

Resource Allocation in DVB-RCS satellite systems

André-Luc Beylot¹, Riadh Dhaou¹ and Cédric Baudoin²

¹ IRIT/ENSEEIH, 2 rue C. Camichel, BP7122, F-31071 Toulouse Cedex 7, France
Beylot@enseeiht.fr, Riadh.Dhaou@enseeiht.fr

² Alcatel Alenia Space, 26 avenue JF Champollion BP1187, F-31037 Toulouse Cedex 1, France Cedric.Baudoin@alcatelaleniast.com

Abstract. This paper compares several approaches for dynamic allocation in geo-stationary networks based on DVB-RCS system. Each Satellite Terminal (ST) regularly sends requests to the Network Control Center (NCC) which in turn allocates resource to the users. Unfortunately, this delayed request-assignment makes the dynamic bandwidth allocation very difficult. Simple mechanisms such as a fixed allocation or requests based on the current size of the terminals' queue are compared to predictive methods based on control theory techniques which have been previously proposed. A lower bound is also derived by considering that the actual size of the buffer can be instantaneously known. It is shown that if the traffic is not really bursty, a fixed allocation which implies lighter signalling mechanisms leads to good results. In bursty traffic conditions, simple mechanisms for which the requests correspond to the actual size of the buffer may lead to the best performance results. *abstract* environment.

Key words: Resource allocation, Modeling and performance evaluation

1 Introduction

DVB-RCS (Digital Video Broadcasting - Return Channel by Satellite) satcom systems provide a shared uplink between various users of terminals. Dynamic allocation mechanisms, taking into account the needs for each terminal, have been investigated to enable an optimal use of the expensive and limited bandwidth. The capacity requests are calculated according to the generated traffic in the terminal. Several modes of calculation are proposed which are based on the size of the buffer information and on the input and/or output rates with possible use of a calculation window taking into account the satellite delay. The allocation requests are then processed in the Network Control Center (NCC). This one carries out a control on the radio resource according to the requests and to their priority. In this paper, we define several methods of calculation of requests and allocations. The allocation loop is characterized by an important delay, a shared access to the resource, a granularity of allocation. The methods aim at decreasing the average latency of the traffic in the buffers while insuring fairness

between satellite terminals. The suggested loops are modelled and evaluated to compare their performance, their complexity and their sensitivity to the traffic type.

2 Dynamic Resource Allocation

2.1 The system

The DVB-RCS standard considers various classes of service on the MAC level and associated allocation methods. Packets have a fixed size.

- CRA: Constant Rate Assignment. This class is not the subject to allocation requests, the capacity is allocated (in slot/frame) at the connection set-up and remains constant. This operation is used for delay sensitive traffics (voice).
- RBDC: Rate Based Dynamic Capacity. This category is used for the traffics with variable flow which can tolerate the response time of the MAC scheduler. The bandwidth is granted on request to run out the instantaneous traffic rate.
- VBDC: Volume Based Dynamic Capacity. Used for the traffics without any time constraints, VBDC requests are expressed in number of slots and satisfied if there are remaining slots.

The duration of a frame T_f is equal to 30ms. An allocation cycle will include propagation delay and various computing times (terminal and NCC) and is set to $d = 23$ frame durations. Thus, the request for frame n must be calculated and sent in frame $n - 23$. In such a context several allocation methods have been proposed.

2.2 Predictive allocation techniques

Several techniques of control were recently applied to communication networks. They lead to performance gains given adequate traffic conditions. The Smith's Predictor (SP) [7] is a popular and very effective long dead-time compensator for stable processes. The main advantage of the SP method is that the time delay is effectively taken outside the control loop in the transfer function. Classical SPs are used to remove potentially destabilising delays from the feedback loop by employing "loop cancellation" technique. The SP has been successfully used in designing congestion control algorithms in TCP [5] and ATM [6] networks. Authors in [4] proposed a DAMA mechanism based on SP controller, making it possible to adapt the load on-demand for satellite networks. Results are limited to the study of simple parameters and to systems which can be described by first order's models with a long delay (the multiple combinations of parameters and the study of systems with models of a higher nature are still open). The SP algorithm can be described as follows. Let $y_i(n)$ denote the allocation request emitted by terminal i during frame n and $u_i(n)$ its buffer size at the beginning

