Radio Connection Management and Signalling Protocols for ATM via Satellites

Bin Fan

Submitted for the Degree of Doctor of Philosophy from the University of Surrey

UniS

Centre for Communication System Research
School of Electronic Engineering, Information Technology and Mathematics
University of Surrey
Guildford, Surrey GU2 5XH, UK

September 2000

© B. Fan 2000
ACKNOWLEDGEMENT

I could not have completed this work without the help and guidance from my two supervisors and many other people to whom I owe a debt of gratitude.

I would like to express my great appreciation and gratitude to Professor Barry G. Evans and Professor Rahim Tafazolli, my supervisors, not only for providing the opportunity for me to accomplish this work but also for their invaluable supervising, unfailing support and constant encouragement during my PhD study. I also thank them for being very kind, patient and understanding when I was being on my learning curve.

I am also grateful for the numerous discussions with Ioannis Mertzanis, Sieamak Vahid and Payam Taaghol on the SECOMS and ASSET projects; Narennhiran Kanagasabapathy and Bilal Ahmed on the satellite channel models; Lloyd Wood and Anthony Sammut on the satellite constellations and intersatellite links; Georgios Sfikas, Ilias Andrikopoulos and Tolga Örs on the BoNes simulation tools. My special thanks also go to John Hibbit, Mauro Fiacco and Stephane Villette for their kind help and assistance with UNIX at the beginning of my PhD study. In particular, I am indebted to Dr. Zhili Sun for reading my transfer report and conducting the examination during my transfer viva.

Special thanks are due to Moira Macmillan, Stephanie Evans, Lakshmi Chennell, Dave Brock, John Hibbit, Adam Kirby, Christopher Clark and Terry Roberts for their help, attention and support in many important and different ways.

My other friends and colleagues, Zhong Fan, Marita Gkizeli, George Aggelou, Nikolaos Dimitriou, Ruben Alvarez, Andreas-Albertos Agius, Hafeez M Aziz, Suji Gunaratne, Shyamalie Thilakawardana, Saied Abedi, Xinjie Yang, Abhaya Sumanasena, Klaus Moessner and many others deserve a pile of thanks for making my study and work here such an enjoyable and pleasant experience to cherish for many years to come.

The Committee of Vice-Chancellors and Principles (CVCP) and the Centre for Communication Systems Research are greatly acknowledged for awarding me the Scholarship/Studentship for my PhD study here.

At last but not at least, I wish to thank my husband, my parents and other family members for their unwavering love, concern and support over the years I have been here.
SUMMARY

The increasing demand for broadband multimedia services and their extension to mobiles has spurred provision via satellites. This is because satellite systems can provide a truly global coverage that cannot be economically realised by either fixed wireline systems or terrestrial mobile systems. Most of next generation satellite systems proposed, such as Astrolink, SPACEWAY and SkyBridge, consider using ATM or ATM based technology over satellite to provide broadband services requiring mobility and higher bandwidth. This thesis therefore concerns future broadband satellite networks that use on-board processing and intersatellite links to extend the terrestrial ATM provision to worldwide satellite systems.

This thesis discusses the air interface connection management and call handling schemes for integrated mobile/fixed ATM-satellite networks. The prime aim is to research an integrated solution that provides effective radio connection management and mobility support whilst maintaining the required QoS at both user terminals and gateway earth stations. We target at minimising the difference in performance between terrestrial ATM and ATM over satellite and providing mobility extension to the ATM protocols, whilst maintaining a high satellite channel efficiency and keeping as little as possible signalling modifications.

In this thesis, an efficient radio connection management scheme, which is designed for a QoS-provisioning transport of ATM traffic over satellite links, and a mobility-enhanced signalling protocol scheme for mobile ATM-satellite networks are proposed. Another large proportion of this thesis is devoted to the optimisations of multiple access and logic link control because these are the major factors that effecting the performance within ATM-satellite integrated systems. As a result, a semi-permanent signalling protocol, a reliability-dependent Selective Repeat Automatic Repeat reQuest (SR ARQ) and an adaptive timer SR ARQ are proposed. In addition to the proposed connection management scheme, a reservation meta-signalling for setting up signalling connections at the user-network radio interface and a mobility-enhanced call handling protocol derived from Q.2931 are proposed. Call control functions such as call routing, location update, paging, handover and authentication are discussed as well.

The proposed signalling protocol architecture provides a protocol reference model for ATM-satellite integrated systems. The verification and demonstration of the advantages of the semi-permanent signalling protocol, which offers a new method to improve the system channel efficiency on signalling connections, have been achieved. The proposed reliability-dependent
SR ARQ protocol provides a novel approach to optimise the transmission throughput to support a variety of traffic types with different QoS requirements in ATM-satellite systems. The proposed connection management scheme effectively manages the air interface connections for ATM services through diversifying connection types, establishing QoS-based connections and deploying an efficient connection mapping and control scheme.

The proposed radio connection management scheme together with the optimised Multiple Access Control (MAC) and ARQ provides a framework of interworking protocols for ATM over satellites. It can also find application in other similar systems that involve the integration of terrestrial protocols and mobile satellites. The research work that has been accomplished herein provides a solution and guidance to the design of signalling protocols for mobile satellite systems to implement ATM technology or indeed other future protocols.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAL</td>
<td>ATM Adaptation Layer</td>
</tr>
<tr>
<td>ABR</td>
<td>Available Bit Rate</td>
</tr>
<tr>
<td>AC</td>
<td>Authentication Centre</td>
</tr>
<tr>
<td>ACH</td>
<td>Acquisition Channel</td>
</tr>
<tr>
<td>ACK</td>
<td>Acknowledgement</td>
</tr>
<tr>
<td>ACKD</td>
<td>ACK with Discarding</td>
</tr>
<tr>
<td>ACKR</td>
<td>ACK with Receiving</td>
</tr>
<tr>
<td>ACTS</td>
<td>Advanced Communication Technologies and Services</td>
</tr>
<tr>
<td>ARQ</td>
<td>Automatic Repeat Request</td>
</tr>
<tr>
<td>ASSET</td>
<td>ACTS Satellite Switching End-to-end Trials</td>
</tr>
<tr>
<td>ATM</td>
<td>Asynchronous Transfer Mode</td>
</tr>
<tr>
<td>ATSR</td>
<td>Adaptive Timer Selective Repeat</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
</tr>
<tr>
<td>B-ISDN</td>
<td>Broadband ISDN</td>
</tr>
<tr>
<td>CAC</td>
<td>Call/Connection Admission Control</td>
</tr>
<tr>
<td>CBR</td>
<td>Constant Bit Rate</td>
</tr>
<tr>
<td>CDTV</td>
<td>Cell delay Variation Tolerance</td>
</tr>
<tr>
<td>CDV</td>
<td>Cell Delay Variation</td>
</tr>
<tr>
<td>CER</td>
<td>Cell Error Rate</td>
</tr>
<tr>
<td>CI</td>
<td>Connection Identifier</td>
</tr>
<tr>
<td>CLD</td>
<td>Changing Link Distance</td>
</tr>
<tr>
<td>CLNS</td>
<td>Connectionless Network Services</td>
</tr>
<tr>
<td>CLP</td>
<td>Cell Loss Priority</td>
</tr>
<tr>
<td>CLR</td>
<td>Cell Loss Rate</td>
</tr>
<tr>
<td>CM</td>
<td>Connection Management</td>
</tr>
<tr>
<td>CMA</td>
<td>Connection Mapping Algorithm</td>
</tr>
<tr>
<td>CMR</td>
<td>Cell Misinsertion Rate</td>
</tr>
<tr>
<td>CMS</td>
<td>Connection Management Scheme</td>
</tr>
<tr>
<td>CPCS</td>
<td>Common Part Convergence Sublayer</td>
</tr>
<tr>
<td>CRC</td>
<td>Cyclic Redundancy Check</td>
</tr>
<tr>
<td>CRT</td>
<td>Conventional Reliable Transmission</td>
</tr>
<tr>
<td>CS</td>
<td>Convergence Sublayer</td>
</tr>
<tr>
<td>CSC</td>
<td>Common Signalling Channel</td>
</tr>
<tr>
<td>CTD</td>
<td>Cell Transfer Delay</td>
</tr>
<tr>
<td>DAMA</td>
<td>Demand Assignment Multiple Access</td>
</tr>
<tr>
<td>DCCH</td>
<td>Dedicated Control Channel</td>
</tr>
<tr>
<td>DL</td>
<td>DownLink</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>DLA</td>
<td>Dynamic Location Area</td>
</tr>
<tr>
<td>DLC</td>
<td>Data Link Control</td>
</tr>
<tr>
<td>DQPSK</td>
<td>Differential Quadrature Phase Shift Keying</td>
</tr>
<tr>
<td>DT</td>
<td>Discarding Timer</td>
</tr>
<tr>
<td>DVE</td>
<td>Distance Varying End-to-end connection</td>
</tr>
<tr>
<td>EFCI</td>
<td>Explicit Feedback Congestion Indication</td>
</tr>
<tr>
<td>EIA</td>
<td>Electronic Industry Association</td>
</tr>
<tr>
<td>EIA/TIA IS-41</td>
<td>EIA/TIA Interim Standard 41</td>
</tr>
<tr>
<td>ETSI</td>
<td>European Telecommunications Standards Institute</td>
</tr>
<tr>
<td>FA</td>
<td>Fixed Assignment</td>
</tr>
<tr>
<td>FDMA</td>
<td>Frequency Division Multiple Access</td>
</tr>
<tr>
<td>FEC</td>
<td>Forward Error Control</td>
</tr>
<tr>
<td>FES</td>
<td>Fixed Earth Station</td>
</tr>
<tr>
<td>FIFO</td>
<td>First In First Out</td>
</tr>
<tr>
<td>FLA</td>
<td>Fixed Location Area</td>
</tr>
<tr>
<td>FLD</td>
<td>Fixed Link Distance</td>
</tr>
<tr>
<td>GBN</td>
<td>Go-Back-N</td>
</tr>
<tr>
<td>GCRA</td>
<td>Generic Cell Rate Algorithm</td>
</tr>
<tr>
<td>GEO</td>
<td>Geosynchronous Earth Orbit</td>
</tr>
<tr>
<td>GFC</td>
<td>Generic Flow Control</td>
</tr>
<tr>
<td>GII</td>
<td>Global Information Infrastructure</td>
</tr>
<tr>
<td>GoS</td>
<td>Grade of Services</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile</td>
</tr>
<tr>
<td>GTW</td>
<td>GaTeWay</td>
</tr>
<tr>
<td>HDLC</td>
<td>High-level Data Link Control</td>
</tr>
<tr>
<td>HEC</td>
<td>Header Error Control</td>
</tr>
<tr>
<td>HEO</td>
<td>Highly Elliptic Orbit</td>
</tr>
<tr>
<td>HLR</td>
<td>Home Location Register</td>
</tr>
<tr>
<td>HO</td>
<td>Handover</td>
</tr>
<tr>
<td>IBRR</td>
<td>In-Band Resource Request</td>
</tr>
<tr>
<td>ICC</td>
<td>Inclined Circular Constellation</td>
</tr>
<tr>
<td>IID</td>
<td>Independent and Identically Distributed</td>
</tr>
<tr>
<td>ISDN</td>
<td>Integrated Services Digital Network</td>
</tr>
<tr>
<td>ISL</td>
<td>Intersatellite Links</td>
</tr>
<tr>
<td>ITU-T</td>
<td>International Telecommunications Union – Telecommunication standardisation sector</td>
</tr>
<tr>
<td>IWF</td>
<td>InterWorking Functions</td>
</tr>
<tr>
<td>IWU</td>
<td>InterWorking Unit</td>
</tr>
<tr>
<td>LA</td>
<td>Location Area</td>
</tr>
<tr>
<td>LAN</td>
<td>Local Area Network</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
</tr>
<tr>
<td>LAP</td>
<td>Link Access Procedure</td>
</tr>
<tr>
<td>LAPB</td>
<td>Link Access Procedure Balanced</td>
</tr>
<tr>
<td>LAPD</td>
<td>Link Access Procedure for the ISDN D channel</td>
</tr>
<tr>
<td>LAPDm</td>
<td>Link Access Procedure for the ISDN Dm channel</td>
</tr>
<tr>
<td>LAPM</td>
<td>Link Access Procedure for Modems</td>
</tr>
<tr>
<td>LEO</td>
<td>Low Earth Orbit</td>
</tr>
<tr>
<td>LES</td>
<td>Local Exchange Services</td>
</tr>
<tr>
<td>LHOM</td>
<td>Local Handover Management</td>
</tr>
<tr>
<td>LIJ</td>
<td>Leaf Initiated Join</td>
</tr>
<tr>
<td>LLC</td>
<td>Logical Link Control</td>
</tr>
<tr>
<td>LM</td>
<td>Location Management</td>
</tr>
<tr>
<td>LU</td>
<td>Location Update</td>
</tr>
<tr>
<td>MT</td>
<td>Mobile Terminated</td>
</tr>
<tr>
<td>MAC</td>
<td>Multiple Access Control</td>
</tr>
<tr>
<td>MCR</td>
<td>Minimum Cell Rate</td>
</tr>
<tr>
<td>MCS</td>
<td>Master Control Station</td>
</tr>
<tr>
<td>MEO</td>
<td>Medium Earth Orbit</td>
</tr>
<tr>
<td>M-F</td>
<td>Mobile user-to-Fixed user</td>
</tr>
<tr>
<td>MF-TDMA</td>
<td>Multi-Frequency TDMA</td>
</tr>
<tr>
<td>MM</td>
<td>Mobility Management</td>
</tr>
<tr>
<td>MO</td>
<td>Mobile Origination</td>
</tr>
<tr>
<td>MSREJ</td>
<td>Multiple SREJ</td>
</tr>
<tr>
<td>MT</td>
<td>Mobile Terminal</td>
</tr>
<tr>
<td>MTP</td>
<td>Message Transfer Part</td>
</tr>
<tr>
<td>NACK</td>
<td>Non-ACKnowledgement</td>
</tr>
<tr>
<td>N-ISDN</td>
<td>Narrowband ISDN</td>
</tr>
<tr>
<td>NNI</td>
<td>Network-Network Interface</td>
</tr>
<tr>
<td>NOC</td>
<td>Network Operation Centre</td>
</tr>
<tr>
<td>NRM</td>
<td>Normal Response Mode</td>
</tr>
<tr>
<td>OAM</td>
<td>Operation And Management</td>
</tr>
<tr>
<td>OBS</td>
<td>On-Board Switch</td>
</tr>
<tr>
<td>OBP</td>
<td>On-Board Processing</td>
</tr>
<tr>
<td>OBRR</td>
<td>Out-of-Band Resource Request</td>
</tr>
<tr>
<td>O/G</td>
<td>Mobile Originated</td>
</tr>
<tr>
<td>OLN</td>
<td>Overlying Network</td>
</tr>
<tr>
<td>P/L</td>
<td>(Satellite) Payload</td>
</tr>
<tr>
<td>PA</td>
<td>Paging Area</td>
</tr>
<tr>
<td>PB</td>
<td>Piggyback</td>
</tr>
<tr>
<td>PC</td>
<td>Polar Constellation</td>
</tr>
<tr>
<td>PCR</td>
<td>Peak Cell Rate</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PCN</td>
<td>Personal Communications Network</td>
</tr>
<tr>
<td>PDH</td>
<td>Plesiochronous Digital Hierarchy</td>
</tr>
<tr>
<td>PDU</td>
<td>Protocol Data Unit</td>
</tr>
<tr>
<td>PER</td>
<td>Packet Error Rate</td>
</tr>
<tr>
<td>PHY</td>
<td>PHYsical layer</td>
</tr>
<tr>
<td>PSP</td>
<td>Permanent Signalling Protocol</td>
</tr>
<tr>
<td>PSVCI</td>
<td>Permanent Signalling VCI</td>
</tr>
<tr>
<td>PTM</td>
<td>Point-to-Multipoint</td>
</tr>
<tr>
<td>PTP</td>
<td>Point-to-Point</td>
</tr>
<tr>
<td>PVC</td>
<td>Permanent Virtual Connections</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature Phase Shift Keying</td>
</tr>
<tr>
<td>RA</td>
<td>Random Access</td>
</tr>
<tr>
<td>RASC</td>
<td>Random Access Signalling Channel</td>
</tr>
<tr>
<td>RASCI</td>
<td>Random Access Signalling Channel</td>
</tr>
<tr>
<td>RCL</td>
<td>Radio Control Layer</td>
</tr>
<tr>
<td>RDSR</td>
<td>Reliability-Dependent Selective Repeat</td>
</tr>
<tr>
<td>RMS</td>
<td>Reservation Meta-Signalling</td>
</tr>
<tr>
<td>RRC</td>
<td>Resource Reservation Channel</td>
</tr>
<tr>
<td>RRT</td>
<td>Relaxed Reliable Transmission</td>
</tr>
<tr>
<td>RS</td>
<td>Reed Solomon</td>
</tr>
<tr>
<td>RT</td>
<td>Retransmission Timeout</td>
</tr>
<tr>
<td>RTD</td>
<td>Round Trip Delay</td>
</tr>
<tr>
<td>Rt-VBR</td>
<td>Real-time Variable Bit Rate</td>
</tr>
<tr>
<td>SAAL</td>
<td>Signalling AAL</td>
</tr>
<tr>
<td>SABM</td>
<td>Set Asynchronous Balanced Mode</td>
</tr>
<tr>
<td>SAR</td>
<td>Segmentation And Reassembly sublayer</td>
</tr>
<tr>
<td>SAT</td>
<td>Satellite-ATM Terminal</td>
</tr>
<tr>
<td>SAW</td>
<td>Stop-and-Wait</td>
</tr>
<tr>
<td>SDH</td>
<td>Synchronous Digital Hierarchy</td>
</tr>
<tr>
<td>SECOMS</td>
<td>SECOMS Satellite EHF Communication for Mobile Multimedia Services</td>
</tr>
<tr>
<td>SMS</td>
<td>Short Message Service</td>
</tr>
<tr>
<td>SONET</td>
<td>Synchronous Optical NETwork</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SPSP</td>
<td>Semi-Permanent Signalling Protocol</td>
</tr>
<tr>
<td>SR</td>
<td>Selective Repeat</td>
</tr>
<tr>
<td>SREJ</td>
<td>Selective Reject</td>
</tr>
<tr>
<td>SSCF</td>
<td>Service Specific Co-ordination Function</td>
</tr>
<tr>
<td>SSCOP</td>
<td>Service Specific Co-ordination Oriented Protocol</td>
</tr>
<tr>
<td>SVC</td>
<td>Switched Virtual Connections</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>TC</td>
<td>Transmission Convergence</td>
</tr>
<tr>
<td>TCH</td>
<td>Traffic Channel</td>
</tr>
<tr>
<td>TCP/IP</td>
<td>Transmission Control Protocol/Internet Protocol</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time Division Multiple Access</td>
</tr>
<tr>
<td>TDD</td>
<td>Time Division Duplex</td>
</tr>
<tr>
<td>TDM</td>
<td>Time Division Multiplex</td>
</tr>
<tr>
<td>TIA</td>
<td>Telecommunication Industry Association</td>
</tr>
<tr>
<td>TRM</td>
<td>Traffic Resource Manager</td>
</tr>
<tr>
<td>UBR</td>
<td>Unspecified Bit Rate</td>
</tr>
<tr>
<td>UL</td>
<td>UpLink</td>
</tr>
<tr>
<td>UNI</td>
<td>User-Network Interface</td>
</tr>
<tr>
<td>UPC</td>
<td>Usage Parameters Control</td>
</tr>
<tr>
<td>UT</td>
<td>Unreliable/Unassured Transmission</td>
</tr>
<tr>
<td>VBR</td>
<td>Variable Bit Rate</td>
</tr>
<tr>
<td>VC</td>
<td>Virtual Connection</td>
</tr>
<tr>
<td>VCI</td>
<td>Virtual Connection Identifier</td>
</tr>
<tr>
<td>VLR</td>
<td>Visitor Location Register</td>
</tr>
<tr>
<td>VP</td>
<td>Virtual Path</td>
</tr>
<tr>
<td>VPI</td>
<td>Virtual Path Identifier</td>
</tr>
<tr>
<td>WZ</td>
<td>Window Size</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION ................................................................................................................... 1

