

An adaptive random-reservation MAC protocol to guarantee QoS for ATM over satellite

T. Örs, Z. Sun and B.G. Evans

Centre for Communication Systems Research,

University of Surrey, Guildford, Surrey GU2 5XH, UK

Phone: +44 1483 259844, Fax: +44 1483 259504

e-mail: {T.Ors,Z.Sun,B.Evans}@ee.surrey.ac.uk

Abstract

In this paper we analyse the performance of an Adaptive Random-Reservation Medium Access Control (MAC) protocol which can support all ATM service classes while providing the required Quality of Service (QoS). Our study focuses on parameter optimisation of the multiple access schemes for ATM over a GEO satellite with on-board processing capabilities, considering various traffic mixes of Constant Bit Rate (CBR), real-time Variable Bit Rate (rt VBR), non-real-time VBR (nrt VBR) and Unspecified Bit Rate (UBR). It is shown that maximum throughput can be achieved by using this access scheme. A TDMA access protocol combining both Random Access and Demand Assignment Multiple Access (DAMA) is particularly suited for a scenario with a high number of terminals and very bursty UBR traffic (e.g. web browsing). UBR sources with short burst length access the slots remaining after the reservation procedure by random access which drastically reduces the slot access delay, at the expense of lower utilisation.

It is shown that the potential user population which can be served is considerably increased by statistically multiplexing bursty traffic over the air interface.

Keywords

ATM, DAMA, Random Access, TDMA, Quality of Service, Satellite, Performance

1 INTRODUCTION

In the recent years, significant progress has been made in the research and standardisation of ATM over terrestrial networks. Whilst optical fibre is the preferred carrier for high-bandwidth terrestrial communication services, satellite systems can play an important role in the B-ISDN. The main strengths of satellites are their fast deployment, global coverage and flexible bandwidth-on-demand capabilities.

After the terrestrial broadband infrastructure will have reached some degree of maturity, satellites are expected to provide broadcast service and also cost-effective links to rural areas complementing the terrestrial network. In this phase satellite networks will provide broadband links to a large number of end users through a User Network Interface (UNI) for accessing the ATM B-ISDN. Portable user terminals are expected to have relatively low average and peak bit rates (up to 2Mbit/s) and the traffic is expected to show large fluctuations. Therefore the access scheme will considerably effect the performance of the system. Furthermore the cost and size of the terminal will have a large impact on the suitability of the satellite solution.

Recent proposals for broadband multimedia satellite systems (Fitzpatrick, 1996), (Fernandez, 1997) are examples for this scenario and increased the attention paid to ATM access over satellite links. The methodology and standards developed for wired ATM networks are ported to the wireless environment, where possible. This will allow seamless integration of terrestrial and satellite networks.

The use of standard ATM protocols to support seamless wired and wireless networking is possible by incorporating a new radio specific protocol sublayer into the ATM protocol model (ATM-Forum, 1996) as shown in Figure 1.

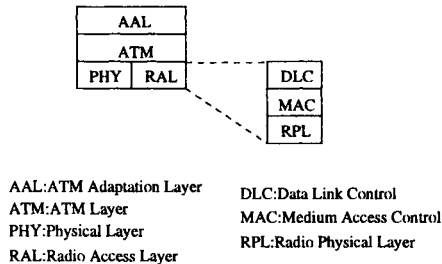


Figure 1 ATM Protocol Model including Radio Access Layer

Considering that satellite communications uses multiple access on a shared medium, a MAC layer, which is not present in traditional ATM networks is needed. The MAC protocol plays a central role as means of accessing the RPL from the ATM layer. The access scheme refers to the physical layer multiplexing technique to a share a common channel among multiple users of possibly multi-services. The problem of statistical multiplexing at the air interface is slightly different to that in the fixed network. In the fixed network the problem is associated with control of bandwidth on an outgoing link from some multiplexing point after buffering has

occurred. It is implicitly assumed that the access links from the source are dimensioned in such a way that they do not impose any constraints on the traffic (e.g. sources can transmit at their peak bit rate). In the air interface the constraint is on the bandwidth available in total to all sources before the buffering/multiplexing point.

