

ATM over VSAT networks using traditional bent-pipe or on-board processing satellites

T. Örs, Z. Sun and B.G. Evans

Centre for Communication Systems Research,

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

Phone: +44 (0)1483 259844, Fax: +44 (0)1483 259504

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

M.H. Hadjitheodosiou

Center for Satellite & Hybrid Communication Networks

University of Maryland, College Park, MD 20742, USA

Phone: +301 405-7904, Fax: +301 314-8586

e-mail: michalis@isr.umd.edu

Abstract

In this paper two VSAT satellite network architectures are proposed to offer access to the B-ISDN and accommodate ATM service needs in an integrated manner, while achieving maximum bandwidth utilisation of the satellite. The first architecture uses a traditional 'bent pipe' repeater satellite. The need for optimization of the channel capacity allocation scheme is discussed, using a preliminary performance analysis for a Dynamic Reservation TDMA system which takes advantage of the flexibility and the statistical multiplexing capabilities of ATM and supports various VBR and CBR services. Some bottlenecks, introduced by the satellite link, in the performance of protocols such as TCP/IP are highlighted. The second architecture uses an OBP satellite with spot beams and cell-switching capabilities. A novel approach, in which statistical multiplexing is performed on-board the satellite, is suggested and results show that a much better satellite bandwidth utilization can be achieved. In this architecture, smaller earth terminals can be used without complex dynamic reservation protocols, but at the cost of a more complex satellite payload.

Keywords

VSAT, ATM, on-board processing, TDMA, TCP/IP

1 INTRODUCTION

Recent developments in multimedia services have created a demand for broadband networks (such as B-ISDN), able to handle diverse traffic characteristics and Quality of Service (QoS) requirements.

Whilst optical fibre is the preferred carrier for high-bandwidth 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 very flexible bandwidth-on-demand capabilities. As new satellites are becoming able to provide broadband services the satellite network configuration and capacity can be gradually increased to match the B-ISDN traffic evolution at each time.

The role of satellites in high-speed networking will evolve according to the evolution of the terrestrial ATM-based B-ISDN. However two main roles for satellites can be identified:

- In the introductory phase, satellites will compensate for the lack of sufficient terrestrial high-bit-rate links mainly by interconnecting a few regional or national distributed broadband networks, usually called 'Broadband Islands'.
- In the mature phase, 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. Satellites are also ideal for interconnecting mobile sites and provide a back-up solution in case of failure of the terrestrial systems.

In the first scenario, satellite links provide high bit rate links between broadband nodes or broadband islands. The interfaces with satellite links in this mode are of the Network-Node Interface (NNI) type. This scenario is characterized by a relatively small number of large earth stations which have a relatively large average bit rate. Since only a small number of larger earth stations is required the cost of the earth station has not a big impact on the suitability of the satellite solution.

The RACE CATALYST project was a demonstrator for this scenario and showed the compatibility of satellite technology with ATM and the terrestrial B-ISDN. The equipment developed during the CATALYST project has been able to interconnect ATM testbeds as well as existing networks such as DQDB, FDDI and Ethernet networks, all using ATM. A detailed explanation of the system design

and performance is provided in (Polese, 1994), (Louvet, 1994), (Hadjitheodosiou, 1994), (Sun, 1996).

In the second scenario the satellite system is located at the border of the B-ISDN and provides access links to a large number of users. This scenario is characterised by a large number of earth stations with relatively low average and peak bit rates (up to 2Mbit/s). The traffic at the earth station is expected to show large fluctuations. Therefore the multiple access scheme will considerably effect the performance of the system. Furthermore the cost and size of the earth station have a large impact on the suitability of the satellite solution.

Very Small Aperture Terminal (VSAT) satellite systems could be used for the second scenario. The main objective is to interconnect a large number of users and make B-ISDN access affordable by lowering earth station cost and providing bandwidth on demand.

2 SOURCE CHARACTERISATION

Source characterisation at the macro level is defining the source traffic characteristics and its QoS requirements. The traffic characteristics of a service are the minimum set of parameters that a user can be expected to declare, to provide the network with as much information as possible to effectively control network traffic and achieve high utilization. The service categories are described in (ATM Forum, 1995).

2.1 Traffic Descriptors and Parameters

The Connection Traffic Descriptors (CTD) play an important role in the preventive control scheme. The Connection Admission Control (CAC) has to consider the CTD in order to allocate the necessary network resources for the connection.

An important issue is the set of traffic parameters to be included in the CTD. The Peak Cell Rate (PCR), Cell Delay Variation Tolerance (CDVT), sustainable cell rate and burst tolerance have been standardized (ITU-T, 1996). Some other widely used traffic parameters have also been proposed: mean rate and mean burst length (Vakil, 1991). In this paper we will use the peak rate, CDVT, mean cell rate and mean burst duration as traffic parameters.

2.2 Source Model

If N identical packet-switched speech connections are multiplexed, (Weinstein, 1978) found that the number of active voice sources i can be modelled very well by a continuous-time birth-death (b-d) process, shown in Figure 1. The inter-arrival time of cells in state i is $1/(i \cdot PCR)$. The parameters used by the model governing the transition rates are: mean burst duration = a^{-1} and mean silence

duration = b^{-1} , which are both exponentially distributed. We showed that this model can be used to represent any on-off type multiplexed homogeneous traffic.

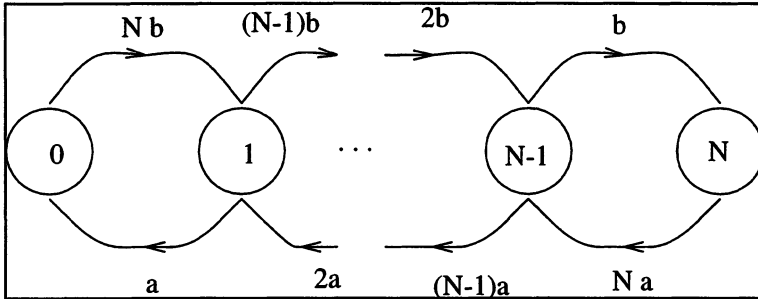


Figure 1 Birth-Death Process Model.

