Dedicated Bandwidth Approach for Channel Allocation in a Multi-Service Up/Down Link of a Low Earth Orbit Satellite Constellation

Rima Abi Fadel[1,2], Samir Tohmé[2]
[2]Ecole Nationale Supérieure des Télécommunications
46, Rue Barrault, 75013, Paris, France
{abifadel, tohme @ inf.enst.fr}
[1]Ecole Supérieure d'Ingénieurs de Beyrouth
Mkalles, Mar Roukoz, BP 11514, Liban
{rima.abifadel@fi.usj.edu.lb}

Abstract: Different users with diverse requirements need to share effciently the Up/Down Link UDL of a Low Earth Orbit LEO satellite constellation. Three types of services are assumed : real time (voice) services with strict constraints over the delay and the bandwidth, non real time (data) services delay tolerant but with bandwidth guarantees requirements and Best Effort services with no guarantees requirements. The scheme proposed consists in dedicating a certain amount of the total bandwidth for serving real time services, while the remaining bandwidth is shared according to a strict priority scheduling scheme between the two data services considered. All services are served as "bursts". Markovian ON-OFF source models are considered, together with a Pareto distribution for data packet lengths with a cut-off at 1502 bytes. The impact of the amount of dedicated bandwidth for real time services on the performance measures of all traffic classes is evaluated.

Key words: Access Networks, Performance Analysis, Satellite Networks.

1. INTRODUCTION

LEO Satellite Constellations are seen as a suitable mean for providing mobile users a global access service to terrestrial networks. Previous work [1,7] studies the connection admission control CAC problem at the air interface level of the satellite constellation using an enhanced trunk reservation technique and considers circuit switching for voice and burst switching for data. In this paper, a complete burst switching scheme is

---

considered. The network considered assumes an integration of three service classes. The highest priority class is the real-time service (voice). Data traffic is separated into two priority classes. The higher priority service has guaranteed bandwidth, and will be referred to as data service, while the lowest priority service called "Best Effort" BE has no guarantees requirements. The scheme proposed considers a division of the total bandwidth into two sets. The first is dedicated for the exclusive use of voice calls. The second is shared between data and BE in a strict priority scheduling manner. A certain boundary separates these two sets. An ON-OFF model is assumed for all sources. All data packet lengths are assumed to be Pareto distributed with a cut-off at 1502 bytes. An active period of the voice service is allocated a unit bandwidth within the dedicated bandwidth set, thus allowing the service of numerous active periods simultaneously. In the second set however, the whole bandwidth is assumed to be equivalent to only one server. The impact of the boundary choice is evaluated.

The "movable boundary" technique has been proposed in [2] for integrating voice calls and data packets. It has been applied in [3] to a broadband network integrating three service types, video, voice and data. [4] proposes to sub-divide network components into multiple sub-components, each having a dedicated bandwidth and carrying only traffic with the same Grade of Service GoS requirement. In [5] and [6], a routing method is developed, based on a base policy, a direct link routing requiring that each call type has a portion of link bandwidth dedicated to it. Priority scheduling, on the other hand, is seen in the literature as one of the simplest forms of scheduling in the context of service differentiation. In[8], static priority scheduling SPS is compared with other scheduling techniques in a General Packet Radio Switch system. In [9], priority is given to real time applications over non real time ones in a WCDMA environment, also calls experiencing handoffs are given priority over newly generated calls. A differentiated service Diffserv model is considered in [10] where flows are partitioned into different classes with different priorities. More general priority assignments than the classical ones are studied in [11]. A non-static priority assignment is also proposed in [12]. Priority scheduling has also been considered in ATM and high speed networks [13]. Delay for CBR traffic is computed in [14]. Finally in [15] a probabilistic priority discipline is proposed in order to avoid the starvation problem.

The paper is organized as follows. In section 2, we introduce the source models and the parameters used for the simulation as well as the admission scheme proposed. In section 3, the impact of the amount of dedicated bandwidth over performance measures of the real time service, for different voice source types, is considered. In section 4, performance measures of data and BE, for different boundary choices and different source types, are evaluated. Section 5 discusses the impact of varying data and BE queue lengths. Section 6 concludes the study.