2 SATELLITE NETWORK ARCHITECTURE

The satellite network architecture for the provision of ATM services to portable terminals employs a satellite with cell switching capabilities (Ors, 1997). Network requirements are for a full meshed point-to-point and point-to-multipoint system. Suitable satellite architectures for meshed terminal networks are expected to employ a spot beam coverage pattern to achieve the high uplink and downlink gain required for mesh connectivity between portable terminals. On-Board Processing (OBP) functions such as switching, channel set-up and multiplexing result in increased complexity on-board the satellite but offer added flexibility and improvement in link performance. The reasons for using OBP functions in this scenario are:

- To maximise the bandwidth utilisation of the satellite.
- To reduce the reservation delay.
- To improve interconnectivity.
- To reduce the ground terminal RF cost.

The 155 and 622 Mbit/s transmission rates conventionally associated with ATM are well above the maximum rates possible with today's portable terminal technology. However, in practice, most individual users will usually require significantly lower traffic rates, especially if there are only a few data or voice terminals located at a remote location. This large number of users with bursty traffic will need a cost-efficient way to communicate between each other and access the ATM/B-ISDN network.

3 FRAMEWORK FOR PROPOSED MAC PROTOCOL

3.1 Design Objectives

The MAC protocol has to be designed to allow statistical multiplexing of ATM traffic over the air interface, especially in the uplink for the independent and spatially distributed terminals. The following design objectives are taken into consideration:

- Maximise the slot utilisation, especially for bursty traffic.
- Guarantee the QoS requirements for all service classes.
- Maximise frame efficiency by minimising overheads.

The minimisation of overheads is not an easy task, especially for ATM which was designed for channels with very good error characteristics (Bit Error Rates around 10^{-10}). To minimise cell loss over the satellite link, channel coding has to be used to make the transmission more robust. A Logical Link Control (LLC) header to facilitate error recovery mechanisms is optional and not in scope of this study. Finally a satellite specific header with satellite routing and wireless resource management fields is added to form a MAC packet.

3.2 Access Schemes

MAC layer access schemes can be typically categorised into four classes: *Fixed Access*, *Random Access*, *Demand Assignment Multiple Access (DAMA)* and *Adaptive Access*. The first three techniques have evolved to meet the needs of constant high traffic with long duration's, sporadic traffic with short to medium duration's, and sporadic traffic with long duration's, respectively (Bohm, 1991). Finally *adaptive access* is used to meet the needs of multiple media which consists of traffic which has all possible traffic patterns mentioned above. Thus to meet the design objectives an Adaptive Access mechanism seems to be the best choice.

3.3 Mapping of ATM Service Classes onto MAC Service Classes

To simplify the conceptual design of the MAC protocol, ATM service classes (ATM-Forum, 1996) can be mapped onto MAC service classes.

Fixed-Rate DAMA is ideal for connections with a constant bit rate such as the CBR service class in ATM networks. Before a connection is set-up, the terminal and satellite negotiate the Quality of Service (QoS) parameters. These QoS parameters determine the characteristics of the connection. Since the parameters will not be modified during the connection, the amount of bandwidth allocated for that connection will not be changed until the connection is terminated. For ATM CBR connections the Peak Cell Rate (PCR) is allocated to the terminal.

Real-time Variable Bit Rate (rt-VBR) services can also be supported with fixed-rate DAMA. For real-time services, the amount of bandwidth assigned to the connection should be close or equal to the PCR to avoid cell delay. The major drawback of this scheme is that a major portion of the bandwidth is wasted when the cell transfer rate is lower than the assigned bandwidth. The major difficulty to employ variable-rate DAMA in ATM over GEO satellite systems is the effect of the large propagation delay (135ms). The computing and negotiation process between the satellite and the terminal may be too long for real-time VBR services and result in unacceptable QoS. The use of variable-rate DAMA for rt-VBR is only possible if the arriving traffic can be predicted one hop delay in advance. Since this is not possible except in some special cases fixed-rate DAMA will be used for rt-VBR.