of the same frame. (K is a gain factor $0 < K < 1$, only integer values for the requests are considered):

$$y_i(n) = \lfloor K(u_i(n) + \sum_{j=n-d}^{n-1} y_i(j)) \rfloor \quad (1)$$

In the present paper, the allocation technique mainly concerns VBDC traffic. Consequently, if the total amount of the requests is lower than the actual number of slots C , all the requests are satisfied. Else, the allocation is proportional to the request.

$$u_i(n) = \begin{cases} y_i(n-d) & \text{if } \sum_k y_k(n-d) \leq C \\ \lfloor \frac{y_i(n-d)}{\sum_k y_k(n-d)} \rfloor + \varepsilon_i(n-d) & \text{otherwise} \end{cases} \quad (2)$$

As results may not be integer values, the remaining slots are randomly and uniformly distributed among active terminals whose request has not been fulfilled. $\varepsilon_i(n-d)$ stands consequently for an additional allocated slot.

A novel approach for dynamic bandwidth allocation (DBA) in satellite networks has been presented in [1–3]. Each ST uses a local adaptive predictor to forecast the future input traffic flow along with a local predictive (receding-horizon) controller to generate a bandwidth request to the NCC. This approach gives good results when the traffic model is well known. The difficulties to adapt the predictor to actual traffic conditions and the complexity of the optimization of the controller have been evaluated [8] but will not be discussed in this paper. Moreover, several papers [13–15] shown that adaptive resource allocation policies may outperform the fixed policies.

2.3 Other allocation techniques

In the present paper, we aim at comparing simple algorithms to the SP method. The first one supposes that the controller knows instantaneously the buffer state of the terminals. Of course, this method cannot be implemented because the necessary delay to indicate packet arrivals is exactly the problem we face. Nevertheless, it leads to optimal performance results to which other techniques may be compared. This method is named “Big Brother” in the present paper. Many algorithms may be chosen to implement fairness principles between STs. When considering infinite buffers, all the work-conserving policies lead to the same value of the mean delay. We did not choose any of them.

The second method which has been considered is a Fixed Allocation. Terminals receive a constant number of slots per frame. This allocation may be done during the set-up period of the terminal. In the present paper, uniform traffic conditions are considered. The number of allocated slots may vary according to the number of active sessions using a simple procedure and with a very light signalling mechanism. For the evaluation, the duration of the allocation period is supposed to be large enough to consider that a terminal is allocated a constant number of slots during a frame period. If it has fewer packets to send, the corresponding

slots will be lost. The principle is that the round-trip delay is so long that it is not efficient to implement solutions at a packet level.

The third method is exactly the opposite of the previous one: a signalling procedure is implemented. The terminals send requests reflecting their current buffer size. This mechanism is named ‘‘Candid Algorithm’’ because no intelligence is required to calculate the size of the requests. The controller collects all those requests, calculates the corresponding numbers of allocated slots and sends back these values.

3 Model and analysis of the system

3.1 Model of the system and traffic assumptions

In the present paper, we do not investigate the actual scheduling of the slots or the exact position of the allocated slots in the frame. Bulk departures are considered at the beginning of the frame during which packets are actually sent. Different traffic models will be considered. For simplicity, we will first assume that packets arrive to the queues according to independent Poisson processes. Interrupted Poisson Processes (IPP) will then be considered in order to show the influence of the burstiness of the input traffic on the performance criteria.