2. SYSTEM AND SOURCE PARAMETERS DESCRIPTIONS

The French Réseau National de Recherches en Télécommunications Satellites Constellation Project [2], proposes a mobile communication access service for mobile users using a LEO satellite constellation. The total bandwidth serving pedestrian users on the uplink is of 72 kbits/s, with a granularity of 2.4 kbits/s. The cellular system that will be considered in this paper is based on cells fixed to the earth. Furthermore, the user mobility is considered as negligible in comparison with the satellite mobility because the typical diameter of a satellite cell is of order of few hundred kilometers, which is large at the space scale of a terrestrial mobile user (pedestrian, car). Thus, the issues associated with the Handover HO will not be considered. Since the bandwidth on the downlink is usually higher than the one on the up link, we focus on the channel allocation on the up link. Three call classes are considered. A voice source is an ON-OFF source. We assume no packet arrivals during the OFF period. And the ON period is served as a burst. Two voice sources types are considered. For both types, ON and OFF times are exponentially distributed with mean times of 3s (respectively 0.352s) for the ON period and 3s (resp. 0.650s) for the OFF period. Since with the parameters given above, thirty communications (N) can be simultaneously established, the number of active voice sources will be assumed to be equal to thirty in the simulation. An ON burst will be allocated the unit bandwidth of 2.4 kbits/s. BOUND, the boundary value represents the number of ON periods that can be simultaneously served. Since BOUND < N, a finite queue will be considered where ON bursts will be entitled to wait for service according to the simple First Come First Serve policy. Web sources are described in the ETSI technical report ETSI TR 101 112. They are modeled as ON-OFF sources with both ON and OFF times exponentially distributed with means for the UDD 8kbits/s model of respectively 12s and 412s. During the ON time a mean of 25 packets are generated with packet inter-arrival times exponentially distributed of mean 0.5s. Packet lengths are distributed following a Pareto distribution with cut-off at 1502 bytes. Throughout the simulations, the same source models are assumed for both data and BE sources. This represents a rather pessimistic case of study, especially for the BE traffic since the latter is usually described with more sporadic models. Other source parameters will be considered in further simulations and will be discussed in the relevant sections. Two queues are considered, one for the data bursts and the other for the BE service. These queues are served according to a strict priority scheduling scheme. A burst will be given the total bandwidth available for data services once admitted into the service, a bandwidth of (N-BOUND)*unit bandwidth. As mentioned above, the service unit is the burst. Figure 1 represents the system.

3.   IMPACT OF THE BOUNDARY CHOICE OVER THE REAL TIME TRAFFIC

In this section, the impact of the choice of the boundary value on the voice traffic performance is evaluated. The queue length is also varied and two source types are considered, the 3s/3s ON/OFF model and the 0.352s/0.65s one. Two performance measures are considered : the burst loss probability and the burst waiting time in queue. A new burst is lost whenever it finds no room in the real time service queue. This study aims to evaluate the benefits of the statistical multiplexing if, instead of considering a pure circuit switched system, the service unit is the burst. Figures 2.a) and 2.b) represent respectively the evolution of the loss probability and of the waiting time for the 3s/3s voice bursts with the number of places in queue m and for different values of the parameter BOUND. Figure 2.a) shows a clear decrease in the burst loss probability for voice as the number of places in queue increase, a decrease that is more significant for smaller values of BOUND. In fact, as BOUND increases, a greater number of bursts can be simultaneously served. A smaller size of the queue is then needed to meet the same loss probability. As for the waiting time in queue represented in figure 2.b), it is observed that the smaller the value of BOUND the bigger the delay. In fact, for higher values of BOUND, the mean arrival rate into the system remains the same while the mean departure rate increases since more channels can be simultaneously used. This leads to a less loaded system and the cases where the queue is full occur less frequently. This means that under the same arrival conditions, the mean number of customers in the system decreases and so does the delay. Another observation is that the delay is more sensitive to m for smaller values of the parameter BOUND. If no queue limitation is done for such systems, the mean customer number that would result is bigger than the chosen queue size. Of course this means that the queue will tend to saturate. If more places were available in queue, the clients would go on filling them, thus increasing the delay, and so on until the queue size becomes quite larger than the mean number of clients expected in the equally loaded infinite queue system.

A comparative study of the two systems with 3s/3s sources and 0.352s/0.65s sources respectively regarding the loss probability for bursts shows that for the same boundary value and the same queue size, the second source type largely outperforms the first one. This can be explained by the fact that the proportion of the ON period regarding the overall activity period is of one half in the first case while it is of almost one third in the second one. This means that the relatively longer OFF periods can be better exploited in this case, thus resulting in a higher statistical multiplexing gain. This is a direct consequence to the fact that for the same peak rate for both source types, the 0.352s/0.65s represents a smaller mean rate. On the other hand,

regarding the waiting delay, much smaller values were noted for this source type as a consequence to the fact that the ON period is shorter.