3 SATELLITE NETWORK ARCHITECTURES

Two satellite network architectures for the support of ATM services will be described. The first architecture uses existing bent pipe repeater satellites and the novelty of the architecture is the use of a dynamic TDMA reservation system for efficient use of the available satellite bandwidth. The second network architecture employs a satellite with cell switching capabilities. The novelty of this architecture is its ability to multiplex and switch all the traffic on-board the satellite avoiding complex reservation protocols and reducing the earth-station complexity and cost. There are currently no satellites in operation with this capability, however this architecture could be implemented when more advanced OBP satellites will be operational.

The 155 and 622 Mbit/s transmission rates conventionally associated with ATM are well above the maximum rates possible with today's VSAT 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 occasionally access the B-ISDN/ATM network. We discuss how VSAT satellite networks could be adapted for these services and investigate some of the problems that need to be resolved.

4 VSAT NETWORKING USING BENT-PIPE SATELLITES

The network architecture using a bent-pipe satellite is shown in Figure 2. It is a mesh network, in which the hub acts as a network management centre for channel allocation and policing, but VSATs that have established a connection talk directly

to each other via the satellite. In this case the propagation delay is a single satellite hop.

We will not discuss the system implementation or protocol layer development for this architecture, but will instead focus on a teletraffic analysis and parameter optimization of the multiple access scheme used.

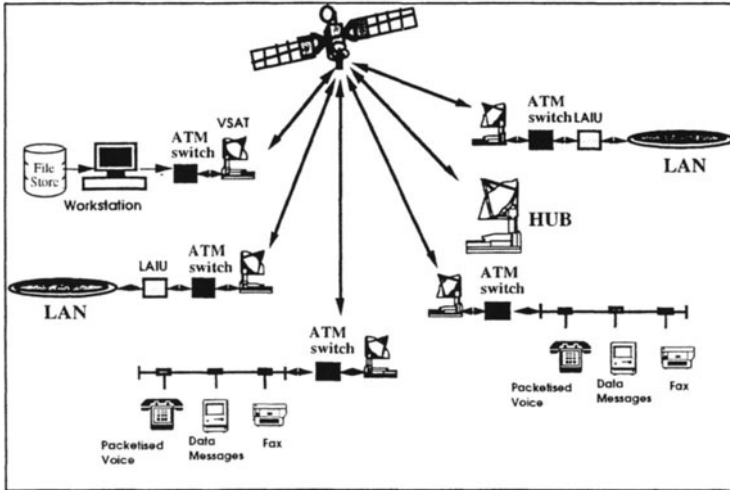


Figure 2 VSAT Mesh Network.

4.1 Satellite Multiple Access

Although many random access/reservation protocols have been developed for VSAT networks carrying bursty data traffic (Raychaudhuri, 1989) conventional contention protocols (e.g. ALOHA-type) cannot be used in this scenario because they cannot give any service guarantees to a user. It could be theoretically possible to use extensions of these for asynchronous traffic (with no guarantees) (Schoute, 1994) but they would not be suitable for mixing various kinds of traffic, in particular delay-sensitive applications.

For CBR and VBR ATM traffic, reservation protocols (e.g. FDMA, TDMA) with static (pre-assigned) bandwidth reservations based on estimated fixed (for CBR traffic) or peak (for VBR traffic) rates could be used, but this would be wasteful and actually defeat the 'bandwidth on demand' advantage: There is therefore a need for a scheme that allows them to share a channel, takes advantage of the 'bandwidth on demand' property of ATM while at the same time provides a compatible interface to the ATM technology.

A pre-assigned scheme could be used to interconnect a small number of stations with well known traffic characteristics (such as the CATALYST network) but not for a large number of users. Dynamic-reservation systems appear to be a better, but more complex solution, but there are still difficulties. A dynamic-reservation Time Division Multiple Access (TDMA) scheme appears to be the best choice, since it

allows the satellite transponder to operate at maximum efficiency. Also, the flexibility of TDMA allows ATM cells to be mapped into the variable length TDMA sub-bursts as the basic unit (slot) of transmission. Satellite systems accommodating both synchronous and asynchronous traffic using TDMA-based techniques have been proposed (Celandroni, 1991), (Zein, 1991). An efficient use of the capacity of the TDMA system can be achieved if we provide a combination of pre-assigned and on-demand capacity in each link, matching the broad mix of VBR and CBR traffic we expect to have.

4.2 A Dynamic TDMA Reservation System

We next focus on a TDMA reservation system consisting of a number of VSATs, a central reservation/control unit (Hub), and a single channel, shared by all active VSATs. The Hub co-ordinates the channel allocation. In a typical VSAT system there could be a number of sources (data terminals, telephones, fax, video) connected to a VSAT via an ATM switch, or the VSAT could act as a gateway for a LAN, in which case a specially developed LAN-ATM Interface Unit (LAIU) is needed to perform the required protocol layer translation (Fig. 2).

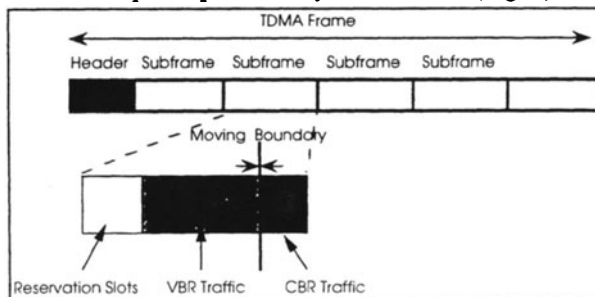


Figure 3 TDMA Frame Structure.