3.2 Models ‘‘Big Brother’’ and ‘‘Candid’’ - Poisson Arrivals

In the case of Poisson Arrivals, the analysis of the two methods can be done through a $M/D^{(H)}/1$ queue with bulk departures where H is the maximal size of the burst. This size is constant and is equal to C in ‘‘Big Brother’’ case and to $\frac{C}{N}$ in ‘‘Candid method’’ (N is the number of terminals). The evolution of such a queue at departure epochs can be described as follows:

$$X_{n+1} = \text{Max}\{X_n - H, 0\} + A_n \quad (3)$$

where X_n is the number of packets in the queue (or in the aggregate queue in the Big Brother case) just before departure and A_n the number of arrivals during frame n . This queue has been extensively studied. Using a z -transform approach, classical solutions consist on finding the roots of the denominator of:

$$X(z) = \frac{A(z)(S(z) - z^H)}{A(z) - z^H} \quad (4)$$

where $X(z)$, $A(z)$, $S(z)$ are the pgf of the stationary queue length, the number of arrivals and the number of departures. The mean number of packets in the queue can be expressed as follows

$$E(X) = \rho + \frac{A''(1) - S''(1)}{2(H - \rho)} \quad (5)$$

ρ is the mean number of arrivals during a frame $\rho = A'(1) = S'(1)$. We propose new bounds of such a term using the following considerations. The first lower

bound consists on observing that the second moment of the number of departures is lower than the second moment of the number of arrivals. It is a lower bound because the number of departures is bounded by the maximal size of the bulk. This bound is reached if the number of departures is exactly equal to the number of arrivals. The second lower bound considers that either 0 or H packets are served. As $S'(1) = \rho$, we get $S''(1) \leq (H - 1)\rho$. An upper bound of the mean queue size is derived by considering that the variance of the number of departures is positive thus, $S''(1) \geq \rho^2 - \rho$. Finally,

$$\text{Max}(\rho + \frac{\rho^2 - (H - 1)\rho}{2(H - \rho)}, \rho) \leq E(X) \leq \rho + \frac{\rho}{2(H - \rho)} \quad (6)$$

In the present paper, we also implemented the very smart solution proposed in [9]. It consists on inverting arrivals and services and considering that Eq. 3 corresponds to the evolution of the response time of a discrete $D/G/1$ queue with, in this case, a service time which is distributed according to a Poisson Process. The mean number of packets in the queue is thus equal to:

$$E(X) = \rho + \sum_{k=1}^{+\infty} \frac{1}{k} \sum_{i=kH}^{+\infty} (i - kH)c_{i,k} \quad (7)$$

where $c_{i,k}$ is the probability that i packets arrives during k frames. The main interest of this method is that these expressions are root-free. To implement this method, we set a maximal value for the infinity. Classical bounds, such as Kingman's formula [10] of $G/G/1$, can also be applied

$$E(X) \leq r + r^2(C_a^2 + C_s^2)2(1 - r)\lambda \quad (8)$$

$r = \frac{\rho}{H}$, λ is the arrival rate of packets, C_a^2 and C_s^2 are respectively the squared coefficient of variation of "arrivals" and "service times" of the $D/G/1$ queue. We can easily find that $C_a^2 = 0$ and that $C_s^2 = \frac{1}{\rho}$. Consequently, $E(X) \leq \rho + \frac{\rho}{2(H - \rho)}$ which is exactly the upper bound that has been derived in Eq. 6.

In order to derive the mean delay, we can subtract from all the previous results the mean number of departure ρ , then apply Little's formula [11] and finally add the mean time between arrival time and the beginning of the following frame. As Poisson arrivals are considered, this parameter is equal to $\frac{1}{2}$ frame duration (a more formal proof can be obtained by deriving the mean number of packets at arrival instant and applying then Little's formula). Thus, the mean delay (expressed in number of frames) is equal to:

$$E(R) = \frac{E(X)}{\rho} - \frac{1}{2} \quad (9)$$

3.3 Models "Big Brother" and "Candid" - Bursty Arrivals

Many models can be considered to take into account the burstiness of the traffic. In the present paper, we consider Interrupted Poisson Process in order to derive

some numerical results. Our aim was mainly to compare the different methods rather than finding exact performance results. As it has been mentioned earlier, if the optimal allocation policy depends to much on the traffic characteristics, an erroneous model of traffic may lead to bad performance results. IPP processes were introduced in order to model ON/OFF traffic with “ON” and “OFF” period duration exponentially distributed (with respective parameter α, β) and Poisson arrivals during the “ON” period (rate γ). The superposition of such processes is not IPP [12]. The performance criteria will thus be derived through the analysis of a $IPP/D^{(H)}/1$ queue for “Candid” method and of a $n - IPP/D^{(H)}/1$ queue in “Big Brother” case.

