

Performance Analysis of Satellite Payload Architectures for Mobile Services

R. A. WYATT-MILLINGTON, Member, IEEE

R. E. SHERIFF

Y. F. HU

University of Bradford
United Kingdom

This paper is concerned with the effects on the network performance of moving parts of what is considered traditionally to belong to the ground segment to on board the satellite. Initially, an overview of geostationary satellite communication systems and payload technology is presented, followed by a description of the network architecture and protocols that are the basis of the simulation models. The results obtained from this testbed are presented before concluding with a discussion of the results obtained.

Manuscript received May 19, 2005; revised November 16, 2005 and April 11, 2006; released for publication June 24, 2006.

IEEE Log No. T-AES/43/1/895029.

Refereeing of this contribution was handled by M. Ruggieri.

Authors' address: School of Engineering, Design & Technology, University of Bradford, Bradford, West Yorkshire, BD7 1DP, UK, E-mail: (r.a.wyatt-millington@brad.ac.uk).

0018-9251/07/\$25.00 © 2007 IEEE

I. INTRODUCTION

Geostationary (GEO) satellites have been used as the space segment within a network since the launch of the Syncom III satellite in 1963. The principles of a GEO satellite were first described by Arthur C. Clarke in 1945 [1], when he proposed the development of a manned object at a fixed point in the sky to broadcast voice communications worldwide. Current GEO systems provide services such as voice and data communication, and broadcast services for TV and radio. GEO satellite systems can be divided into two categories: global and regional systems. INMARSAT,¹ which has been one of the major global players in satellite communications for many years, is an example of a global GEO system operator that provides both voice and data services via its INMARSAT-3 satellites. The INMARSAT GAN (global area network) system provides ISDN services at data rates up to 64 kbit/s. In the near future, the proposed broadband GAN (BGAN) [2], using the INMARSAT-4 (I-4) satellites, will provide Universal Mobile Telecommunications System (UMTS) and International Mobile Telecommunications-2000 (IMT-2000) compliant services, delivering broadband multimedia data to mobile users at data rates up to 432 kbit/s. The first I-4 satellite was successfully launched on March 11, 2005, and took over digital service provision from one of the INMARSAT-3 satellites at the end of May 2005.² The Indonesian-based Asia Cellular System (ACeS) [3] and the Saudi Arabian-based THURAYA Satellite System [4] are examples of regional GEO satellite systems that use the nonproprietary geo-mobile radio (GMR) specifications developed jointly between the respective companies and the European Telecommunications Standards Institute (ETSI) and the Telecommunications Industry Association (TIA) of the United States. The THURAYA satellite capacity is also leased by INMARSAT for its regional BGAN service, which is an intermediate step towards the BGAN service. The regional BGAN service is available within the coverage area of the THURAYA-1 satellite and offers data rates of up to 144 kbit/s via a satellite modem that is around the size of a notebook computer.

The aim of this paper is to present results of an investigation into what the effects are on the performance of a GEO satellite network when portions of the ground segment functionality are placed within the space segment as part of the on-board processing (OBP) of a regenerative satellite. The majority of deployed communication satellites have a payload that contains, at the most, on-board switching (OBS) at the IF of the transparent

¹<http://www.inmarsat.com/>

²<http://countdown.inmarsat.com/news/00016875.aspx/>

transponder. The use of full regenerative satellites is not very common due to two inherent technical issues, as well as for economic constraints given the longer lead-time required for the development of these complex payloads. The first technical issue is the difficulty in dealing with problems that arise within a complex payload once the satellite is in space, and the second is that a regenerative payload using current space-hardened technology is locked onto one modulation scheme and access technique [5].

It is this second problem that is of greater concern to satellite system designers as with the current lifetime of satellites, and expected development lead-time, the system could be locked into the technology for anything up to 25 years, which in communication terms is a couple of generations of technology. However, as integrated circuit manufacturers improve their programmable digital devices to a level where such devices can withstand the harsh space environment, the use of reprogrammable, regenerative satellites becomes a more feasible option. This will allow the advantages of regenerative payloads to be exploited.

Indeed the advantages of reconfigurable computers in space, as well as the potential problems were identified by Bergmann and Dawood [6] in their work in the Cooperative Research Centre for Satellite Systems (CRCSS) in Australia. On December 14, 2002, an experimental reconfigurable payload was launched into space as part of the Australian scientific mission satellite FEDSAT [7]. The core of this payload was a radiation-hardened Xilinx field programmable gate array (FPGA). In July 2003, the payload “healed” itself by detecting, analysing, and repairing a fault caused by space radiation, with no need for human intervention [8], thereby proving that reconfigurable computers may have a future in the development of regenerative payloads, and hence for the possibility of moving traditional ground segment functionality on board the satellite.

The ideas presented here for moving what is traditionally considered ground segment functionality into the space segment are looking forward to a time when such satellites are a feasible alternative to the current generation of satellites. In particular, the issues of additional power requirements of the processing on-board, as well as the additional mass, will require close attention. This paper looks at constant bit rate (CBR), circuit-switched traffic, but the approach could equally well be applied to variable bit rate (VBR), packet-switched traffic. In this case, the greater complexity of the resource management algorithms due to the need to continually change the time slot allocation would need to be considered.

This paper is divided into a number of sections, with Section II exploring in more detail the satellite communication systems, in terms of payload technology, and the role of satellites. Section III

presents the GMR standards used for the research presented here, with Section IV examining the network architectures that are simulated. Section V gives more details about the simulation, with the results being presented in Section VI. Finally, some conclusions are drawn in Section VII.

II. SATELLITE COMMUNICATION SYSTEMS

A. Payload Technology

The definition of a satellite payload given by [9] is:

...the system on-board the satellite, which provides the link for the communication signals path.

In the past, this communication link was between two ground stations, but the more advanced payloads of today can provide for the interconnectivity between large numbers of mobile users, either directly, or via ground stations. This extra functionality allows for the needs of a more mobile population, rapidly changing traffic demands, and varying operational scenarios.

The fundamental payload functions are to receive and filter the uplink signals, and to provide frequency conversion and amplification of the signals for retransmission on the downlink. There are two basic types of payload: transparent, where the original data is not recovered, although some OBS may take place; and regenerative, where it is recovered and OBP is used. The complexity of the OBP can vary from a simple switch to much more complex systems. However, as a general rule of thumb, the more complex the on-board processing, the more features that can be offered in a flexible manner [10].

Reference [5] introduces the need for OBP to provide multimedia services via Ka-band satellites to small user terminals all over the world. Two advantages of satellite systems with OBP are:

- 1) improved link quality as compared with transparent systems,
- 2) ability to provide direct interconnection between terminals via use of OBS—reduces latency in link and hence improves quality of service.

Many of the papers published in the field of OBP have tended to concentrate on the use of OBS in satellites, particularly in terms of asynchronous transfer mode (ATM) switching within the satellite payload [11–14]. As well as the ability of OBS to provide direct interconnection between terminals, it can also enhance user access via selective landing—where the gateway closest to the required destination is chosen, thereby avoiding possible network congestion within the terrestrial network. A design for an experimental multimedia on-board switch developed for the European Space Agency (ESA) is presented in [13], along with an ATM-based

implementation. However, the presented architecture can be extended to MPEG (Motion Picture Experts Group), DVB (digital video-broadcasting) or label-switched IP-based systems.

Reference [15] examines the use of OBS within an otherwise transparent payload in the context of a satellite with multiple spot beams. This type of system can provide, on top of the normal advantages of a system with OBS, VBR services and bandwidth-on-demand. An example of the use of an on-board switch is within the THURAYA³ system to provide the functionality for single-hop terminal-to-terminal calls. The control data still passes to the gateway, but the traffic data is switched on the satellite between the two users. This reduces the perceived delay between the actual speech starting and it being heard by the listener. The THURAYA satellites also use digital beam-forming (DBF) to produce the spot beams. The use of DBF allows reconfiguration of the beams within the coverage area and dynamic allocation of resources to allow for traffic "hot-spots."

The use of OBP in current and planned commercial satellites in EUTELSAT is examined in [16], which presents the SKYPLEX OBP from its earliest conception to future use in multimedia satellites to provide services such as dynamic bandwidth allocation, and on-board connectivity. SKYPLEX was the first multimedia OBP module in a commercial satellite when it was brought into service on the Hot BirdTM 4 satellite in 1998 and is being continually developed to enhance the service provided by the Hot BirdTM constellation.

A possible protocol architecture for an ATM-based satellite system is given in [11]. This article examines the payload architecture within such a system in terms of the split in the resource management and control functionality between the space and ground control segments. The OBS provides full connectivity between any uplink spot beam to any downlink spot beam, thus enhancing the access capabilities of the system. The main focus of the paper was on the effects of such a protocol architecture on the performance of the system, especially during call set-up, which is one focus of the research carried out for this paper. Another proposed ground/space split in terms of the medium access control (MAC) sublayer is given in [17], where the scheduler part of the MAC is moved on-board with the remaining elements on the ground. Various possible advantages of such a scheme are presented.