A VSAT with a message to transmit, sends its request with the message size to the hub, which will then let the terminal know when to start the message transmission. The TDMA frame (Figure 3) is divided into Reservation slots and Data slots. A Reservation slot is that period of time in which terminals report their requests to the reservation unit. There are only a few Reservation slots available and a VSAT selects one at random without knowing whether another station is using the same slot. There is a possibility of collision, and if, after a time-out, the VSAT does not receive an acknowledgement of its request, it waits a random time and re-transmits it. The acknowledgement from the Hub to VSATs is transmitted in TDM mode over a separate, faster channel.

Data slots represent the part of the frame in which a terminal can transmit its message. In every frame there are many data slots and the Hub will assign enough sequential data slots to a particular successful request. A data slot is assigned to at most one terminal and therefore there is no possibility of collision. The available

data slots can be reserved for the duration of a CBR connection or allocated on a burst-by-burst basis to VBR services. We can see that the frame partition (i.e. number of slots allocated to reservation, CBR and VBR traffic) needs to be adjusted to the traffic load and mix we have, and clearly an adaptive allocation using a moving boundary would represent a more efficient (but much more complex to implement!) use of the capacity, if the traffic load changes frequently with time. Unless the number of reservation slots/frame is carefully adjusted the result would be either low capacity utilization and long delays (too many reservation slots, less capacity available for information transmission) or network backlog (too few reservation slots resulting in successive collisions and high delay).

A possible implementation scenario would be to allocate a portion of the capacity to each VSAT according to its mean traffic rate (derived for the load of all the sources connected to that station) and its QoS requirements, and allow it to request additional capacity using the reservation slots or by piggy-backing the request with the data sent in the data slots. In this way, a service with strict low delay/jitter requirements (e.g. voice or video) could use the allocated capacity, while a less demanding service (e.g. data transmission) could take advantage of the capacity reservation, thus maximizing the efficient use of the channel. The actual procedure might be very delicate and details need to be investigated in greater length for a realistic implementation. Each VSAT must monitor its traffic arrival rate and the network load (e.g. from the time it takes to establish a reservation, number of successive collisions) and be able to block new users if there is a danger of exceeding a certain threshold.

4.3 CBR and VBR Throughput Performance Analysis

For the analysis we consider a network consisting of 10 VSATs, with a total channel capacity of 2 Mbit/s. Unless otherwise stated we assume an error-free channel and a mesh configuration, where each satellite hop takes approx. 0.27 sec. Part of the capacity can be assigned to each VSAT based on expected load, while the remaining capacity could be left for dynamic reservations. For example, if we expect to have a mean overall traffic rate of 0.1 Mbit/s per VSAT and there are 10 active VSATs (with similar traffic loads), we can allocate 50% of the channel capacity (1Mbit/s) for this by dividing it equally amongst the stations. The remaining 50% is available for dynamic reservations every time the load at a particular VSAT exceeds the allocated capacity.

Figure 4 (a) and (b) show the net throughput vs traffic load for CBR and VBR traffic and various capacity allocations, assuming the total channel capacity is available (no fixed *a priori* assignments). For example a 75% capacity allocation in Fig. 4 (a) would mean that 75% of the channel capacity is allocated to CBR traffic and 25% available for VBR services. A burstiness (peak/mean traffic rate) of 200 is assumed in the VBR case. We can see from these results that in both

cases the channel capacity appears to be a limiting factor (graphs reach a knee after which throughput remains constant), but the effects are more profound in the case of VBR traffic, because of the delay in establishing successive reservations. However, for sufficient channel capacity (100% allocation for either CBR or VBR), the maximum achieved throughput is higher in the VBR case, because statistical multiplexing enables a better channel utilization. Clearly, a capacity allocation based on the expected mix of VBR/CBR services would offer the best performance.

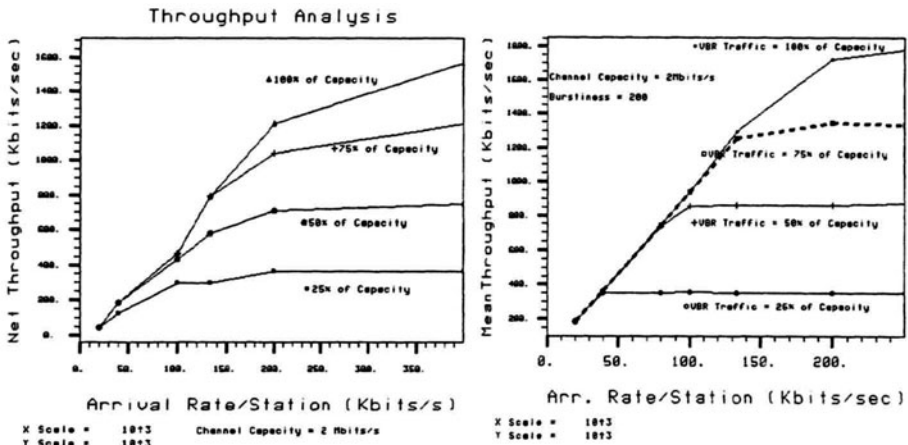


Figure 4 (a) Throughput CBR Traffic (b) Throughput VBR Traffic.

