven levels have been used (L=11): level 1 means $\beta_1 = 0$, level 2 means $\beta_2 = 0.1$, level 3 means $\beta_3 = 0.2$, and so on, up to level 11 that means $\beta_{11} = 1$. A simplified notation will be used in the following. The specific value of β will be explicitly indicated. For instance, the probability of having a degradation level 3 at station 1 will be indicated as $p^{(1)}_{\beta_3=0.2}$.

Station 1, 2, 3 (i=1, 2, 3)	$p^{(1)}_{\beta_l=1} = \begin{cases} 1 & \text{if } l=11 \\ 0 & \forall l \neq 11 \end{cases}$

Table 1. Channel parameters.

Test 1	$p^{(4)}_{\beta_l=1} = \begin{cases} 1 & \text{if } l=11 \\ 0 & \forall l \neq 11 \end{cases}$
Test 2	$p^{(4)}_{\beta_l=0.9} = \begin{cases} 1 & \text{if } l=10 \\ 0 & \forall l \neq 10 \end{cases}$
....	....
Test 10	$p^{(4)}_{\beta_l=0.1} = \begin{cases} 1 & \text{if } l=2 \\ 0 & \forall l \neq 2 \end{cases}$
Test 11	$p^{(4)}_{\beta_l=0} = \begin{cases} 1 & \text{if } l=1 \\ 0 & \forall l \neq 1 \end{cases}$

Table 2. Test definition - Degradation levels, station 4 (i=4).

Concerning this part the results of CSFL-ABASC and of CAP-ABASC strategies are exactly the same. Each test has a specific degradation with probability 1; the differences between CSFL-ABASC and CAP-ABASC cannot result. CSFL-ABASC is the label used in the following. Table 3 contains the bandwidth allocations. It may be compared with the allocation performed by RPA and OC-ABASC, which do not vary with the degradation, being independent of it. The minimum bandwidths that have been computed are: 12 mpb for stations 1 and 2, 7 mpb for stations 3 and 4.

	St 1	St 2	St 3	St 4
RPA	23	21	10	10
OC-ABASC	20	20	12	12
CSFL-ABASC				
Test1	20	20	12	12
Test2	19	20	11	13
Test3	19	20	11	14
Test4,5	19	19	11	15
Test6	18	18	10	18
Test7	17	17	10	20
Test8	15	16	9	24
Test9,10,11	24	25	15	0

Table 3. Bandwidth allocations (measured in mpb).

It is important to note that the algorithm tries to allocate the bandwidth to the station affected by fading. When the degradation is so serious that any allocation would be useless (Tests 9-11), no bandwidth is allocated to the degraded station, which is substantially neglected by the algorithm. The call blocking probability is 3.12% for stations 1, 2 and 4.38% for station 3. All the strategies (RPA, OC-ABASC, CSFL-ABASC) provide the same results. The call blocking probability of station 4 is shown in Fig. 2.

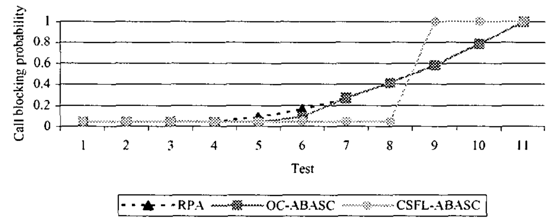

Fig. 2. Call blocking probability, station 4.

CSFL-ABASC, as well as CAP-ABASC, which provides the same performance in these tests, keeps the call blocking probability under the threshold (5%) also for relevant degradations. CSFL-ABASC maintains a value of 4.38% up to the Test8 ($p^{(4)}_{\beta_l=0.3} = \begin{cases} 1 & \text{if } l=4 \\ 0 & \forall l \neq 4 \end{cases}$). The performance improvement is paid by an increase in the packet dropping probability of the non-degraded stations. Fig. 3 shows the packet dropping probability for stations 1 and 2. A similar behavior has been measured, not reported, for station 3. The bandwidth is taken from the stations not affected by fading without interfering with the guaranteed traffic. It is important to observe the increasing packet dropping probability from Test3 up to the peak of Test8 in Fig. 3, where much bandwidth is given to the degraded station. At the same time, the packet dropping probability of station 4 is strongly reduced (Fig. 4).