B. Role of Satellite

The current trend in the mobile-satellite industry appears to be in serving smaller niche markets, rather

³www.thuraya.com

than directly competing with the terrestrial cellular networks. Many of the networks envisaged at the start of the nineties were non-GEO solutions to provide personal mobile communications as a single system, whereas now, at the start of the 21st century, the technological pendulum appears to have swung back towards the GEO solution providing personal mobile communications in cooperation with terrestrial systems. Satellite systems are unable to compete with the low call tariffs available on terrestrial systems, and so are now seen as extending the coverage of existing terrestrial mobile systems.

Satellite communications are particularly useful in areas of the world where there is no cellular coverage. For instance, mountaineers use satellite phones to keep in touch (see [3] for an example of pictures sent via the ACeS system from a team climbing K2 in the Himalayas). Satellite phones can also be useful for people who travel the world, and may not be able to use the local terrestrial network.

Satellites, particularly at K-band, can provide broadband access to areas where it is expensive to lay optical fibres to provide broadband access. One interesting use is in tele-medicine, such as when field hospitals send medical information to a base hospital to obtain an accurate diagnosis and/or treatment.

All of the above uses for satellite communications are in niche markets for which satellite communications can solve problems in communications that may be difficult or expensive to overcome otherwise.

III. GEO-MOBILE RADIO SPECIFICATIONS

The GMR specifications define the first standardized satellite mobile communication systems. Until recently, mobile-satellite system solutions have tended to be propriety, and no two systems have been alike, which made global roaming difficult. The GMR specifications define a GEO satellite system that reuses much of the GSM specifications, and hence allows good integration (at inter mobile switching centre (MSC) level) with GSM networks. There are two sets of GMR specifications.

1) *GMR-1 Family*: A general description can be found in [18]. This uses L-band links between the mobile stations and the satellite (spot beams), and Ku-band or C-band feeder links between the gateways and the satellite (regional beam). The satellite operations control centre (SOC) is attached to a gateway. Fig. 1 shows the network architecture for a GMR-1 system.

2) *GMR-2*: A general description can be found in [19]. This uses L-band links between the mobile stations and the satellite (spot beams), and C-band feeder links between the gateways and the satellite (regional beam). The satellite control facility (SCF)

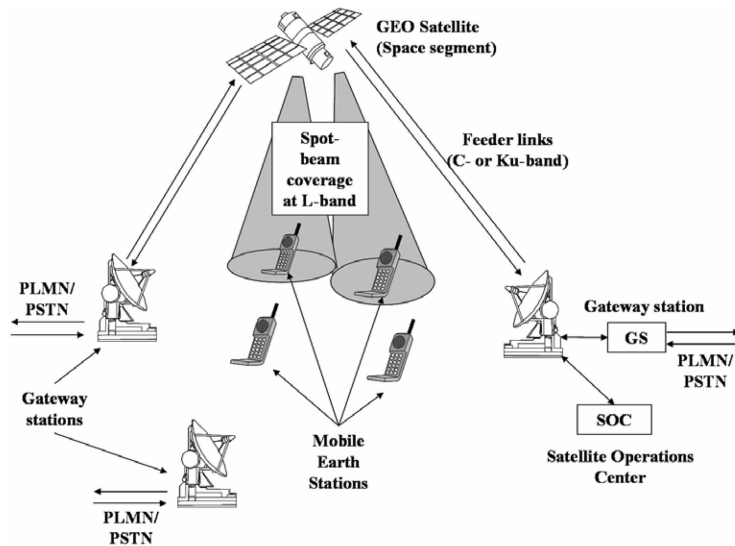


Fig. 1. GMR-1 network architecture [20]—basic architecture of GMR-1 based satellite network with operations centre collocated with gateway.

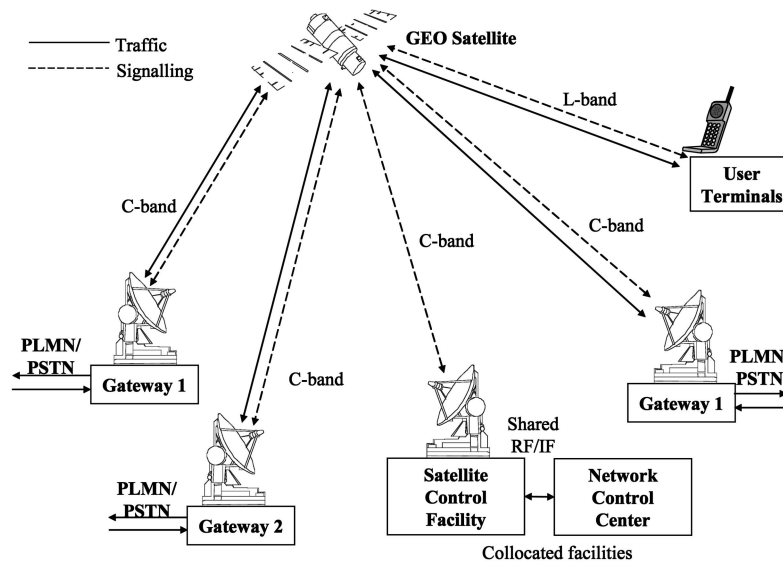


Fig. 2. GMR-2 network architecture [21]—basic architecture of GMR-2 based satellite network with operations centre separate from gateway.

and the network control centre (NCC) are collocated as a separate entity with its own antenna. Fig. 2 shows the network architecture for a GMR-2 system.

Table I shows the general features of GSM, and the two sets of GMR standards. As can be seen GMR-2 is closely related to GSM, whereas GMR-1 has some major differences on the radio interface and in the physical layer. In particular, GMR-1 uses a different physical channel from either GSM or GMR-2, where each physical channel occupies a number of timeslots within a 24 slot frame at a specific frequency, as opposed to the case in GSM/GMR-2 where a physical channel occupies a single timeslot within an 8 slot frame at a specific frequency.

THURAYA uses the GMR-1 family of specifications, whereas ACeS uses the GMR-2 family of specifications. The concept behind the GMR specifications is to extend the GSM Phase 2 services into areas where there is little or no terrestrial coverage by using a GEO satellite (in case of ACeS) or a set of GEO satellites (as for THURAYA, which has two satellites in orbit as of June 2003). The GMR specifications have been designed so that there is a strong resemblance to GSM at the upper protocol layers. In fact, the A-interface between the gateway/base station and the MSC is identical in both GMR and GSM, which in turn implies that the MSC and databases such as the home/visited location register (H/VLR) in GMR are the same as those used in GSM.

TABLE I
Comparison of GSM and GMR Standards

	GSM	GMR-1	GMR-2
Radio Interface			
Multiple Access Modulation	FDM/TDMA GMSK	FDM/TDMA Normal: $\pi/4$ -CQPSK BACH: BPSK	FDM/TDMA Forward: OQPSK Return: GMSK
RF channel spacing	200 kHz	31.25 kHz	200 kHz (forward link) 4 × 50 kHz (return link)
Timing			
Frame size	4.615 ms	40 ms	4.615 ms
Timeslots/frame	8	24	8
Bits/timeslot	156.25	78	156.25
Gross data rate	270.83 kbps	195.6 kbps	270.83 kbps (forward) 67.71 kbps (return)
Physical channel definition	RF channel + timeslot	RF channel + starting timeslot + number of timeslots	RF channel + timeslot
Basic Voice Traffic	Full-rate—net data rate of 13 kbit/s	TCH3—3 timeslots—net data rate of 2 kbit/s	Quarter-rate—net data rate of 3.6 kbit/s
SDCCH	Net data rate of 0.782 kbit/s	Net data rate of 1.15 kbit/s	Net data rate of 0.2625 kbit/s
Power Control			
Measurements	MS (forward link) BTS (return link)	MS (forward link) GTS (return link)	MS (forward link) GTS (return link)
Decision	BSC	MS (forward link) GSC (return link)	GSC
Implementation	MS (return link) BTS (forward link)	MS (return link) GTS (forward link)	MS (return link) GTS (forward link)
Operating Frequency Bands			
Mobile Link		L-band	L-band
Feeder Link		C-band or Ku-band	C-band

The use of many of the same network components as in GSM also allows for easy integration between the GSM and GMR networks: the GMR MSC can be seen as another MSC within the network. Both operating systems use a dual-mode mobile terminal, which supports both GSM900 and ACeS (at 1700 MHz). The terminal automatically selects the GSM service when available, thus minimizing costs to the user.

Following the development and publication of the GMR-1 and GMR-2 specifications, the GMR working group within the ETSI Satellite and Earth Stations (SES) technical committee, with the support of THURAYA developed the GEO-mobile packet radio service (GMPRS) specifications. These were published as Release 2 of the GMR specifications in March 2003 and add general packet radio service (GPRS) functionality to the existing GMR-1 standards [20]. Future evolution may see the addition of the EDGE (enhanced data rates for GSM evolution) functionality as a 3G solution for GMR satellite systems, enabling higher data rates, without changing the radio interface, other than the modulation schemes which can be done via a software upgrade.

IV. NETWORK ARCHITECTURE AND FUNCTIONS

A. System Architecture

Fig. 3 shows the system architecture of the three payload architectures investigated in this paper:

- 1) the bent-pipe scenario, where all the network functionality is part of the ground segment, as shown in Fig. 3(a);
- 2) the GTS-onboard scenario, where the functionality of the gateway transceiver station (GTS) is part of the space segment, as shown in Fig. 3(b);
- 3) the GSS-onboard scenario, where the functionality of both the GTS and the gateway station controller (GSC), which together form the gateway subsystem (GSS), is on-board the satellite, as shown in Fig. 3(c).