To conclude this section, an important result is that the choice of the boundary will be subject to the voice source nature as well as the performance measures required for this strict constrained class of service. This choice also has to consider the data services behavior.

4. IMPACT OF THE BOUNDARY CHOICE OVER DATA AND BEST EFFORT

Different loads for the data and the BE classes are considered. This first approach to the problem evaluates the effectiveness of the head of the line HOL scheduling scheme considered. Figures 3.a) and 3.b) represent respectively the loss probability and the waiting time in queue for data bursts, such as WEB bursts. These figures show that for a given number of Web active sources, the loss probability of Web bursts is very slightly affected by the number of active BE sources. The slight increase noted is due to the non-preemptive priority scheme. The delay increase with the number of BE active sources is more visible since the residual BE service time is a component of the waiting time of Web bursts. Simulation results not represented here show that for a given number of BE clients, loss probability as well as waiting time in queue for BE clients significantly increase with the number of Web active sources, due to the strict priority scheme chosen. This leads to a starvation problem with high loss probabilities (around 0.5) and large waiting time in queue (around 40s for a service time around 3s) that get more and more critical as the number of Web active clients increases. Figures 4.a) and 4.b) represent the performance measures for the data service when BOUND is varied. An important observation is that for higher values of BOUND, an additional increase affects more significantly the performance measures of the lower classes of service. Only the Web performance curves are represented, but simulation shows that the same conclusions hold for the BE service. In fact, BE is in an even worse situation than what is described for Web when BOUND increases, especially when the load of the Web class increases. Other simulations were run in order to evaluate the impact of varying the source models for data and BE sources. The same mean rate of 266.415 bits/s per source is considered, while the peak rate is varied. Considered values are 64 and 144 kbits/s. An increase of the data service with guarantees loss probability with the increase of the peak rate of the sources is noted. Even though the mean rate is the same, a rapid queue saturation problem may arise when different sources are in their ON state, simultaneously transmitting at their peak rate. The performance degradation is however not drastic since the mean rate is relatively small. Another

consequence of this small value is that the mean waiting delay of data is almost unaffected with the peak rate variation.

5.   IMPACT OF DIFFERENT QUEUE LENGTHS FOR DATA AND BE

In this section, data and BE queue lengths are varied. Two values are considered, 5 and 10 places. A "place" in queue is assumed to be large enough to receive a whole burst. Figure 5.a) shows that data loss probability is very sensible to the data queue length. It considerably decreases, for any amount of BE active communications, as the data queue increases. The price to pay, however, is the important increase in the waiting delay for this service, as shows figure 5.b), especially when more active sources compete for the system resources. Increasing the BE queue length however has almost no effect over the performance of data since absolute priority is affected to the latter. Other simulation results show that, for a light load associated with the data service, increasing the data queue length does not affect the performance of BE bursts. On the other hand, the waiting time of BE bursts is significantly penalized when data bursts represent a significant load since BE bursts can only be served when the data queue evolves to an empty state. It can, as the number of active BE sources increases, attain an order of a minute. Its increase however with the BE queue length is far less important than what is observed when the data queue length increases, while leading to better blocking conditions. The data queue should thus be carefully dimensioned since it affects the data service and can contribute in the establishment of a starvation situation for BE bursts. Other simulations show that if a cutoff at 10 kbytes is chosen, a drastic degradation of all performance measures is noted.

6.   CONCLUSION

In this paper, a dedicated bandwidth scheme, reserving an amount of bandwidth for voice, is combined with a HOL mechanism that controls access to the remaining bandwidth of the lower priority data and BE bursts. Simulations show that the boundary choice is function of the voice and data source types, as well as of the number of simultaneously active sources of each class. The HOL mechanism applied to the data services shows that the number of active sources of each type is to be controlled in order to guarantee Quality of Service parameters for each service class, in accordance with the system parameters, that is the amount of available bandwidth and buffering space. The dimensioning problem must be carefully performed since various, and conflicting parameters should be considered, and serious problems such as the severe starvation problem should be excluded.
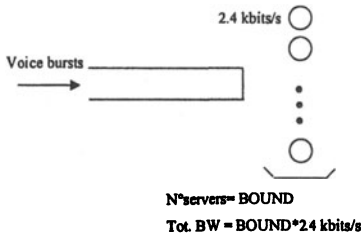
2.4 kbits/s