A scenario where fixed-rate DAMA is efficient for rt-VBR services is when the terminal can multiplex traffic from multiple services. In this case the aggregate

traffic can be approximated as a constant cell flow by using a small amount of shaping.

VBR services which are not time sensitive can use burst reservation. In this case cells can be buffered in the terminal till the required bandwidth is reserved and in case the queue exceeds a certain threshold more bandwidth can be requested. Thus using variable-rate DAMA the bandwidth of a connection can be adjusted according to the change of the data transfer rate.

No numerical commitments are made for the UBR service class and this service category is intended for non-real time applications. UBR services could be supported by variable-rate DAMA. However the fact that this service class has the lowest priority (because no commitments to cell loss or cell delay are made) has to be considered. We propose that UBR could transmit data directly to the unoccupied MAC slots without reservation. The unreserved slots are broadcasted on the downlink to be accessed by random access. This is particularly appealing for bursty interactive services with short duration, for which the long slot reservation delay is unacceptable.

4 THE RANDOM-RESERVATION ADAPTIVE ASSIGNMENT PROTOCOL

The TDMA frame of the adaptive assignment protocol is divided into Request slots, Control slots, Reservation slots and Random Access slots, as shown in Figure 2. The protocol is based on the proposals by (Bohm, 1993), (Celandroni, 1991), (Zein, 1991) with modifications to achieve the design objectives for multi-service networks. Prioritised queuing of requests on-board the satellite is the most important difference from previous proposals.

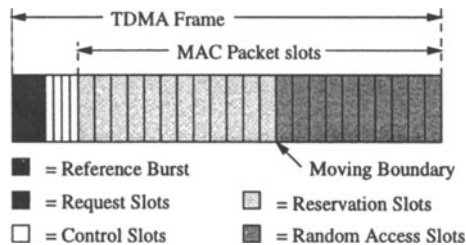


Figure 2 Uplink Frame Structure

A Request slot is that period of time in which terminals report their requests to the on-board Radio Resource Management (RRM). There are only a few Request slots available and a terminal selects one at random without knowing whether another station is using the same slot. If more than one terminal selects the same request slot, a collision occurs and terminals have to retransmit after waiting for a mean retransmit waiting time determined by the collision resolution algorithm. The MAC

protocol ensures that the collision probability stays low. The reason for using request slots is because ATM networks support different services which have different loss and delay requirements.

If a single request was received for a request slot (successful request), the on-board RRM module tries to allocate the necessary Reservation slots. If no Reservation slots are available, the request can either be blocked (called blocking probability) or queued. We propose to queue successful requests in a prioritised queue so that the terminal does not need to compete with other terminals for a request slot again. By queuing successful requests, Reservation slots can be allocated according to their priorities.

Once the Reservation slots are reserved (successful reservation), an acknowledgement is transmitted to the terminal in Time Division Multiplex (TDM) mode on the downlink frame. Reservation slots represent the part of the frame in which a terminal can transmit its message after a successful reservation. In every frame there are many reservation slots and the on-board RRM module will assign reservation slots to a particular successful request according to its priority. A reservation slot is assigned to only one terminal and therefore there is no possibility of collision.

On the other hand Random Access (RA) slots represent the part of the frame in which terminals can transmit without the need of making a reservation. The slots available for RA are broadcasted in the downlink frame. This part is for services which don't want to wait for the lengthy reservation procedure. In RA mode it is not possible to guarantee a certain QoS to users although the protocol will try to minimise the number of collisions to maximise throughput by using an adaptive collision resolution algorithm. RA should only be used by UBR sources with relatively small burst length since RA traffic is allowed to reserve slots and have to contend for each MAC packet. An improvement in performance for RA is possible by using the UNR+ service class, where a minimum bandwidth is guaranteed for UBR services (Goya, 1996)

Unless the number of request slots per frame is carefully adjusted the result would be either low capacity utilisation and long delays (too many request slots, less capacity available for information transmission) or network backlog (too few request slots resulting in successive collisions and high delay). The number of request slots should be fixed for system behaviour where the number of collisions can be controlled by broadcasting a message in the downlink that services with lower priority should not send/resend requests till the collisions have been resolved. Our analysis has shown that two request slots provide adequate performance for an average call holding time of longer than six seconds. However when the number of collisions can't be controlled (very short average call holding time) new request slots can be added by reducing the number of control slots.