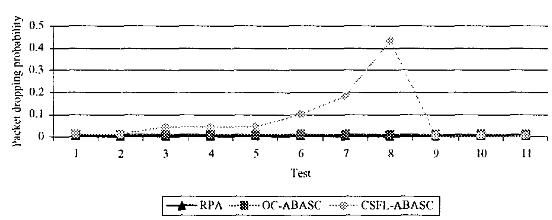

Fig. 3. Packet dropping probability, station 1 and station 2.

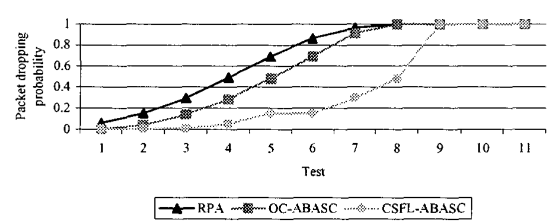

Fig. 4. Packet dropping probability, station 4.

The second part of the results is dedicated to the analysis of the performance by varying the probability of having a degradation level ($p^{(i)}_{level}$), which is the other parameter to model the fading. The results, in this case, show the comparison between CSFL-ABASC and CAP-ABASC. The results of RPA and OC-ABASC were aimed at showing the adaptability of CSFL and CAP to the channel status and to help understand their behavior. The variation of the parameter $p_{level}^{(i)}$ allows focusing on the different performance offered by the two fading-dependent schemes.

The parameters have been kept the same as in the previous case. Two levels of fading have been used (L=2): β_1 and β_2. The former has been set to 1 for all the tests. The latter has been varied but, for the sake of simplicity, only the case $\beta_2 = 0.4$ is shown in the following. It has been selected because it is very meaningful to evidence the difference between the two schemes. Table 4 summarizes the probability of degradation of the first three stations. Table 5 lists the status of the fourth station.

	$p_{\beta_1=1}^{(i)}$	$p_{\beta_2=0.4}^{(i)}$
Station 1, 2, 3 (i=1, 2, 3)	1	0

Table 4. Probability of degradation, station 1, 2, 3.

	$p_{\beta_1=1}^{(4)}$	$p_{\beta_2=0.4}^{(4)}$
Test 1	1	0
Test 2	0.99	0.01
Test 3	0.98	0.02
Test 4	0.95	0.05
Test 5-6-7-8-9-10-11-12	0.9-0.8-0.7-0.6-0.5-0.4-0.2-0	0.1-0.2-0.3-0.4-0.5-0.6-0.8-1

Table 5. Probability of degradation, station 4.

The allocations performed have been reported in Table 6, for the two algorithms. The values corresponding to the two schemes are separated by the symbol "-" and CAP-ABASC is reported in italics to simplify the identification. The minimum bandwidth required $C_{min}^{(i)}$ is also shown.

	St 1	St 2	St 3	St 4
Minimum bandwidth ($C_{min}^{(i)}$)	12	12	7	7
CSFL - *CAP*				
Test1,2,3	20 - *20*	20 - *20*	12 - *12*	12 - *12*
Test4	18 - *20*	18 - *20*	10 - *11*	18 - *13*
Test5	18 - *19*	18 - *19*	10 - *11*	18 - *15*
Test6,7,8,9,10,11	18 - *18*	18 - *18*	10 - *10*	18 - *18*
Test12	17 - *17*	17 - *17*	10 - *10*	20 - *20*

Table 6. Bandwidth allocations (measured in mpb).

As stated in the description of the algorithms (section III), a minimum probability of having degradation is sufficient for CSFL to allocate much bandwidth to the degraded station. The Tests from 2 to 5 are exemplary. In Test2 a probability 0.01 of having a degradation is enough to allocate 18 mpb to station 4. CAP-ABASC, as clear from Table 6, is less influenced by the probability and the bandwidth allocation is smoother than in CSFL. The effect should be clear from the results. The call blocking probability is always below 5% for the first three stations. In more detail, it corresponds to 3.12%, for stations 1 and 2, and to 4.38% for station 3, both for CSFL and CAP. Fig. 5 shows the call blocking probability for station 4. Both methods are below the 5% threshold.