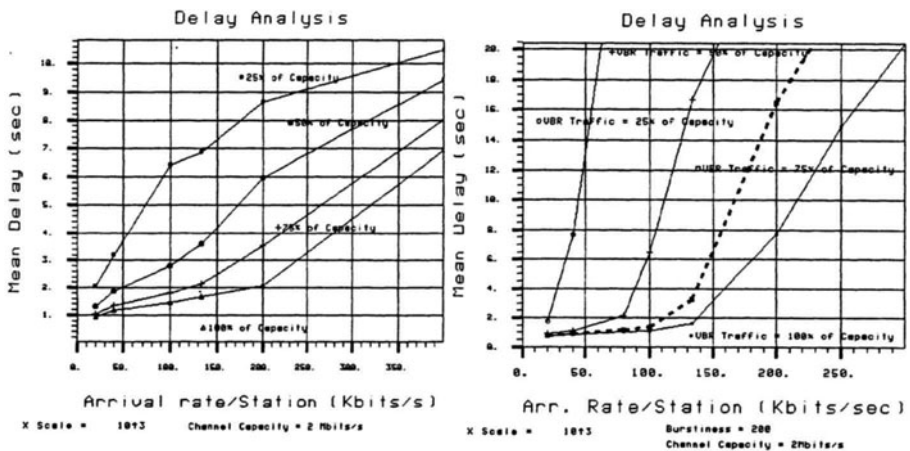


Figure 5 (a) Delay CBR Traffic (b) Delay VBR Traffic.

Combining Fig. 4 (a) and (b) with a delay performance analysis Fig.5 (a) and (b) for specific services could determine the capacity partition required to guarantee a particular QoS, based on the traffic load.

4.4 TCP/IP Throughput Performance

An essential service a VSAT system offering LAN interconnection will be required to support is file transfer, and TCP/IP is one of the most common and robust protocols for this application. To see the effect of the satellite delay and channel errors on the protocols performance a simplified implementation of a VSAT system consisting of 10 stations and using the Reservation TDMA scheme described earlier is simulated, for both star and mesh configurations. Figure 6 shows the throughput as a function of the channel capacity for both configurations, for the error-free channel case.

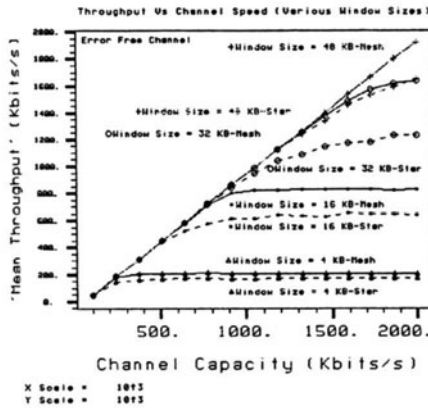


Figure 6 TCP/IP Throughput vs Capacity for Error Free Channel.

It is obvious that the window size of the TCP transport protocol is a limiting factor. With a window of 48 Kbyte, which is the maximum window the SunOS implementation of TCP allows, the maximum achievable throughput is 1.9 Mbit/s (which corresponds to 1.5 Mbit/s net throughput) for the mesh case and 1.6 Mbit/s (1.25 Mbit/s net) for a star system, at the maximum channel speed of 2 Mbit/s. For smaller windows we can see that increasing the channel speed does not increase the throughput and leads to low capacity utilization. Also, for smaller windows we can see an almost linear relationship between capacity and throughput until the saturation point is reached. A star configuration (with double-hop delays) has a much lower throughput than a mesh configuration, especially as the window limitations become less significant, so there is a clear advantage if a mesh architecture is used, as we would expect.

It is important to note that a mesh configuration would need very hard recovery algorithms for successful operation, especially in the presence of channel errors. This could have a heavy impact on the system efficiency and offset some of the advantage for mesh systems deduced from Figure 6.

For a system with such a high propagation delay, the channel Bit Error Rate (BER) would have a significant effect on the systems throughput, especially in the

case of larger window sizes. The effects of burst errors in a satellite environment can be prevented by interleaving at the ground terminal (Lunsford, 1995) but at the cost of introducing buffering delay.

5 VSAT NETWORKING USING OBP SATELLITES

In this architecture there is no need for a central reservation/control unit on ground (Hub). Network requirements are for a full meshed point-to-point and point-to-multipoint system. Suitable satellite architectures for meshed VSAT networks are expected to employ a spot beam coverage pattern to achieve the high uplink and downlink gain required for mesh connectivity between VSATs. 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 maximize the bandwidth utilization of the satellite.
- To improve interconnectivity.
- To reduce the ground terminal RF cost.

5.1 Satellite Multiple Access

The need and consequent emphasis on low-cost ground terminals in this architecture eliminates a full-band TDMA. It would require extremely accurate timing at each terminal because of the large bandwidth, if a large number of users are to be supported. Full-band TDMA also requires a very high burst rate which demands high transmit power from each VSAT. These characteristics translate directly into expensive equipment. In MF/TDMA a group of VSATs in the same uplink beam with low traffic requirements can share the use of a low bit rate (e.g. 2 Mbit/s) TDMA carrier. VSAT users with higher throughput can use full 2Mbit/s TDM transmission. The frame efficiency is modest because of the need to provide burst preamble fields for bit- and frame-time recovery as well as for carrier recovery. An exact frame efficiency can be calculated only when the lengths of these required fields are known. In this study, we assume an information frame efficiency of 95%.

Uplink transmission uses a 24 ms MF-TDMA frame with 51 time-slots as shown in Figure 7. Each carrier can support maximum 50 terminals. Within a time slot the data is transmitted in portions of 111 bytes. One such portion is called a Frame Unit (FU). Each FU contains two ATM cells (96 bytes of information payload), resulting in a carrier data capacity of 32kbit/s. The choice of two ATM cells (and the frame size) was a trade-off between delay and link efficiency. 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 specified by the telephone service (400ms) taking

into account the satellite propagation delay and the delay introduced by the terrestrial B-ISDN. For four ATM cells per FU the frame size is nearly 60ms. On the other hand, as the frame unit size is reduced the link efficiency is reduced because the FU header becomes increasingly significant compared to the FU size.