The protocol stacks for the above scenarios are shown in Fig. 4. These protocol stacks are based on the GMR protocols, which in turn are based on the terrestrial GSM protocols. It is worth noting that this paper concentrates on circuit-switched traffic, as the provision of reliable wireless voice traffic is still an issue in many areas of the world, and circuit-switched services are the predominant method of supplying reliable voice traffic.

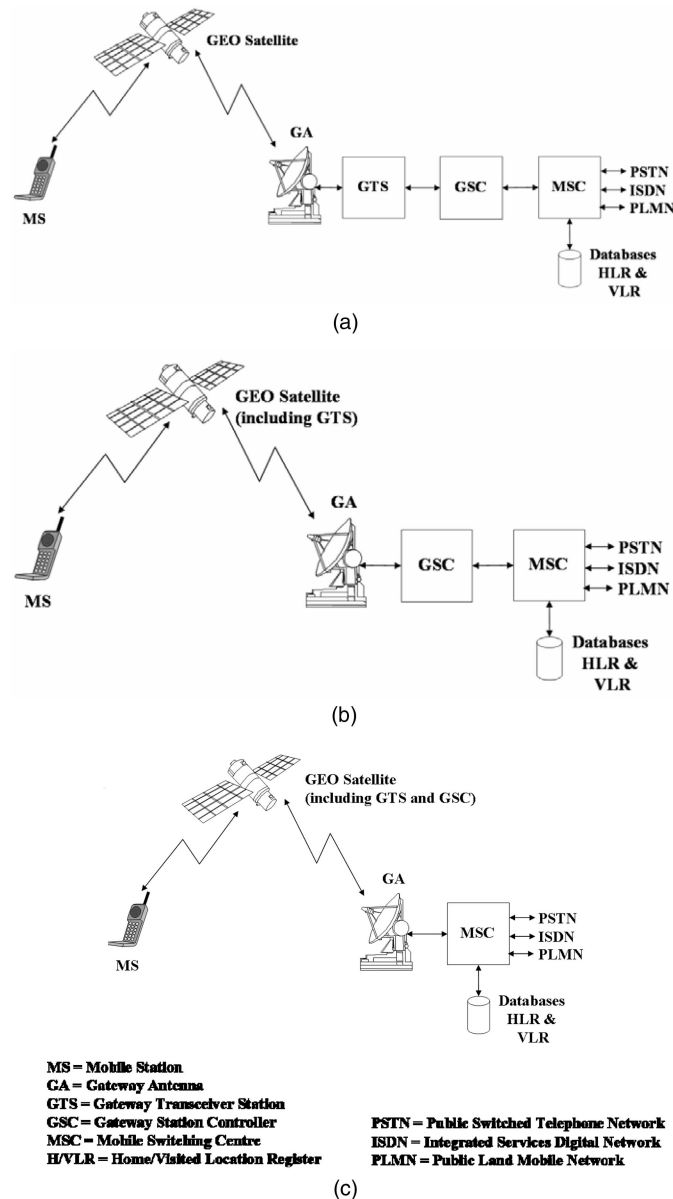


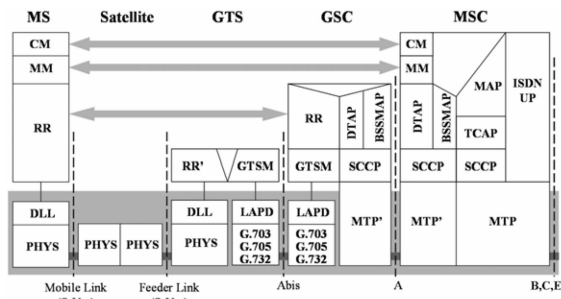
Fig. 3. System architecture. (a) Bent-pipe scenario: all gateway functionality on ground. (b) GTS on board scenario: part of gateway functionality in satellite. (c) GSS on board scenario: all of gateway functionality in satellite.

Table II gives a brief description of each of the layers or sublayers within the protocol stacks that were modelled in the simulation. Each layer/sublayer has its own particular responsibility within the system, although there is interdependence between the layers. Section V gives more details on the specific simulation model used for the results presented in this paper.

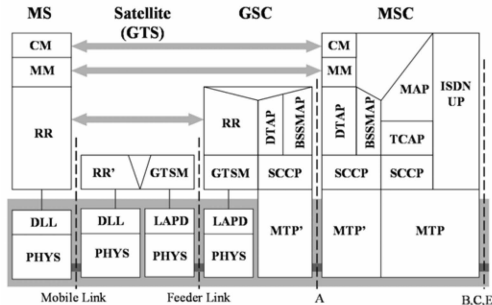
B. Call Set-Up Procedure

Fig. 5 shows the signalling flows for the mobile-terminating call-setup procedure. In this procedure, the call is initiated by the network, which detects an incoming call for the mobile. The MSC, upon receiving notification of this incoming call, requests the data relating to the called mobile station

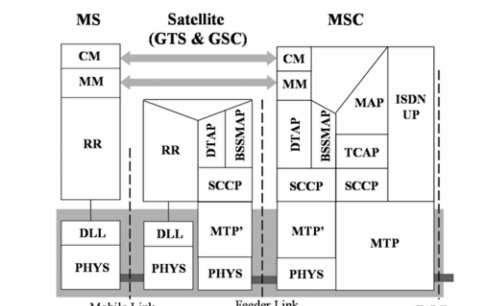
(MS) from the VLR. The requested data is transmitted (via GSM mobile application part (MAP)) from the VLR to the MSC as a page primitive. This primitive is received by the call control (CC) sublayer within the MSC, which then initiates the set-up of a mobility management (MM) connection, by transmitting a paging request message to the gateway. The gateway initiates the establishment of a radio resource (RR) connection, by passing this message to the called MS via the paging channel (PCH). In order to establish an RR connection, a dedicated channel has to be assigned on the air interface to the connection between the MS and the GTS. The call set-up then proceeds with the channel request message from the MS to the GSC, which requests the assignment of a standalone dedicated control channel (SDCCH). The details of



(a)



(b)



(c)

CM Connection Management
 MM Mobility Management
 RR Radio Resource Management
 DLL Data Link Layer
 GTSM Gateway Transceiver Station Management
 BSSMAP Base Station Mobile Application Part
 DTAP Direct Transfer Application Part
 SCCP Signaling Connection Control Part
 TCAP Transaction Capabilities Application Part
 MTP Message Transfer Part
 MAP Mobile Application Part
 UP User Part

Fig. 4. Protocol stacks for the three scenarios. (a) Bent-pipe scenario: as for GMR-based systems. (b) GTS on board scenario: new physical layer on S-Abis link. (c) GSS on board scenario: new physical layer under MTP' on S-A link, so non-GSM-standard MSC.

the assigned SDCCH are transmitted to the MS in an immediate assignment message, which triggers the immediate assignment procedure as detailed in the next section. The SDCCH is then used to authenticate the MS via an exchange of messages between the MM sublayers in the MS and the MSC, and also the VLR (for authentication keys). Following the authentication procedure, ciphering is applied to the dedicated channel, and following this, an exchange of messages between the CC sublayers in the MSC and the MS occurs to establish the CC connection. The *SETUP* message from the MSC to the MS indicates to the MS CC that a CC connection needs to be established. The return of a *CALL CONFIRMED* message to the MSC indicates to the network that the CC connection

is now established. The final stage of the procedure is to inform the called user of the call via the *ALERTING* and *CONNECT* messages. If the user accepts the call, then the MS returns a *CONNECT ACK* message to the network, which passes the message to the calling party, and a data exchange ensues.

The call set-up procedure relies on two RR procedures, as detailed in the next section.

C. Radio Resource Management

In terms of the call set-up procedure discussed above, the main responsibility of the RR sublayer is to assign and then establish the dedicated channels on the air interface. Initially a dedicated control channel is established for the early stages of the call set-up procedure, including the authentication and ciphering procedures. The procedure used for the assignment and establishment of the SDCCH is shown in Fig. 6(a). The procedure is initiated after the reception of the *PAGING REQUEST* by the MS, when it requests the assignment of the channel via the *CHAN REQUEST* message on the random access channel (RACH). The GSC assigns the SDCCH, activates the channel in the GTS and then informs the MS of the assignment via the *IMMEDIATE ASSIGNMENT* message on the access grant and paging channel (AGCH/PCH). The MS then initiates the channel establishment using the *DL_EST_REQ* primitive, which contains the response message (*PAGING RESPONSE* in case of mobile-terminating call set-up). Once the channel has been established using the data link layer procedure, the MM sublayers in both MS and MSC are informed, and the MSC initiates authentication procedures.