1.1 Background ........................................................................................................................................... 1

1.2 Motivation .............................................................................................................................................. 4

1.3 Research Objective ............................................................................................................................... 6

1.4 Research Approach ............................................................................................................................... 6

1.5 Original Work & Major Achievements ................................................................................................. 7

1.6 Thesis Structure ...................................................................................................................................... 9

CHAPTER 2 MOBILE SATELLITE SYSTEM CHARACTERISTICS AND ATM OVERVIEW ....................................................................................................................... 11

2.1 Mobile Satellite System Characteristics ............................................................................................ 11
   2.1.1 Satellite Constellation .................................................................................................................... 11
       2.1.1.1 Types of Constellations .............................................................................................................. 11
       2.1.1.2 Constellation Design ................................................................................................................ 12
       2.1.1.3 Current Satellite Constellations ................................................................................................. 17
   2.1.2 Satellite Multi-beam Coverage ...................................................................................................... 18
   2.1.3 Intersatellite Links .......................................................................................................................... 18
   2.1.4 Satellite Channel ............................................................................................................................ 20

2.2 ATM Introduction ................................................................................................................................. 23
   2.2.1 ATM Protocol Layers ................................................................................................................... 24
       2.2.1.1 Physical Layer .......................................................................................................................... 24
       2.2.1.2 ATM Layer ............................................................................................................................. 25
       2.2.1.3 ATM Adaptation Layer ............................................................................................................ 25
   2.2.2 ATM Traffic Classes and QoS Provisioning .................................................................................. 26
   2.2.3 Traffic Management ......................................................................................................................... 27
   2.2.4 Statistical Multiplexing and ATM Connection ............................................................................. 28
   2.2.5 Call Admission Control .................................................................................................................. 28
   2.2.6 ATM UNI Signalling ....................................................................................................................... 28
       2.2.6.1 Meta-signalling ....................................................................................................................... 29
       2.2.6.2 Q.2931 Protocol ...................................................................................................................... 30

2.3 Conclusion ............................................................................................................................................. 33

CHAPTER 3 ISSUES OF ATM OVER SATELLITE ......................................................................................... 35

3.1 Introduction ........................................................................................................................................... 35

3.2 Integration Difficulties .......................................................................................................................... 35
5.2.1 Frame Structure ................................................................. 80
5.2.2 Request Methods ................................................................. 82
5.2.3 Scheduling Algorithm ......................................................... 83
5.2.4 Service Adaptability ............................................................ 83

5.3 Optimisation of MAC Protocol for Signalling Traffic ......................................................... 84
5.3.1 Semi-permanent Signalling Protocol ......................................................... 84
5.3.2 Evaluation Approach ............................................................... 87
5.3.3 Performance Parameters .......................................................... 88
5.3.4 Simulation Results of the Performance in an Error-Free Channel .................................. 89
  5.3.4.1 Simulation Results and Discussions .............................................. 90
  5.3.4.2 Simulation Discussion ................................................................. 94
5.3.5 Simulation Results of the Performance with the Presence of Channel Fading ............. 95
  5.3.5.1 Simulation Results and Discussion .............................................. 95
  5.3.5.2 Simulation Discussion ................................................................. 99
5.3.6 Conclusion ........................................................................ 99

CHAPTER 6 OPTIMISATION OF ARQ PROTOCOL ................................................................. 101

6.1 Introduction ........................................................................ 101

6.2 SR ARQ Protocols ................................................................ 102
  6.2.1 Protocol Description .............................................................. 102
  6.2.2 Packet Numbering ................................................................. 103
  6.2.3 Protocol Efficiency ................................................................. 104

6.3 Reliability-Dependent SR ARQ Protocol .............................................. 105
  6.3.1 Scheme Description ................................................................. 105
    6.3.1.1 Relaxed Reliable Transmission ............................................... 107

6.4 Adaptive Timer SR ARQ ............................................................ 109
  6.4.1 Problem Identification ............................................................. 110
  6.4.2 ATSR ARQ Protocol Description ............................................. 111
  6.4.3 Evaluation Method ................................................................. 112
  6.4.4 Simulation Results and Discussion ........................................... 114

6.5 The Architecture of ARQ Protocol for ATM-satellite Systems ......................... 120

6.6 Conclusions ........................................................................ 125

CHAPTER 7 A QOS-PROVISIONING CONNECTION MANAGEMENT SCHEME .... 127

7.1 Introduction ........................................................................ 127

7.2 System Model ..................................................................... 128

7.3 The Proposed Connection Management Scheme ....................................................... 129
  7.3.1 Connection Definition and Diversification ..................................... 129
  7.3.2 Connection Mapping Scheme .................................................. 130
    7.3.2.1 Mapping the Connection Mode ............................................ 131
LIST OF FIGURES

Figure 1-1 ATM-satellite integration issues.................................................................4
Figure 1-2 Thesis structures..........................................................................................9
Figure 2-1 Major elliptical orbital parameters............................................................13
Figure 2-2 Orbital parameters in a geocentric-equational coordinate system.............13
Figure 2-3 Satellite look angle (a) Elevation...............................................................14
Figure 2-4 Satellite look angle (b) Azimuth.................................................................14
Figure 2-5 Street of coverage method...........................................................................16
Figure 2-6 Number of satellites N versus angular coverage [BES78].........................16
Figure 2-7 Spherical triangle method..........................................................................17
Figure 2-8 A typical multi-beam satellite’s spotbeam pattern....................................18
Figure 2-9 Iridium’s inter-plane and intra-plane ISL..................................................19
Figure 2-10 Iridium’s ISL configuration. (N: North; S: South).................................20
Figure 2-11 Zorzi’s first-order Markov model............................................................22
Figure 2-12 ATM signalling interfaces (TE: Terminal Equipment).........................24
Figure 2-13 B-ISDN protocol reference model..........................................................24
Figure 2-14 ATM signalling protocol stack...............................................................29
Figure 2-15 Peer-to-peer communications across the UNI (SAP: service access point) 30
Figure 2-16 Signalling procedure for point-to-point call set-up..................................31
Figure 2-17 Point-to-Multipoint call set-up...............................................................32
Figure 2-18 Leaf joins an active LIJ call.................................................................32
Figure 2-19 Leaf joins an inactive LIJ call...............................................................33
Figure 3-1 Handover rate per user versus satellite spotbeam....................................38
Figure 3-2 Handover rate per user versus mean terminal arrival rate.......................38
Figure 3-3 Length of intra-plane ISL (ISL96) and inter-plane ISL (ISL94) during one orbit period....40
Figure 3-4 Major ATM-satellite integration issues.....................................................41
Figure 3-5 Performance comparison between slotted Aloha and pure Aloha [PAH95] 43
Figure 3-6 Throughput performance of slotted Aloha in the presence of channel fading with reference to the maximum allowed packet repeat times. 45
Figure 3-7 Delay performance of slotted Aloha in the presence of channel fading with reference to the maximum allowed packet repeat times..................................................45
Figure 3-8 Throughput performance of slotted Aloha in the presence of channel fading with reference to the backlog timeslots.................................................................46
Figure 3-9 Delay performance of slotted Aloha in the presence of channel fading with reference to the backlog timeslots.................................................................47
Figure 3-10 Throughput efficiency comparisons between ARQ protocols..................51
Figure 3-11 Delay and throughput of SR ARQ versus average bit error rate on a satellite channel...52
Figure 3-12 Delay and throughput of SR ARQ versus window size on a satellite channel. 54
Figure 3-13 Multi-level location areas.......................................................................56
Figure 3-14 Concepts of FLA and DLA.................................................................57
Figure 3-15 Relationship between paging area and location area............................58
Figure 4-1 Two approaches for ATM-Satellite interconnection (MT: mobile terminal) 64
Figure 4-2 Interworking between SECOMS with B-ISDN (IWU: interworking unit).....65
Figure 4-3 ATM-satellite system architecture.........................................................66
Figure 4-4 Options of ATM protocol architecture....................................................70
Figure 4-5 Proposed ATM-satellite protocol reference model....................................70
Figure 4-6 Network interfaces..................................................................................71
Figure 4-7 ATM-satellite protocol architecture—Control plane...............................73
Figure 4-8 ATM-satellite protocol architecture—User plane......................................73
Figure 4-9 ATM-satellite protocol architecture—Mobility plane...............................74
Figure 4-10 SECOMS channel frame structure.......................................................75
Figure 4-11 Extended ATM packet structure for transmission on mobile satellite link....77
Figure 5-1 The basic frame structure of the reservation/demand MAC protocol........81
Figure 5-2 Major message elements for RRC, RAC, FC and TCH channels................82
Figure 5-3 Illustration of the PSP protocol...............................................................85

xiv
Figure 8-8 Mobile originated call set-up ................................................................. 156
Figure 8-9 Mobile terminated call set-up .............................................................. 157
# List of Tables

Table 2-1 Commonly used constellation parameters. ......................................................... 12  
Table 2-2 QoS parameters.................................................................................................. 27  
Table 2-3 Traffic descriptors............................................................................................ 27  
Table 3-1 System parameters related to the constellation.................................................. 40  
Table 4-1 End-to-end connection supported in ATM-satellite network. (S: Support, X: Not support.) 68  
Table 4-2 Definitions and protocol functions of ATM-satellite network interfaces...................... 72  
Table 5-1 Simulation parameters....................................................................................... 95  
Table 6-1 Four intermediate satellites’ position at t = 0...................................................... 113  
Table 6-2 Simulation parameters ..................................................................................... 115  
Table 6-3 Comparison of packet loss ratio........................................................................ 119  
Table 7-1 Connection mapping scheme.............................................................................. 135
Chapter 1 Introduction

1.1 Background

Today in the telecommunication field, three major trends can be observed: the evolution towards fixed broadband networks, the development of the third generation terrestrial mobile communications and the emergence of mobile satellite communications.

The evolution of the fixed network from the Integrated Services Digital Network (ISDN) to the Broadband ISDN (B-ISDN) aims to provide all existing and emerging telecommunication services within one network. After initial efforts to bring about a single infrastructure for voice, data and video transmissions under the umbrella of ISDN had failed due to the essential limitation of circuit switching, a new technology, Asynchronous Transfer Mode (ATM), that enables a packet-switched, bandwidth-on-demand services for video and data transmissions, was developed. ATM technology has been the key technology in turning the B-ISDN vision into reality in the last decades. The prospect of ATM and B-ISDN depends on commercial aspects such as the availability of reasonably priced equipment, the consumers’ interest in the new services and the costs of maintaining and upgrading existing systems. However, it is widely envisaged that ATM will play a major role in the future telecommunications [PRY93].

Telecommunications has also rapidly extended in the mobile area. Cellular radio technology has enjoyed explosive growth worldwide to meet the demand for personal mobility that cannot be provided by the fixed network. This expansion has been stimulated both by technological advances and by increased demands for mobile services. Advances in semiconductor technology allow low cost, lightweight and portable units to come onto the market, fuelling the increase of land mobile services. Along with voice service, mobile systems also have extended data transmission capability to offer features such as fax, short message service and paging. Since the first cellular mobile telephone networks were introduced in the early 1980s, growth in the number of subscribers has accelerated rapidly in both the most developed and developing countries. Furthermore, progress from the second-generation to third-generation mobile systems that offer higher capacities, more services, better performance and higher cost-effectiveness, has occurred.

However, both fixed and terrestrial mobile systems cannot economically provide services in such areas as seas, mountains and forests without using satellites, which can extend coverage to form a truly global coverage. The first satellite constellation consisting of three satellites
was proposed by Arthur C. Clarke in 1945 and the first satellite SPUTNIK was successfully launched into orbit in 1957 by the former Soviet Union. Satellite communication technologies thenceforth have become part of people's everyday life. The old generation Geosynchronous Earth Orbit (GEO) satellite systems such as INTELSAT and INMARSAT have been in operation for about three decades and have been used for many purposes including weather forecasting, navigation, reconnaissance and communications [COM97].

In the 1980's, mobile-satellite service technology advanced from initial concepts to practical system design and service implementation. Since then there has been implementation of regional and domestic GEO mobile satellites. Due to the long propagation delay associated with GEO satellites, Low Earth Orbit/Medium Earth Orbit (LEO/MEO) satellites have recently attracted much interest for their low-delay transmission features and characteristics that can better support real-time services. LEO/MEO satellites have a relatively short lifetime and disadvantage of complicated satellite tracking and operation due to dynamic satellite movement. Both GEO and MEO/LEO have their own distinctive advantages and both will contribute to future global communications.

In recent years, driven by converging technological and economic forces and demands for global access, connectivity and more service types, the competition to build global satellite networks to deliver mobile services has been intense [STU97]. A number of LEO and MEO systems such as IRIDIUM (now deceased) and GLOBALSTAR have been developed, and are being conceived by other telecom players such as ICO-Teledesic to provide space-based telecommunications.

Motorola's IRIDIUM system was conceived in 1987, and is the very first private global wireless communication system that utilises small handheld terminals to provide voice services along with fax, paging and low rate data transmission services over satellite. The IRIDIUM system was completely deployed in May 1998. It was expected to provide cellular-like services in the areas where terrestrial cellular service is unavailable, or where the Public Switched Telephone Network (PSTN) is not well developed. IRIDIUM represents one of the largest ever global service launches. Its launch implies the advance of the personal satellite communications.

Unfortunately, IRIDIUM was finally deceased after two years since its first launch. Telecommunication analysts pointed out a variety of factors including poor marketing campaign, limited in-building usage, high cost and undesirable handsets (large, heavy, expensive and short battery life). IRIDIUM also suffered from the threats of the continued roll out of improved cellular GSM networks and the introduction of the third generation mobile
systems. It is perceived that one of the big design strategy failures is that IRIDIUM only offered enough bandwidth for voice and paging services but failed to offer the adaptability towards the next generation broadband satellite systems [ROB00]. However, the latest telecom space race between companies like INMARSAT, ICO-Teledesic and Globalstar is even more ambitious than IRIDIUM’s plan to provide the broadband telecom services of much higher capacity and wider breadth.

The strong demand for a broadband communication system that will provide voice, data, internet, video and other existing and emerging multimedia services to anyone anywhere with worldwide connectivity and mobility has spurred the need for an integration of terrestrial and satellite networks. Satellite networks are expected to interface with terrestrial networks at high data rates and also provide networking access to a variety of users directly. Research and development on satellite and terrestrial network integration are increasing at an unprecedentedly fast pace.

Within satellite networks, the current technology trend is to use Internet Protocols (IP) and ATM to carry a variety of services.

The consideration of using IP is because of the growth of Internet networks and the increasing usage of the Internet services on the ground. TCP/IP was originally designed for data transmission. The advantages of simplicity, scalability and universality were the keys of its success. Visioning the multimedia communication future, extensive research and development in modification and enhancement of TCP/IP is being pursued to satisfy new and emerging needs such as supporting voice, video-conferencing and providing security and mobility [JAI98].