Voice bursts

N°servers= BOUND
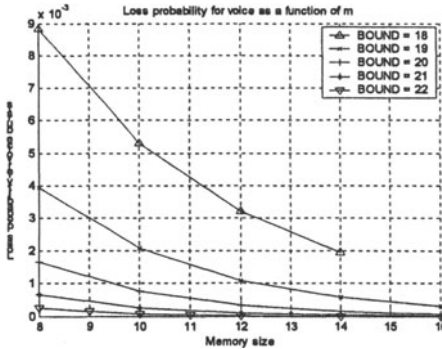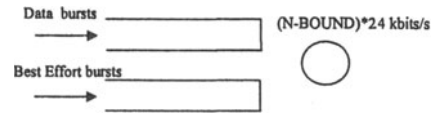Tot. BW = BOUND*24 kbits/s

Figure 1.a) Dedicated channels for voice.

Data bursts

(N-BOUND)*24 kbits/s

Best Effort bursts

Figure 1.b) Head of the Line scheduling for data service with guarantees and Best Effort service.

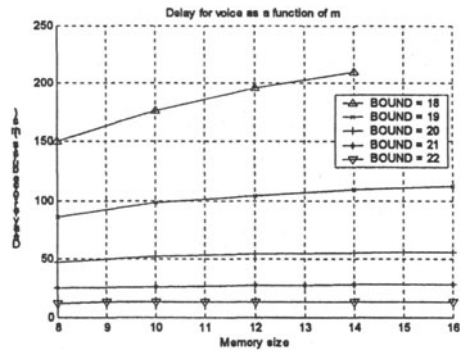*Figure 2.a)* Voice bursts loss probability as a function of m and BOUND.

Figure 2.b) Voice bursts' waiting time as a function of m and BOUND.
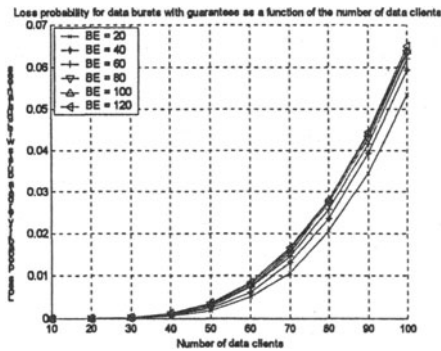
Figure 3.a) Web bursts loss probability.
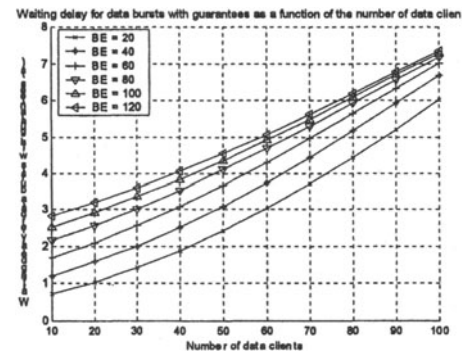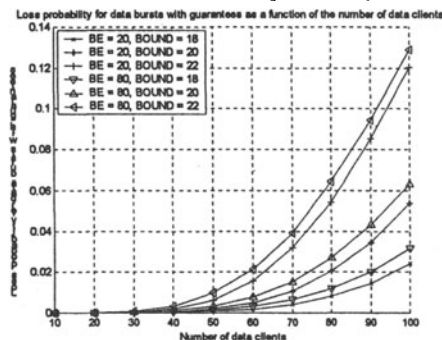

Figure 3.b) Web bursts' waiting time

Figure 4.a) Web bursts loss probability for different values of BOUND.
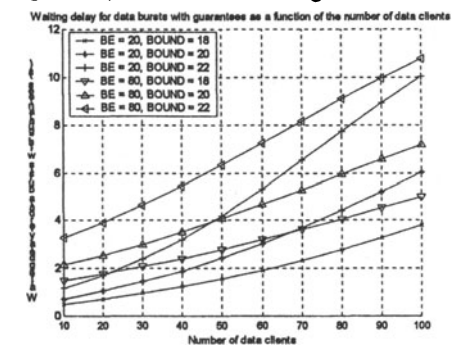

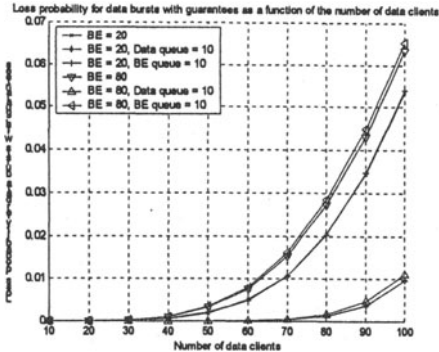*Figure 4.b)* Web bursts' waiting time in queue for different values of BOUND.