The requests for dynamic slot allocation are done using the Control Slots which are assigned, on a round-robin basis to all terminals which request the variable-rate DAMA MAC class. The number of control slots is set to eight to minimise the frame overhead.

The satellite frame introduces a constant delay equal to the frame length, on the cells of a stream connection. Therefore the selection of the frame size should be small enough to satisfy the delay limit of real-time services (400ms) (ITU-T,1996) taking into account the satellite propagation and processing delays and the delay introduced by the terrestrial B-ISDN.

The MAC packet slot period has been chosen to support a 32 kbit/s CBR stream and corresponds to one frame unit of 384 un-coded information bits every uplink frame. This results in a frame period of 11.9 ms to transmit 84 ATM cells per second using AAL5. There are 64 MAC packet slots to support 2.048 Mbit/s of traffic per spot-beam on the uplink. The actual uplink transmission rate is higher due to ATM and MAC layer overheads.

5 ANALYSIS OF ADAPTIVE RANDOM-RESERVATION MAC PROTOCOLS

5.1 CBR and nrt-VBR Analysis

By queuing requests when no reservation slots are available (buffer size=100) and by using an adaptive collision resolution algorithm the blocking and collision probabilities can be minimised for a traffic mix of CBR and nrt-VBR. Thus this section will mainly focus on the performance of the system in terms of throughput and reservation delay. One of the main evaluating factor of various access protocols is the normalised information throughput defined as:

$$\frac{\text{throughput}}{\text{capacity}} = \frac{\text{the number of occupied MAC slots in a frame}}{\text{the number of MAC slots in a frame}}$$

Nrt-VBR connections first have to reserve the initial capacity by using a Request slot and can then use a Control Slot assigned to them in round-robin fashion for other requests. Thus there are two different reservation delay values. The delay for the initial capacity request using the request slot is called 'Call Reservation Delay' (or only Reservation Delay) and the delay for the dynamic bandwidth reservations using the control slot is called 'Burst Reservation Delay'. The slot reservation is relinquished during the silence periods of the source. In this way the air interface bandwidth is shared between multiple sources achieving statistical multiplexing on the air interface.

The offered traffic load of the system can be calculated by multiplying the average number of calls originating per unit time (λ) with the mean call holding time (h). Furthermore a source may occupy more than one slot according to its PCR which has to be taken into account:

$$\text{Offered traffic load (Erlang)} = (\lambda \cdot h \cdot \text{PCR}) / (32 \cdot 10^3) \quad (1)$$

For VBR source the call holding time has to be divided by the burstiness (β) to calculate the offered load. The average call holding time for all services is assumed to be one minute.

The normalised load is defined as:

$$\text{Normalised Load} = \text{Offered Traffic Load} / \text{Number of MAC slots} \tag{2}$$

Each nrt-VBR source represented by an on-off source model shown in Figure 3 where a^{-1} is the mean burst period and b^{-1} is the mean silence period which are both exponentially distributed. The burstiness is defined as:

$$\beta = \frac{\text{Peak Rate}}{\text{Mean Rate}} = \frac{a^{-1} + b^{-1}}{a^{-1}} \tag{3}$$

The VBR mean burst period is 100ms and the PCR during the burst is 64 kbit/s.