However, TCP/IP was developed without taking into consideration its performance over long-delay satellite links, with the result that extensive efforts are now underway to extend its functionality with added features for the satellite environment such as larger TCP windows, selective acknowledgement and better congestion avoidance [NET00].

Compared to IP, ATM was designed to handle real time voice and video, in addition to more conventional computer data. It is capable of providing services to all traffic types. It has advantages of offering a wide range of QoS guarantees and high bandwidth utilisation by deploying the bandwidth on demand. Because of these advantages, ATM is considered to play a major role in the future telecommunications.
The development of ATM inspired the need to create integrated systems connecting ATM and satellites for the next generation satellite telecommunications [AKY97]. Most of the next generation broadband satellite systems proposed, such as Astrolink, SPACEWAY, SkyBridge, EuroSkyWay, Celestri, consider using ATM or ATM based technology over satellite to provide services requiring mobility or/and higher bandwidth. As Internet and private intranet applications will also be carried over integrated ATM-satellite network, some research projects are carried out to investigate the IP over ATM-satellite networks [GOY99]. In these proposed satellite systems, ATM is used to provide at least packet transport layer services.

Furthermore, the new advanced satellite systems, provided with on-board switching, processing and intersatellite links, provide a better environment for extending ATM technology. Seeking a seamless integration of terrestrial ATM and satellites whilst providing a required QoS, has therefore become a major topic for the current research and development.

1.2 Motivation

As the user mobility and air interface are to be merged into an integrated system, areas for immediate considerations in ATM via satellite include many new aspects. From a signalling protocol design point of view, these issues can be divided into two categories, the radio connection management related issues and mobility enhancement related issues, as shown in Figure 1-1.

![Figure 1-1: ATM-satellite integration issues](image)

The radio connection management related protocols are expected to be placed below the ATM layer to provide radio connection establishment, maintenance and release services to the
ATM layer and to ensure that the ATM layer required QoS can be satisfied. Protocol issues such as wireless multiple access, logic link control, connection management and resource allocation need to be addressed. Meanwhile, the satellite channel efficiency is also required to be maintained. This requires that the radio connection management related protocols, such as Multiple Access Control (MAC) and Logic Link Control (LLC), need to be optimised to adapt to different ATM services types.

As many proposed satellite systems use ATM as the packet transport protocol, QoS provisioning for the ATM layer is especially important. A successful seamless integration of terrestrial ATM and satellites relies on a successful maintenance of QoS comparative to that of terrestrial ATM, which offers a wide range of quality levels from guaranteed, best effort, to non-guaranteed QoS [BHA97]. Terrestrial ATM succeeds in achieving a high protocol efficiency by taking advantage of the low bit error fibre transmission media and using techniques such as statistical multiplexing, traffic contract and Virtual Path Identifier/Virtual Connection Identifier (VPI/VCI) switching. But it does not support user mobility and has some protocol function simplicities such as no link-by-link error control performed for 48 byte cell payloads at the physical and ATM layer. In satellite systems, factors such as satellite channel impairments, user mobility, long propagation delay and LEO/MEO constellation dynamics degrade the system performance at the radio interface. These factors impact on the transportation of ATM traffic over satellites and on the connection management related functionalities at the air interface. There is thus a difference in performance between terrestrial ATM and ATM over satellite. Such discrepancies can however be resolved by provision of new connection management scheme at the radio interface.

The main mobility related protocol functions are incorporated into the higher layer of the ATM protocol. They are mainly required to provide mobility and security services. Some additional signalling may be added to serve the mobile specific services such as short message. Issues such as location update, handover, call routing and call establishment need to be addressed.

In order to address the above-mentioned issues, in this work, we investigate a fully integrated, mobility enhanced signalling protocol architecture for ATM-satellite networks, which provides mobility support to the user terminals. Our focus will be on designing an air interface connection management scheme and optimising the MAC and LLC protocols in order to flexibly handle different ATM services types and achieve satellite channel efficiency. On the LLC layer, we optimise the Automatic Repeat reQuest (ARQ) that has direct impact on the QoS of ATM layer.
1.3 Research Objective

Based on the new requirements of satellite-ATM network in this research, we target at proposing a mobility-enhanced and fully integrated signalling and protocol architecture for ATM-satellite networks and designing an efficient radio connection management scheme (CMS) for provisioning a QoS-guaranteed transport of ATM traffic over satellite links. Because CMS utilises services provided by MAC and LLC protocols, MAC and LLC protocols firstly need to be optimised in order to achieve a truly effective CMS and satellite channel efficiency. Four basic criteria, provisioning of required QoS, maximum system capacity, system implementation simplicity and minimum modification of the ATM, are the main considerations during this work.

The eventual system should enable the network to handle different ATM services according to their service types and QoS requirements and keep unused bandwidth on a connection to a minimum. The designed air interface ATM-satellite signalling and protocols should be made efficient and harmonised with the ATM protocol. Extended mobility functions and satellite link related MAC and LLC sub-layers need to be incorporated into the existing ATM layers. Call control functions need to be promoted from the existing ATM signalling. Signalling implementation issues including routing, location management, handover and signalling procedures need also to be addressed [JAI97].

1.4 Research Approach

A survey was first conducted to study protocol issues and protocol modification/enhancement proposals on the integration of ATM and satellites in the wireless related, e.g. [CHI94][RAY94][DOS95][ACA96][AYA96][RAY96][UME96][WAK96][JAI97][KOT97]. Building on these research contributions and taking into account special features of satellite systems, we proposed a fully integrated and flexible ATM-satellite signalling protocol architecture.

Given that the MAC and LLC protocols are the two major protocols contributing to the overall system performance and provide services to radio connection management, we thus first investigated and optimised these two protocols separately. The MAC protocol (TDMA based) is optimised to be adaptive to different service types and to use assignment-on-demand to achieve a maximum channel utilisation and low access delay. The optimisation is considered for both user data and signalling traffic. The ARQ protocols of the LLC layer have direct impacts on the QoS of ATM layer. We therefore optimise the ARQ scheme to be
service- and reliability-dependent in order to meet the QoS parameters of cell transfer delay and cell loss ratio and to enhancement of the error control performed by Head Error Control (HEC) and ATM Adaptation Layer (AAL) so that serious cell discarding on ATM layer can be avoided. It was also modified to overcome the low channel efficiency problem encountered in an end-to-end connection routed through intersatellite links (ISL).

Based on the optimised MAC and ARQ protocol, we then investigated the radio connection management (CM) schemes that could provide a QoS-guaranteed ATM traffic transmission over satellite channels. A CM scheme is mainly affected by four factors, e.g. the physical channel, the MAC protocol, the logic link control protocol and the resource allocation scheme. These four factors have usually been studied separately in the previous research work. However, herein we study them together as this is most appropriate for an effective ATM traffic transportation scheme over satellite. The performance objectives for the ATM traffic transmission over satellite were studied first, then a novel radio connection management scheme was proposed. The scheme can compensate QoS degradation caused by the mobile radio environment and enables the network to map an ATM service onto a suitable satellite connection with minimum bandwidth provisioning. It was then optimized to achieve specific performance objectives of cell loss rate (CLR), cell transfer delay (CTD) and bandwidth efficiency in the ATM-satellite network.

In addition to the proposed connection management scheme, we have investigated the connection control and the call handling protocols. Our interests were especially on the establishment of both signalling connection and traffic connection. In order to efficiently set up a connection, the meta-signalling protocol specified in ATM was modified into reservation meta-signalling for an ATM-satellite system. As the user mobility and air interface design are important in an integrated system, we thus studied the call-handling issues in ATM-satellite, which included authentication, location update, paging, handover routing, incoming delivery outgoing call routing. This enabled us to outline the call set-up procedures for both mobile originated and mobile terminated calls.

1.5 Original Work & Major Achievements

This thesis demonstrates the following original work and major achievements:

- Proposal of a new and fully integrated ATM-satellite signalling protocol architecture that provides mobility for ATM users.
• Validation of the application of demand assignment MAC scheme into signalling traffic and proposal of a semi-permanent signalling protocol. Proposal of a service-adaptive MAC protocol.

• Proposal of a service-adaptive ARQ protocol. Proposal of a reliability-dependent SR ARQ protocol to support a variety of traffic types with different QoS objective, especially those delay-constrained services.

• Study for the first time the impact of distance changing end-to-end connections on performance of the Selective Repeat (SR) ARQ protocol, with particular reference to the connection routed via ISL of a LEO/MEO satellite constellation.

• Proposal of an Adaptive Timer SR ARQ (ATSR ARQ) protocol to be used on a distance-changing connection for improved throughput performance.

• Proposal of a QoS-provisioning radio connection management scheme for ATM via satellite, taking into accounts the influence of all four factors: the fading channel, MAC protocol, ARQ protocol and resource allocation.

• Proposal and verification of a Reservation Meta-signalling (RMS) for setting up signalling connections at the air interface in an ATM-satellite system.

During this PhD work, the author has contributed to two European Advanced Communications Technologies and Services (ACTS) projects, Satellite EHF communications for multimedia-mobile services (SECOMS) and ACTS Satellite Switching End-to-end Trials (ASSET).

SECOMS (10/96-12/97) aimed to define the system architecture and elements and develop related technologies for the future advanced mobile multimedia satellite services operating in Ka- (20-30 GHz) and EHF- (40-50 GHz) bands. The original work input to the SECOMS project includes modelling and performance analysis of multiple access and call control signalling.

ASSET (03/98 – 12/99) was a follow-up to the previous SECOMS project and its derived commercial initiative, EuroSkyWay. It proposed the implementation of a constellation of GEO satellites with the capability of multi-protocol interworking to provide broadband multimedia services to low-cost fixed, portable and mobile terminals. The original work contributed to this project includes connection management scheme design, evaluation and consolidation.
Part of the work presented in this thesis has been input into the above two projects. These works can be found in the project technical reports.

A list of publications and technical reports related to this PhD work and the work done for SECOMS and ASSET projects are given in the last section of this thesis.

### 1.6 Thesis Structure

The structure of this thesis is illustrated in Figure 1-2.

![Figure 1-2 Thesis structures.](image)

Chapter 2 provides an introduction to the characteristics of mobile satellite systems and ATM technology. We focus mainly on the features of specific importance and relevance to the work presented in this thesis. Chapter 3 presents the major challenges posed by a seamless integration of mobile satellite and ATM and analyses the impacts on QoS and ATM performance. Integration requirements and key integration tasks of the signalling protocols are addressed. Based on the discussion presented in chapters 2 and 3, we introduce a flexible and fully integrated signalling protocol architecture for the mobile ATM-satellite in chapter 4. In order to meet the QoS requirements specified in the ATM layer, the optimisation of the MAC protocol is presented in chapter 5 and ARQ protocol in chapter 6 respectively to make them more adaptive to different service types supported by ATM. Utilising the optimised...
MAC and ARQ protocol, we further propose, in chapter 7, a QoS-provisioning connection management scheme in order to map the individual ATM services onto a suitable connection that can provide the required QoS with a minimal channel bandwidth. Finally, we present a call control signalling scheme in chapter 8 focusing on call routing techniques and call establishments for both traffic connections and signalling connections. Chapter 9 gives an overview of the work presented in the thesis, highlights the major conclusions and possible future work.
Chapter 2 Mobile Satellite System Characteristics and ATM Overview

The mobile satellite system characteristics and an overview of ATM technology are presented in this chapter. We focus mainly on those features that are of specific importance and direct relevance to the work presented in this thesis covering constellation, channel model, ATM Q.2931, ATM service types and QoS requirements. The aim of this chapter is to introduce the background and environment in which this research was carried out.

2.1 Mobile Satellite System Characteristics

2.1.1 Satellite Constellation

2.1.1.1 Types of Constellations

There are three types of satellite constellations using different orbits, GEO, MEO and LEO. Each orbit may be applicable to a specific application. Altitudes of the orbits in GEO, MEO and LEO constellations are chosen by taking into account the radiation belts around the earth. Such radiation can damage solar panels and electronics on board the satellites and reduce the power generating capacity of the solar arrays [MAR97].

GEO satellites placed at 35,786 km above the earth remains quasi-stationary relative to a fixed spot on the earth. A constant coverage area can therefore be provided by GEO satellites across the visible portion of the earth. This was the original consideration behind the establishment of first-generation satellite communications using GEO. However, due to long propagation delays associated with GEO (around 250ms from transmitter to receiver), LEO/MEO have recently attracted much interest for their advantage of low-delay transmission to support improved real-time services.

LEO refers to the orbit between 500km and 1,500km and MEO between 10,000km and 12,000km. These types of orbits reduce the propagation delay as compared to GEO. But LEO and MEO satellites do not appear stationary to earth stations as their orbital periods are less than 24hrs e.g. LEO around 120mins and MEO around 1-2hrs. The nature of these satellite dynamics is the major disadvantage of LEO/MEO compared to GEO. The LEO/MEO constellations complicate ground station tracking and operation as the satellites are dynamic.

According to the orbital inclination, a MEO or LEO constellation can be either Polar Constellation (PC) or Inclined Circular Constellation (ICC). Polar constellations consist of a number of orbits that have a common intersection at the polar axis with an inclination angle around 90°. ICC has an inclination angle of between 90° and 0°.
2.1.1.2 Constellation Design

**Satellite Position**

In order to describe the satellite orbit positions and the configurations of the satellite constellations in the space, orbital parameters are commonly used [EVA99][RIC95]. These parameters are given in Table 2-1. Their definitions are explained in the Figure 2-1, Figure 2-2, Figure 2-3 and Figure 2-4.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Denotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semi-major axis</td>
<td>a</td>
</tr>
<tr>
<td>Semi-minor axis</td>
<td>b</td>
</tr>
<tr>
<td>Eccentricity</td>
<td>$e = \sqrt{a^2 - b^2} \over 2$</td>
</tr>
<tr>
<td>Inclination angle: the angle between the plane of the orbit and the equatorial plane measured at the ascending node in a northward direction.</td>
<td>$i$</td>
</tr>
<tr>
<td>Right Angle of Ascending Node: the angle between the direction of the Vernal Equinox (X-axis in Figure 2-2) and the ascending node.</td>
<td>$\Omega$</td>
</tr>
<tr>
<td>Argument of perigee: the angle in the orbital plane between the line of the nodes and the perigee of the orbits.</td>
<td>$\omega$</td>
</tr>
<tr>
<td>True anomaly</td>
<td>$\theta$</td>
</tr>
<tr>
<td>Elevation angle</td>
<td>$\sigma$</td>
</tr>
<tr>
<td>Azimuth</td>
<td>A</td>
</tr>
<tr>
<td>Range</td>
<td>d</td>
</tr>
<tr>
<td>Radius of the earth</td>
<td>R</td>
</tr>
<tr>
<td>Orbit period</td>
<td>T</td>
</tr>
<tr>
<td>Total number of satellites in the constellation</td>
<td>N</td>
</tr>
<tr>
<td>Number of orbital planes</td>
<td>P</td>
</tr>
<tr>
<td>Number of satellites per plane</td>
<td>Q</td>
</tr>
<tr>
<td>Satellites Altitude</td>
<td>h</td>
</tr>
</tbody>
</table>

Table 2-1 Commonly used constellation parameters.
Parameters such as satellite orbital period, velocity and position are of particular importance. The satellite orbital period for a satellite in a circular orbit is given by

$$T = \sqrt{\frac{4\pi^2}{\mu}} \frac{(R+h)^3}{R}$$  \hspace{1cm} (2-1)$$

and the satellite velocity \( V \) is given by

$$V = \sqrt{\frac{\mu}{(R+h)}}$$  \hspace{1cm} (2-2)$$
where the $\mu$ is the gravitational parameter. It is also important to determine the look angle of a satellite from a fixed earth station or a mobile terminal. The satellite position, relative to the observation point, is commonly specified by two parameters, the satellite azimuth and elevation angle [WIN98]. The meanings of these two parameters are shown in Figure 2-3 and Figure 2-4.

![Figure 2-3 Satellite look angle (a) Elevation.](image)

![Figure 2-4 Satellite look angle (b) Azimuth.](image)

The elevation angle is the angle that a satellite makes with the tangent at the specified points on the earth. The satellite elevation angle is given as [RIC95].
\[ \sigma = \arctan \left( \frac{\cos(\psi) - \delta}{\sin(\psi)} \right) \]  \hspace{1cm} (2-3) 

where \( \delta = \frac{R}{R + h} \), \( \psi \) is the coverage angle.

The azimuth of a satellite from a given point is the angle that the satellite direction makes with the direction of true north measured in the clockwise direction. The azimuth can be obtained as

\[ A = \left\lfloor \arctan \left( \frac{\tan(\phi_e)}{\sin(\theta_e)} \right) \right\rfloor, \quad \phi_e = \phi_s - \phi_e \]  \hspace{1cm} (2-4) 

where \( \theta_e \) and \( \phi_e \) are the latitude and longitude of the observation point, respectively. \( \phi_s \) is the longitude of the sub-satellite point.

The line of sight distance between the observation point and the satellite is called the Range \( d \) of a satellite, expressed as

\[ d = \sqrt{R^2 + (R + h)^2 - 2R(R + h)\cos(\psi)} \]  \hspace{1cm} (2-5) 

**Satellite Constellation Design**

Recent constellation design focuses on designing large LEO/MEO constellations to achieve a global coverage for real-time personal communications. Such a constellation has advantages of a low propagation delay and a small size terminal as a result of reduced requirement on gain and power of the terminal. These advantages make the idea of provision of real-time services with handheld telephones more straightforward.