We will consider the embedded Markov Chain at departure epochs, by considering the number of packets in the queue and the number of active source(s). The number of arrivals during a frame period A_n depends on the number of active sources. For “Candid” mechanism, we thus analyzed the continuous process corresponding to the number of packets generated since the beginning of the period $N(t)$ and the state of the source $E(t)$. The transient analysis of this process have been previously been performed [12] using a z -transform approach $\Phi_i(z, t) = \sum_k Pr[E(t) = i, N(t) = k]z^k$. In the present paper, we prefer a Laplace-Transform approach.

Let $P_j(k, i, t) = Pr[N(t) = k, E(t) = i | E(0) = j]$ and $P_j^*(i, k, s)$ denote its Laplace transform. We easily get the initial values:

$$\begin{cases} P_0^*(0, 1, s) = \frac{\alpha}{\tau} \\ P_0^*(0, 0, s) = \frac{s+\beta+\gamma}{\tau} \end{cases} \begin{cases} P_1^*(0, 1, s) = \frac{\alpha+s}{\tau} \\ P_1^*(0, 0, s) = \frac{\beta}{\tau} \end{cases} \quad (10)$$

where $\tau = s^2 + (\alpha + \beta + \gamma)s + \alpha\gamma$. The following terms can be recursively described as follows: $P_j^*(k, 0, s) = \frac{\beta\gamma}{\tau} P_j^*(k-1, 1, s)$ and $P_j^*(k, 1, s) = \frac{\gamma(s+\alpha)}{\tau} P_j^*(k-1, 1, s)$. Let s_1 and s_2 denote the poles of the denominator τ . After some calculations, those Laplace transforms can be inverted. They are of the following forms:

$$P_j(k, i, t) = \sum_{m=1}^{k+1} \frac{t^{j-1}}{(j-1)!} (u_{k,i,m} \exp(+s_1 t) + v_{k,i,m} \exp(+s_2 t)) \quad (11)$$

Coefficients $u_{k,i,m}$ and $v_{k,i,m}$ are recursively computed [8]. A numerical method has been implemented to derive the values of those functions in $t = T_f$ and we only consider a finite number of terms. Let $R_{j,i,k}^1 = P_j(k, i, T_f)$. The numerical solution leads to the steady state probabilities (number of packets, state of the source) $x_{k,j}$ and thus the mean number of packets just before departure is equal to $E(X) = \sum_k k(x_{k,0} + x_{k,1})$. Using the same argument as in the Poisson arrival case, we can again apply Eq. 9.

In Candid method, we have to study a $n - IPP/D^{(H)}/1$ queue. We nearly applied the same method. The number of arrivals during a frame period duration given the number of active sources are recursively derived by using previous

equations.

$$R_{j,i,k}^{m+1} = \sum_{p=0}^1 \sum_{q=0}^k R_{j,i,q}^m R_{0,p,k-q}^1 \left(1 - \frac{i}{m}\right) + R_{j-1,i,q}^m R_{1,p,k-q}^1 \frac{i}{m} \quad (12)$$