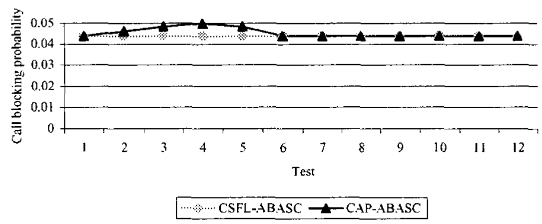

Fig. 5. Call blocking probability, station 4.

The packet dropping probability for stations 1 and 2 is shown in Fig. 6. The same quantity for stations 3 and 4 is reported in Fig. 7 and Fig. 8, respectively. Even if the call blocking probability is a little bit higher in CAP-ABASC (Fig. 5), but always lower than the threshold, the packet dropping probability is always lower in CAP-ABASC (see Fig. 6 and Fig. 7, Test 2 to Test 5). It means that the bandwidth is better allocated in CAP-ABASC without penalizing the guaranteed traffic, which keeps the constraint on the call blocking probability. The packet dropping probability of station 4 is the same for both schemes (Fig. 8). The same behavior may bè observed in the last two graphs, where the average call blocking probability (Fig. 9) and the average packet dropping probability (Fig. 10) are shown for the overall system. The constraint on the guaranteed traffic is always maintained but the performance of the best effort traffic is drastically improved by CAP-ABASC. The effect is amplified for more serious fading (smaller values of β).

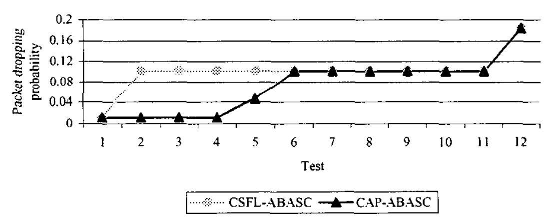

Fig. 6. Packet dropping probability, station 1 and 2.

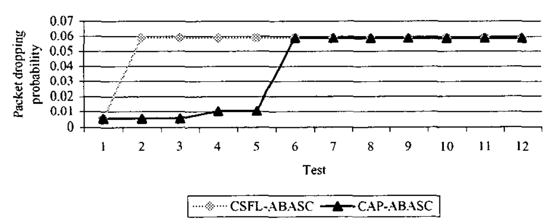

Fig. 7. Packet dropping probability, station 3.

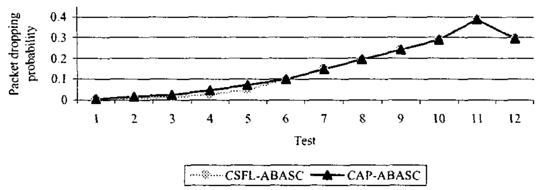

Fig. 8. Packet dropping probability, station 4.

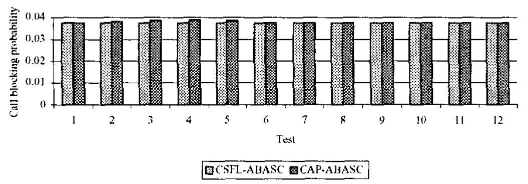

Fig. 9. Average call blocking probability.

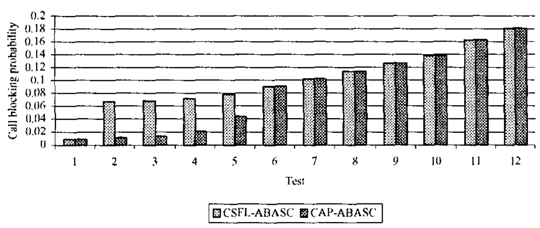

Fig. 10. Average packet dropping probability.

V. CONCLUSIONS

The paper has presented two bandwidth allocation algorithms to support multimedia traffic in a satellite environment (CSFL-ABASC and CAP-ABASC). Two types of traffic have been considered: a guaranteed traffic, which needs precise Quality of Service (QoS) requirements, and a "best effort" traffic. N earth stations, connected through a geostationary satellite, compete for the inbound bandwidth. A station, which has the role of master, manages satellite network resources and allocates the bandwidth to the other stations. The scheme proposed is aimed at keeping a given constraint on the call blocking probability of the guaranteed traffic and at minimizing the packet discarding probability of the best effort portion. A minimum grade of service for both traffics involved even in case of degradation of the satellite channel should be offered. The results reported have shown that the two strategies allow to maintain the required QoS for the guaranteed traffic at all stations and to improve the performance concerning non-guaranteed traffic in the stations where there is channel degradation. CAP-ABASC allocates the bandwidth more precisely and allows a stronger reduction of the packet dropping probability of the best effort traffic without penalizing the guaranteed traffic.