The objective of the constellation design is to optimise the orbital geometry so as to provide the required earth coverage while minimising the number of satellites. In designing the constellation for a Personal Communications Network (PCN); service areas, types of communication services and the QoS to be provided plus network issues have to be taken into consideration. Two commonly used constellation methods [NAR98] are briefly described below.

**Street of Coverage Method**

The street of coverage method is used to optimise constellations that have equally inclined orbits with an equal distance between intra-plane satellites, and place an equal number of satellites into each orbit. Each satellite moves around the earth and their track on the earth forms a “coverage strip”. When all
the satellites are designed to move around the earth towards the same direction and their plane separation is small enough, a global coverage can be formed by the tracks made by all these individual satellites orbits. The concept of the "street of coverage" method is shown in Figure 2-5. Using this method, Beste optimised single visibility coverage for the whole earth as well as for regions extending from the poles to any arbitrarily chosen altitude [BES78]. The number of satellites versus orbital altitude for whole earth coverage using Beste's approximation equation is shown in Figure 2-6.

Figure 2-5 Street of coverage method.

Figure 2-6 Number of satellites N versus angular coverage $\psi$ for a global coverage [BES78].

Spherical Triangle Method
The spherical triangle method divides the earth's surface into a number of spherical triangles [WAL73][BAL80]. The triangles are formed by linking the sub-satellite points of a plane to its two adjacent sub-satellites points of the neighbouring planes, shown in Figure 2-7. The centre point of such a triangle is \( R_{ijk} \). The method first searches for the maximum \( R_{ijk} \), \( R_{ijk}(\text{max}) \), in all the spherical triangles over an orbital period of time. Then it searches all the \( R_{ijk}(\text{max}) \) values for all different inclinations and initial phase angles. The constellation giving the smallest \( R_{ijk}(\text{max}) \) would be the optimised constellation. For a double coverage, the optimisation should be performed provided that another sub-satellite point is enclosed in each spherical triangle. The further details can be found in the references mentioned above.

![Figure 2-7 Spherical triangle method.](image)

The choice of different types of constellations will depend on the desired service area and specific applications such as regional, global, real time and etc.

### 2.1.1.3 Current Satellite Constellations

In the race to bring the first successful global satellite system and offer services to large remote areas on the earth, many satellite systems have been either developed or are under development. The major systems include Iridium, GlobalStar, ICO and Teledesic. All these systems are intended to offer a variety of services covering voice, data, paging and fax with Teledesic extending services to broadband. The major system design parameters including the constellation parameters are given in Appendix A.
2.1.2 Satellite Multi-beam Coverage

In order to increase capacity with a fixed allocation of radio spectrum additional frequency reuse is needed and thus in turn means increased satellite antenna gain and number of spot beams. The typical satellite spotbeam pattern for Iridium is shown in Figure 2-8 [WIN98].

The deployment of multi-beam antennas, however, has its disadvantages. It imposes the requirement of handovers between the spotbeams, and the frequency of required handover could be very high in a LEO constellation. For example, Iridium has 48 spotbeams and each footprint covers an area with 4600km in diameter. Because the antenna spotbeams move at a speed of 6.6km/s relative to the earth, the handovers occur once every 1.5 minutes or so. The high frequency of handovers complicates the systems design and affects the QoS.

2.1.3 Intersatellite Links

New advanced satellite systems such as Iridium and Teledesic employ intersatellite links to form the connectivity of the LEO/ICO-satellite network. Intersatellite links bring a major benefit to transport the long-distance traffic over reliable and high capacity space connections [WER97C]. This reduces the reliance on terrestrial connections and can decrease the communication delays. The presence of ISLs also introduces flexibility in routing and avoids the need for visibility of both user and gateway by each satellite in the constellation [WER97C].
Intersatellite links are the direct connections between satellites in a constellation. There are two types of intersatellite links, intraplane ISLs connecting successive satellites within the same orbit plane and interplane ISLs connecting satellites in adjacent co-rotating orbit planes [WOO95]. Interplane ISLs generally require antenna steering whereas intraplane ISLs can be maintained with fixed antennas. The length of intraplane ISLs is constant, whilst that of interplane ISL's varies with time and reaches its maximum at the equator. Intraplane ISLs can be permanently maintained whilst interplane ISLs have to be deactivated in Polar regions. This is because the line-of-sight of interplane ISLs could be interrupted by the earth.

As an example shown in Figure 2-9, the Iridium constellation used polar orbits with 11 satellites in each orbit. Iridium provided two intra-plane connections to satellites forward and backward and two inter-plane connections to satellites in each neighbouring plane. Because the Iridium LEO satellites circulate the earth at a constant speed, the coverage area of a LEO satellite changes continuously. The visibility period of a satellite, $T_v$, defined as the maximum time duration that a terminal residing in the coverage region of a satellite can directly communicate with that satellite, is typical around 10 minutes [UZU97A]. ISLs are consequently switched on and off during operation. Due to relative velocities and Doppler shift, there are only three lines of interplane ISLs near to the equator as shown in Figure 2-10 [WOO95].
In the systems that employ intersatellite links, end-to-end connections are expected to be set up by using a set of intersatellite links. The traffic is expected to be routed in space for the maximum duration possible. Routing is determined by users' geographical locations and certain optimisation criteria such as minimum hop or minimum cost at call set-up [UZU97B]. One distinct feature of this kind of connection is that the connection length varies with the time during the active period. Besides the continuous changes caused by co-rotating interplane ISLs, discontinuous changes caused by the deactivation of implemented interplane ISLs in the polar region also presents discrete-time contributions to the connection, which necessitates connection handover. Connection handover could result in updating the connection partially or in setting up a whole new route in some cases.

2.1.4 Satellite Channel

Channel fading factors can be characterised by path loss, Additive White Gaussian Noise (AWGN), Doppler effect, multipath and shadowing. Due to these factors, the signal power could be seriously attenuated and the signal envelope could be badly distorted. The distortion can lead to a high bit error rate in the transmitted digital signal and degradation in the service quality to an unacceptable level. An accurate channel model is therefore important for assessing the communication system QoS performance.

With the development of the packet switching technique, one of the key technical areas is to determine the effect of channel fading on network performance. QoS of user applications is closely related to the performance of the packet transmission over radio channel, which is controlled by a data link protocol. A set of packet-level parameters such as Cell Loss Ratio (CLR), Cell Error Rate (CER) and Cell Transfer Delay (CTD) are used in ATM to characterise the performance of ATM layer and an end-to-end connection [ATF96]. Therefore, a packet-level fading channel model is needed for realistic simulation and analysis of the QoS parameters and network performance.
Correlated Fading Channel Model

Extensive research work has been performed to model and represent the dynamic channel behaviour. The channel model can generally be divided into two categories, one an analogue model and the other a digital model. The analogue model attempts to represent the channel behaviour at signal level for DSP level simulation. Two widely used analogue models are the narrowband analogue model proposed by Lutz [LUT91] and the wideband channel model proposed by Turin [TUR90]. Digital models, such as Gilbert channel model [GIL60] and Lutz two-state channel model [LUT91] are generally used for analyses. They attempt to represent the channel behaviour at a discrete time. However, these digital models are often based on bits or symbols, they cannot be efficiently used to obtain the packet statistics of the network performance evaluation. Hence these models are not very suitable for block level performance analyses. Michele Zorzi has proposed a block-level two-state Markov model to represent bursty fading channels, which is suitable for network performance evaluation [ZOR95]. This model that has been used in our work is briefly described herein.

Zorzi verified that a first-order Markov model for data transmission on flat fading channels is sufficient accurate enough to reproduce the channel memory. This model treats the IDD (independent and identically distributed) model as a special case. It has been found that the proposed model gives significantly different results from the previously proposed IDD model used in the block transmission wherein block errors are assumed independent and identically distributed. On the other hand, Zorzi found that the results were relatively insensitive to the memory of the Markov model. A first-order model is thus sufficient to represent the channel memory and a higher-order model is therefore not considered. Zorzi has derived the packet success/failure process over a correlated flat fading channel and provided us with a technique to evaluate packet transmission in a more realistic channel fading condition.

Zorzi’s work has verified the accuracy of a first-order binary Markov model for the success/failure process of data blocks under the assumption that the considered data rate is relatively high (hundreds of kbits/s), so that the duration of a packet of hundreds of bits is smaller than the coherence time. This implies a narrowband channel. He has considered a frequency non-selective (flat) multipath fading channel, which is modelled as a multiplicative complex function, $\alpha(t)$, and is adequately described as a random process.

The packet success/failure process is modelled as the outcome of a comparison of the instantaneous signal-to-noise ratio to a threshold value, $SNR$. If it is above the threshold, the packet is successfully decoded with probability 1; otherwise, the packet is lost with probability 0. If $F$ is the value of the fading margin, which is defined as a margin of extra signal strength built into the system to compensate for transitional fading conditions, the instantaneous signal-to-noise ratio (taking into
account the effect of the fading) is given by $SNR,F\alpha^2(t)$. Hence, the binary process that describes packet success/failure on the channel, $\beta_n$, can be obtained by quantization of the squared magnitude of the complex Gaussian description with the threshold $1/F$, i.e.,

$$\beta_j = \begin{cases} 0 & \text{if } V_j^2 > 1/F \\ 1 & \text{if } V_j^2 \leq 1/F \end{cases}$$

where $V_j = |\alpha(jT)|$ is the amplitude of the fading envelope at time $jT$; $T$ is the packet duration, and "1" stands for a packet failure. The above success/failure process on a mobile radio channel is described by a first-order Markov model shown in Figure 2-11.

![Figure 2-11 Zorzi’s first-order Markov model.](image)

The parameters of the Markov model can be determined on the basis of the fading model and the characteristics of the communication scheme. The transition probability matrix that describes the channel is given by

$$M_r = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}$$

where $p$ and $1-q$ are the probabilities that the packet transmission in slot $j$ is successful, given that the transmission in slot $j-1$ is successful or unsuccessful, respectively. Given the matrix $M_r$, the channel model is completely characterised. In particular, the steady-state probability, $P_E$, a probability when a packet error occurs due to fading and noise, is

$$P_E = \frac{1-p}{2-p-q}$$

Also, note that $(1 - q)^{-1}$ represents the average length of a burst of errors, which is described by a geometric random variable. The parameters of the above Markov model can be derived from

$$P_E = 1 - e^{-\mu_F}$$
Mobile Satellite System Characteristics and ATM Overview

and

\[ q = 1 - \frac{Q(\theta, \rho \theta) - Q(\rho \theta, \theta)}{e^{\mu F} - 1} \]  

(2-11)

where

\[ \theta = \sqrt{\frac{2 f \rho}{1 - \rho^2}} \]

(2-12)

\[ \rho = J_0(2\pi \int_T f) \]

\rho is the correlation coefficient of two successive samples of the complex amplitude of a fading channel with Doppler frequency \( f_d \), taken \( T \) seconds apart, \( f_n T \) is the normalised Doppler bandwidth, and

\[ Q(x, y) = \int_y^\infty e^{-\frac{(x^2+y^2)}{2}} I_0(xw)wdw \]  

(2-13)

is the Marcum Q function, wherein \( I_n \) is the modified Bessel function of the first kind and of zeroth order.

Note that the correlation properties of the fading process depend only on \( f_n T \). When it is small (e.g. < 0.1), the process is very correlated ("slow" fading); on the other hand, for a large value of \( f_n T \) (e.g. < 0.2), successive samples of the channel are almost independent ("fast" fading). For high data rates (i.e. small \( T \)), the fading process can typically be considered to be slow fading for typical speeds of moving mobile terminals, so that the dependence between transmissions of consecutive data packets is non-negligible.

2.2 ATM Introduction

Two standardisation organisations, the ITU-T (International Telecommunications Union – Telecommunication Standardisation Sector) and the ATM Forum, are currently involved in ATM standards. ETSI (European Telecommunications Standards Institute) has adopted a policy of producing ATM standards compatible with those of ITU-T.

ATM is designed for B-ISDN network. It consists of four types of interfaces, Network-Network Interface (NNI), User-Network Interface (UNI), Private UNI (P-UNI) and Private NNI (P-NNI), as shown in Figure 2-12 [ONV95]. ATM is a connection-oriented technology capable of supporting all classes of traffic in one transmission. User QoS requirements are declared and negotiated at the connection set-up and will be guaranteed in the following communication by all the intermediate switches between end users.
2.2.1 ATM Protocol Layers

The protocol reference model of B-ISDN is shown in Figure 2-13. It consists of three planes, control, user and management. The control plane handles all the connection management related functions, addressing and routing. The user plane transmits end-to-end user information between communication parties. The network Operation And Management (OAM) functions are provided by the management plane. All three planes use the common low layers: physical layer, ATM layer and ATM Adaptation Layer (AAL).

![Figure 2-13 B-ISDN protocol reference model.](image)

2.2.1.1 Physical Layer

ATM was originally designed for high speed and low Bit Error Rate (BER) wireline transmission media such as Synchronous Optical NETwork (SONET) and Synchronous Digital Hierarchy (SDH) that are most common ones. These two standards use optical fibre as the transmission medium. Coaxial cable and twisted pair can also be used for ATM but the bit rates achieved are lower than those achieved via optical fibres.
The ATM physical layer comprises two sub-layers, the physical medium (PM) and the transmission convergence (TC) sub-layers. The PM sub-layer takes care of the transmission of the bits across the physical medium. The TC is more ATM dependent and deals with the functions necessary to support the ATM cell type data transfer. These functions include, for example, Header Error Correction (HEC) generation, cell delineation, transmission frame generation and recovery functions [ONV95].

2.2.1.2 ATM Layer

To keep up with high-speed transmission links, the ATM layer function is simplified. The ATM layer only copes with cell header related functions and layer management. User information is carried in ATM cells that are fixed 53 bytes with a 5 bytes header. The header contains Virtual Path Identifier and Virtual Connection Identifier (VPI/VCI) used by the ATM switch to route incoming cells to the next link. A VPI/VCI is allocated for a virtual connection by the switch when the connection is set up, and it remains unchanged for the entire lifetime of the connection.

ATM layer buffers incoming and outgoing cells, handle various traffic management functions such as cell loss priority marking, generic flow control and monitoring of the transmission rate. The eight-bit HEC can correct single bit error in the header. No link-by-link based error control is performed for 48 byte cell payload at the physical layer and ATM layer. Similar cell formats are employed by the UNI and NNI. The only difference is that the Generic Flow Control (GFC) field does not appear in the NNI cell format and it is used to extend the VPI field in NNI.

2.2.1.3 ATM Adaptation Layer

The ATM Adaptation layer (AAL) performs segmentation and reassembly services, which map user/control management Protocol Data Units (PDU) into ATM cells, and adapt different services to the ATM network. The AAL layer enhances the services provided by the ATM layer to meet the requirements of specific applications. These applications could be user applications or control management functions.

A variety of services to be transported over ATM layer are classified into four types, each of which has its own specific requirement for AAL. Different AALs are therefore defined according to three basic parameters: timing relationship between source and destination, bit rate and connection mode.

AAL 1 supports Constant Bit Rate (CBR) services. To support the CBR services, service data unit, timing information and data structure information are transferred between source and destination. Functions for source clock recovery and control of cell delay variation are also included. AAL 1 also offers an indication of the erroneous information.

AAL 2 is used mainly for Variable Bit Rate (VBR) services with timing requirements.
AAL 3/4 supports VBR services that do not require a timing relationship between the source and the destination. This is recommended by ITU for transferring data services that are sensitive to loss, but not to delay. One of the most important functionalities of AAL 3/4 is that it provides either assured or non-assured peer-to-peer operational procedure to both message and stream modes. The AAL 3/4 performs the error control via the cyclic redundancy code (CRC) field and cell retransmission. Flow control is also offered between endpoints.

The service of AAL 5 is similar to that of AAL 3/4 except for no multiplexing supported. AAL 5 is modified from AAL 3/4. It is used to provide a service with less overhead and better error detection for signalling and data services [OTH95]. Other than CRC-10 used in AAL 3/4, AAL 5 uses CRC-32.

**Signalling AAL**

The SAAL is the service specific signalling adaptation layer implemented on top of ATM layer at the control plane. It provides reliable transport of signalling messages in an ATM network. The SAAL consists of SSCF (Service-Specific Coordination Function) and SSCOP (Service-Specific Connection-Oriented Protocol). The SSCOP performs functions similar to a classic link control protocol. SSCF is defined as Q.2130 to interpret the Q.2931 protocol at the UNI and Q.2140 to interpret the MTP-3 protocol at the NNI.

**2.2.2 ATM Traffic Classes and QoS Provisioning**

The ATM forum has defined a set of traffic classes and their parameters. The defined ATM traffic classes include Constant Bit Rate (CBR), real-time Variable Bit Rate (rt-VBR), non-real time Variable Bit Rate (nrt-VBR), Available Bit Rate (ABR) and Unspecified Bit Rate (UBR). CBR and rt-VBR require certain QoS guarantee on Cell Loss Rate (CLR), Cell Transfer Delay (CTD) and Cell Delay Variation (CDV) from the network. Nrt-VBR, ABR and UBR are not delay sensitive services. However, nrt-VBR requires a guarantee for the CLR. ABR requires a "best effort" approach on bandwidth allocation. UBR service does not offer any service guarantees. A user is free to send any amount of data up to a specified maximum while the network makes no guarantees at all on the cell loss rate, delay or delay variation that might occur.