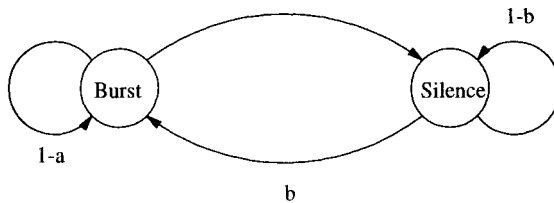


Figure 3 The on-off source model

First the results for only nrt-VBR are presented to show the advantage of multiplexing bursty traffic over the air interface. Simulations are carried out for traffic with burstiness of 5,10 and 20. Note that for all the simulations the number of active terminals is equal to the call arrival rate. The confidence interval for all simulations is 95% and not shown on the graphs for the neatness of the results.

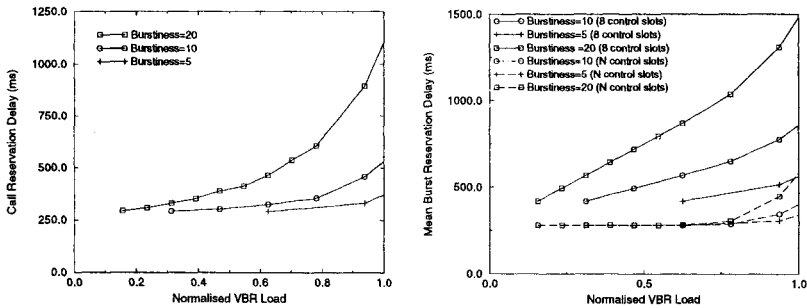


Figure 4 (a) Call Reservation (b) Burst Reservation Delay for nrt-VBR services with various burstiness values.

The simulation results for this scenario show that a throughput of up to 0.95 can be achieved. The reservation delay remains reasonably low, for low burstiness values. As the burstiness is increased so does the mean reservation delay (Figure 4) as more sources need to be multiplexed on the air interface to achieve high utilisation.

Choosing a low number of control slots to minimise frame overheads results in increased burst reservation delay as can be seen in Figure 4 (b). The increase in burst reservation delay become more visible as the number of active terminals

increases. If each source is assigned an individual control slot then there will be large frame overhead for N sources while achieving very low burst reservation delays. Note that N (number of terminals) is equal to the call arrival rate which can be calculated from equations (1)-(3). Changing the burstiness for a fixed load changes the mean rate and hence the number of terminals. Since the terminal is expected to buffer bursts at the PCR for the burst reservation delay period, a reduced burst reservation delay results in smaller buffer requirement for the terminal, to avoid cell loss due to buffer overflow.

Next the effects of CBR traffic on nrt-VBR traffic is investigated by fixing the nrt-VBR load at 0.6 and varying the CBR load. The required CBR PCR is assumed 32kbit/s. The increase in the call and burst reservation delay as a function of CBR load is shown in Figure 5. The increased delay is due to the fact that as the total offered load is increased, it exceeds the system capacity and hence request have to be buffered which results in longer delays till bandwidth allocation can be made. Again the burst reservation delay is minimised by allocating a control slot for each VBR source as shown in Figure 5 (b). For eight control slots the burst reservation delay increases since control slots are assigned in round-robin fashion to terminals which can only request bandwidth when the control slot is assigned to them.

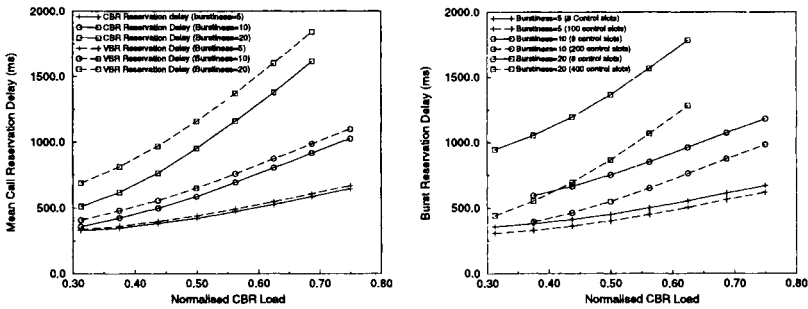


Figure 5 (a) Mean Call Reservation Delay (b) Mean Burst Reservation Delay

The achieved throughput is similar for various burstiness values as shown in Figure 6, due to the low number of collisions and blockings.

