Performance of multiple access protocols in geo-stationary satellite systems

Heba Koraitim, Samir Tohmé †

Marouane Berrada, Americo Brajal ‡

† Ecole Nationale Supérieure des Télécommunications

46, rue Barrault 75634, Paris-France,

Tel.:33 1 45817449, Fax.:33 1 45891664

e-mail: koraitim, tohme@res.enst.fr

‡ Laboratoires d'Electronique Philips S.A.S.

22, avenue Descartes - BP 15 - 94453 Limeil Brevannes Cedex France,

Tel.:33 1 45106700, Fax.:33 1 45106960

e-mail: berrada, brajal@lep.research.philips.com

Abstract

Two packet multiple access schemes, the DQRAP and the ARRA are modeled and evaluated in the geostationary satellite environment. A new protocol is afterwards proposed joining the advantages of both studied schemes, and more adapted to interactive multimedia applications over satellite uplinks. The Generalized Retransmission Announcement Protocol, GRAP, re-groups the immediate access by contention at low loads, and the reservation access for higher loads to achieve a better channel efficiency. Simulation results illustrate an improved throughput/delay characteristics and a higher protocol stability. Enhanced versions of the protocol are also proposed and evaluated to further improve its efficiency, with reasonable additional complexity.

Keywords

VSAT, satellite networks, multiple access protocols, packet switching

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35388-3_42](https://doi.org/10.1007/978-0-387-35388-3_42)

1 INTRODUCTION

Satellites systems have been recently rediscovered, to complement terrestrial networks in providing a worldwide access to the evolving multimedia services.

Multiple access protocols constitute one of the most important aspects that largely determine the performance of communications networks. They define the means by which a subscriber can establish the contact to the network and hence gain access to its resources. Network subscribers are then competing in a pre-established manner to share the resources of a specific link. Many families of protocols have been proposed in the literature in different network contexts (Lee *et al.* 1983), (Lee *et al.* 1984), (Raychaudhuri *et al.* 1987), (Wong *et al.* 1991) and (Mohammed *et al.* 1994). The features of these protocols differ according to the type of the considered network and its topology.

Satellite networks have specific characteristics which largely influence the features of the multiple access scheme to be adopted. From one side, the inherent broadcast feature of satellite links enables the broadcast of information to all covered users at almost the same time. However, the large round-trip propagation delay characterizing the geostationary satellite environment imposes some limitations in protocol design.

Two multiple access schemes are studied and analyzed in this work, the Distributed Queueing Random Access Protocol (DQRAP), and the Announced Retransmission Random Access (ARRA) protocol. We have decided to compare these two as they rely on the same basic idea of allowing contention access at low loads and switching to reservation access at higher loads. Besides, both protocols separate new arrivals from retransmitted messages to avoid excessive conflict and collisions. In DQRAP (Xu *et al.* 1993, Koperda *et al.* 1995), this is done by blocking new arrivals from entering the system if a collision resolution cycle is in progress. In ARRA, on the other hand, new arrivals are allowed to transmit in a separate part of the frame called the common mini-slot pool (CMP) (Raychaudhuri 1985).

These two protocols differ however in the collision resolution principle applied in each case. While DQRAP adopts a tree-based approach (Capetanakis 1979) for resolving collisions between reservation requests, ARRA protocol randomly arbitrates retransmission requests over the total number of frame slots.

DQRAP was originally proposed for terrestrial hybrid/fiber coaxial networks. We have therefore adapted the access algorithm to the satellite environment before modeling and examining the performance of both protocols on an uplink satellite channel. A new protocol, the Generalized Retransmission Announcement Protocol GRAP, is then proposed which further enhances the channel efficiency and achieves a better system stability at higher loads.

In the next section, the two protocols, DQRAP and ARRA are briefly described, where the DQRAP adaptation to the satellite environment is particularly emphasized. The two protocols are modeled and their simulation results

are presented. In section 3, the GRAP protocol is introduced, and detailed, explaining the modifications added to enhance its performance. Section 4 reveals the protocol models developed to extract the performance evaluation results which demonstrate the behavior of GRAP and E-GRAP and compare it to other simulated protocols.