For every traffic class there exist a number of parameters that characterise both the traffic and the QoS requirements. These parameters are given in Table 2-2 and Table 2-3. The requested parameters have to be defined by the user when a connection set-up is requested.
2.2.3 Traffic Management

For the success of ATM, it is important that it provides a good traffic management for both bursty and non-bursty sources based on the type of the traffic and the quality of service desired. Before a connection can actually be allocated to a user, a traffic contract has to be established to provide guaranteed QoS [ATF96]. Users can specify a set of parameters to describe their traffic characteristics and declare their desired QoS. These QoS and traffic description parameters are given in Table 2-2 and Table 2-3 respectively. These QoS parameters have to be agreed by or negotiated with the network during connection set-up. The network uses these parameters to perform connection admission control to ensure the QoS for existing and incoming users.

<table>
<thead>
<tr>
<th>QoS Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CER</td>
<td>Cell Error Ratio</td>
</tr>
<tr>
<td>SECBR</td>
<td>Severely Errored Cell Block Ratio</td>
</tr>
<tr>
<td>CLR</td>
<td>Cell Loss Ratio</td>
</tr>
<tr>
<td>CTD</td>
<td>Cell Transfer Delay</td>
</tr>
<tr>
<td>CDV</td>
<td>Cell Delay Variation</td>
</tr>
<tr>
<td>CMR</td>
<td>Cell Misinsertion Rate</td>
</tr>
</tbody>
</table>

Table 2-2 QoS parameters.

<table>
<thead>
<tr>
<th>Traffic Descriptors</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>Peak Cell Rate</td>
</tr>
<tr>
<td>SCR</td>
<td>Sustained Cell Rate</td>
</tr>
<tr>
<td>BT</td>
<td>Burst Tolerance</td>
</tr>
<tr>
<td>MCR</td>
<td>Minimum Cell Rate</td>
</tr>
<tr>
<td>CDVT</td>
<td>Cell Delay Variation Tolerance</td>
</tr>
</tbody>
</table>

Table 2-3 Traffic descriptors.

In order to protect the established connection, the network monitors and controls the traffic to ensure users to remain with their negotiated parameters. This is called usage parameter control. Upon detecting a non-conforming source, the network can drop or delay violating cells whilst warning the source user.

Congestion control lies at the heart of the general problem of traffic management. The Generic Cell Rate Algorithm (GCRA) for flow control and Explicit Feedback Congestion Indication (EFCI) for congestion control are recommended by the ATM Forum to shape the traffic. When the network is badly congested, cells with lower priority will be discarded, and end users will be informed to slow down the transmission [ONV95].

The ABR service can adapt its source rate to changing network conditions. The key feature of the ABR traffic class is its flow control method that allows the free link capacity to be dynamically used.
The flow control makes ABR well suited for the transmission of cell loss sensitive data. The ATM Forum has put forward a rate-based flow control method, wherein the transmission rate of the source is controlled. The rate-based approach uses end-to-end rate-based control. Another flow control is a credit-based system, wherein the traffic is controlled per link according to the number of buffers available in the next switch [JAI96].

### 2.2.4 Statistical Multiplexing and ATM Connection

Statistical multiplexing is one of major advantages of ATM. ATM adopts statistical multiplexing aimed at achieving the best bandwidth utilisation whilst QoS requirement of delay and loss for different types of real time and non-real time traffics are also satisfied during periods of congestion. However, the benefit arises only if there are enough uncorrelated sources to be multiplexed [PRY95].

An ATM connection is called a virtual connection (VC). There are two main types of virtual connections, Permanent Virtual Connections (PVC) and Switched Virtual Connections (SVC). PVC is pre-configured and SVC is dynamically established. A virtual connection could be point-to-point or point-to-multipoint.

### 2.2.5 Call Admission Control

CAC's completion is restricted to the call set-up phase and to the call re-negotiation phase. CAC operates at the UNI (User-Network Interface) and at the NNI (Network-Network Interface). CAC allocates buffers and bandwidth to a VP (virtual path) according to a traffic descriptor, QoS requirements, capacity of the network, and measurements of the actual traffic load in the different network nodes, as an option.

An efficient CAC algorithm should attain high utilisation of the transmission link by means of statistical multiplexing, whilst maintaining the QoS requirements of all existing connections and therefore avoiding congestion. The efficiency of a CAC scheme is measured by its success in meeting the requirements defined for all traffic management functions including efficiency, simplicity, flexibility and robustness. In particular, since CAC runs at call set-up time, time consuming and complex CAC algorithms are not appropriate for real services.

### 2.2.6 ATM UNI Signalling

The two main candidates for the signalling protocol over the UNI are the Q.2931 standardised from the ITU-T and the UNI-signalling from the ATM-Forum. The desired protocol is to be capable of dealing with call control and call routing functions in support of set-up, maintenance and release of connections.
The ATM-Forum UNI Signalling Specification (version 4.0) is based on Q.2931, but has several variations of Q.2931. The ITU-T specifications do not allow a leaf to join an ADD PARTY Connection without the intervention of the root. The ATM-Forum UNI establishes procedures for the leaf to initiate an ADD PARTY request, which is handled by the network. The root is not involved in this operation. UNI 4.0 also provides for an ATM group address, which is a collection of ATM end systems.

The ATM signalling protocol stack is shown in Figure 2-14. As Q.2931 does not possess any error-compensation mechanism, a special SSCOP (service specific co-ordination oriented protocol) layer for signalling is built on top of AAL5 (ITU-T signalling protocol stack) to provide a highly reliable call control signalling message transmission. Signalling SSCF (service specific co-ordination function) maps the particular requirements of UNI signalling to the requirements of the ATM layer. CPCS (common part convergence sublayer) and SAR (segmentation and reassembly sublayer) provide an uninsured information transfer but detect the error message.

```
<table>
<thead>
<tr>
<th>Q.2931</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSCF</td>
</tr>
<tr>
<td>SSCOP</td>
</tr>
<tr>
<td>CPCS</td>
</tr>
<tr>
<td>AAL5</td>
</tr>
<tr>
<td>SAR</td>
</tr>
<tr>
<td>ATM</td>
</tr>
<tr>
<td>PHY</td>
</tr>
</tbody>
</table>
```

Figure 2-14 ATM signalling protocol stack.

2.2.6.1 Meta-signalling

To support a wide variety of services, ATM requires a number of different signalling procedures to handle different functionality and to be used in different circumstances. Thus, the B-ISDN signalling channels are not permanently defined. UNI signalling connection set-ups utilise two methods in ATM. For a point to-point configuration, a dedicated signalling channel with default value VCI=5 for Permanent Signalling VCI (PSVCI) is used. For a point-to-multipoint case, signalling connection is set up on demand by a meta-signalling protocol.

Meta-signalling is defined to establish, check and release user-to-network signalling channels in ITU-T recommendation Q.2120 [ITU95]. It is a simple protocol with 5 messages. All the commands are given space in one ATM cell, each via a reserved ATM channel with VCI=1 and default VPI=0. All the meta-signalling messages are of the same size. Meta-signalling is invoked by the user side.
following a call establishment request. It could also be invoked by the network to check the existence of on-going signalling connection.

2.2.6.2 Q.2931 Protocol

In order to support point-to-point dynamic connection set-up, connection re-negotiation and connection teardown, and the Q.2931 protocol running on top of SAAL is used. This signalling protocol allows connections to be set up with characteristics that meet the needs of both the users and the network. It consists of the protocol and messages that are exchanged between an ATM end user and the network. The peer-to-peer communication across the UNI is shown in Figure 2-15.

![Diagram of Peer-to-peer communications across the UNI](image)

*Figure 2-15 Peer-to-peer communications across the UNI (SAP: service access point).*

**Point-to-Point Call set-up Signalling**

UNI signalling takes place between end stations and the network. Signalling messages provide the network with enough information to characterise the source and to locate the destination UNI. An overview of the sequence of events and the message exchanged to establish a point-to-point connection at UNI is shown in Figure 2-16.
Point-to-Multipoint Call Set-up Signalling

Point-to-multipoint (PTM) call processing required by multimedia applications, such as videoconferencing, is supported in ATM. Supporting point-to-multipoint calls is a distinguishing feature of ATM. Current PTM signalling specifications include ITU-T Q.2971 recommendation and PTM signalling in ATM forum UNI signalling 4.0. Point-to-multipoint call set-ups adopt ‘add party’ or ‘leaf initiated join’ methods [ATF96].

Add Party Capability

ITU-T recommends the “add party” procedure shown in the Figure 2-17 to set up PTM connections. A stem connection is first set up using point-to-point signalling between a calling user and a first party user. The rest of the parties, which are called leaves, join the group upon the “add party” request sent by the root. This procedure only allows the root of a connection to join the leaves to the connection.
Leaf Initiated Join (LIJ) Capability

Because of the drawback of the add party procedure, the ATM forum has recommended the "leaf initiated join" to allow one or more leaves to join the connection without intervention from the root.

LIJ has Network LIJ and Root LIJ two operation modes. In the network, LIJ procedures are different when a leaf joins an active LIJ call from when a leaf joins an inactive LIJ call. Leaves joining active LIJ and inactive calls are shown in Figure 2-18 and Figure 2-19, respectively.

Figure 2-17 Point-to-Multipoint call set-up.

Figure 2-18 Leaf joins an active LIJ call.
2.3 Conclusion

From this introduction of the basic characteristics of mobile satellite systems and ATM technology, it will be appreciated that mobile satellite communication technology is considerably different from fixed wire line telecommunication technology. The satellite constellation and radio transmission environments are also different from the transmission environments for which ATM is designed. Therefore, an integration of ATM and mobile satellites would be expected to raise numerous issues and problems. The integration issues and problems will be discussed in the next chapter.
Chapter 3 Issues of ATM over Satellite

Having discussed the characteristics of satellite systems and ATM technology, we address major issues of a full integration of ATM and satellites in this chapter. The integration difficulties and emerged QoS provisioning problems are firstly discussed. We then analyse their impacts on the extension of ATM and the requirements on signalling protocol for an ATM-satellite integrated system. Considering that a radio connection management is the key to the success of ATM over satellite, we particularly investigate connection management related aspects, such as multiple access schemes, logic link control and connection control scheme, concerning extension and enhancement of these signalling protocol to transporting ATM traffic in mobile satellite channels.

3.1 Introduction

The satellite constellation and radio transmission environment are considerably different from the transmission environments for which ATM is designed. Unsurprisingly, an integration of ATM and mobile satellites raises numerous issues and challenges. In this chapter, we address major issues of a full integration of ATM and satellites and identify major tasks that need to be tackled.

3.2 Integration Difficulties

As mentioned in the previous chapter, ATM technology succeeds in achieving high bandwidth utilisation and offering a wide range of quality levels by employing techniques such as fast packet switching, VPI/VCI routing, statistical multiplexing and traffic management. However, since ATM does not support wireless access and user mobility, the desired integration suffers following major difficulties when using ATM in a satellite environment. Many issues are identified and discussed in the following sections.

3.2.1 Channel Impairment

ATM was originally designed for a high speed and low BER wireline transmission media. For example, optical fibre transmission systems have a BER of around $10^{-9}$, and digital circuits are expected to provide end-to-end network BER in the order of $10^{-9}$ to $10^{-10}$ [FLU95]. Due to these high quality transmission media, no link-by-link based error control is performed for 48 byte cell payload at physical and ATM layer, and detection for misinserted and lost cells and means for dealing with cell delay variations are placed at a higher layer. These simplifications enable ATM to keep up with high-speed transmission links.
But satellite systems suffering from atmospheric loss, multipath fading and shadowing have a poor transmission quality. A typical satellite link might operate with a BER between $10^{-3}$-$10^{-6}$ [PRO89] [FEH91]. The BER performances present a big contrast with the BER performance via optical fibres. The achievable BER in mobile links necessitates more complex error protection and recovery schemes to achieve acceptable service BERs. For such poor quality of mobile links, the simplicities made in ATM protocol design could incur high packet error and packet loss resulting in network performing a complex error protection and recovery procedure. Extra bandwidth and processing resource will be consumed and user QoS such as Cell Loss Rate (CLR), Cell Transfer Delay (CTD) and Cell Delay Variation (CDV) will be degraded. The impacts also necessitates a careful consideration of the selection frequency, transmit and receive power levels, and appropriate coding mechanisms.

### 3.2.2 User Mobility

In order to provide users with roaming capability, new signalling and protocol mechanisms, such as location management, paging, handover and database management, have to be incorporated into the ATM protocol. This leads to a more sophisticated signalling protocol than for pure ATM. More signalling traffic will be generated and more channel resource will be required. This situation will be worsened when channel fading occurs causing signalling messages to be retransmitted repeatedly. As a consequence, performance of call processing is degraded and call-dropping probability is increased.

Since the handover rate, the mobility related signalling efficiency and generated signalling load are dependent on the frequency of users’ roaming rate from one cell to another, a mobility model is widely used to calculate cell boundary crossing rates and location area crossing rates. This mobility model gives the boundary-crossing rate as [THO98],

$$M = \frac{\rho v L}{\pi}$$

(3-1)

where $\rho$ is area population density, $v$ is average moving speed and $L$ is length of perimeter of the area.

On the basis of this mobility model, the handover rate of a satellite system can be calculated. Satellite handover occurs when in-call terminals cross the spotbeam area boundary. Assuming that $M_{sp}$ is the spotbeam boundary crossing rate, then we have
Let $E_{term}$ be the probability of terminal being in a call when it crosses a spotbeam area boundary, then the handover rate $P_{ho}$ can be derived as,

$$P_{ho} = M_{sp} E_{term}$$  \hspace{1cm} (3-3)$$

Assuming that the call arrival rate of a user terminal is Poisson distributed with mean $\lambda$, then the inter-call arrival time $t_i$ will be exponentially distributed $1/\lambda$. Assuming that the call holding time of a user is also exponentially distributed with mean call holding time $t_h$. Then an on-off model can be used to represent the user initiated call behaviour with “on” standing for a user in call and “off” standing for a user in idle state. Therefore, the mean time $T_{in}$ that a user is in a call is given as

$$T_{in} = \frac{t_h}{\left( t_h + \frac{1}{\lambda} \right)}$$ \hspace{1cm} (3-4)$$

We can then derive the handover rate per user as,

$$P_{ho} = \left( \frac{\rho v L}{\pi} T_{ho} \right) / N$$  \hspace{1cm} (3-5)$$

where $N$ is the user population in the spotbeam. Let $A_s$ be the size of the spotbeam and $R$ be the radius of the spotbeam, then the equation can be further expressed as,

$$P_{ho} = \frac{2 \pi R}{A_s} \frac{t_h}{\left( t_h + \frac{1}{\lambda} \right)}$$ \hspace{1cm} (3-6)$$

A typical LEO/MEO spotbeam size is between $2 \times 10^5$ and $2 \times 10^6 \text{ km}^2$ [WIN98]. Given the mean call holding time of 5 minutes and user call arrival rate 0.5 calls per hour, the handover rate per user versus spotbeam size is shown in Figure 3-1. The handover rate versus call arrival rate is given in Figure 3-2.
Location update (LU) rate can also be calculated. Location update has two different cases. The first case is the location update that occurs within the same location area such as periodic location updates. This type of location update is called Registration. The second case is Location Area (LA) crossing updates. If we assume $F_{in}$ is the fraction of registration and $F_{out}$ is the fraction of the LA crossing rate, then we have
LA crossing updates will not be performed when inter-spotbeam handover occurs. It is only performed at the boundary when the user is not in a call. The number of LA crossing location updates $P_{lu}$ can be calculated as

$$P_{lu} = M_{sp} (1 - E_{term})$$

So the LA crossing updates per user is

$$P_{lu} = \frac{M_{sp} L (1 - T_{in})}{\pi} / N$$

$$P_{lu} = \frac{2vR}{A_s} \left[ 1 - \frac{t_h}{t_h + \frac{1}{\lambda}} \right]$$

### 3.2.3 Satellite Constellation Dynamics

New advanced satellite system constellations equipped with on-board switching and intersatellite links, provide a better environment for extending ATM technology. However, satellite constellations bring with them two major problems. The first is that large constellations, especially GEO constellations, have a long propagation delay. The typical roundtrip propagation delays of different constellation types are shown in Table 3-1 [NAR98]. This long delay can significantly increase the latency in the feedback mechanism of some signalling and protocol systems. Signalling protocols, especially congestion control and flow control, which require a fast feedback, may become inefficient. The other problem is that the continuously moving satellites of the MEO/LEO satellite constellations complicate the system signalling protocols by adding moving spotbeams, distance-changing connections and Doppler shifts. Thus frequent connection handovers and satellite handovers are inevitable. Due to the satellite movement, the terminal handover also occurs frequently even though the mobile terminal may be stationary. The typical terminal and satellite handover frequencies experienced in LEO and MEO systems are shown in Table 3-1.

For those systems that route the traffic via space connections formed by ISLs, the changing link distance is an additional complication. The variations of intra-plane and inter-plane ISL link distance in the Iridium constellation are shown in Figure 3-3 [WER97B]. The dynamic ISLs have negative impacts on the performance of the data link transmission.
Table 3-1 System parameters related to the constellation.

<table>
<thead>
<tr>
<th></th>
<th>LEO</th>
<th>MEO</th>
<th>GEO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude (Km)</td>
<td>500-1500</td>
<td>-10000</td>
<td>35786</td>
</tr>
<tr>
<td>Round-trip delay (ms)</td>
<td>6.6-20</td>
<td>-134</td>
<td>500</td>
</tr>
<tr>
<td>Satellite handover frequency (min/handover)</td>
<td>5-10</td>
<td>5-10</td>
<td>-</td>
</tr>
<tr>
<td>Stationary terminal handover frequency (min/handover)</td>
<td>1-2</td>
<td>2-4</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 3-3 Length of inter-plane ISL (ISL96) and inter-plane ISL (ISL94) during one orbit period.