4 Results

4.1 Analytical results - Poisson Arrivals

In the present work we consider a high load $\rho = 0.8$ (for lighter traffic conditions, predictive methods are less useful). The frame duration T_f is equal to 30ms. The maximum aggregate bit rate is set to 2Mbps which nearly leads to a maximal number of slots per frame equal to $C = 140$. In figure 1, we represented the mean access delay as a function of the number of terminals N . The case $N = 1$ corresponds to “Big Brother” mechanism. We compared the different analytical methods (upper bound, lower bound, numerical solution) to simulation results. The numerical method leads to excellent results. They are nearly equal to those obtained using a discrete event simulation whose confidence intervals are extremely low. The upper bound is quite accurate. The load is high, it corresponds to traffic conditions for which Kingman’s formula is known to be accurate. The difference between this upper bound and simulation results is between 3% and 20%. The lower bounds are also accurate; the one considering that the variance of the number of departures is equal to the variance of the number of arrivals leads to good results when the number of terminals is low. In this case, the number of slots allocated to a terminal is high and consequently, most of the packets arrived during frame n will be emitted in frame $n + 1$. The other lower bound is exact when each terminals are allocated one slot per frame and leads to inconsistent results when the number of terminals decreases.

4.2 Comparison of the Allocation methods, Poisson Arrivals

We will compare the fixed allocation method, the SP technique (for various values of K), and the lower bound under Poisson arrivals. The standardized load of the system is 0.8 and we vary the number of terminals. SP leads to worse results than a fixed allocation especially when the number of terminals is large. With regard to the SPs parameter setting, we note that larger values of K lead to better results. The value $K = 1$ was considered, even if it constitutes an extreme value (the series of requests diverge). We can note that for this value of K the difference with the fixed allocation method is lower. We are inclined to say that if the system is heavily loaded, all the resources will be very intensely used and that devoting them to a terminal or another does not change the global result. As Poisson arrivals are considered, the lower bound is nearly equal to 0.5 (a packet has to wait at least for the beginning of the following frame) and does not depend on the number of terminals (we did not plot a curve ; it corresponds to the fixed technique with $N = 1$ terminals). We can finally think that these

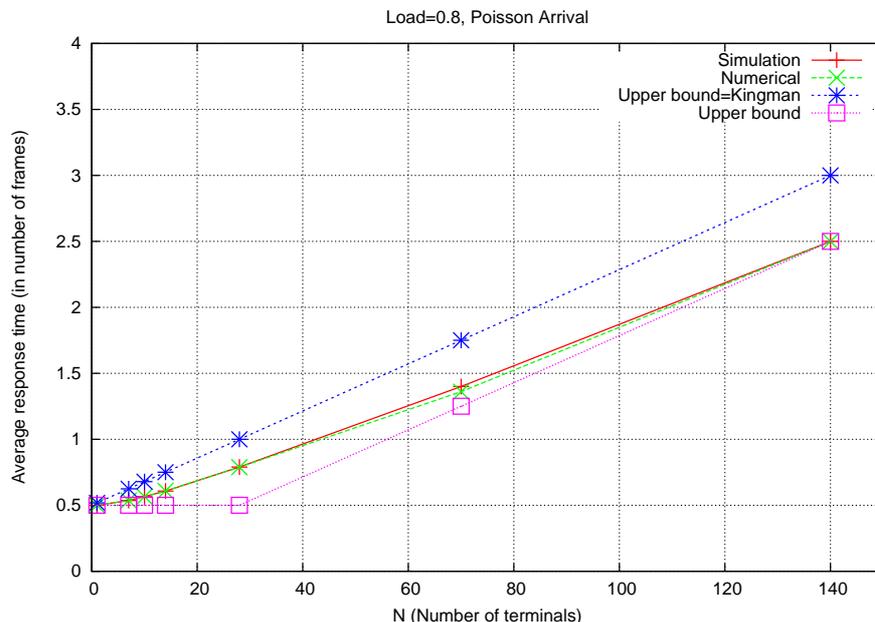


Fig. 1. Analytical results: Mean Access Delay (in number of frames) as a function of the number of terminals, Poisson Arrivals

good results are due to Poisson arrivals which are not sporadic enough compared to real traffic. Poisson sources result in a rather good regularity in arrivals flows, which could correspond to an unfavourable configuration for the SP. Under a sporadic traffic, in particular with larger durations of activity, SP may have time enough to stabilize itself and thus to outperform a fixed allocation.