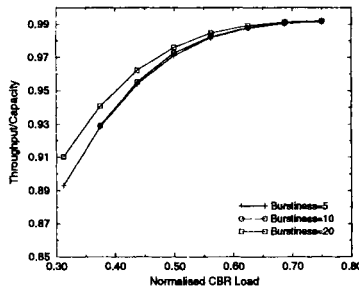


Figure 6 Total System Throughput vs Normalised CBR Load

5.2 CBR, VBR and UBR Analysis

The novelty of the adaptive MAC protocol is its ability to support lower priority UBR traffic in both reservation and random access mode according to the terminals burst duration.

The UBR load is calculated in the same way as for VBR traffic and the throughput of UBR for a pure-reservation system, (simulation parameters shown in Table-1) in a scenario with a high number of UBR sources with very bursty traffic ($\beta=200-5000$) is shown in Figure 7(a). The required PCR for UBR sources is assumed 32 kbit/s and the UBR burst period is varied from 13ms (1 cell) to 416 ms. As it can be seen the amount of carried traffic remains unacceptably low, even for low to medium load due to the high number of collisions, in particular for high burstiness values where the burst period is short and a high number of terminals are admitted to achieve high utilisation. The pure-reservation MAC throughput increases for traffic with longer burst duration (or lower burstiness). Only for burstiness values lower than 200 can a throughput higher than 25% of the frame capacity be achieved.

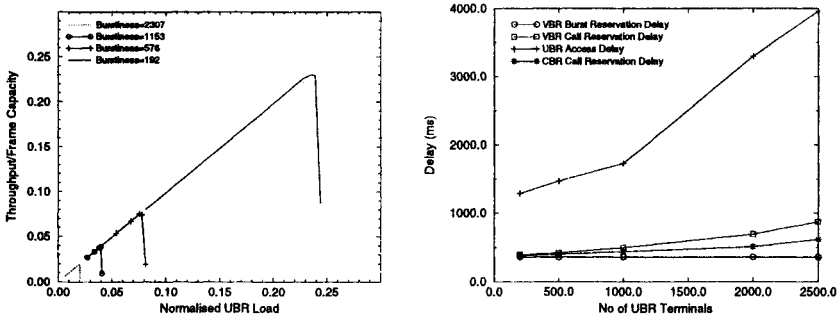


Figure 7 (a) Throughput vs normalised UBR load (b) Reservation Delay vs Arrival Rate (per minute)

Since the offered UBR traffic load is dependent on the burstiness of the traffic Figure 7 (b) shows the delay as a function of the UBR terminal numbers (which is also the number of arrivals per minute). As it can be seen, UBR traffic is increasing the call reservation delay of higher priority CBR and VBR. Since all the terminals content for the same reservation slots, UBR traffic increases the collision probability and the access delay.

To improve the UBR throughput and the access delay we propose that UBR sources access the MAC slots remaining after the reservation procedure by Random Access (RA). This way the lengthy reservation procedure is avoided and the number of collisions reduced. The reservation delay of terminals using reservation (CBR and VBR terminals) is unaffected by RA terminals.

The UBR throughput using RA is limited by 36.8 % of the available RA capacity (theoretical Slotted Aloha limit). This throughput is higher than the pure-

reservation throughput for very bursty traffic ($\beta=500-5000$). Our simulations showed that in order to minimise RA collisions and access delays, the UBR load has to be kept around 25% of the available RA capacity. The throughput of UBR vs access delay is shown in Figure 8(a). Note that the normalised UBR load is found by dividing the offered load by the number of available RA slots. Since ATM is connection-oriented, the system can refuse new UBR connections to keep the UBR load at acceptable levels. However as the number of available RA slots changes every frame there may be instants of decreased throughput till the collision resolution algorithm optimises system performance by reducing the RA collisions. Figure 8(b) shows the increased collision probability vs UBR load when the number of available RA slots is reduced by 1/3 on the next frame.