3.3 Integration Issues

The various factors mentioned in section 3.2 provide reduced performance between terrestrial ATM and satellite ATM and pose a barrier to their integration. These factors lead to the problem of the QoS provisioning in the ATM-satellite networks and also raises requirements to design effective signalling protocols to overcome these difficulties in order to maintain a comparative QoS to that of ATM. The major requirements appear at the radio interface on the packet data transportation, resource management and mobility management related protocols. These include considerations of the header compression, Multiple Access Control (MAC), Logic Link Control (LLC), network security, mobility management, call admission control, traffic management and call control and connection control. These issues are shown in will be individually addressed in the following sections. There also exist other physical layer issues, which are not to be dealt with in this work.
3.3.1 **Header Compression**

As has been mentioned before the satellite channel is an expensive resource in terms of bandwidth and one of the issues to be addressed in satellite ATM systems is the high percentage of headers in the extended satellite ATM packet. Apart from 10 percent header overhead inherent in the small-sized ATM \((5/(5+48))\) cells, extra header has to be added to provide error protection for the physical layer information. The extra size of the header added by layers below the ATM layer is dependent on the specific protocols used. Clearly, header compression techniques are required to reduce the bandwidth taken by the header information.

3.3.2 **Multiple Access Control**

MAC protocols are used to allocate communication channels to independent, competing users by regulating the transmission of packets. The task is to manage the bandwidth resource to achieve a high access throughput and low access delay. The selection of a suitable MAC technique is one of the major issues in the mobile satellite ATM system.

3.3.2.1 **Multiple Access Protocols**

MAC protocols can be grouped into four classes [PAS97]: fixed assignment, random access, demand assignment (centrally controlled and distributed controlled), adaptive schemes. Fixed assignment and random assignment protocols are simple and require minimal scheduling. Fixed-assignment techniques, such as time division multiple access (TDMA) and frequency division multiple access (FDMA), incorporate permanent channel assignments for individual users. These schemes perform well when each user transmits a steady flow of messages, however, in bursty traffic applications, they are inefficient.
**Random Access**

Despite the drawback of reduced channel throughput, random access protocols have become a fundamental part of mobile communication systems. Slotted Aloha is commonly used in conjunction with other scheme in recent MAC proposals. It provides an easy way for users to share the system resource. In a random access protocol, a transmitting node always transmits at the full rate of the channel. As it is a contention-based protocol, packets sent to a channel could end up in collision. Retransmission with a randomised backlog delay is therefore required to resolve the collision and maintain the packet success probability at the next transmission. The randomised delay is usually chosen from a uniformly distributed time range.

The performance of slotted aloha is characterised by channel throughput $S$, delay $T_d$ and system stability. The channel throughput is defined as the average number of successful packet transmissions per timeslot duration. The major problems with slotted ALOHA are the low achievable throughput and possible instability. The long propagation delay and channel fading encountered in mobile satellite systems are the major factors in degrading its performance.

The methods, S-G analysis, has been used to analyse slotted Aloha [TAS86]. $G$, the traffic intensity or the total traffic “offered” to the channel, is defined as the number of packet transmissions attempted per packet time. S-G analysis assumes an infinite user population and Poisson distributed traffic intensity. In other words, the probability that $K$ packets are generated during a given packet time obeys a Poisson distribution with a mean of $G$ packets, that is,

$$P(k) = \frac{G^k e^{-G}}{k!}$$  \hfill (3-7)

Then the throughput $S$ can be expressed as a function of $G$,

$$S = Ge^{-G}, \quad S_{\text{max}} < 1/e$$  \hfill (3-8)

Its performance in comparison with pure Aloha is plotted in Figure 3-5. The throughput for slotted Aloha peaks at $G = 1.0$, where $S = 1/e$. It is double that of pure ALOHA. It is worth noting that at the point of peak throughput, 37% of the slots carry successfully transmitted packets whilst the same percentage of slots are empty, therefore about 26% are in collision.
[PAH95]. If the network operates at higher traffic loads, the numbers of empty and successful slots both decrease sharply and the number of collisions increases rapidly.

![Graph showing the performance comparison between slotted Aloha and pure Aloha](image)

Figure 3-5 Performance comparison between slotted Aloha and pure Aloha [PAH95].

From Figure 3-5, we can see that in response to the traffic increase beyond the point of $G=1$, the throughput $S$ will decrease. This reduction in throughput means that there are fewer successful packet transmissions and more collisions. This in turn means that the number of retransmissions grows, further increasing both the backlog of messages to be transmitted and the traffic load $G$, consequently decreasing the throughput $S$. Thus there is likelihood of an undesirable scenario that the operating point keeps moving to the right and the throughput eventually goes towards zero. This is referred to as channel saturation and the system operation will then be in an unstable condition.

Slotted Aloha cannot be considered a very efficient scheme due to its low throughput. However, its throughput can be slightly improved in the presence of signal capture effects. In the conventional analysis of slotted Aloha, it is assumed that all collided packets are destroyed in each collision. But in the analysis where the presence of signal capture is considered, some of the packets involved in a collision will survive. In an ideal situation, one packet survives all collisions with any $k$ interfering packets. This case is referred to as Perfect Capture. It offers an upper bound on the throughput of slotted Aloha with capture, given by [PAH95]

$$S = 1 - e^{-G}$$

(3-9)
However, practical capture probability is less and is a function of modulation and coding, distribution of user terminals, packet length and signal-to-noise (SNR) power ratio.

The average access delay can also be calculated. If we define $T_p$ as delay from the transmitter to the receiver, $T_s$ duration of the slot, $k$ ratio of randomisation delay to $T_s$, then average packet delay in an error-free satellite channel can be expressed as [HA90]:

$$T_d = T_p + \frac{3}{2}T_s + \left( e^G - 1 \right) \cdot \left[ T_p + \frac{(k + 2)}{2}T_s \right]$$

(3-10)

When the channel fading is encountered, the performance experienced by the slotted Aloha decreases considerably. Figure 3-6 to Figure 3-9 reveal its performances on a satellite channel suffering from fading. No capture effect is assumed. The parameters used are given as,

- Satellite propagation delay from transmitter to receiver: 250 ms.
- Mean packet arrival rate: 113.2 packets/s.
- Timeslot duration: 0.00265s
- Guard time period: 1.325E-4s.
- Backlog timeslots: 5
- Packet length: 640bits
- Retransmission timer: 0.27s.

It can be seen that the access delay and channel throughput are seriously affected by the channel condition. When a high packet error rate is encountered, all the active users that access the network will experience a high call blocking probability. The low channel throughput can block the normal system operation and affect the system stability. The mean access delay shown in Figure 3-7 falls into the range between 350ms to 800ms. Such performance is very poor for real time services. This indicates that the slotted Aloha random access cannot be directly adopted into an ATM-satellite system. Note that Figure 3-6 and Figure 3-7 give results for a system load of 0.3. The performance becomes worse when the channel load is increased. In the case that the channel is heavily loaded, the network has to monitor the channel condition and return the feedback information or control command to the terminals to recover from any further system performance degradation.

Increasing the number of retransmit times for a collided packet or a lost packet is one way to improve the packet success. However, it has been demonstrated that this method is not very effective in the presence of channel fading because the increased repeats of the lost or collided packet also add more traffic load onto the channel.
A collided packet must be retransmitted. But before it is retransmitted again, a random delay, usually called backlog delay, is applied to reduce the possibility of a second collision. This delay is usually chosen in terms of number of timeslots, that is, the terminal chooses to wait a random number of timeslots. The random number is chosen from a pre-defined range and the chosen number of timeslots is called backlog timeslots.
Without the presence of channel fading, the parameter of backlog timeslot has a clear influence on the access throughput. However, when the channel fading is involved, the influence of the backlog delay becomes insignificant. This can be seen from Figure 3-8. The parameter of backlog timeslots has less impact on the access delay compared to the impact on the throughput. As shown in the Figure 3-9, the chosen backlog timeslots in the simulation are 3, 5 and 7. These small values of backlog timeslots make no much difference on the access delay performance in both good and bad channel conditions.

From the evaluation results, we can derive the following conclusions. Firstly, the channel fading is a major factor that seriously affects the access performance when the bit error rate is beyond 1.0E-4 (channel coding is not considered here). Secondly, the system performance and stability rely greatly on the input traffic load. Access load control is necessary to prevent the channel from instability. Thirdly, the random access alone is not suitable for satellite systems with long propagation delays.

Clearly, slotted Aloha cannot offer high channel efficiency for satellite systems, or short access delays for the delay-constrained real-time services. The deployment of such a protocol on satellite has thus to be made with care. Therefore, the slotted Aloha access scheme is usually used in conjunction with other schemes.

Figure 3-8 Throughput performance of slotted Aloha in the presence of channel fading with reference to the backlog timeslots.
Issues of ATM over Satellite

Figure 3-9 Delay performance of slotted Aloha in the presence of channel fading with reference to the backlog timeslots.

Demand Assignment

Demand assignment techniques are usually complex but achieve higher bandwidth utilisation and perform well under high loads. Demand assignment involves two stages: a reservation stage followed by an assignment stage. Control of the reservation and transmission stages can be either centralised or distributed. With distributed control, users determine their actions but base their actions on the broadcast information. With centralised control, users submit their requests to a centralised control function and base their actions on the network instructions. Many recent demand assignment proposals are based on demand assignment multiple access (DAMA) proposed by Zein et al [ZEI92], such as ASCA (adaptive satellite-channel assignment) reservation protocol proposed by Sagawa and Okada [SAG95], MB/R-DAMA (Movable Boundary/Random-DAMA) proposed by Bohm et al and CFDAMA (Combined Free/Demand Assignment Multiple Access) proposed by Le-Ngoc and Mohammed [LEN93].

Adaptive Schemes

Adaptive schemes are designed to handle situations involving a combination of traffic types or a time-varying mixture. To accommodate a combination of traffic types, channels can be partitioned into several sections, each operating under its own protocol. Adaptive protocols can provide good performance over a wide range of conditions. The access scheme adapts itself smoothly to network load fluctuations.
3.3.2.2 MAC schemes for ATM-satellite

Existing satellite multiple access and bandwidth allocation techniques are generally oriented towards specific application or traffic type. They cannot efficiently handle a large number of telecommunication services with different traffic characteristics [ZEI92]. To extend ATM to satellite systems, a new MAC protocol has to be developed to expand statistical multiplexing from fixed network to satellite air interfaces. The desired MAC protocol must be able to satisfactorily handle the different ATM services (CBR, VBR, ABR and UBR) with their radically different source characteristics and QoS requirements. Additionally, the protocol must be flexible enough to allow the maximum number of users to access the satellite channels simultaneously, and robust enough to cope with the burst-errors caused by the mobile satellite environment, long propagation delays and multi-path fading and shadowing.

In the literature, many combined adaptive MAC schemes have been proposed to address this problem. The recent MAC proposals for satellite systems fall into categories of centralised and combined adaptive techniques. They attempt to design a MAC that can adapts itself smoothly to service types and network load fluctuations. It has been widely accepted that a reservation-based demand assignment MAC is a good solution. The protocol can satisfy service requirements for different traffic classes and high channel utilisation because of its built-in reservation mechanism.

Amongst the proposed schemes, the most attractive one is CFDAMA (Combined Free/Demand Assignment Multiple Access) proposed by Le-Ngoc and Mohammed [LEN93]. This proposal is an enhanced DAMA scheme. This scheme can achieve high utilisation whilst delay performance still remains competitive. According to this scheme, the scheduler first allocates the channels on the basis of reservation requests pending in the queue. In the absence of such a request, it freely assigns the remaining channels to the terminals. There are three strategies for a terminal to send its request, fixed assigned (FA), random access (RA) and piggyback (PB). The CFDAMA protocol therefore is further divided into CFDAMA-FA, CFDAMA-RA and CFDAMA-PB. A combined CFDAMA-RA with CFDAMA-PB could be selected as a promising candidate for the access scheme for signalling traffic.

3.3.3 Logic Link Control

Relatively low transmission speed and high bit error rate are the main disadvantages of satellite links. Although the end-to-end error control is performed by the ATM-Adaptation Layer (AAL), they are unable to achieve the required error performance in ATM-satellite systems and may incur serious cell discarding in ATM layer. Therefore the Logic Link
Issues of ATM over Satellite

Control (LLC) scheme is expected to perform error control on the radio link, expectedly with minimal cell transfer delay and delay variation, to enhance the error control performed by the Head Error Control (HEC) and AAL layers. This is to ensure that serious cell discarding on the ATM layer can be avoided.

The major task of LLC is to achieve a reliable and high throughput transmission on the satellite channel whilst maintaining the user required QoS, such as cell loss rate and cell transfer delay. This requires that the new LLC protocol to have a strong error control power and adaptability to different ATM service types. A key step towards achieving this lies in the improvement of the error control function to be realised by Forward Error Control (FEC) coding and Automatic-Repeat-Request (ARQ) error control functions. This is because the performance of these error control functions is prone to channel interference resulting in poor link efficiency and channel throughput on the mobile radio links. Moreover, the mobility of LEO constellations also causes a significant degradation in the performance of ARQ due to the distance-changing connection routed via ISLs and frequent satellite handovers [WAR94].

There is no doubt that a combined FEC/ARQ scheme is a good error control option for the future ATM-satellite system [CAI97]. However, as far as the ATM layer QoS is concerned, ARQ is the major factor that determines the QoS parameter of Cell Loss rate, Cell Transfer Delay and Cell Delay Variation. In this work, we will concentrate on the investigation of the ARQ protocol scheme for ATM via satellite. A discussion of the ARQ protocol, focused on a selective repeat ARQ protocol, is therefore particularly relevant and is the subject of the following section.

3.3.3.1 ARQ Protocol

ARQ aims to provide the error recovery function to satisfy the reliability requirements of no packet loss, no duplication and first-in-first-out at the data link and network level. A reliable and robust protocol with reasonable throughput efficiency and modest buffer requirement is the common goal of designing ARQ protocol design.

ARQ is realised by packet retransmission. The three basic components of ARQ are: (1) erroneous and loss packet detection, (2) informing senders of erroneous and lost packets, (3) retransmission of erroneous and loss packets. Detection of the erroneous and loss packets is performed by checking the order of sequence number of the received packets. Informing senders of missing or erroneous packets is performed by sending positive acknowledgements for the correctly received packets or negative rejects for the missing or erroneous packets.
In accordance with different methods of packet retransmission, ARQ splits into three categories: Stop and Wait (SAW), Go-Back-N (GBN) and Selective-Repeat (SR). The SAW protocol represents the simplest ARQ procedure and was implemented in early error-control systems. This scheme is simple but inherently inefficient because of the idle time spent waiting for an acknowledgement of each transmitted packet. The basic GBN ARQ scheme is; whenever a received packet is detected in error, the receiver also rejects the next N-1 received packets, even though many of them may be error free. The retransmission of many error-free packets makes the GBN quite ineffective for satellite systems with large round-trip delays and expensive bandwidth.

The SR ARQ protocol is operated in such a way that only the erroneous and lost packets are retransmitted. The transmitter keeps the newly arrived packets in a buffer before they are transmitted, and removes them when the receiver correctly receives them. The receiver acknowledges reception of the packet by an acknowledgement feedback message if the packet is correctly received. In the absence of the feedback, the packet will be retransmitted immediately when the end-to-end timer expires. Packets received out of sequence at the receiver have to be re-sequenced to make sure that packets passed to the upper layer are in order. Thus, sufficient buffer capacity is required since one erroneous packet can cause successive packets to remain in the buffer.

Logic link protocol performances are characterised by packet transfer delay and throughput efficiency. The packet transfer delay between two points in the network is defined as the elapsed time from which the first bit of a packet leaves the source point to the time the last bit arrives at the destination point. Throughput efficiency \( \eta \) is defined as the average fraction of time that the protocol uses to transmit new packets. During the rest of the time, the protocol waits for ACKs or for timeout, or retransmits copies of old packets. Two aspects seriously affect ARQ performance, channel fading condition and receiver memory capacity.

Research has shown that amongst these three ARQ protocols, SR offers significant benefits for long distance connections in terms of high link efficiency under the condition of infinite buffer at the receiver [LIN84]. But this high link efficiency is achieved at the expense of extensive buffering requirements and much complex logic at both transmitter and receiver. However, SR ARQ can still outperform the other two schemes in the finite buffer case provided that the buffer overflow is properly handled. Transmission window control is a commonly used control method.
Figure 3-10 gives a comparison of the throughput efficiency of the three ARQ protocols for window size 8. These results are produced based on the ARQ throughput analyses presented in [WAD91].

![Throughput efficiency comparison between ARQ protocols](image)

Figure 3-10 Throughput efficiency comparisons between ARQ protocols.

The above figures clearly show that SR ARQ offers much higher throughput than GBN and SAW. In addition to the throughput, SR ARQ is more robust than the other two protocols in counteracting the high packet error rate presented by the fading channel. The throughput of SR ARQ improves more quickly than GBN and SAW with the increase of channel condition. With presence of server fading, SR ARQ demonstrates a stronger robustness and maintains a much higher throughput than GBN and SAW.