REFERENCES

[1] D.E. Brooks, C. Buffinton, D.R. Beering, A.W. William, D. Ivancic, M. Zernic, D.J. Hoder, "ACTS 118x Final Report, High Speed TCP Interoperability Testing", NASA/TM - 1999-209272, July 1999.

[2] R.T. Gedney, R. Schertler, F. Gargione, *The Advanced Communications Technology Satellite*, SciTech Publishing, Mendham, NJ, 2000.

[3] D. Adami, A. Conti, D. Dardari, M. Marchese, L. S. Ronga, "CNIT-ASI Project 'Integration of Multimedia Services on Heterogeneous Satellite Networks': An Overview", *Proc. 5th Ka-Band Utilization Conf.* , Taormina, Italy, Oct. 1999, pp. 219-226.

[4] V. Birmani, "Resource Allocation for Ka-Band Broadband Satellite Systems", CSHCN MS 99-6, Master of Science Thesis, 1999; also in http://www.isr.umd.edu/TechReports/CSHCN/1999/.

[5] L.J. Ippolito, "Propagation Consideration for Low Margin Ka-Band Systems", *Proc. 3rd Ka-Band Utilization Conf.*, Italy, Sept. 1997, pp. 161-167.

[6] R.T. Gedney, T.A. Coney, "Effective Use of Rain Fade Compensation for Ka-Band Satellites", *Proc. 3rd Ka-Band Utilization Conf.*, Italy, Sept. 1997, pp. 153-160.

[7] R.J. Acosta, "Rain Fade Compensation Alternatives for Ka-Band Communication Satellites", *Proc. 3rd Ka-Band Utilization Conf.*, Italy, Sept. 1997, pp. 145-152.

[8] R. Bolla, F. Davoli, "Control of Multirate Synchronous Streams in Hybrid TDM Access Networks", *IEEE/ACM Trans. Networking*, vol. 5, no. 2, pp. 291-304, April 1997.

[9] R. Bolla, F. Davoli, S. Ricciardi, "A Hierarchical Control Structure for Multimedia Access Networks", *Proc. IEEE Internat. Conf. Commun. (ICC '99)*, Vancouver, Canada, June 1999, pp. 1320-1325.

[10] R. Bolla, F. Davoli, M. Marchese, "A Bandwidth Allocation Strategy for Multimedia Traffic in a Satellite Network", *Proc. IEEE Globecom2000*, San Francisco, CA, Nov. 2000, pp. 1130-1134.

[11] G. Maral, M. Bousquet, *Satellite Communications Systems - Systems, Techniques and Technology*, III Ed., John Wiley & Sons, Chichester, 1998.

[12] B. Tsybakov, N.D. Georganas, "On Self-Similar Traffic in ATM Queues: Definition, Overflow Probability Bound, and Cell Delay Distribution", *IEEE/ACM Trans. Networking*, vol. 5, no. 3, pp. 397-409, 1997.

[13] I. Metzanis, G. Sfikas, R. Tafazolli, B.G. Evans, "Protocol Architectures for Satellite ATM Broadband Networks", *IEEE Commun. Mag.*, vol. 37, no. 3, pp. 46-54, March 1999.

[14] B. Tsybakov, N.D. Georganas, "Overflow Probability in ATM Queue with Self-Similar Input Traffic". *Proc. IEEE Internat. Conf. Commun. (ICC'97)*, Montreal, Canada, 1997.

[15] S. Ghani, M. Schwartz, "A Decomposition Approximation for the Analysis of Voice/Data Integration", *IEEE Trans. Commun.*, vol. 42, pp. 2441-2452, July 1994.

0-7803-7097-1/01/$10.00 ©2001 IEEE