An early approach to the assignment of the traffic channel (TCH) is taken in this paper, with it being assigned after the call request is acknowledged with the *CALL CONFIRMED* message. Early assignment is used due to the inherent delay of the satellite air interface, which means that if a late assignment (off-air call set-up) approach is used, where the TCH is assigned only after the mobile subscriber has accepted the call, then there would be a large delay between the user accepting the call, and the point when voice traffic could be transmitted. The assignment procedure is shown in Fig. 6(b), and, as can be seen, is initiated by the network MM sublayer. Again, the GSC assigns the channel, and activates it in the GTS, before informing the MS RR sublayer of the assigned channel using the *ASSIGNMENT COMMAND* message via the SDCCH. The first stage of the procedure in the MS involves the local release of the SDCCH, before the TCH is established as in the immediate assignment procedure, creating an associated control channel (ACCH) along with the TCH. When the *DL_EST_IND* primitive is received by the GTS RR, it performs the local release of the SDCCH

TABLE II
Description of Responsibility of Layers in GSM/GMR Systems

(Sub)Layer	Responsibility
Connection Management (CM) Layer 3	Call Control (CC) entity: Mobile-originating and mobile-terminating call establishment; in-call modifications, call release, and call reestablishment if a call is lost for any reason. Requires active MM connection.
Mobility Management (MM) Layer 3	Mobility of user terminals, such as informing the network of location and providing user identity confidentiality. Requires active RR connection.
Radio Resource Management (RR) Layer 3	Assignment, allocation and administration of radio resources, the acquisition of system information from the broadcast control channel (BCCH) and the cell selection procedure.
Data Link Layer (DLL) Layer 2	Transparent transfer of messages between protocol entities of Layer 3, via the appropriate logical channels. Three channel types: broadcast channels (forward link only), control channels (common and dedicated) and traffic channels (TCH).
Physical Layer Layer 1	Implements error detecting and error correcting codes, handles modulation of signal onto appropriate frequency. Maps logical channels from Layer 2 onto physical channels.

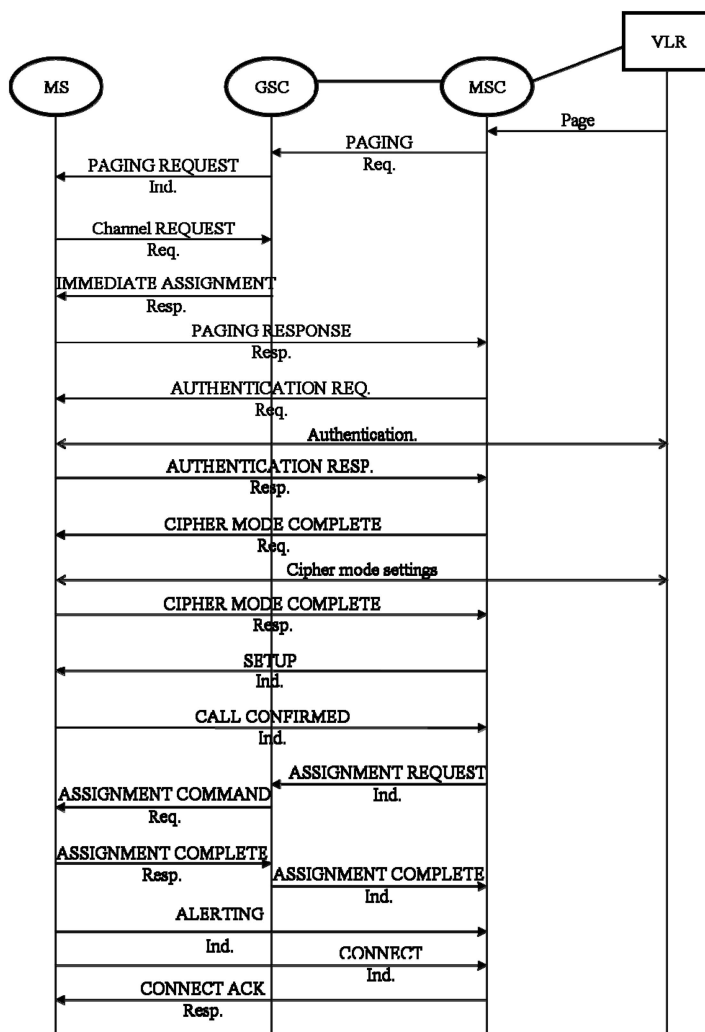


Fig. 5. Signaling flow for mobile-terminating call set-up procedure.

on the network side. The major difference with the immediate assignment procedure is that the response message (ASSIGNMENT COMPLETE) is not transmitted via the DL_EST_REQ primitive, but is transmitted via a normal acknowledged data primitive

once the channel is established. It is important to note that within the RR sublayer, the majority of the procedures are terminated in the BSC RR, but a few are handled by the BTS RR, including ciphering, assembly of channel measurements from the MS and

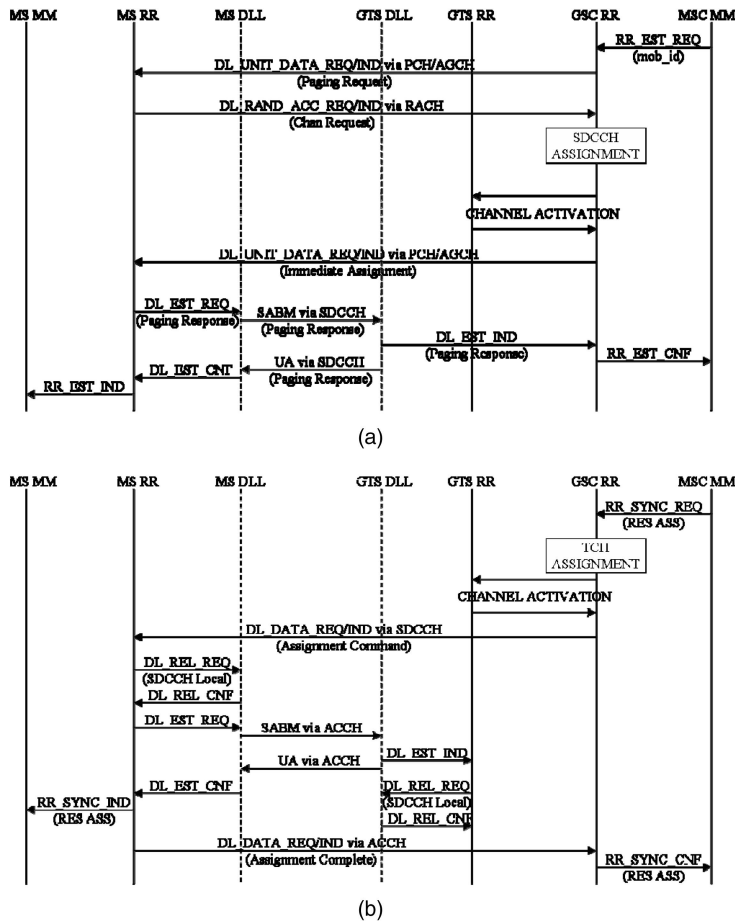


Fig. 6. RR management procedures involved in call setup. (a) Immediate assignment procedure used to set up initial dedicated control channel for transmission of control (signaling) data between mobile and gateway/MSC. (b) Assignment procedure: change of dedicated channel to TCH, thereby allowing user voice/data to be transmitted across the air interface.

transceiver, and their transfer to the BSC (possibly with processing in the BTS) and power control commands from the BTS to the MS.

V. SIMULATION

A. Simulation Approach

The analysis of the network performance of the three scenarios was performed using models developed in OPNET modeler. This is a widely used tool within the satellite research community for simulating networks involving satellites [21, 22]. It divides the simulation model into three domains: network, node, and process, each of which corresponds to an entity within the generic network architecture as shown in Fig. 7. In order to develop simulation models for this paper, selected functionality from the protocol stacks shown in Fig. 4 was implemented as shown in Fig. 8. Each physical component was modelled as a node, with the protocol (sub)layers within the component modelled as OPNET processes. The links between the satellite and the ground components used the OPNET radio model, using the allocated channel frequencies

and bandwidth from the GMR specifications, and modulation curves (bit error rate (BER) versus E_b/N_0) that take into account the effect of the channel coding and interleaving applied to the data (obtained from SIMULINK models of the physical layers of the systems). All the messages defined in [23]–[25] were implemented as packets within OPNET. The packet formats are as specified in the GSM/GMR specifications with mandatory fields only; the optional and conditional fields are ignored. The satellite channel was taken to be the narrowband Lutz channel model [26] as shown in Fig. 9. The Markovian model basically represents an on-off channel where the channel is either in a good state (transmission possible) or a bad state (no transmission possible). In the simulation, for ease of simulation, a Bernoulli process was used to simulate this, where the probability of a zero (bad state in these simulations) is the time-shadowing factor of the Lutz channel A . The time-shadowing factor is defined in (1), where D_g is the mean duration of the good state and D_b is the mean duration of the bad state

$$A = \frac{D_b}{D_g + D_b}. \quad (1)$$

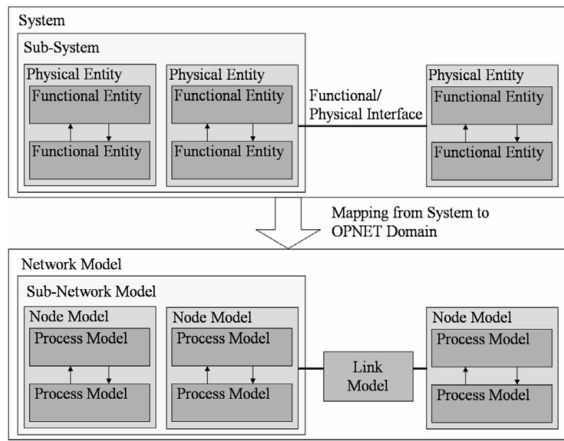
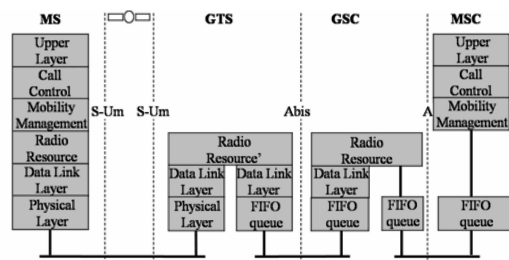
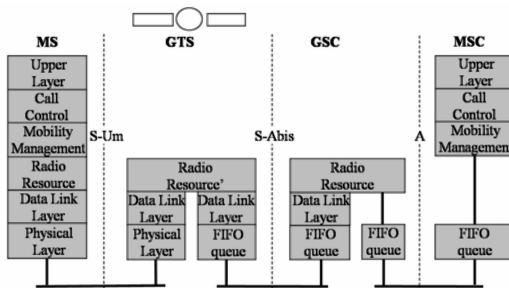