### 3.3.3.2 ARQ Protocol for ATM-Satellite System

However, when applying the SR ARQ scheme to mobile satellite connections, poor performance is still a remaining problem to be solved. For instance, the evaluation results of SR ARQ protocol from the SECOMS project is shown in Figure 3-11 and Figure 3-12 and demonstrates that, for a service transmitted on a GEO satellite channel using SR ARQ protocol, the transmission quality is vulnerable to channel fading, small window size and insufficient channel bandwidth assignment. (Note that the channel coding scheme is not considered here.)
Issues of ATM over Satellite

Figure 3-11 Delay and throughput of SR ARQ versus average bit error rate on a satellite channel.
The above figures show the delay and throughput performance of the SECOMS data link layer varying with the average bit error rate in the channel. The evaluation is obtained with an assumption of end-to-end delay of 250ms because the SECOMS project is based on a GEO satellite constellation. The transmitter and receiver buffer capacities are assumed to be 64kbits. A cell bit rate traffic arrives at the transmitter buffer with average bit rate 3kbps and the channel transmission rate is 24kbps. Considering the possible buffering delay at the receiver and transmitter, the packet retransmission timer is set to be 680ms.

The high burst error rate is the major factor that causes a low throughput and long delay on the data link layer. This factor has significant impact on the cell transfer delay parameter of those delay constrained real time services such as video services. It is found that the bit error rate of 1.0E-4 is a critical point for the protocol performance. The delay and throughput degrades quickly when the bit error rate goes above 1.0E-4, but they can remain at a stable performance level when the bit error rate is below 1.0E-4. However, it can be observed that the performance remains at a low throughput of 0.125. The reason is the significantly influence of the high bit error rate and increase of the retransmission numbers.
Issues of ATM over Satellite

The window size is a system design parameter that is determined by many factors such as system capacity, services supported and packet transfer delay constraints. The mean packet transfer delays and throughput versus different window size at different bit error rates are shown in Figure 3-12, in which the delay and throughput demonstrate poor performance at a small window size. Due to satellite systems' long round-trip delay and possibly long end-to-end delay, a small window size can become a bottleneck to the traffic transmission. As a consequence, the buffer may overflow and packet dropping will occur. For some data services with stringent requirements on cell loss levels for instance, the desired QoS may not be maintained for this reason.

The window size should be large enough to allow the terminal to transmit at a full rate that enables the efficient use of all the assigned channel bandwidth. Otherwise, the allocated channel resource will be wasted and the user's QoS will not be guaranteed. However, a larger window size requires a larger buffer capacity at the receiver. A compromise may be needed to balance these requirements.

Although there have been extensive discussion of the development of efficient ARQ schemes for the satellite channel, little work is done to discuss the ARQ protocol for mobile ATM-satellite systems. ARQ protocol design becomes increasingly difficult as ATM supports many different services with different data rates, burstiness and QoS requirements.
By identifying and quantitatively studying the trade-offs and impacts that arise in the satellite environment, our primary interests lie in the study of SR ARQ for its improvement in providing improved performance of delay-constrained data traffic over bursty fading channels; its adaptability to support a variety of traffic types with different QoS requirements and its trade-offs between throughput and delay. The protocol parameters to be considered are retransmission timeout, window size, transmitter and receiver buffer size.

3.3.4 Network Security

ATM-satellite systems involve a wireless transmission media and mobility and the network has to provide security services and management to achieve two goals. The first is to protect user services’ confidentiality and data integrity to counteract security threats such as loss or corruption of information and forgery [FED98]. The second is to protect the network against unauthorised access. In the GSM system, authentication, encryption and user identity protection are the major security functions used to provide security services to the users. As a consequence, security services such as authentication, access control, data integrity confidentiality and scrambling capability has to be provided.

3.3.5 Mobility Management

Since there is no terminal mobility functions provided in ATM, a mobility related protocol has to be incorporated into the ATM protocol harmoniously with minimum signalling modifications. Broadly, mobility management can be categorised into radio mobility and network mobility. Radio mobility mainly refers to the handover process and network mobility mainly refers to the location management, which includes location update and paging [TAB97].

With the increased mobile user population, user subscription data and traffic generated by location update and paging are rapidly growing. Efficiently managing the user database and reducing the signalling load becomes a major target. In addition, call-level QoS also needs to be supported to minimise the call-dropping rate caused by handover failure in the ATM-satellite system. Handover occurs when the communication connection is temporary unavailable due to heavy shadowing, the terminal approaching the boundary of the spotbeam or the satellite moving out of the serving area. In the following we will discuss each of these aspects.
3.3.5.1 Location Management

Location Management (LM) is used to enable the system to know the current location of a powered-on mobile station and ensure that the incoming calls can be delivered to mobile users. It is essentially based on a user’s mobility and his incoming call rate characteristics. LM methods implemented in existing location management standards, Electronic Industry Association/ Telecommunication Industry Association (EIA/TIA) Interim Standard 41 (IS-41) and Global System for Mobile Communications (GSM), make use of the concept of ‘location area (LA)’ to track a user’s location. Location update (LU) procedures are used to update user location databases when the user crosses the LA boundary.

Location Update

The key issues in LM lie in the definition of the location area and the method used to update the user’s location information. An overview of recent LM approaches for the terrestrial mobile systems has been given in [TAB97]. These approaches or schemes can be divided into memoryless schemes and memory-based schemes. Memory-based schemes are intelligent types. A design of this type of LM scheme has been motivated by the fact that current systems act as a memoryless processor and repeat a number of actions that can be otherwise avoided if predicted. Provided that prediction is very efficient, it will greatly reduce the LU signalling traffic and increase successful paging probability at an early paging stage. Therefore it improves the signalling efficiency. The proposed memory-based LM schemes are mainly based on the observation and statistics of user behaviour. Examples shown in Figure 3-13 are Multilevel LAs [HU95], which define hierarchical level of LAs with different system coverage.

![Figure 3-13 Multilevel location areas.](image)

Instead of using pre-designed multilevel LA sizes, other approaches such as [TAB95] have been proposed to adjust the LA size dynamically according to each user’s mobility and LU
rate. This scheme may reduce signalling traffic but at the price of increasing computation load of a system.

An LEO satellite system has larger sized spotbeams than the terrestrial cells. But the motion of the satellites leads to a fast movement of the footprints and spotbeams. Due to a high boundary crossing rate in LEO systems, implementing individual user-based location update schemes could cause a high location update rate resulting in a heavy computation load and location update signalling load despite the low user mobility. GEO satellite systems, on the other hand, have large spotbeams that cover a large area, and the user's roaming frequency between spotbeams are relatively low, thus a periodic location update could be inefficient.

The location update methods for satellite systems are mainly divided into fixed location area methods (FLA) and dynamic location area methods (DLA) according to the research work performed at the University of Surrey [MEB95]. The FLA method is purely based on the rate of the boundary crossing. But in the DLA method, the location area boundary is time-varying and changes with certain conditions. Its LA is defined as the area centred at the position where the terminal receives its last incoming call.

![FLA and DLA concepts](image_url)

**Figure 3-14 Concepts of FLA and DLA.**

It is found that large overlaps occur in the DLA method and a larger number of location updates is required in the DLA method than the FLA method. Based on the above observation, a pre-defined FLA method, which is either single level or multilevel LA, is perceived suitable for satellite systems.
Paging Strategies

A paging strategy is much dependent on the definition of the Paging Area (PA) that is related to the location area definition. The location area can be divided into a number of PAs with their size no larger than the location area. The size and number of paging areas are decided by the statistics of the user mobility. Generally, paging can be divided into following categories:

- **One-step paging**: This scheme pages a user in all PAs of the location area at the same time. This scheme has the least paging delay but generates most paging traffic.

- **N-step paging**: This scheme prioritises PAs into groups according to the user mobility profile. A user will be paged first in the PAs with the highest apparent probability. If a user cannot be found at the first step, n (n≥2) steps will be followed until the user is found or the paging time limit is reached.

An optimal paging strategy for the ATM-satellite will be N-step paging with a high success rate at the first step. This requires that the network can predict the user’s position with high accuracy. However, this is not an easy task. In addition, because of the trade off between the location update and the paging, selection of one-step or N-step paging is also decided by the system requirement and traffic condition.

![LA and PA](image1)

(a) Paging area has the same size as the location area.

![PA and LA](image2)

(b) Paging area is smaller than the location area.

Figure 3-15 Relationship between paging area and location area.
3.3.5.2 Handover

Handover is the switching of an on-going call to a different spotbeam or cell. Handover has to be addressed in order to offer a high insurance of link continuity and call level service quality. Otherwise, it can cause a shorter break in the communication or a call drooping. There exists some issues to be addressed for handover, including handover initialisation, handover routing and handover resource management.

It is obvious that satellite movement causes more frequent handover occurrences. For this reason the system performance depends strongly on the handover performance. Handover must maintain the continuity and integrity of calls at a reasonable quality level to ensure the user required QoS. The handovers are likely to be network initiated in satellite-ATM systems due to the following advantages:

- The network has accurate information of the history of the channel quality and has advantage to make a reasonable prediction of the channel quality in the next interested time period, especially in an LEO system, where certain percentage of handovers are attributed to the satellite handover.

- The network has more information of its resource and possibly neighbouring spotbeams’ resource, and it can start to look for a new handover target immediately whilst it instructs the mobile terminal to prepare a handover.

In this case, the mobile terminal is required to submit the results of the link measurement report to the network regularly or upon request by the network. However, a suitable handover scheme needs further study.

3.3.6 Call Admission Control

Call admission control (CAC) is executed at the call set-up phase and the call re-negotiation phase to decide whether to accept or reject a new connection or an ongoing connection requesting for reallocation. Conventional call admission control in satellite systems is primarily based on available channel resources, first-come-first-served principles and a complete sharing of channel resource by different services. This simple call admission control may cause heavy congestion and result in degrading the quality of existing calls in an ATM-satellite system.

The CAC algorithm to be developed is expected to perform call admission on the basis of service class, PDR (Peak Data Rate) and Cell Loss Ratio whilst taking the existing
resource, required QoS, congestion conditions and network efficiency into account. Especially, since CAC runs at call set-up time, time consuming and complex CAC algorithms are not appropriate for real systems.

In the case of satellite networks, channel fading could seriously degrade the throughput performance of a connection, extra bandwidth may be needed to protect the user required QoS. This extra bandwidth demand needs to be addressed by a CAC although the amount of extra bandwidth is difficult to predict. In addition, the influence due to the increasingly occurring handovers should also be taken into account.

3.3.7 Traffic Management

In the ATM network, the basic traffic management specified is the Call Admission Control and Usage Parameters Control. In addition to these basic control mechanisms, congestion control, traffic shaping and flow control are also used. In the ATM-satellite system, these control mechanisms should also be supported in order to meet the QoS objectives and maintain the stability of the network operation. These procedures operate according to the service categories and with their traffic parameters.

The ATM-satellite network periodically monitors the occupancy status of each uplink and downlink carrier and broadcasts the measured status in the downlink carriers. The downlink carrier load status can be nominal, overloaded or congested. The ground terminals need to implement a proper congestion avoidance policy within short delay durations.

3.3.8 Call Control and Connection Control

To accommodate terminal mobility, an additional wireless control needs to be incorporated into the ATM network protocol. The new signalling control must be flexible enough to control functions common to both mobile and fixed ATM and to utilise sufficiently the fixed ATM network protocols. It is required to support control plane functions related to resource control and management that are specific to establishing, maintaining and releasing a wireless link over satellite. Besides this, it is expected to support mobility management functions including terminal authentication and registration, location update, power measurements and control, and handover. Signalling connection set-up is to request both VPI/VCI and channel resource to establish a signalling channel over the satellite. An efficient connection management scheme is equally important to provide services in support of set-up, maintenance and release of connections in order to establish full network connectivity.
Signalling procedures devoted to the connection control functions require a dedicated system database in order to store proper system information, which supports call control signalling.

A further discussion of call control signalling is presented in Chapter 8.

3.4 Conclusion

In this chapter, we have discussed the major issues of ATM via satellite. The focus has been on the connection management related protocols, such as multiple access control, ARQ protocol, location update and handover. It has been demonstrated that there are two major problems to be tackled in order to achieve an integrated ATM-satellite system. The first is the connection management techniques that have to be developed to provide a QoS-guaranteed transportation of ATM traffic on a mobile link over the satellite in order to ensure the ATM layer QoS. The second is an effective mobility management scheme needed to ensure that ATM-satellite systems can provide a fast user location and smooth handovers. In this work, we will focus on an investigation of the first problem.
Chapter 4 ATM-Satellite Integrated Network Protocol Architecture

Based on existing research contributions and taking into account distinguished features of satellites, we propose a fully integrated, efficient and flexible signalling and protocol architecture for ATM via satellite. We delineate the proposed integration architecture. The considered system integration scenario, proposed protocol reference model and signalling protocol architecture are presented. In addition, system connectivity, interworking function and layer protocol functions are discussed.

4.1 Introduction

An integrated ATM-satellite system needs efficient networking protocols to provide connectivity and mobility services with appropriate QoS. The system needs to provide connectivity services at least to the following call types in support of call/connection set-up, release and management.

- Fixed ATM network user originated calls to mobile ATM-satellite user (F-M).
- Mobile ATM-satellite user originated calls to Fixed ATM network user (M-F).
- Mobile ATM-satellite user originated or terminated calls to another Mobile ATM-satellite user (M-M).

The mobile ATM terminals are considered to be able to operate in dual mode. They can be directly connected to the fixed ATM network via a cable, as well as connected to satellites via a front-end device that performs the interworking function. The mobility services need to provide ATM user terminals with a roaming capability in the service area. The provided mobility should be completely transparent to the ATM network protocol's switching point to which a user is connected.

In this chapter, we propose a fully integrated and flexible signalling and protocol architecture for ATM via satellites. The proposed architecture is developed from the ATM protocol. The mobility functions and MAC, LLC functions are incorporated into the existing ATM layers. The proposed ATM-satellite protocol reference mode and signalling protocol architecture are also presented. In addition, system connectivity, interworking functions and layer protocol functions are discussed.
4.2 Integrated System Architecture

4.2.1 Interconnection Approaches

Provisioning of B-ISDN services over satellite is a common target of many on-going projects and emerging commercial mobile satellite communication systems. From the system protocol design point of view, ATM-Satellite interconnection solutions can be categorised into indirect interconnection and direct interconnection. Differences between these two approaches are illustrated in Figure 4-1.

![Diagram of indirect and direct interconnection]

(a) Indirect interconnection.

(b) Direct interconnection.

Figure 4-1 Two approaches for ATM-Satellite interconnection (MT: mobile terminal).

Indirect interconnection is a simple way to provide B-ISDN services over satellite. The satellite system only needs to provide a transmission path between the InterWorking Functions (IWF) and B-ISDN service is provided transparently to the network. As the core system is independent of external networks, it is very flexible for the network to have an efficient protocol optimised to the satellite environment only. There are few issues on the merge of the two different protocols. On the other hand, this approach may cause a heavy packet header, complex interworking function and heavy processing load due to the complete protocol interpretation between mobile terminals, satellite network and ATM network protocols.
Satellite EHF Communication for Mobile Multimedia Services (SECOMS) network is an example of such an indirect service provisioning system. At the university of Surrey, the research group worked on SECOMS from 1995-1998. Figure 4-2 illustrates its network architecture. SECOMS uses a limited number of GEO satellites working at Ka and EHF bands, capable of providing a wide range of broadband mobile services. The system operates as an underlying core access network and provides a means to transparently transfer both signalling and data traffic over satellites [MER97]. The external network protocol is terminated at the gateway stations and specific satellite terminals. The external protocol is translated by an IWF.

Direct interconnection aims to provide a fully integrated system connecting ATM and satellite. In this approach the ATM protocol is enhanced to accommodate mobility and wireless medium related functions. Modification to the ATM protocol is required to be minimal. The mobility compliant ATM terminals can set up connections across the satellite network using most of its own call control/connection protocols. Services are provided directly by the satellite system to be connected to the backbone ATM network. This scheme can be more efficient than the former approach. It is therefore chosen in realising the goal of a fully integrated ATM-satellite network.

### 4.2.2 Integrated Networking Scenario

A network architecture of the ATM-satellite integrated system is proposed in Figure 4-3, where the ATM-satellite system network with mobile ATM-satellite terminals and backbone ATM networks is shown. The system is designed to support satellite connections to the mobile terminals. The satellite network is seen as a part of the B-ISDN. But all the mobility specific signallings are terminated at a special gateway earth station. This particular earth
station performs all the interworking functions. The network elements of this ATM-satellite architecture are described in the following sections.

![ATM-Satellite System Architecture](image)

**Figure 4-3 ATM-satellite system architecture.**

### 4.2.2.1 Network Elements

**SAT (Satellite-ATM Terminal)**

The SAT is a dual mode terminal with mobility. It can be connected to the fixed ATM network by a cable. It can also be connected to satellites via front-end equipment which implements the interworking function between the satellite system and the ATM network. The terminal types considered include portable terminals and handheld terminals. The SAT supports the enhanced User-Network Interface (UNI) signalling Q.2931+M, which is incorporated with mobility management protocols.

**MCS (Master Control Station)**

It is assumed that a satellite earth control station is always integrated within an MCS. The MCS is thus in charge of tracking, monitoring and controlling the satellites. The MCS provides the overall control of the satellite network resources. This centralised node is responsible for allocating the connection identifiers and assisting in allocating frequencies and channels to the SAT. The MCS also deals with call handling and call routing. The main
protocol control functions of the MCS are to manage procedures such as synchronisation, authentication, call admission and connection handling and routing. The MCS has full visibility of the network configuration and connections, both active and those being set-up. Information supporting authentication and connection management must be available. Thus the MCS will be equipped with dedicated databases in order to store proper system information and support all procedures described above.