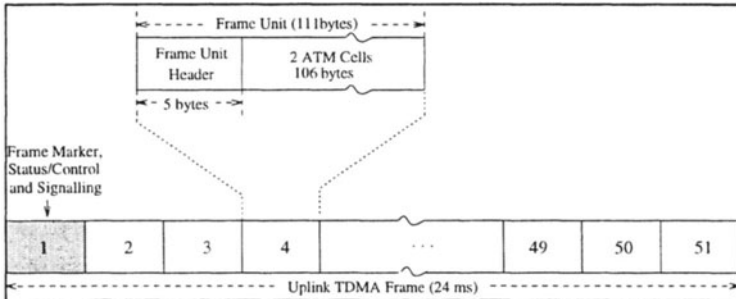


Figure 7 Uplink Frame Structure.

The uplink transmission rate per terminal is 37 kbit/s due to the FU overhead. Downlink transmission uses a 24 ms TDM frame with a transmission rate of 16 Mbit/s. The system used for performance evaluation was a scaled down design with 16 carriers and 16 beams. Note that cell switching resulted in a low data transmission frame efficiency of 81.6% (51 slots of 32 kbit/s on a 2Mbit/s carrier). However, as will be shown later the actual throughput will be higher.

5.2 Ground Segment

The functions of the VSAT earth station are shown in Figure 8. The User Network Interface (UNI) constitutes the boundary between the terrestrial system and the satellite network. The arriving cells are controlled by the Usage Parameter Control (UPC) for their conformance with the traffic contract. Only the peak rate is controlled for CBR and ABR traffic. However for VBR sources also the mean rate is controlled using a dual Leaky Bucket (LB) configuration. The UPC parameters for the ABR type traffic are adaptive (dynamic LB) and changed according to the explicit rate feedback signal from the satellite. The cell scheduling module is responsible for scheduling CBR, VBR, ABR and UBR traffic. CBR and VBR connections have priority and maximum burst utilization is achieved by detecting the silence periods of VBR voice connections. ABR traffic has higher priority than UBR. The cell processor also performs a number of functions. First the processing function sorts the cells by VP and assigns new VPIs. Next, the ATM cells are interleaved in blocks of two. Each packet called Frame Unit (FU) and shown in Figure-7 contains two cells destined for the same downlink beam. This simplifies the on-board processing. Then VPs are mapped into satellite packet addresses for on-board routing and FU headers are assigned. The packets are then scrambled before modulation for transmission. The procedure for the destination VSAT earth station primarily consists of the procedure above in reverse order.

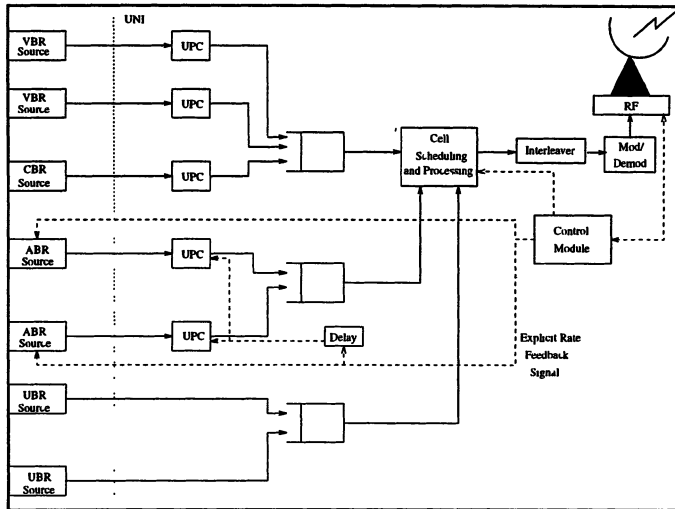


Figure 8 VSAT Earth Station Functions.

5.3 Space Segment

The most important element of the space segment is the on-board processor with cell-switching capabilities for satellite ATM networking. Several concepts for on-board switching have been proposed. However, these systems have been mainly targeted at applications supporting a predominance of voice traffic, and therefore employ circuit-switched techniques. Examples are the NASA Advanced Communication Technology Satellite (Wright, 1992) and the European Space Agency (ESA) time-space-time prototype (Fromm, 1992). With the introduction of new services, the question arises of how to offer better satellite resource utilization in a mixed traffic (voice, data and video) environment.

We propose to statistically multiplex the traffic on board the cell-switching satellite for maximum bandwidth utilization. On the uplink the satellite is accessed via 16 Uplink Groups (UG). Each UG comprises of 16 carriers with a transmission rate of 2.048 Mbit/s. The payload configuration of this system is shown in Figure 9. On board the satellite the 16 carriers are first demodulated using a Multi-Carrier Demodulator (MCD) for low power and low weight implementation. The input processor processes the FUs extracting the cells. Then all cells from the 16 carriers are multiplexed discarding empty cells. The output of the multiplexing buffer (and the input processor) is 16 Mbit/s, which is the cell switch rate of one port of the OBP cell switch. In this scenario maximum utilisation of the burst slots is not required (considering that VSATs don't generate much traffic) for maximum bandwidth utilization of the satellite, by concentrating the traffic in the sky.

This introduces some issues for the design and analysis of the satellite architecture. Many considerations previously the concern only of the ground segments now shift to the space segment. The on-board processor allocates

bandwidth on demand and performs statistical multiplexing. This essentially changes the nature of the satellite from a deterministic system to a stochastic system. In a stochastic system, the arriving traffic is random and statistical fluctuations may cause congestion, where cell loss due to buffer overflow might occur. Thus, it is necessary to incorporate traffic and control mechanisms to regulate the input traffic.