2 PROTOCOLS IN THE SATELLITE CONTEXT

Due to the large round-trip propagation delay encountered in satellite links, the channel is always structured in the form of frames, whose length can vary to cover a part of, or the entire propagation time. In other words, the frame duration may be one or a multiple of channel frame durations. For both DQRAP and ARRA, the frame is divided into a number of equal length slots, where each slot is further divided into a data transmission slot (DS) and a number of control mini-slots (MS).

2.1 Satellite DQRAP

The frame structure of the DQRAP protocol is shown in figure 1. The frame duration is equal to the two-way round-trip propagation time from a user terminal to the network control center (NCC) and back again to the terminal. The status of the system is broadcasted to all terminals within the satellite

Figure 1 DQRAP frame structure for satellite links

coverage by the NCC. Each active terminal keeps track of the channel status by listening to the feedback information of the NCC. A distributed queueing discipline is maintained by each terminal to memorize the status of each slot, and keep track of its own position in the queues. This discipline functions by storing the status information of each slot in two global queues, the transmission queue, TQ , and the collision resolution queue, RQ .

When a terminal generates a packet, it will listen to the information concerning the queues status of the upcoming slot after the packet arrival. If both queues are empty (slot 1 in figure 1), the packet will be transmitted in DS together with a reservation request transmitted in a randomly selected MS of the same slot.

If the DS is successful, the corresponding reservation will be canceled. Otherwise, if the DS collides, while the MS is successful, the terminal then enters the (TQ) of the corresponding slot on the frame, and keeps track of its position in the distributed queue to transmit its data in the same slot of the next frame after the round-trip propagation delay.

However, if both, the DS and MS suffer collisions (more frequent at higher loads), the terminal will be placed in the *RQ*, in the same position with all other terminals colliding in the same MS of the corresponding DS. Colliding requests are organized in the *RQ* by the order of the position of their MS in the slot (Collisions in MS 1 enter in first position, while those in MS 2 enter in the second position).

The first group in the *RQ* will retransmit only their request in the same slot, after randomly selecting another MS. Request collisions in a certain slot are hence resolved according to a tree algorithm by progressively dividing colliding groups in mini-slots into smaller groups. Collision resolution trials are then separated by the round-trip propagation delay (or the frame duration), as they are always attempted in the same slot of the frame.

Whenever a packet arrives in a terminal and learns, by the feedback information, that the *RQ* of the upcoming slot is empty, while the *TQ* is not (slot 2 in the figure), the terminal transmits only a contending reservation request in one of the MSs of this slot. If however, the *TQ* is empty, while the *RQ* is not (slot 3), or if both are non-empty (slot 4), the packet will be held in an arrivals queue in the terminal until a slot on the frame is detected with an empty *RQ*.

An occupied *RQ* indicates that a collision resolution cycle is in progress in this slot, and hence new arrivals are prohibited from entering the system. This measure serves to decouple new arrivals transmissions from retransmissions, thus increasing the protocol stability.

2.2 Satellite ARRA

In contrast to the DQRAP protocol, which was originally proposed for terrestrial networks, the ARRA protocol initially targeted broadcast channels with a large round-trip propagation delay. The frame structure of ARRA is shown in figure 2, where the round-trip delay consists of a multiple of frame durations.

The frame is divided into a number K of slots and each slot in turn, is divided into a data slot (DS) and K mini-slots (MS). Similar to the DQRAP

protocol, the length of an MS is much smaller than the length of a DS. Both, the DS and the MS, have the same roles as in DQRAP for transmitting the data packet and the reservation requests respectively. An additional field of K MSs constitute the first part of the frame to serve as a Common Mini-slot Pool (CMP).

Figure 2 ARRA frame structure for satellite links

When a packet is generated, the terminal will wait for the beginning of the next frame to read the feedback information transmitted by the NCC concerning the frame. If there are free slots available for contention on this frame, it will choose a free slot at random and transmit its packet in the DS field of the slot. A reservation request is also simultaneously transmitted in one of the MSs of the same slot, to advertise the need for reservation if the data message collides. The MS within the slot is randomly chosen and the selected MS number indicates the slot in which retransmission will take place in the following frame, after the round-trip propagation delay. Each terminal is then *announcing* a reservation in case of collision.