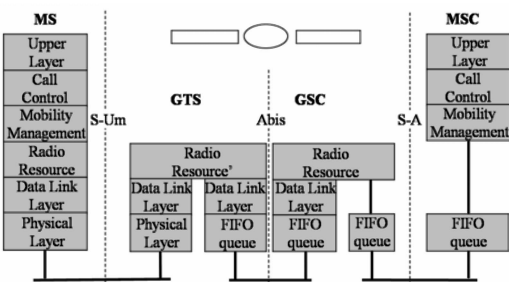
Fig. 7. Mapping of generic system architecture to OPNET architecture: basis for creating network simulations within OPNET modeler.



(a)



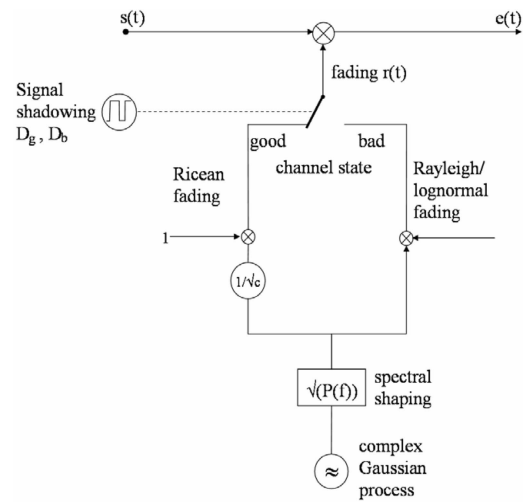
(b)



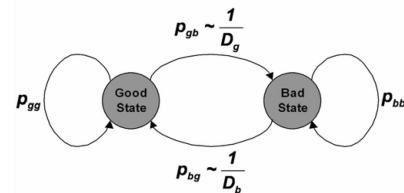
(c)

Fig. 8. Simulated protocol stacks for three scenarios.

Functionality that is included in simulation models can be clearly identified, along with position of functionality within network for each scenario. (a) Bent-pipe scenario: note use first-in first-out (FIFO) queues on Abis and A interfaces. (b) GTS on board scenario. (c) GSS on board scenario.



(a)



(b)

Fig. 9. Lutz narrowband model of land-mobile satellite channel. Top figure gives fading signal that can be applied to transmitted signal, bottom figure shows Markovian representation of model.

(a) Satellite channel model: generated fading characteristics.

(b) Two-state Markov model shows possible transitions within system depending on mean durations of good and bad states.

B. Simulation Scenarios

Network models were developed for the three network configurations for both GMR-1 and GMR-2 based networks, as well as a baseline GSM model (for comparison). Using these models the following two series of results were obtained.

1) Single user scenario: all dedicated channels available all the time. This stage was used to look at differences between GSM, GMR-1, and GMR-2 systems.

2) Multi-user scenario: the availability of dedicated channels defined by an external data file that emulates the occupation of channels by other users. This enables the number of call failures during call set-up (call blocking) and during a call (dropped call) to be measured. This allowed the call blocking and dropped call probabilities to be calculated.

In the first scenario, the starting time of the call set-up request was varied, to simulate a mobile terminating call set-up request at various points within the timers of the system. The position of the request can affect when data is transmitted over the interface between the MS and the GTS given the mapping of the logical channels to the physical channels.

In the second scenario, only the GMR-2 network scenarios were considered. This scenario allowed for an investigation of the effect of two factors on the call blocking (no call set-up) and dropped call probabilities, viz.: the number of available SDCCH, which can be 2, 4, 6, or 8 channels; and the satellite availability, as determined by the time-shadowing factor for the Lutz channel.

VI. RESULTS

It should be noted that all results in this section are theoretical and represent worst case scenarios based on timer values in specifications and low data rate links.

A. Single-User Scenario

For the single-user scenario results, the delays of various subprocedures within the mobile-terminating call set-up procedure, as well as the overall call establishment delay, were collected. The following measurements were made within the OPNET model.

1) The delay of the RR connection ESTablishment (RR EST in): equivalent to paging and immediate assignment procedures as shown in Fig. 6(a) measured in the network MM sublayer;

2) The delay of the AUTHentication and CIPHERing procedures (AUTH & CIPH)—measured in the MM sublayer of the network;

3) The delay between transmission of SETUP and reception of CALL CONF messages (SETUP & CALL CONF)—measured in the CC sublayer of the network;

4) The delay in the RESource ASSignment (RES ASS): equivalent to the assignment procedure shown in Fig. 6(b)—measured in the MM sublayer;

5) The delay in the CALL ESTablishment (CALL EST)—measured in the upper layer of network.

Fig. 10 shows the collection points for these statistics within the call set-up procedure. The statistics were gathered for each of the seven network configurations 1) GSM, 2) GMR-1 bent-pipe, 3) GMR-1 with GTS on-board, 4) GMR-1 with GSS on-board, 5) GMR-2 bent-pipe, 6) GMR-2 with GTS on-board, and 7) GMR-2 with GSS on-board. The GSM results were used as baseline results to look at the effect of the satellite on the operation of the networks.

Fig. 11 shows a comparison of the average delays obtained for the seven network configurations. Overall, these results show that for the GMR-1 based scenarios, the effect of placing portions of the traditional ground segment functionality is very noticeable in the call set-up procedure. When the GTS functionality is moved on board, the average

call set-up time decreases by around 1 s from that of the bent-pipe scenario. There is a further significant decrease in average time taken to establish the call when the full GSS functionality is on board the satellite of around 3 s between the GTS and GSS scenarios, which leads to an overall decrease of around 4 s from the bent-pipe scenario.

In the case of the GMR-2 based scenarios the difference is less significant. Indeed for the GTS on-board scenario there is actually a small increase (around 0.3 s) in the average call set-up time over that for the bent-pipe scenario. There is a small decrease in corresponding value for the GSS on-board scenario of around 0.5 s compared with the standard network. However, it is worth noting that the average call set-up times for the GMR-2 based bent-pipe and GTS on-board scenarios are lower than those for the corresponding GMR-1 based scenarios, due to the differing physical layers. This means that signals tend to have a longer processing delay in GMR-1 than in GMR-2.

The average amount of delay contributed by each procedure varies considerably depending on the network configuration. As would be expected, the GSM delays are all the smallest, although the transmission delay within the physical layer makes a much more significant contribution to the delays for GSM—as shown with the SETUP and CALL CONF delay. The AUTH & CIPH delays for the GMR-1 scenarios are much greater than those for the GMR-2 scenarios because there will be a longer delay before the next SDCCH in GMR-1 as it occurs once every eight frames (40 ms frame), as opposed to every 34 frames (4.615 ms frame) in GMR-2. Particularly for the bent-pipe case, each message arrives at the physical layer very soon after the last SDCCH was received, therefore it has to wait around seven frames (280 ms) for GMR-1, but less than eight frames for GMR-2 (36.92 ms). Therefore, although the SDCCH messages incur an interleaving delay of 124.605 ms in GMR-2, this is still less than the delay incurred waiting for the channel in GMR-1, which also includes a transmission delay of at least 40 ms as the message is sent in two consecutive frames. Also, there will be segmentation of the messages in the data link layer (DLL) in GMR-1 as the maximum number of bytes per DLL message for GMR-1 is much less than that for GMR-2 (seven as opposed to eighteen for SDCCH). This will cause further delays in the transmission of the messages, as each segment will occupy one SDCCH burst.

A similar factor causes the differences between GMR-1 and GMR-2 for the SETUP and CALL CONF delays, although only two messages are involved here, so the difference is much less pronounced.