4.3 Influence of the burstiness, IPP arrivals

Let us now compare the results obtained with fixed allocation, SP, Candid and “Big Brother” methods under bursty traffic conditions. We consider again a normalized load of 0.8. IPP processes are characterized by three parameters: the mean duration of “OFF” periods $\frac{1}{\alpha}$, the mean duration of ON periods $\frac{1}{\beta}$ and the arrival rate during “ON” period γ . Thus, $\rho = \frac{N}{C} \gamma \frac{\alpha}{\alpha + \beta}$ γ is to packets per frame during active periods. We took several values of the duration rate of active periods β and we deduced the duration rate of silence periods α as a function of the number of terminal N . Results are depicted in Fig. 3 and 4. In Fig. 3, the average duration of “ON” periods is short (1 frame). The difference between the allocation mechanisms and the lower limit remains reasonable. The fixed allocation gives better results than the SP; the performance is degraded when we increase the number of terminals. The “candid method” gives the worst results but these results are less sensitive to the number of

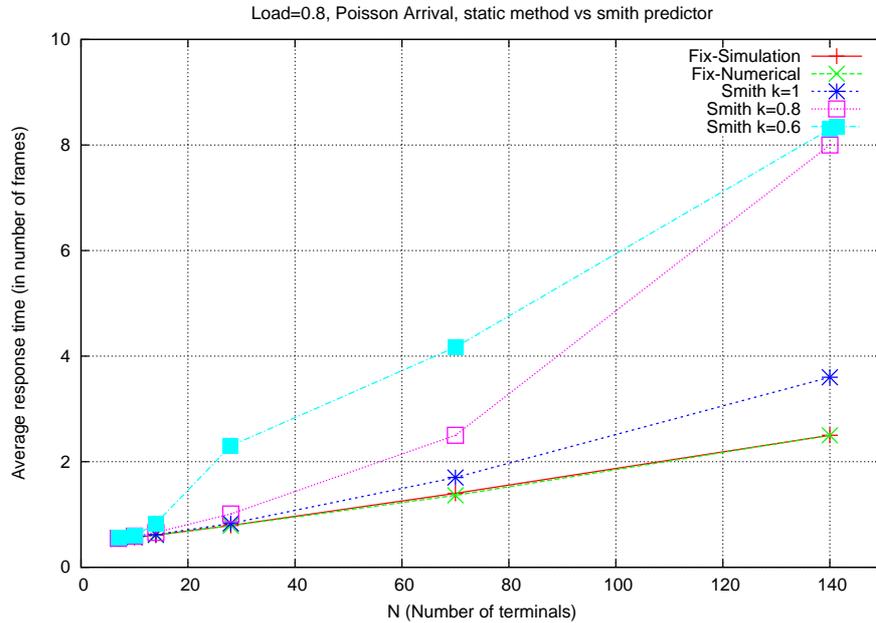


Fig. 2. Mean Access Delay as a function of the number of terminals, Poisson Arrivals

terminals. In this configuration, the best results obtained with the SP correspond to the value $K = 1.0$ for the same reasons invoked when considering Poisson arrivals. We also note that the durations of the “ON” periods remain short compared to the supposed reactivity of the SP.

In Fig. 4, the mean duration of the “ON” periods is set to 100 frames. The results derived from the different allocation methods are now very far away from the lower bound. The system becomes extremely weak; the periods of overload are very long. Even in the “big brother” case, the performance deteriorates. The performance is clearly degraded when we increase the number of terminals with the fixed allocation whereas dynamic methods allow on a limited horizon a more efficient absorption of traffic bursts. SP and candid methods are less sensitive than the fixed allocation on the number of terminals. If we consider $N = 10$ terminals, all the methods nearly lead to the same results. Starting from $N = 14$, we observe for the first time the interest of the SP compared to the fixed allocation method that does not work properly. In this configuration, in the fixed allocation scheme, terminals are allocated 7 slots in each frame, which does not make it possible any more to absorb instantaneously peaks of traffic. The benefit of the absence of signalling is reduced. The optimal value of K is $K = 0.6$. The SP is better (except for $K = 1.0$) than the fixed allocation... but in these cases, “Candid” method leads to better performance results.