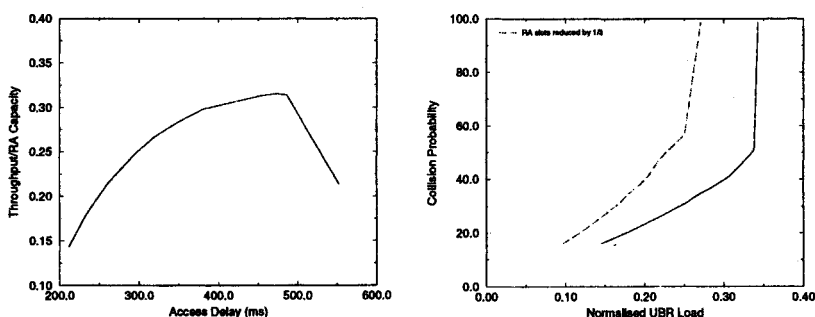


Figure 8 (a) Throughput vs Access Delay (b) Collision Probability vs Normalised UBR Load

The maximum number of UBR terminals which can be supported can be found as:

$$\text{Maximum number of UBR sources} = \text{Normalised Load} \cdot \beta \cdot \text{number of RA slots}$$

For a normalised UBR load of 0.25 a burstiness of 4500 and 10 RA slots, 11250 UBR terminals can be supported. For pure-reservation this number was around 3000 because of the increase in collisions.

6 CONCLUSIONS

As user demands become more complex, satellite networks are expected to support a much wider range of services. As satellites will play an important role in the deployment of ATM networks, we addressed the optimisation of the capacity allocation scheme, using performance results for an adaptive MAC scheme.

It has been shown that considerable improvements in delay performance and satellite bandwidth utilisation are possible if next-generation satellite technology (OBP) and an adaptive MAC is used. The traditional demand-assignment scheme using a ground terminal as control station has two major drawbacks: long set-up and reservation time and limited channel utility. Both are due to the long

propagation delay of the satellite link. Both disadvantages can be improved by processing channel requests in the satellite to allocate frame slots.

The mapping of ATM service classes to MAC classes and the use of a prioritised request queue provides the QoS differentiation required by ATM networks. Furthermore the advantages of a Random-Reservation MAC scheme for a scenario with a high number of very bursty users was illustrated. The user population which can be supported was shown to be much higher than with a pure-reservation scheme while also achieving higher throughput and lower access delay for low to medium UBR traffic load.

7 REFERENCES

- ATM-Forum (1996) *Proposed Charter, Work Plan and Schedule for a Wireless ATM Working Group*, 96-0712/PLEN.
- ATM Forum (1996) *Traffic Management Specification Version 4*, af-tm-0056, April.
- Bohm, S. Elhakeem, A.K. and Murthy, V.K.M. (1993) Analysis of a movable boundary random/DAMA accessing technique for future integrated services satellites, pages 1283-1289, IEEE Globecom
- Celandroni, N. and Ferro, E. (1991) The FODA-TDMA Satellite Access Scheme: Presentation, Study of the System and Results., *IEEE Trans. Comms.*, COM-39(12), pages 1823-1831.
- Fernandez, R. (1997) The Ka-band Quest Continues, Via Satellite, March.
- Fitzpatrick, E.J. (1996) Hughes Spaceway: Wireless Interactive Broadband Service, Second Ka-Band Utilization Conference, September 24-26.
- Guerin R. and Heinanen J. (1996), UBR+ Service Category Definition, ATM-Forum 96-1598
- ITU-T (1996) *B-ISDN ATM Layer Cell Transfer Performance*, Draft 4R Rec.I.356.
- Ors, T. Sun, Z. and Evans, B.G. (1997) A Meshed VSAT Satellite Network Architecture using an On-Board ATM Switch, IEEE IPCCC, pages 208-215.
- Zein, T. Maral, G., Tondriaux, M. and Seret, D. (1991) A Dynamic Allocation Protocol for a Satellite Network Integrated with B-ISDN., *Proc. 2nd ECSC*, pages 15-20.