The RES ASSIGN delay of GMR-2 is much greater than that of GMR-1 due to the different

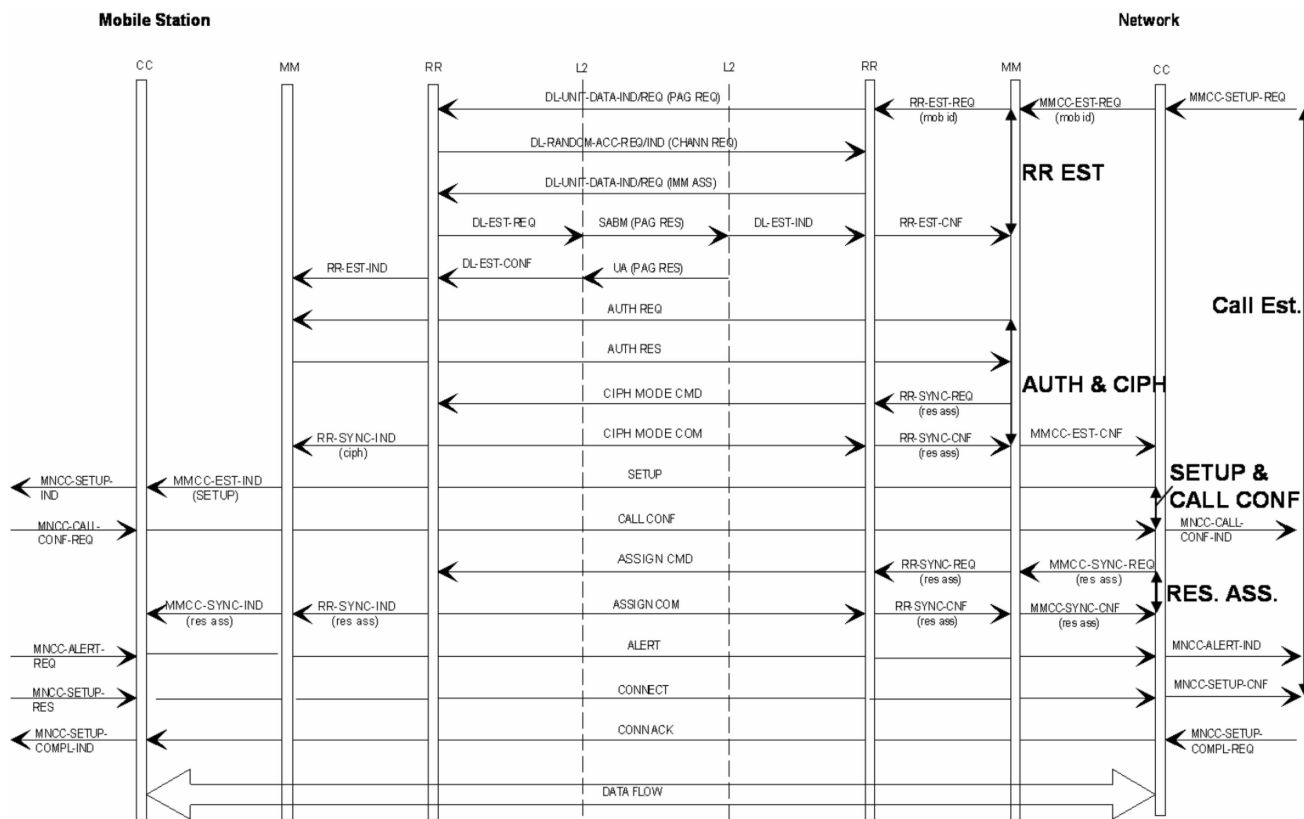


Fig. 10. Collection points for single user scenario statistics shows which delays are measured in gateway and between which messages delays are measured.

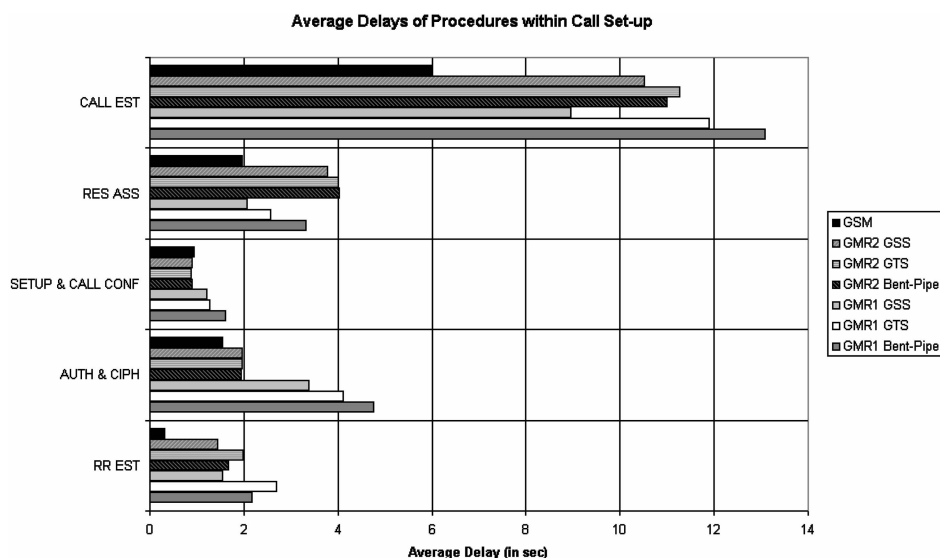


Fig. 11. Comparison of average procedural delays for all scenarios.

associated control channels used in the two systems. GMR-1 uses the fast associated control channel (FACCH), in which the messages are transmitted in the next frame, as it steals slots from the TCH. GMR-2 uses the slow associated control channel (SACCH), which has assigned frames in which it can be transmitted, and is also interleaved across four SACCH frames, which adds to the amount of transmission delay incurred by a message.

B. Multi-User Scenarios—Call Failures

1) *Overview:* Call failure is an important factor within a satellite communications system, as it is a visible measure of performance for the consumer. A user would very quickly become dissatisfied if either calls could not be made via his/her telephone, or calls were being continually dropped. There are two probability measures that can be used to describe

the performance of a network with regard to call failures. Call blocking probability is a measure of the probability of a call being blocked during the call set-up procedure (in other words the probability of a failure during set-up). Probability of a dropped call is a measure of the probability of a successfully connected call being dropped before the user ends it.

The call blocking probability is affected by the availability of dedicated channels (SDCCH and TCH) during an attempted call set-up, and by the satellite channel availability (channel in a good fade state rather than a bad fade state). The probability of a dropped call is only affected by the satellite channel availability, as a dedicated channel has already been assigned to the call. The effect of these factors on the two probability measures is the aim of this set of results.

In order to obtain the probabilities it is necessary to record in the MSC upper layer model (calling user emulator) three values, namely the total number of call attempts (N_a), the number of successful call set-ups (N_s), and the number of dropped calls (N_d).

From these three values, the call blocking probability (P_{cb}) and probability of a dropped call (P_{cd}) can be calculated as shown in (2) and (3)

$$P_{cb} = \left(\frac{N_a - N_s}{N_a} \cdot 100 \right) \% \quad (2)$$

$$P_{cd} = \left(\frac{N_d}{N_s} \cdot 100 \right) \% \quad (3)$$

For each simulation run, the values of N_a , N_s , and N_d were recorded at the end of the simulation as statistics. The two probability values could then be calculated and recorded for each run.

2) *Call Blocking Probability*: Two factors have an effect on the call blocking probability: the number of available SDCCH, and as a consequence the number of available TCH (one TCH occupies space of two SDCCH); and the satellite availability. In these simulations, the number of SDCCH is set via the appropriate attributes in the MS, GTS, and GSC nodes, whereas the availability of the satellite channel is represented by the Lutz time-shadowing factor A . This factor defines the probability that the channel is in the bad state (Rayleigh fading), as opposed to the good state (Rician fading). For instance, if A is 0.01 then 99% of the time the channel is in the good state, i.e., data can be transmitted via the radio link.

Fig. 12 shows the call blocking probability results as affected by the number of SDCCH available with a fixed time-shadowing factor of 0.002, and by the satellite availability with six SDCCH available. Fig. 12(a) shows that that the call blocking probability is very closely related to the number of available SDCCH. Additional information on the reasons for the blocked calls can be seen in Table III, which details the number of expiries of timers T_2.1 (no