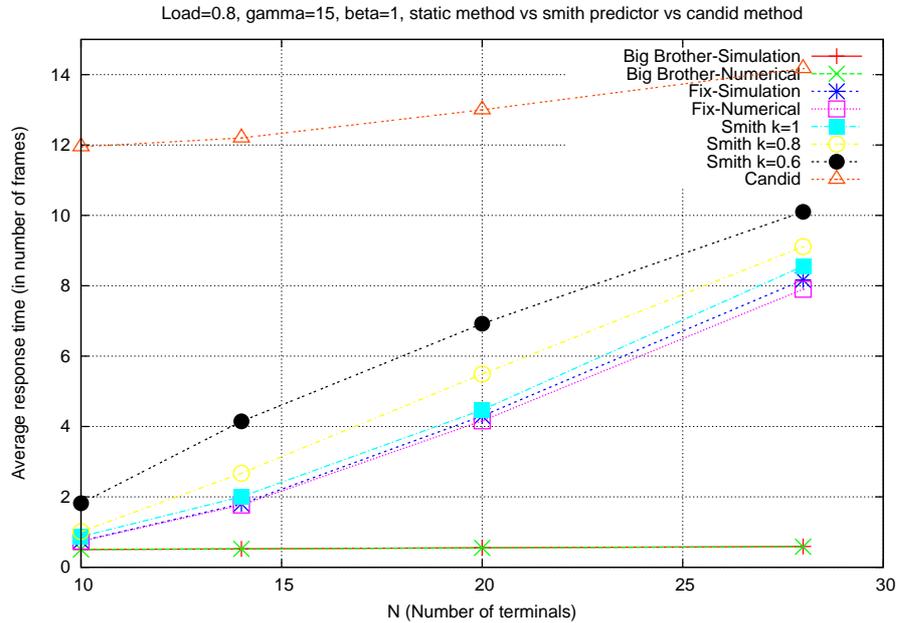


Fig. 3. Mean Access Delay as a function of the number of terminals, IPP arrivals

5 Conclusion

In GEO Satellite systems, dynamic resource allocation is a difficult problem because of the very important delay which separates requests from their response and of the strong potential variability of injected traffics. The Smith's predictor has been extensively studied in communication networks. If traffic variations or delays are low, this method can be interesting. Nevertheless, fixed allocation or multi-step allocation appears more effective in the satellite context. Taking into account of a predictor of traffic is undoubtedly important. However, our experiment shows that if the traffic has not the same nature as the modelled traffic, results are questionable. In our experiments, it appears that for not very sporadic traffics, the fixed allocation is the most efficient (because it does not generate significant signalling's overhead and delay) and that for much more sporadic traffics, a "naive" method aiming at emitting the current size of the buffer absorbs more efficiently traffic variations in particular over long durations. In conclusion, the methods based on the Smith's predictor may possibly be better than simpler fixed allocation methods when the systems are heavily loaded or when the durations of the burst periods are large. However, if these periods are really long, some signalling could make it possible to switch from a step of fixed allocation to another one. It could be advisable to consider, in further analysis, real traffic traces and to focus also on throughput and fairness.