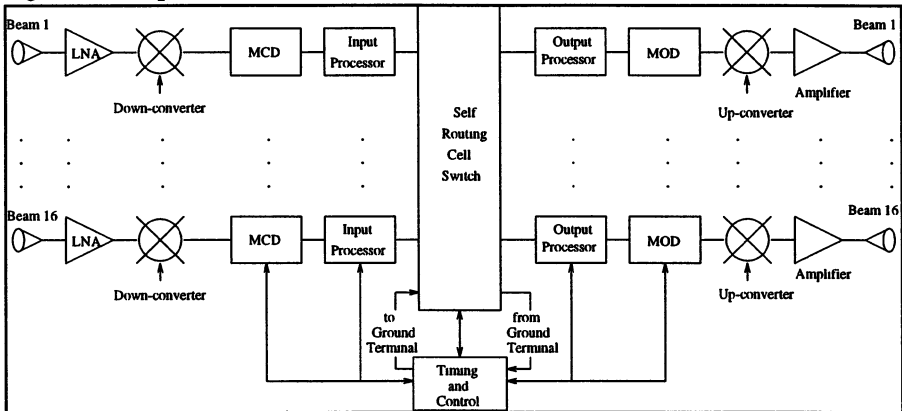


Figure 9 Payload Block Diagram.

5.4 Traffic and Congestion Control

Providing the desired QoS for various traffic categories in ATM networks is not an easy task. The design of a suitable ATM traffic and congestion control is the most important challenge for the success of an ATM-based B-ISDN. Therefore it has been the subject of vigorous research over recent years.

Various control mechanisms have been proposed for ATM networks. These can be classified into two categories:

- Preventive control techniques, which attempt to prevent congestion by taking appropriate action before congestion occurs. A preventive control mechanism consists of the *connection admission control* (call level) and *usage parameter control* (cell level).
- Reactive control techniques (burst level), which are used to recover from a congested state.

Connection Admission Control (CAC)

The CAC decides whether or not a connection can be accepted. Network resources are allocated according to the traffic contract which is negotiated between the user and the network. Parameters which form this contract are the Traffic Descriptors, QoS requirements, Conformance Definition and the Service Category. A connection request is accepted only when sufficient resources are available to establish the call through the whole network at its required QoS and maintain the

agreed QoS of existing calls. This also applies to re-negotiation of connection parameters within a given call.

Calculation of CLR for N Identical Multiplexed Sources

A variety of CAC algorithms have been proposed. The aim is to use an algorithm that is simple (in terms of processing and storage requirements) and efficient (to allow statistical multiplexing gain).

The fluid-flow approach (Anick, 1982) gives accurate results for the superposition of N identical ATM sources on an ATM link compared to other CAC algorithms (Sykas, 1992). The fluid-flow model considers information is continuous, as if it were a fluid. There is no packetization of information in cells which means that only the burst scale is considered, resulting in considerable computational simplification while providing accurate results if the multiplexing buffer is large enough to absorb cell scale fluctuations.

Usage Parameter Control (UPC)

UPC is defined as the set of actions taken by the network to monitor and control traffic in terms of conformity with the agreed traffic contract at the user access. The main purpose is to protect network resources (in particular the satellite link capacity) from misbehaviour that could affect the QoS of other established connections. The Leaky Bucket (LB) is generally agreed to achieve the best performance compromise of the mechanisms studied for UPC (Onvural, 1994).

Reactive Congestion Control

Although preventive control tries to prevent congestion before it actually occurs the satellite system may experience congestion due to multiplexing buffer or switch output buffer overflow. In this case, where the network relies only on the UPC and no feedback information is exchanged between the network and the source, no action can be taken once congestion has occurred.

Rate-based control was recommended for ABR services, where information about the state of the network is conveyed to the source through special control cells called Resource Management (RM) cells (ATM-Forum, 1995). Explicit Rate (ER) indication is used in our system, in which the satellite notifies the source of the exact bandwidth share it should be using in order to avoid congestion. Various algorithms have been proposed for the calculation of the fair bandwidth share per connection (Jain, 1995). The source model used for the ABR service is the persistent source which always sends with the maximum permitted rate. Different simulations (Barnhart, 1995) have shown that this model imposes the heaviest constraints on the network and is therefore very appropriate for testing the throughput and cell loss of the ABR service.

The switch can determine congestion either by measuring the traffic arrival rate or by monitoring the buffer status. We will use the former method where the on-board switch measures the current load by counting the number of cells received

during a fixed averaging interval. Based on the known capacity of the link, the switch can determine whether it is overloaded or underloaded. If an overload state is detected the activity of the source is reduced. After entering the normal state the source activity is gradually increased to avoid oscillation. The aim of the rate control is to keep the utilisation at a target level with minimum cell loss. The choice of the target utilisation (for example $85 \pm 5\%$) has an important impact on the system performance as will be shown in the next Section. Whenever the link utilisation is in the given range the link is said to be operating efficiently.

5.5 OBP System Performance

Whatever assumptions are made on the traffic entering the VSAT, the multiplexing and multiple access functions in the up-links alter its statistical characteristics. In order to evaluate the performance of the OBP buffers we chose to use the fluid-flow approach, adding the contribution to the cell loss caused by the different natures of the traffic arriving on-board. On-board the satellite two buffers for rt- and nrt-traffic are used.

Only the rt-VBR and ABR traffic categories are considered since they represent integrated transport of real-time and non-real time traffic flows. CBR traffic and UBR traffic can easily be integrated in the traffic management scheme. The bandwidth assignment to CBR services is the peak rate and traffic control of the peak rate is not difficult. Thus CBR traffic will only have the effect of background traffic. Handling UBR traffic introduces the issue of protecting the QoS objectives of the ABR connections.

The equilibrium probability of i sources being on P_b , is given by the binomial distribution and C is defined as the capacity how many active sources can be supported on the link without reaching the overload state.

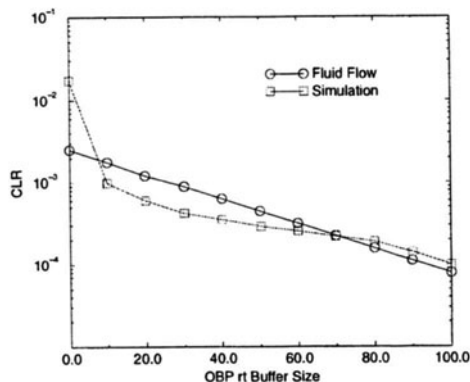


Figure 10 Simulation and Theoretical Results for 50 Voice Sources.