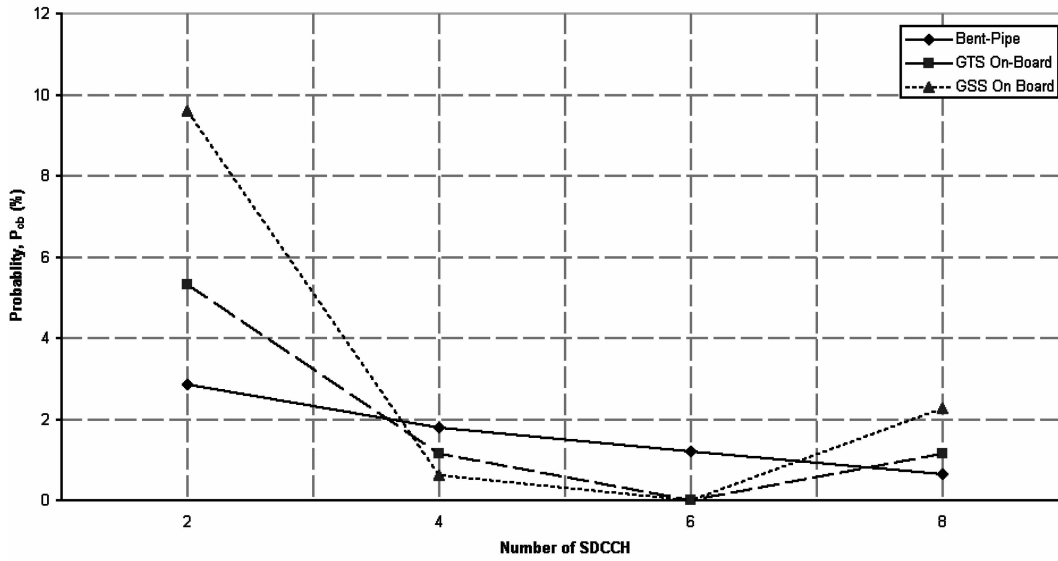
SDCCH) and T310 (no TCH) for the three scenarios. For all the scenarios, the greatest probability of a call being blocked is when there are only two SDCCH available, where the most likely cause of a call being blocked is the nonavailability of a SDCCH, as all timer expiries are T_2.1. On the other hand, when eight SDCCH are available, it is more likely that the call will be blocked by the nonavailability of a TCH, as the timer expiries are all T310. Therefore, the probability is both dependent on the number of available SDCCH and the number of available TCH. As can be clearly seen in Fig. 12(b), the call blocking probability for time-shadowing factors up to $A = 0.01$ is similar for all three scenarios—and relatively low (less than 2% of calls are blocked). For time-shadowing factors higher than this value, the call blocking probability for the GTS on-board scenario rises steeply to give a probability of more than 50% of calls being blocked for 90% channel availability ($A = 0.1$). The bent-pipe and GSS on-board scenarios have similar results for a factor of 0.05 (around 7.5% of calls are blocked), but then the bent-pipe scenario increases steeply to approximately 40% of calls being blocked at $A = 0.1$. The GSS on-board scenario has the best results at a factor of 0.1, as only around 15% of calls are blocked.

3) *Probability of a Dropped Call*: The probability of a dropped call is affected by the availability of the satellite channel, as the most common cause of a satellite call being aborted is a fade event caused either by shadowing or rain attenuation. In the case of a fade, the satellite channel becomes very noisy and the BER rate is too high to be accurately corrected by the error correcting codes on the radio link. This means that the effective connection is lost, so the call is aborted.

Fig. 13 shows the measured results for the three network configurations. As can be seen, the probability of a dropped call is less than 10% for shadowing factors up to 0.01 (channel available 99% of the time). The probability rises steeply after this point, with over 80% of calls being dropped on the bent-pipe system, 50% being dropped in GTS on-board scenario and 40% being dropped in GSS on-board scenario when the satellite channel availability decreases to 90% of the time ($A = 0.1$).

4) *Analysis*: Overall these results show that the movement of traditional ground-segment functionality on-board a satellite only noticeably effects the network performance in terms of call blocking and call dropping probabilities at lower levels of satellite availability. At a satellite availability of 99% or greater, the GSS on-board and GTS on-board scenarios have slightly lower call blocking and call dropping probabilities than the bent-pipe scenario. It is at higher time-shadowing factors that the effects become prominent, where the GSS on-board scenario demonstrates a marked improvement over

Call Blocking Probability
Time-Shadowing Factor, $A = 0.002$



Call Blocking Probability
Number of SDCCH = 6

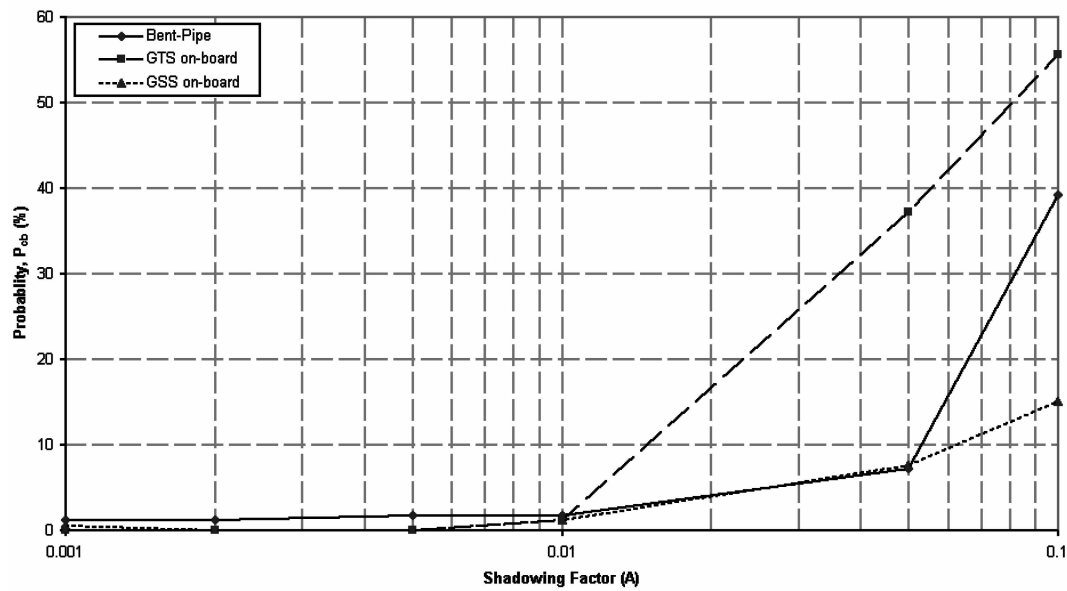


Fig. 12. Call blocking probability results showing how two factors can effect how likely it is that a call set-up request will not succeed (call blocked) either due to nonavailability of dedicated channels or due to loss of transmission on satellite link because of shadowing or fading of satellite channel. (a) Effect of number of SDCCH at fixed Lutz time-shadowing factor. (b) Effect of satellite availability as measured by Lutz time-shadowing factor at fixed number of SDCCH.

the bent-pipe scenario. The GTS on-board scenario, however, performs considerably worse than the other two, and again shows that this scenario is not really optimal.

There is obviously a compromise to be reached over the number of available SDCCH and the number of possible TCH in order to have the best chance of a call being accepted in the various GMR-2 based scenarios. This does depend on the other

traffic within the gateway, and the use of dynamic resource allocation is a topic for future investigation to optimize the channel assignments within a gateway. The satellite availability has a noticeable effect on the call blocking and call dropping probabilities. A satellite availability of 99% or greater (i.e., satellite is only not available 1% or less of time) gives a reasonably high chance that a call will be accepted, and once accepted will not be dropped. At lower

Probability of a Dropped Call
Number of SDCCH = 6

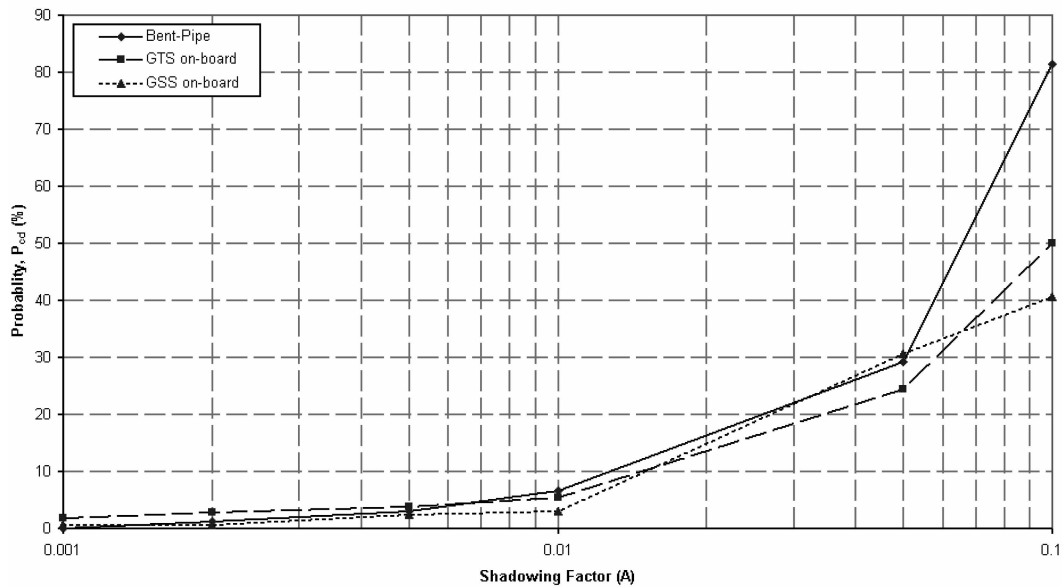


Fig. 13. Probability of dropped call results showing effect of satellite link availability (as measured by Lutz time-shadowing factor) on chances of call being dropped before user has finished.

TABLE III
Number of Timer Expirations

Number of SDCCH	Bent-Pipe		GTS on Board		GSS on Board	
	T.2_1	T310	T.2_1	T310	T.2_1	T310
2	4	0	7	0	14	0
4	0	0	0	1	1	0
6	0	2	0	0	0	0
8	0	1	0	2	0	3

availability values, the increase in the probability of a call being dropped or blocked in the first place is substantial. However, due to the fact that fewer signals are transmitted across the feeder link when the entire gateway functionality is on board, the probabilities in this scenario are lower than those for the scenario when all the gateway functionality is on the ground. The scenario where only part of the gateway functionality is on board the satellite has a greater risk of a call being dropped at lower satellite availability values as more messages are transmitted via the feeder links, so are affected by the satellite availability. The call dropping probability is less in this scenario, but this is due to the smaller number of successful calls in the first place.