If no free slots are available on the frame, a terminal will only transmit a reservation request in one of the MSs of the Common Mini-slot Pool (CMP) at the beginning of the frame, to reserve the corresponding slot number in the upcoming frame. The feedback information regarding the status of the frame and the available slots for contention is broadcasted by the NCC before the beginning of each frame.

No special measures are adopted in this protocol to resolve collisions between retransmissions, except for the fact that each back-logged terminal in a collision will randomly choose a different slot each time by announcing in the corresponding MS number. This randomizes the retransmission trials over the whole available set of free slots on the frame.

2.3 Performance of DQRAP and ARRA

The two protocols were modeled in a VSAT star network configuration, where a satellite with on-board processing capabilities ensures the NCC access control functions. A Poisson model was considered to model the traffic generated by an infinite number of VSAT user terminals.

DQRAP protocol was simulated with a frame duration equal to the round-trip propagation delay (0.27 sec.), consisting of 108 slots, which is also the number of interleaved DQRAP engines. On the other hand, a frame duration of 20 ms was considered for ARRA protocol, with 8 slots per frame ($K = 8$). This frame duration and number of slots per frame were chosen as a compromise, between the frame length and the overhead introduced by the mini-slots. Both, DQRAP and ARRA, have the same slot length, and were tested for the same uplink channel capacity.

DQRAP protocol was modeled for two values of MS, the first equals three, while the second equals eight. The latter value was particularly considered in order to compare its performance to that of ARRA, when they both have the same number of MSs per slot. The additional overhead introduced by the MSs is not counted for in the results, as we are only considering the throughput of the data slots.

We have noticed, as illustrated in figure 3, that the throughput of the ARRA saturates at a value of 0.425. This is due to the excessive number of collisions taking place between retransmissions and newly arriving reservation requests, in the limited number of mini-slots of the CMP. As the network loading exceeds 0.43, an unstable state is reached where collisions are multiplying and the system output saturates. The throughput/delay characteristics of the ARRA protocol are shown in figure 4.

Figure 3 ARRA throughput

Figure 4 ARRA characteristics

The performance of DQRAP with the two values of MS (3, 8) is presented in figure 5. Obviously, the DQRAP protocol largely outperforms the ARRA

protocol in terms of maximum channel throughput, which reaches 0.98. This is due to the complete separation of the transmission and collision resolution processes characterizing DQRAP.

Figure 5 DQRAP throughput/delay characteristics

The protocol is also stable, over a wide range of loading conditions, due to the application of the tree-based collision resolution scheme. It is noticed that increasing the number of MSs per slot from 3 to 8 has a slight effect on reducing the average data delay at lower channel loads, but has no significance at higher loads, as the two curves coincide at a load of 0.9.

3 GENERALIZED RETRANSMISSION ANNOUNCEMENT PROTOCOL

The Generalized Retransmission Announcement Protocol, GRAP, combines some aspects of the previously described protocols to develop an access scheme adapted to the satellite characteristics, and at the same time achieving an acceptable efficiency and stability.

The frame structure of GRAP is shown in figure 6. Similar to the ARRA, the frame is divided into K slots, and each slot consists of a data slot field (DS) and K mini-slots (MS) field. The frame duration is chosen such that the

overhead introduced by the K MSs length is very small compared to the DS length.

Figure 6 GRAP frame structure

3.1 GRAP Access Procedure

New arrivals When a user terminal becomes active, it will monitor the down-link stream, to read the status of the upcoming frame and the set of available slots (SAS) broadcasted by the NCC. At the beginning of the following frame, if there are free slots available for contention, the terminal will transmit its data message, accompanied by an anticipating retransmission announcement placed in a randomly chosen MS of the same slot. The MS number in the slot indicates the slot number that the terminal requires to reserve in the next frame in case of collision.

Retransmissions If an active terminal receives a feedback indicating that its packet has collided, it will search the SAS for the slot number of its announced reservation. If the slot number is not included in the SAS, it assumes that its reservation has been successful and waits for the assigned slot to transmit its data message.