In order to verify the accuracy of the fluid-flow approach for the calculation of the on-board multiplexing buffer cell loss, simulations were carried out for 50 packet voice sources. The traffic parameters for the voice source are PCR=32kbit/s, mean burst duration=0.352 ms and mean silence duration=0.65 ms. Figure 10 shows the theoretical and simulation results. As it can be seen the theoretical results underestimate the cell loss for small buffer sizes where cell scale fluctuations can't be absorbed. As expected for larger buffer sizes the fluid flow model overestimates the CLR slightly. However it is interesting to note that as the buffer size exceeds 70 cells the theoretical results become less accurate, which is due to the satellite uplink. We will therefore use theoretical results with Simulation results are generally more accurate in capturing the system details. All simulation results presented in this work have a 95% confidence interval which are not shown for the neatness of the figures. The normalised capacity (C/N) for a certain CLR and a rt-buffer size of 50 if 800 voice sources (maximum number which can be supported by one UG with 16 carriers), 400 or 50 voice sources are multiplexed, is shown in Fig.11. It can be seen that for $N=800$ an effective bandwidth of $0.4375 \cdot 32 \text{ kbit/s} = 14 \text{ kbit/s}$ is sufficient for a CLR of $1.45 \cdot 10^{-9}$. This is 43.75% of the peak rate resulting in an throughput increase by 228% compared to circuit switching (where the peak rate is assigned to each connection).

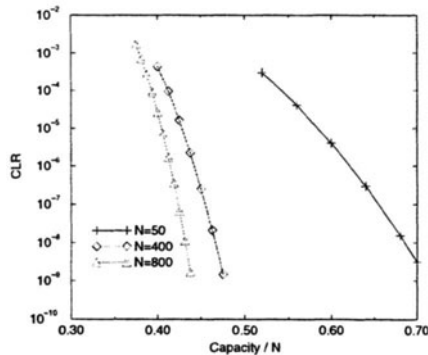


Figure 11 CLR vs C/N .

The remaining bandwidth (16Mbit/s-bandwidth allocated to rt-sources) can be allocated to UBR and/or ABR sources by taking advantage of the silence periods of the voice source. In the worst case when 800 terminals are transmitting voice and also want to transmit ABR traffic the fair share for each terminal is 6 Kbit/s. In order to observe the effects of the target utilisation (u) on cell loss for different buffer sizes we performed simulations for 800 voice sources and a total bandwidth of 16 Mbit/s. The MCR for each ABR source connection request was assumed zero so that the bandwidth assigned to each source was the same. A buffer size of 50 cells is used for higher priority real-time traffic to guarantee their QoS. The propagation delay to all earth stations was assumed identical (260 ms).

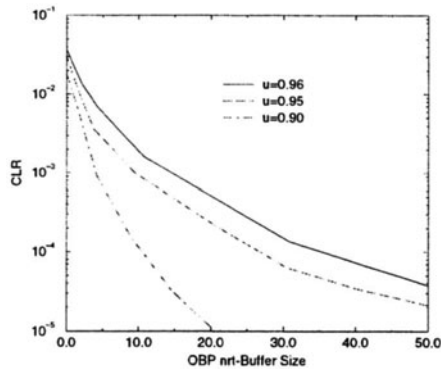


Figure 12 CLR vs nrt buffer size for various target utilisation values.

If the target utilisation level is set high more link bandwidth is assigned to each ABR connection and the rate of the sources is only reduced when the target utilization (u) is exceeded. Figure 12 shows the CLR of ABR traffic as a function of the nrt-buffer size for various target utilisation limit values. As can be seen, the cell loss increases very sharply as the target utilization increases. For 96% target utilization, the maximum bandwidth assigned to ABR sources was 6.3 Mbit/s. Looking at Figure 11, it can be seen that $800 \cdot 14 \text{ kbit/s} = 11.2 \text{ Mbit/s}$ of the link bandwidth has to be assigned to voice sources for a CLR of $1.45 \cdot 10^{-9}$. For such a CLR only 4.8 Mbit/s has to be assigned to ABR sources (86% utilization). Thus the system target utilization would be around this value for stable operation. The required total buffer size to provide the necessary CLR is 100 cells (5.3 Kbyte).

The use of a shared buffer for the multiplexer could provide improved performance at the expense of additional complexity and is for further study.

6 CONCLUSIONS

As user demands become more complex, VSAT satellite networks, which have so far been successful in providing mainly packet data services, are expected to support a much wider range of services. As satellites will play an important role in the early deployment of ATM networks, we have discussed two different satellite-based architectures (for the short and long term) that could be used to provide various ATM services and addressed some of the problems that need to be resolved.

The need for optimisation of the capacity allocation scheme, when using bent pipe satellites was discussed using some preliminary performance results for a dynamic reservation system. Some bottlenecks in the performance of protocols such as TCP/IP introduced by the satellite link were highlighted. We have seen

that the link capacity and the window size could have severe implications on the systems performance.

To provide full ATM connectivity via satellite using bent-pipe satellites will require development of more dynamic and efficient satellite access scenarios.

Considerable improvements in delay performance and satellite bandwidth utilization are possible if next-generation satellite technology (OBP) is used. The traditional demand-assignment scheme using a ground terminal as control station has two important drawbacks: long set-up time (about 500ms caused by two hops assuming a negligible processing time at the control station) and limited channel utility. Both are due to the long propagation delay of the satellite link. Both disadvantages can be removed by processing channel requests in the satellite to allocate frame slots. An on-board cell-switching satellite with spot beams is considered in architecture 2.