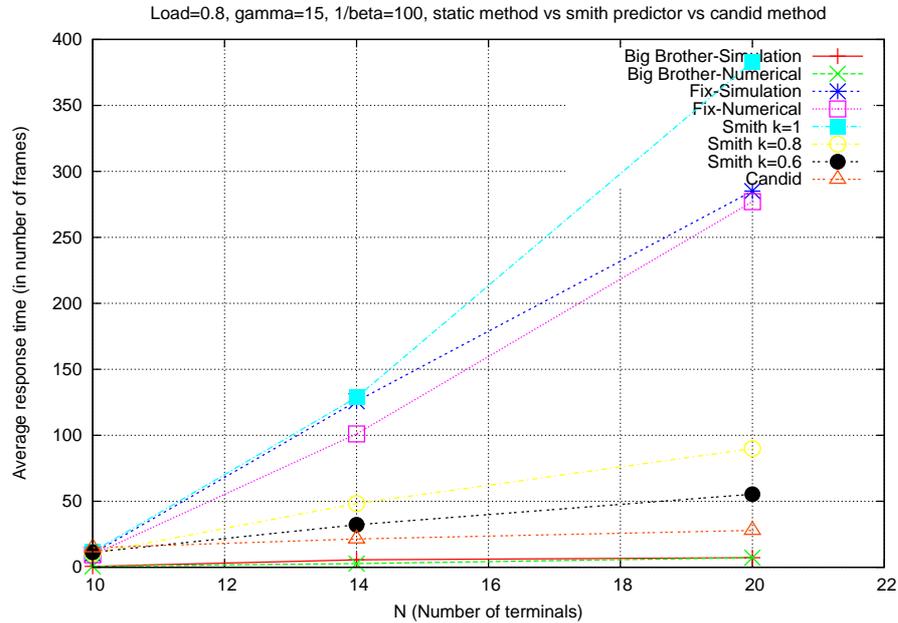


Fig. 4. Mean Access Delay as a function of the number of terminals, IPP arrivals

References

1. Chisci L, Fantacci R, Pecorella T : Multi-terminal dynamic bandwidth allocation in GEO Satellite Networks. In: IEEE VTC04 Spring (2004) 2797-2801.
2. Chisci L, Fantacci R, Pecorella T : Predictive bandwidth control for GEO satellite networks. In: IEEE ICC 2004 (2004) 3958-3962.
3. Chisci L, Fantacci R, Francoli F, Pecorella T : Dynamic Bandwidth Allocation via Distributed Predictive Control in Satellite Networks. In: IEEE ISCCSP04 (2004) 373-376.
4. Delli Priscoli F, Pietrabissa A : Load-adaptive bandwidth-on-demand protocol for satellite networks. In: IEEE ICDC02, Las Vegas (2002) 4066-4071.
5. S. Mascolo S : Congestion control in high-speed communication networks using the Smith principle. Automatica, Vol. 35 (1999) 1921-1935.
6. Mascolo S, Di Sciascio E, Grieco A : End-to-End Congestion Control and Bandwidth Estimation in High Speed ATM Networks. In: International Conference ITI 2001, Pula, Croatia (2001) 57-62.
7. Smith O.J.M : Closer control of loops with dead time. Chemical Engineering Progress. Vol. 53 (1957) 217-219.
8. Beylot AL, Dhaou R : Optimisation de boucles d'allocation. Alcatel Space Grant Final Report (2006).
9. Janssen A, van Leeuwarden J : Analytic computation schemes for the discrete-time bulk service queue. Queueing Systems. Vol. 50 Springer (2005) 141-163.
10. Kingman J : On the Algebra of Queues. Methuen, London (1966).
11. Little J : A proof for the queuing formula : $L = \lambda W$, Op Research Vol. 9 (1961) 383-387

12. Fischer W, Meier-Hellstern K : The MMPP cookbook. *Perf Eval.* Vol. 18 (1992) 149-171
13. Alagoz F, Vogcic B-R, Walters D, Altrustamani A, Pikholtz R-L : Fixed versus adaptive admission control in direct broadcast Satellite networks with return channel. *IEEE Jour. on Sel. Areas in Comm.* Vol. 22 (2004) 238-249.
14. Baglietto M, Davoli F, Marchese M, Mongelli M : Neural approximation of open-loop feedback rate control in satellite networks. *IEEETrans on Neural Net.*, Vol. 16 (2005) 1195-1211.
15. Celandroni N, Davoli F, Ferro E, Gotta A : Networking with multi-service GEO satellites: Cross-layer approaches for bandwidth allocation. *Int. Jour. of Sat. Comm. and Net.*, Vol. 24 (2006) 387-403.