On the contrary, if the announced slot is included in the SAS, it means that the reservation has not been successful due to collisions or other errors. The terminal then retransmits a reservation request in a randomly chosen MS of an already reserved slot, since reserved slots are not accompanied by anticipating announcements. In this case, the MSs associated with the set of reserved slots (SRS) play the role of the CMP field in the ARRA protocol and replace it, hence reducing the overhead associated with the protocol. Furthermore, the

number of MSs now available for request retransmissions is a multiple of the CMP, and this multiple is equal to the number of slots included in the SRS.

This has the advantage that at high loads, when many slots are reserved, there is more space to retransmit reservation requests, and at low loads, more free slots will be available for contention.

3.2 Flow control

In order to limit the risk of repeated collisions at high network loads, a flow control mechanism relying on a back-off delay is introduced in the terminals' access procedure. Before retransmitting a collided request, the terminal will have to wait for a certain amount of delay, before attempting the retransmission. This delay is calculated as a function of the number of repeated collisions to which the terminal has been exposed and is directly proportional to it.

Another control aspect is envisaged whenever the network load goes down, and free slots are again available on the frame, where colliding requests are allowed to recycle in the system as new arrivals. The recycling feature is offered to requests that have exceeded a certain number of repeated collisions, to limit the number of conflicts in MSs and hence improve the protocol stability.

The number of repeated collisions, after which retransmissions are recycled as new arrivals in the system, is a parameter of the protocol that has to be adjusted for optimal operating efficiency. The recycling aspect not only reduces the retransmission load in the MSs, but also reduces the retransmission delay and helps in re-partitioning the load between new arrivals and retransmissions.

3.3 NCC operation

The NCC contains all the control and monitoring functions to organize and regulate the GRAP procedure. The uplink frame is completely received before any analysis and processing can be started. The NCC then treats the received frame and analyzes it before sending the feedback to all active terminals on the down-link.

1. **Frame analysis:** The NCC first searches the DS fields of the frame slots to detect successful and collided messages. It then proceeds with the detection of collisions in the MSs. Successful reservations associated with successful contention messages are canceled, while those associated with collided messages are memorized, together with successful retransmission requests in the MSs of the reserved slots in the SRS.

The NCC then allocates slots to successful reservations, and eliminates some of those who had a *virtual collision* (those who place their reservation

request in the same MS number of different slots). The elimination process can be done based on a priority or simply a random principle.

The advantage of having a number of MSs equal to the number of slots on the frame, is the reduction of the amount of information carried by each MS. Hence, the NCC is not obliged to specify the terminal ID associated with each reservation. A terminal can just place a signal in the MS and wait to know whether the corresponding slot is included or not in the SAS.

2. **Feedback formulation:** The set of available slots (SAS) is afterwards formulated, including all non reserved slots on the frame. On the down-link, this SAS is broadcasted, with all other feedback information concerning the reserved slots and their position on the frame. Each terminal then receives the result of its transmission trial, after the propagation and the NCC processing delays (which can consist of one or a multiple of the up-link frame duration).

3.4 Enhanced GRAP

In order to further enhance the performance of the protocol, a queueing system is added in the NCC to store virtually colliding reservation requests. The expression *virtual collision* designates the conflict occurring between terminals placing their reservation requests in the same MS number of two or more different frame slots. Although these requests can be correctly received at the NCC, their intention to reserve the same slot on the next frame causes the NCC to hold the reservation for only one terminal and cancel the others.

The queueing system then tends to overcome this weakness by preserves the reservations for all *virtually colliding* requests. Two queueing disciplines can be envisaged. The first one comprises a single distributed queue for the whole frame, while the second comprises a number K of queues, one for each slot on the frame.

(a) Single reservation queue

In this scenario, only one reservation queue is maintained by the NCC to store virtually collided requests, for which no place can be found on the upcoming frame, and postpone their allocations for future frames. This can be modeled as shown in figure 7.

The single queue follows a priority discipline, with the higher priority given to requests that have suffered more collisions and hence, more delay. At the NCC, these requests are scheduled at the beginning of the frame and are completely served before new arriving requests. If there remains some place afterwards on the frame, new requests can be accommodated.