In our proposal maximum utilisation of the burst slots using complex dynamic reservation protocols, which introduce high delay, is not required. Maximum bandwidth utilization of the satellite is achieved by statistical multiplexing the traffic on-board the satellite. Greater connectivity, improved spectral efficiency, improved use of the switch capacity and greater EIRP per VSAT terminal are the main advantages of using a state-of-the-art satellite system. These advantages enable new services while potentially reducing the cost of earth stations.

However development of an OBP cell-switched satellite communication network introduces some issues, such as traffic and congestion control, which have been addressed in this paper. The used CAC algorithm and choice of target utilization are the main factors affecting the system performance. Even slight overdimensioning of the system results in high cell loss and thus unacceptable performance. It has been shown that rt and non-rt traffic can be integrated and increases in throughput up to 228% achieved, for a very low CLR.

We conclude that the use of VSATs for ATM services is possible, but careful system design and dimensioning is very important to provide the required quality of service.

7 REFERENCES

- Anick,D. Mitra,D. and Sondhi, M.M. (1982) Stochastic Theory of Data-Handling Systems with Multiple Sources. *Bell System Technical Journal*, 61(8)
- ATM-Forum (1995) ATM-Forum, Traffic Management Specification Version 4.0, 95-0013R8
- Barnhart A.W. (1994) Baseline Model for rate-control simulations", ATM Forum 94-0399
- 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), 1823-1831

- Fromm, H.H. and Bella, L. (1992) System architecture and design criteria for the ESA OBP development. *9th Int. Conf. on Digital Sat. Com.*, 57-62
- ITU-T (1996) Traffic Control and Congestion Control in B-ISDN, *Rec.I.371*
- Hadjitheodosiou, M.H. Komisarczuk, P. Coakley, F. Evans, B. and Smythe, C. (1994) Broadband Island Interconnection via Satellite-Performance Analysis for the RACE II-CATALYST Project. *International Journal of Satellite Communications*, V.12, 223-238
- Jain, R. (1995) Congestion Control and Traffic Management in ATM Networks: Recent Advances and A Survey. *Computer Networks and ISDN Systems*
- Louvet, B. and Chellingsworth, S. (1994) Satellite integration into broadband networks, *Electrical Communications*
- Lunsford, J. Narayanaswamy, S. Chitre, D. and Neilbert, M. (1995) Link Enhancement for ATM over Satellite Links. *IEE 10th International Conference on Digital Satellite Communication*, 129-136
- Onvural, Raif O. (1994) *Asynchronous Transfer Mode Networks: Performance Issues*, Artech House
- Polese, P. Mort, R. and Combarel L. (1994) Satellites in UMTS and B-ISDN: status of activities and perspectives. *Electronics and Communication Engineering Journal*, 297-303
- Raychaudhuri, D. and Joseph K. (1987) Ku-Band Satellite Data Networks using VSATs-Part1: Multi-access Protocols. *Int. J. Sat. Comms.*, 195-212
- Schoute, F.C. Awater, G.A. and Giesbers, B.M.P. (1994) Aloha Performance Models. *IEEE GLOBECOM*
- Sun, Z. Ors, T. and EVANS B.G. (1996), Interconnection of Broadband Islands via Satellite-Experiments on the RACE II CATALYST Project, *IEEE Globecom Workshop on Transport Protocols for High-Speed Broadband Networks*
- Sykas, E.D. Paschalidis, I.Ch. and Vlakos, K.M. (1992) Congestion Avoidance in ATM Networks, *IEEE INFOCOM*
- Vakil, F. and Saito, H. (1991) On Congestion Control in ATM Networks. *IEEE LTS*, 55-65
- Weinstein, C.J., (1979) Fractional speech loss and talker activity model for TASI and for packet-switched speech. *IEEE Trans. on Communications*, V.COM-26, 1253-1257
- Wright, D. and Balombin, J. (1992) ACTS System Capabilities and Performance, *AIAA*, 1135-1146
- 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*, 15-20

8 BIOGRAPHY

Tolga Örs received his BSc degree in Electrical Engineering from Istanbul University, Turkey and the MSc degree in Satellite and Telecommunication Engineering from University of Surrey, England. He is currently pursuing his Ph.D. at the University of Surrey working on optimization and performance evaluation of ATM traffic control functions for terrestrial and satellite networks. His research interests include traffic characterization, performance evaluation and traffic & congestion control in ATM networks.

Michael Hadjitheodosiou received the M.A. degree from the University of Cambridge, U.K., the M.S. degree from the University of California, Irvine and the Ph.D. degree from the University of Surrey, U.K. all in electrical engineering, in 1989, 1991 and 1995 respectively. He was a visiting fellow at the Canadian Government Communications Research Center (1995-96). He is currently with the Center for Satellite & Hybrid Communication Networks, University of Maryland, College Park. His research interests include multi-access protocols, performance optimisation of satellite and mobile networks and ATM over next generation broadband satellite systems.

Zhili SUN received his BSc from Nanjing University, China and his MPhil and PhD from Lancaster University, UK. He is currently a Senior Research Fellow in the Centre for Communication Systems Research, University of Surrey, UK. From 1989 to 93 he was a postdoctoral researcher, in the Telecommunication Group, Queen Mary and Westfield College, University of London. He is now working on an EPSRC project "B-ISDN implementation via satellite-ATM island interconnection" and also as a co-ordinator of the THESEUS project within the European ACTS programme. He teaches MSc and industrial short courses on ATM and Broadband ISDN.

Barry Evans received the BSc and PhD degrees in Electrical Engineering and Microwave Systems from the University of Leeds in 1965/1968. He was Telecom Lecturer-Reader in Telecommunications Systems at the University of Essex from 1969-1983. In 1983 he was appointed to the Alec Harley Reeves Chair of Information Systems Engineering at the University of Surrey, and in 1990 became the first Director of the post-graduate centre for Satellite Engineering Research which houses some 100 researchers. Barry Evans is editor of the International Journal of Satellite Communications and author of three books and over 200 journal papers.