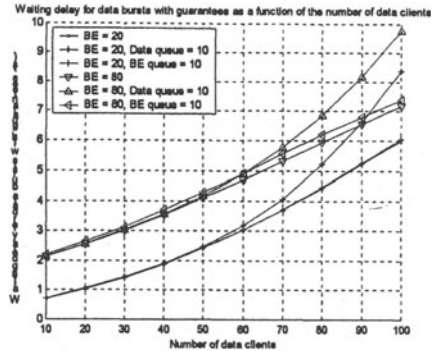*Figure 5.a)* Web bursts loss probability for different queue lengths.

Figure 5.b) Web bursts' waiting time in queue for different queue lengths.

REFERENCES

[1] R. Abi Fadel, S. Tohmé. *Connection Admission Control CAC and Resources Allocation RA on the Up/Down Link UDL in a Low Earth Orbit LEO Satellite Constellation*. Proc. IEEE Int. Symp. Comp. and Commun., Jul.2001.

[2] P. Zafiropoulo. *Flexible Multiplexing for Networks Supporting Line-switched and Packet-Switched Data Traffic*. ICCC 1974.

[3] G. Stamatelos, J. Hayes. *An integrated system for video, voice and data communications*. Proc. IEEE SUPERCOMM/ICC'92, Vol. 1 , 1992.

[4] H. Ji, J. Hui, E. Karasan. *GoS-Based Pricing and Resource Allocation for Multimedia Broadband Networks*. Proc. IEEE INFOCOM 1996, Vol. 3 , 1996.

[5] A. Kolarov, J. Hui. *Least Cost Routing in Multiple-Service Networks*. Proc. IEEE INFOCOM, Toronto, Canada, June 1994.

[6] A. Kolarov, J. Hui. *Least Cost Routing in Multi-Service Networks*. Proc. IEEE INFOCOM 1995, Vol. 1 , 1995.

[7] R. Abi Fadel, S. Tohmé. *Connection Admission Control CAC and Differentiated Resources Allocation RA in a Low Earth Orbit LEO Satellite Constellation*. Proc. of the IFIP Networking 2002 Conf., May 2002.

[8] Q. Pang, A. Bigloo, V. Leung, C. Scholefield. *Service Scheduling for General Packet Radio Service Classes*. IEEE Wireless Communications and Networking Conf., vol.3, 1999.

[9] M. Kazmi, P. Godlewski, C. Cordier. *Admission Control Strategy and Scheduling Algorithms for Downlink Packet Transmission in WCDMA*. IEEE Vehicular Techn. Conf., vol.2, 2000.

[10] S. Wang, D. Xuan, R. Bettati, W. Zhao. *Providing Absolute Differentiated Services with Statistical Guarantees in Static Priority Scheduling Networks*. Proc. Seventh IEEE Symp. Real-Time Technology and Applications 2001.

[11] S. Wang, D. Xuan, R. Bettati, W. Zhao. *Providing Absolute Differentiated Services for Real-Time Applications in Static Priority Scheduling Networks*. Proc. IEEE INFOCOM 2001, Vol. 2 , 2001.

[12] K. Siriwong, R. Ammar. *QoS Using Delay-Synchronized Dynamic Priority Scheduling*. Proc. IEEE Int. Symp. Comp. and Commun., Jul.2001.

[13] C. Li, R. Bettati, W. Zhao. *Static Priority Scheduling for ATM Networks*. Proc. 18th IEEE Real-Time Systems Symp., 1997.

[14] K. Iida, T. Takine, H. Sunahara, Y. Oie. *Delay Analysis for CBR Traffic under Static-Priority Scheduling*. IEEE/ACM Transactions on Networking, vol.9, no.2, April 2001.

[15] Y. Jiang, C.K. Tham, C.C.Ko. *A probabilistic Priority Scheduling Discipline for Multi Service Networks*. Proc. IEEE Int. Symp. Comp. and Commun., Jul.2001.