Since the remaining free slots on the frame may not correspond to the originally reserved ones for some waiting requests, the NCC must include the

Figure 7 Single queue E-GRAP model

reserved slot number in its feedback information for each terminal, to notify it of its new reserved slot.

The advantage of this approach is the overall reduction of the mean message delay at higher loads, since the priority is proportional to the number of repeated collisions. A tradeoff has then to be made between the encountered mean delay and the maximum channel efficiency, since more feedback information has to be sent on the down-link.

(b) Multiple reservation queues

The second scenario envisaged to avoid virtual collisions is the introduction of a number of waiting queues to store successful reservation requests. This number is equal to the number of slots per frame K , and hence a waiting queue is envisaged per slot as illustrated in figure 8.

Figure 8 Multiple queues E-GRAP model

When several terminals transmit their reservation requests in the same mini-slot number of two different slots, all the successful requests are retained and stored in the queue corresponding to the targeted slot. Their position in the queue may either be random, follows their position on the frame, or be decided following a certain priority discipline.

The extra feedback information concerning the reserved slot number is not needed in this approach. This is however, at the expense of the possibility of suffering more delay, waiting for the same slot in future frames, while there may be other free slots on the present frame.

The two approaches explained above help in reducing the number of re-transmissions and hence, possible future collisions. Therefore they confine the delay variation to a smaller value than that experienced by the simple GRAP.

4 GRAP AND E-GRAP PERFORMANCE

Figure 9 VSAT terminal model

Simulation models have been developed for GRAP and E-GRAP to examine their performance in the same VSAT network configuration, previously described to test DQRAP and ARRA.

Simplified block diagrams for the VSAT terminal model and the NCC model are shown in figures 9 and 10. A frame duration of 20 ms and $K = 8$ slots per frame have been maintained for the system.

Figure 11 illustrates the throughput/delay characteristics of the GRAP and the two versions of E-GRAP with a single and multiple queues. The GRAP achieves a channel throughput of 0.95 for a mean data delay that approaches 2 seconds. It is clear that, introducing the queueing discipline has improved

Figure 10 NCC model

Figure 11 GRAP and E-GRAP characteristics

the performance of GRAP by increasing the maximum channel throughput to 0.98, at the same time of reducing the mean data delay in the system. E-GRAP with a single queue further reduces the mean delay at high loads, compared to the multiple queues E-GRAP, due to the complete statistical multiplexing aspect on which it is based. In this version of the protocol then, the number of MS per slot can be different from the number of slots per frame, since a terminal can be reserved any slot on the frame.

Figure 12 Performance comparison between DQRAP and E-GRAP

Figure 13 GRAP delay distribution

Figure 14 E-GRAP single queue delay distribution

Figure 15 E-GRAP multiple queues delay distribution

The additional delay introduced by the multiple queues E-GRAP at higher loads, is due to the fact that each terminal must wait for a specific slot that corresponds to the number of MS in which it has previously placed its reservation. The advantage of this technique however, is that the overhead introduced by the MSs in the slot is reduced due to the transmission of less control and feedback information.

Compared to the DQRAP protocol, as indicated in figure 12, both GRAP and E-GRAP outperform DQRAP, in terms of mean delay, up to a load of 0.95. As the load exceeds the latter value, the improvement introduced by E-GRAP is significant where it attains a throughput value of 0.98 for a delay of only 1.4 seconds. This demonstrates the E-GRAP stability at very high loads.

The delay distribution of GRAP, E-GRAP with single queue and E-GRAP with multiple queues, is measured and illustrated in figures 13, 14, and 15 respectively for a 0.8 loading factor. The E-GRAP is shown to achieve a better delay distribution, where the delay density function is confined to 1.3 seconds at 0.8 load, compared to around 3 seconds for the GRAP.

The distribution of the mean delay encountered by data packets in the DQRAP protocol, at the same load (0.8), is given in figures 16 and 17 for 3 and 8 MS respectively. The distributions for both MS values have little difference between them, but they are more dispersed compared to those of GRAP and E-GRAP. The dispersion in the delay values is spread over a wider range reaching a value of 5 seconds. This is due to the tree algorithm applied to resolve collisions of reservation requests in DQRAP.