VII. CONCLUSIONS

The results presented above have shown that there are some advantages in placing parts of the gateway on board the satellite in terms of call set-up delay and the probability of a call being blocked or dropped. One of the major findings is that it is better

to move the whole gateway on board (GSS on board scenario) as opposed to just some of the gateway, particularly when aiming to minimize the chances of a call being blocked. When the whole gateway is moved on board, the worst case call set-up time is considerably improved (4 s less than for bent-pipe scenario) for GMR-1 based systems. For the GMR-2 based systems, the effect is more noticeable in terms of the call-blocking probability, where placing the entire gateway on board the satellite reduces the call-blocking probability significantly at lower levels of satellite availability.

Moving functionality may improve call set-up times, and the chance of a call set-up succeeding and the subsequent call not being dropped. However, there are other factors that mitigate against such a transfer of functionality—in particular the issue of reliability. The trade-off between performance and reliability is a constant factor for designers and developers of satellite networks, and indeed of virtually any telecommunications network. The design of a telecommunications system always involves a series of compromises in order to provide the best available service to the user.

REFERENCES

- [1] Clarke, A. C. Extra terrestrial relays. *Wireless World*, (1945), 305–308.
- [2] Franchi, A., Howell, A., and Sengupta, J. Broadband mobile via satellite: INMARSAT BGAN. In *IEE Colloquium: Broadband Satellite: The Critical Success Factors: Technology, Services & Markets*, 2000, 23/1–23/7.

- [3] Watson, J., and Barton, M.
The ACeS air interface and system experience.
In *4th European Workshop on Mobile and Personal Satellite Communications* (EMPS 2000), 2000, 11–15.
- [4] Noerpel, A., and Goodings, R.
Geo-mobile satellite system air interface overview and performance.
In *4th European Workshop on Mobile and Personal Satellite Communications* (EMPS 2000), 2000, 6–10.
- [5] Eckhardt, G.
Guest editorial: On-board processing.
Space Communications, **17**, 1–3 (2001), 1–3.
- [6] Bergmann, N., and Dawood, A.
Reconfigurable computers in space: Problems, solutions and future directions.
In G. Brebner, K.-H. Brenner, H. E. Gindy, and H. Schmeck (Eds.), *Schlo Dagstuhl Seminar Report 00261: Dynamically Reconfigurable Architectures*, 2000, 4–5.
- [7] Bergsman, P.
Xilinx FPGA blasted into orbit.
Xcell Journal, 46 (Summer 2003). Available:
http://www.xilinx.com/publications/xcellonline/xcell46/xc.pdf/xc_satellite46.pdf.
- [8] SpaceDaily (July 24 2003)
Computer, heal thyself.
SpaceDaily. [Online]. Available:
<http://www.spacedaily.com/news/robot-03m.html>.
- [9] Evans, B. G.
Satellite Communication Systems (3rd ed.).
London: The Institution of Electrical Engineers, 1999.
- [10] Rosenburg, C., and Montpetit, M.-J.
Editorial: Special issue on the mini-conference:
Broadband satellite networking at IFIP networking 2000, Paris, May 2000.
International Journal of Satellite Communication Systems, **19**, 5 (2001), 425–427.
- [11] Mertzanis, I., Sfikas, G., Tafazolli, R., and Evans, B. G.
Satellite-ATM networking and call performance evaluation for multimedia broadband services.
International Journal of Satellite Communication Systems, **17**, 2 & 3 (1999), 107–127.
- [12] Yoshimura, N., and Kadowaki, N.
A simulation study for the performance of an on-board ATM switching scheme for broadband satellite communications network.
Space Communications, **17**, 1–3 (2001), 137–141.
- [13] Buerkle, W., and Trefz, M.
On-board switching architectures for multimedia satellite systems.
Space Communications, **17**, 1–3 (2001), 215–229.
- [14] Noeldeke, C. M.
Air interface and payload architecture for GEO multimedia communications satellites.
Space Communications, **17**, 1–3 (2001), 49–58.
- [15] Ors, T., and Rosenburg, C.
Providing IP QoS over GEO satellite systems using MPLS.
International Journal of Satellite Communication Systems, **19**, 5 (2001), 443–461.
- [16] Fenech, H. T., Pujante-Cuadrapani, A., and Lance, E.
EUTELSAT multimedia satellites.
Space Communications, **17**, 1–3 (2001), 145–164.
- [17] Le-Ngoc, T., and Elshabrawy, T.
Broadband satellite access for interactive multimedia services.
Space Communications, **17**, 1–3 (2001), 35–48.
- [18] GMR-1 01.202 v1.1.1
GEO-Mobile Radio Interface Specifications; Part 1: General Specifications; Sub-part 3: General System Description.
Mar. 2001, (ETSI TS 101 376-1-3 v1.1.1).
- [19] GMR-2 01.202 v1.1.1
GEO-Mobile Radio Interface Specifications; Part 1: General Specifications; Sub-part 3: General System Description.
Mar. 2001, (ETSI TS 101 377-1-3 v1.1.1).
- [20] GMPRS-1 03.064 v2.2.1
GEO-Mobile Radio Interface Specifications (Release 2); General Packet Radio Service; Part 3: Network specifications; Sub-part 22: Overall description of the GMPRS radio interface; Stage 2; GMPRS-1 03.064.
Mar. 2005, (ETSI TS 101 376-3-22).
- [21] Chan, P. M. L., Wyatt-Millington, R. A., Sheriff, R. E., Hu, Y. F., Conforto, P., and Tocci, C.
Design and evaluation of signaling protocols for mobility management in an integrated IP environment.
Computer Networks, **38**, 4 (2002), 517–530.
- [22] Kandus, G., Svigelj, A., and Mohoric, M.
The impact of different scheduling policies on traffic class dependent routing in intersatellite link networks.
International Journal of Satellite Communications and Networking, **22**, 5 (2004), 533–546.
- [23] GSM 04.08 v4.23.1
Digital cellular telecommunications system (Phase 2); Mobile radio interface; Layer 3 specification.
Oct. 1999, (ETS 300 557 v4.23.1).
- [24] GMR-1 04.008 v1.1.1
GEO-Mobile Radio Interface Specifications; Part 4: Radio interface protocol layer specifications; Sub-part 8: Mobile Radio Interface Layer 3 Specifications.
Feb. 2001, (ETSI TS 101 376 4-8 v1.1.1).
- [25] GMR-2 04.008 v1.1.1
GEO-Mobile Radio Interface Specifications; Part 4: Radio interface protocol specifications; Sub-part 8: Mobile Radio Interface Layer 3 Specifications.
Mar. 2001, (ETSI TS 101 377-4-7 v1.1.1).
- [26] Lutz, E., Werner, M., and Jahn, A.
Satellite Systems for Personal and Broadband Communications.
Berlin, Germany: Springer-Verlag, 2000.



Rosemary A. Wyatt-Millington (M'99) gained her M.Eng. in electrical and information sciences from the University of Cambridge in 1997, her M.Sc. in radio frequency and microwave engineering, and her Ph.D. in satellite communications from the University of Bradford in 1998 and 2004, respectively.

Following her M.Sc., she spent a year working as a guest researcher at the National Institute of Standards and Technology (NIST) in Gaithersburg, MD. Following her time at NIST, she worked for three years as a research assistant in the Department of Electronics and Telecommunications at the University of Bradford, followed by another two years as a teaching fellow in the School of Engineering, Design & Technology (EDT) at the same university. She was appointed as a Lecturer in EDT in September 2004. Her major research interests include the use of standard terrestrial protocols in satellite networks, IP mobility, and radio resource management within satellite systems.

Dr. Wyatt-Millington is a chartered engineer and a member of the Institute of Engineering and Technology in the UK.



Ray E. Sheriff received a B.Eng. in electrical and electronic engineering from the University of Leeds, UK in 1986 and his Ph.D. from the University of Bradford in 1995.

In 1991, he took up a lectureship with the University of Bradford, UK, and was appointed to the post of Professor of Electronic Engineering in 2002. His research interests include network protocol design and performance evaluation in the fields of mobile and satellite communications.

Y. Fun Hu received her 1st class B.Sc. honours degree in mathematical sciences and Ph.D. in information systems engineering from the University of Bradford, UK, in 1984 and 1988, respectively.

She is currently Professor of Wireless Communications Engineering in the School of Engineering, Design and Technology (EDT) at the University of Bradford, where she leads the Mobile Networks and Applications Group of the Mobile and Satellite Communications Research Centre (MSCRC). Since 1992, she has been actively involved in the development of 3G mobile communications systems through participation in numerous high profile EU framework programme projects. Her research interests include wireless network protocol design, mobility management, radio resource management, network security, wireless sensor networks and technologies. She was a UK delegate of the EU COST 253 (1996–2000), COST 256 (1997–2001) and COST 272 Actions (2001–2005). She was an executive member of the IEE Electronics and Communications Divisions Professional Network Group on Satellite Systems and Applications (2000–2002) and has been a member of the Technical Advisor Panel of the same professional network group since 2002.