The results indicate hence, that protocols applying tree-based collision resolution algorithms are stable at high loads, but at the expense of additional delay and delay variation. Besides, the DQRAP protocol introduces more complex functions in the user terminals, since each terminal have to keep track of the state of each of the transmission queue TQ , and the collision resolution queue RQ for each slot on the frame.

A back-off delay applied in the GRAP protocol improves the delay performance over that of DQRAP, with much less complexity. The enhanced versions of the protocol, with a queueing discipline, further enhances the protocol performance, at the expense of managing one or multiple distributed transmission queues.

Figure 16 Delay distribution of DQRAP with 3 mini-slots

Figure 17 Delay distribution of DQRAP with 8 mini-slots

5 CONCLUSIONS

We have modeled and analyzed two packet multiple access protocols in the satellite environment, the DQRAP and the ARRA protocol. The first is based on a tree collision resolution algorithm, which proceeds in parallel to the transmission process of packets in two separate distributed queues for each slot. The second simulated protocol relies on a simple retransmission policy to transmit reservation requests.

A new protocol, the GRAP, was proposed as a compromise between both techniques. It relies on the principle of complete separation between new arrivals and retransmissions. New arrivals access the contention sub-frame, while the retransmissions contend in the mini-slots of the reservation sub-frame. The protocol, however, preserves the right for retransmissions to recycle as new arrivals when a certain number of collisions is encountered.

The frame structure of the ARRA has then been adopted for the proposed GRAP, after removing the CMP field, and the distributed queueing policy of the E-GRAP was inspired by the DQRAP protocol. GRAP and E-GRAP are hence less complex than DQRAP as the whole system can keep track of just one queue (the reservation queue), while they exhibit a more complex feature due to the recycling aspect. This can be easily handled by VSAT terminals.

GRAP was found to outperform both DQRAP and ARRA, where the enhanced versions achieve a better performance and stability at very high loads. E-GRAP exhibits the lowest delay (and hence delay variation), which makes it better adapted to data traffic with real-time requirements, and multimedia applications, such as interactive on-demand services over Internet.

REFERENCES

- Lee, H. W. and Mark, J. W. (1983) Combined random/reservation access for packet switched transmission over a satellite with on-board processing: Part I - global beam satellite. *IEEE Trans. on Commun.*, COM-31 (N0. 10):1161-1171.
- Lee, H. W. and Mark, J. W. (1983) Combined random/reservation access for packet switched transmission over a satellite with on-board processing: Part II - multi-beam satellite. *IEEE Trans. on Commun.*, COM-32 (N0. 10):1161-1171.
- Raychaudhuri D. and Joseph, K. (1987) Ku-band satellite data networks using very small aperture terminals-part 1: Multi-access protocols. *Int. Journ. of Sat. Comm.*, 5:195-212.
- Wong, E. W. M. and Yum, T. S. (1991) A controlled multiaccess protocol for packet satellite communications. *IEEE Trans. on Commun.*, vol. 39 No. 7:1133-1140.
- Mohammed, Jahangir I. and Le-Ngoc, Tho (1994) Performance analysis of combined free/demand assignment multiple access CFDMA protocol

- for packet satellite communications. *Proceedings of ICC'94 Conference*, 869–873, New Orleans, 1994.
- Xu, Wenxin and Campbell, Graham (1993) A Distributed Queueing Random Access Protocol for a broadcast channel. *Proceedings SIGCOMM'93*, 270–278, Ithaca, N.Y., USA.
- Koperda, Frank and Lin, Bouchung and Collins, D. Jason (1995) A proposal to use XDQRAP for the IEEE 802.14. *Distribution IEEE 802.14 working group*, July 1995.
- Raychaudhuri, Dipankar (1985) Announced retransmission random access protocol. *IEEE Trans. on Commun.*, COM-33, No. 11:1183–1190.
- Capetanakis, John I. (1979) Tree algorithms for packet broadcast channels. *IEEE Trans. on Information Theory*, Vol. IT-25(No. 5):505–515.