**HLR / VLR (Home Location Register/Visitor Location Register)**

HLR and VLR are the concepts adopted from GSM for subscribers' databases. HLR holds the subscriber's information relevant to the provisioning of telecommunications services, as well as the current location of the subscriber. As a physical machine, an HLR is typically a standalone computer, without switching capability but able to handle hundreds of thousands of subscribers [MOU92]. Another function of the HLR is the management of security data for the authentication of subscribers.

A VLR is always assumed to be integrated with an MCS in the system. The VLR is the second subscriber database. It is responsible for temporarily storing subscription data for those subscribers currently situated in the services area of a corresponding MSC.

**GTW (GaTeWay)**

The GTW station is a Fixed Earth Station (FES) with interworking functionality to support inter-connectivity to the external ATM networks. One of its tasks is to separate the signalling information relevant to the call control from the signalling information relevant to the Mobility Management (MM). It is a switch responsible for fetching the location information and routing the call towards the MSC through which the subscriber can obtain services at any instant.

**Constellation and Payload**

The satellite constellation used for integration could be a GEO, MEO or LEO constellation. It is assumed that the satellite is equipped with on-board processing switching capability. A major on-board processing is the traffic resource allocation, which is managed by a traffic resource manager. Intersatellite links are assumed to be used to route the connection through the sky. A satellite has a direct connection to the MCS, SAT and GTW. An MCS is connected to other MCS's via the satellite.
TRM (Traffic Resource Manager)

The traffic resource manager represents a centralised entity, residing on-board, that manages the satellite resources on demand. The TRM is responsible for assigning resources to an connection and is also in charge of implementing the On-Board Switch (OBS) control that allows setting up, updating and releasing the inputs/output connections [SEC97D].

The TRM needs to have a full visibility of the commitment status of the system resources, both in uplink and in downlink. The TRM will be equipped with dedicated databases, keeping the following information about status of the overall up-link and down-link resources, connection parameters such as identifier, current status, service category, resources assigned and resource holding time.

On-Board Switch

The OBS represents a component of system switching architecture capable of performing the routing of elementary traffic channels from any uplink to any downlink carrier.

4.2.2.2 Networking Connectivity

The ATM-satellite system has to offer connectivity services in support of ATM network protocols and ATM end users with mobility capability and appropriate QoS requirements. The system needs to provide full network connectivity services to the SAT, fixed ATM users and Network Operation Centre (NOC). The end-to-end connections considered to be supported are given in Table 4-1. These connections can be pre-defined or dynamically set up. The network should be capable of providing connectivity services to a full range of applications including those to emerge in the future.

<table>
<thead>
<tr>
<th>Connection Type</th>
<th>Point-to-point</th>
<th>Point-to-multipoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAT ↔ SAT</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>SAT ↔ ATM end user</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>SAT ↔ HLR</td>
<td>S</td>
<td>X</td>
</tr>
<tr>
<td>SAT ↔ NOC</td>
<td>S</td>
<td>X</td>
</tr>
<tr>
<td>GTW ↔ GTW</td>
<td>S</td>
<td>X</td>
</tr>
<tr>
<td>MCS ↔ MCS</td>
<td>S</td>
<td>X</td>
</tr>
<tr>
<td>MCS ↔ NOC</td>
<td>S</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 4-1 End-to-end connection supported in ATM-satellite network. (S: Support, X: Not support.)
4.3 Network Signalling Protocol Architecture

4.3.1 Protocol Reference Model

The ATM-satellite protocol reference model is expected to incorporate the mobility and wireless access functions into standard ATM protocol reference model. There are three approaches to define the ATM-Satellite integrated protocol architecture.

The approach shown in Figure 4-4 (a) almost completely separates radio connection and mobility-related signalling from ATM signalling. This approach is similar to an indirect ATM-satellite integration. Because the ATM protocol runs on top of a satellite specific protocol, the ATM data and signalling traffic are completely encapsulated in the satellite specific packets. Hence approach (a) ends up with signalling inefficiency caused by heavy header and increased processing loads.

The approach in Figure 4-4 (b) incorporates all radio and mobility-related signalling into a top layer. This approach is inefficient in bandwidth utilisation. The approach simply implies that all the logical channels at the radio interface must have the timeslot equivalent to that of the encoded packet, which consists of a radio related header and ATM cell. However, some logic channels, such as paging, random access and reservation channels, may not require such a large packet size and high transmission speed. But the bandwidth cannot be tailored into different sized timeslots in this approach. Therefore, this proposal is not considered flexible enough to achieve high bandwidth utilisation.
Based on the above discussion, we propose the protocol reference model shown in Figure 4-5. This protocol is made according to the principle of isolating radio connection-related signalling from fixed ATM-related and from mobility-related signalling. The mobility specific functions, such as location update and security service, are implemented at top layer in conjunction with Q.2931. The mobility enhanced Q.2931 is thus called Q.2931^{M}. Some other mobility specific signalling, such as handover and paging, are performed at the Radio Control Layer (RCL). Logic Link Control (LLC) and MAC layer are provided below the ATM layer to ensure reliable packet transmission.

This protocol retains the advantage of the ATM protocol. At the same time, it retains certain flexibility for managing the radio connection and resources. The RCL is responsible for radio connection set-up, release and quality management.
4.3.2 Interfaces

The signalling interfaces have been identified to provide reference points for protocol design. The interfaces are given in Figure 4-6. These interfaces and their protocol functions are summarised in Table 4-2.

The interface A between the SAT terminal and the MCS via the satellite is considered as a User Network Interface (UNI), which performs most of the ATM UNI signalling functions. There is an argument that the UNI interface should be between the SAT and GTW because the GTW has the IWF and acts mostly as an ATM switch. However, because the MCS is a centralised node that has full control of the network resource and administration, the relationship between the MCS and the GTW is a master-slave relationship. The GTW follows the MCS’s instructions to establish and clear satellite connections. It is also noted that the GTW is a switch and it handles its own processing and buffering capacity. For this reason, it can be regarded as an independent peer switch to the MCS station. These two considerations make it difficult to decide the protocol function of interface B, which is between the MCS and GTW via satellite. However, there are two other considerations. Firstly, the UNI and Network-Network Interface (NNI) are very similar to each other. Secondly, if we assume that the GTW is always attached to an ATM switch, the GTW can then be regarded as a mobility enhanced ATM switch, named as ATM+. The relationship between ATM+ and MCS is virtually a peer-to-peer relationship. On the basis of these two points, the interface B is defined as an NNI interface and so is the interface F.
ATM-Satellite Integrated Network Protocol Architecture

<table>
<thead>
<tr>
<th>Interface Name</th>
<th>Involved Function Entities</th>
<th>Interface Type</th>
<th>Interpretation to the B-ISDN Interface</th>
<th>Function Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>SAT-MCS</td>
<td>Virtual</td>
<td>UNI</td>
<td>Q.2931 signalling, mobility control, radio connection control</td>
</tr>
<tr>
<td>B</td>
<td>MCS-GTW</td>
<td>Virtual</td>
<td>NNI</td>
<td>NNI signalling, mobility control, radio connection control</td>
</tr>
<tr>
<td>C</td>
<td>SAT-P/L</td>
<td>Physical</td>
<td>-</td>
<td>Resource management and radio transmission/reception control.</td>
</tr>
<tr>
<td>D</td>
<td>MCS-P/L</td>
<td>Physical</td>
<td>-</td>
<td>Resource management and radio transmission/reception and payload control.</td>
</tr>
<tr>
<td>E</td>
<td>GTW-P/L</td>
<td>Physical</td>
<td>-</td>
<td>Resource management and radio transmission/reception control.</td>
</tr>
<tr>
<td>F</td>
<td>GTW-ATM</td>
<td>Physical</td>
<td>NNI</td>
<td>NNI signalling, IWP function, mobility management.</td>
</tr>
<tr>
<td>O</td>
<td>SAT-GTW</td>
<td>Virtual</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4-2 Definitions and protocol functions of ATM-satellite network interfaces.

Interfaces C, E and D are between the SAT/GTW/MCS and payload. They co-ordinate the lower layer functions related to resource requests, allocation and release plus some others and have physical layer and MAC layer signalling communications.

4.3.3 Network Protocol Architecture

Based on the reference interfaces defined above, a network signalling protocol architecture is proposed. A design of the ATM-satellite signalling protocol architecture mainly addresses three signalling issues. They are:

- the communication between functional entities in support of mobility procedures,
- the routing of signalling messages,
- the radio connection management related features such as logic link control, MAC and handover.

Control Plane

Figure 4-7 shows the protocol architecture of the control plane. In this control plane, the radio connection set-up, maintenance and release functions are completed by the signalling exchanges amongst SAT, MCS and GTW. But the RCL layer signalling requests the services from the MAC layer in which signalling interacts between SAT, MCS, GTW and TRM. The resource allocation related signalling is terminated at the satellite TRM.
It is seen that the ATM and AAL layer protocols of the user terminals comply with the ATM standards. Q.2931™ protocol has mobility protocols that are implemented within the satellite network and are terminated at the IWF of GTW. Because the UNI lies logically between the SAT and the MCS, MCS Q.2931™ handles the call handling functions meanwhile.

**User Plane**

The protocol architecture of the user plane is shown in Figure 4-8. User plane information transmissions are mainly conducted between the SAT and GTW, based on the LLC layer. The encapsulated ATM packets rather than ATM cells are transmitted on the virtual connection between SAT and GTW. The encapsulated ATM packet is actually an ATM cell with a header added from the LLC and RLC layer. The user plane ATM layer transmission also relies on services from the MAC layer to adjust the user’s short term resource requirements and resource allocation, especially in the case of semi-permanent connections. The AAL layer peer connection lies directly between the SAT and the fixed ATM end-user.
Mobility Plane

In addition to the user and control planes, a mobility management (MM) plane is also introduced. The MM plane is specifically dedicated to the handling of procedures concerning personal mobility. We assume that the VLR resides with the MCS and an interface is provided between the MCS and HLR for the MCS to request necessary subscribers' information. The mobility related function is actually managed by the MCS, HLR and GTW. The GTW filters messages submitted by the Q.2931 and generates security and user location related mobility signalling messages. It co-operates with the HLR and MCS to deliver incoming calls to called mobile users. The MCS terminates the mobility related signalling from the SAT and obtains subscribers' database information from the VLR or HLR. When a mobile user requests to change its subscription database, the MM layer signalling will enable the MCS to set up an end-to-end connection between MCS and HLR.

![Diagram of ATM-satellite protocol architecture](image)

**Figure 4-9 ATM-satellite protocol architecture—Mobility plane.**

### 4.3.4 Layered Protocol Functions

The layered protocol functions decide services that can be provided from each layer to a higher layer.

#### 4.3.4.1 Physical Layer Protocol Functions

The major functions of the physical layer include formatting and transmission/reception of uplink TDMA and downlink TDM frames, FEC encoding and decoding, modulation and demodulation, etc [POD98]. Commonly used FEC schemes for satellite systems are Reed Solomon (RS) and convolutional coding. The QPSK is the usually used modulation scheme.
The satellite radio link channels are usually organised into several frame levels e.g. frame, multi-frame and super-frame. These frames are further organised to different logical channels for transmission of different signalling and data traffic. The full transmission rate at the physical layer is mainly determined by the system bandwidth.

The satellite bandwidth is a limited resource. Typical carrier frequencies for current fixed point-to-point satellite services are 6/4 GHz (C band) and 14/12 GHz (Ku band). Traditional C and Ku-band transponder bandwidths are 36 MHz and 72 MHz [ZHA97]. To support multimedia services, the need for a larger bandwidth has pushed the satellite systems to move up to a high frequency of 30/20 GHz Ka-band. For example, the ACTS project, SECOMS, (and its follow-up commercial system EuroSkyWay) operates at Ka-band and EHF band (40/50 GHz). Its proposed frame structure is shown in Figure 4-10. The full transmission rate can reach up to 2.048Mbps on the uplink and 32.768Mbps on the downlink. The forward error control encoder provides Reed Solomon (RS) (80, 64) encoding and differential encoding of all frame units.

4.3.4.2 Multiple Access Layer Protocol Functions

The MAC layer deals mainly with resource, multiplexing and collision resolution related functions. These functions include:

- Initial access,
- Out-of-band channel reservation,
- Management of transmission modes,
- Recording and updating assignment information of each active connection,
- Resolution of collision and retransmission,
• Multiplexing and De-multiplexing.

The terminal uses an initial access when it generates a new call and tries to establish a new connection via the satellite network. During the communication, it also performs out-of-band channel reservation when extra bandwidth is needed.

An in-depth study of multiple access schemes and the MAC scheme proposed for ATM-satellite system will be presented in chapter 5.

4.3.4.3 Logical Link Control Layer Protocol Function

The LLC can provide a reliable on the radio connection. It performs error detection and correction using CRC (Cyclic Redundancy Check). Those corrupted packets that cannot be corrected by CRC will be recovered by ARQ retransmission. In addition, LLC is also designed to provide a relaxed reliable transmission in ATM-satellite system. A modified selective repeat ARQ scheme (SR ARQ), named as the reliability-dependent SR ARQ discussed in Chapter 6, is used at this layer to cope with inefficiency of the conventional SR ARQ. The proposed protocol also has an adaptive timer mechanism. The major logic link layer functions include:

- Satellite packet header encapsulation/extraction,
- Error detection and correction, ARQ,
- Cell sequence ordering,
- Detection of transmission and operational errors on a link connection,
- Recovery from detected errors,
- Flow control.

The packet structure at the LLC layer is assumed to be similar to the High-level DLC (HDLC) packet structure but contains different header information. The packet structure is shown in Figure 4-11. The contained information elements should include Connection Identifier (CI), Radio control layer Message Type (RMT), In-Band Resource Request (IBRR), the Nth Packet Retransmission (NPR), Packet Sequence Control (PSC), ATM cell and CRC. CI uniquely identifies one logical link between source and destination. RMT indicates whether the carried payload is LLC layer signalling message or user data. IBRR is used to reserve uplink resource and NPR gives the number of times that the current packet has been transmitted from the sender. PSC field is the same as the control field in HDLC frame. It mainly includes send-packet sequence number and receive-packet sequence number.
4.3.4.4 Radio Connection Control Layer Protocol Functions

The radio connection control layer performs connection management functions to set up, release and maintain radio connections for signalling or data traffic. It is also in charge of controlling requirements and release of the channel resource. It determines the protocol operation mode and operation parameters of the LLC and MAC layer for each application according to its service types, QoS requirements and sends operation instructions to the LLC and MAC layer. It should be noted that this layer has some mobility related functions such as paging and handover. One of the reasons to locate these functions at the RCL layer is the aim to provide a smooth handover and continuous radio connection to the ATM layer. A list of the RCL layer function is given below:

- Connection Control,
- Resource and traffic Management,
- Paging process,
- Handover,
- Attach/ Detach procedure.

An in-band channel reservation function is provided at this layer for the SAT to send a request to prolong allowed transmission time or to request extra channel resources. This is performed by appropriately setting the IBRR request field inside the extended satellite ATM packet. The IBRR is a field with a length of few bits to keep the packet header light.

4.3.4.5 Q.2931\textsuperscript{M1} Layer Protocol Functions

The Q.2931\textsuperscript{M} is responsible for providing call control services with set-up, maintenance and release of either a mobile originated or a mobile terminated call, which can be a point-to-point call or a point-to-multipoint call. Q.2931\textsuperscript{M} also provides the mobile ATM user terminal mobility services completely transparent to the ATM switching point to which the user is connected.

\footnote{Q.2931\textsuperscript{M} is a mobility enhanced Q.2931 protocol.}
For call control aspects, Q.2931*M utilises most of the Q.2931 protocol function and should provide the following services:

- Call set-up and release,
- Call parameters updating,
- Call management procedure,
- Related resource and traffic management,
- Routing management.

In addition to the Q.2931 call control function, the mobility management provided in Q.2931*M has a relatively independent control function specifically dedicated to handling procedures concerning the personal mobility so that the mobility management signalling can be separated from the call control signalling. The mobility management functions considered include:

- Security,
- Authentication,
- Location Update,
- Registration.

4.4 Conclusion

In this chapter, we have proposed a new and fully integrated ATM-satellite signalling protocol architecture, which accounts for mobility of ATM users. The proposed integration scenario, network connectivity, protocol reference model and layered protocol functions have been presented. This architecture will act as a general basis for the overall system design. Further studies on multiple access control and logic link control protocol, as well as connection management and call control, will be presented in the following chapters. We will focus on the optimisation of these protocols to provide QoS-provisioning transportation of ATM traffic over mobile satellite links and the enhancement of ATM signalling protocols to include the mobility and radio resource management.
Chapter 5 Optimisation of MAC Protocol

In this chapter, we introduce a service-adaptive MAC protocol. The protocol is based on a reservation-based demand assignment scheme. The major advantage of this scheme is that it offers high channel efficiency. We also propose a semi-permanent signalling protocol for conveying signalling traffic. Applying dynamic resource assignment instead of conventional fixed resource assignment to signalling traffic is the essence of this scheme. The focus of this chapter is on exhibiting performance evaluation and optimisation results of the semi-permanent signalling protocol in order to demonstrate its advantages of bandwidth efficiency over other schemes.

5.1 Introduction

In order to achieve an efficient connection management scheme to support different ATM services with a minimal provisioning of satellite bandwidth, the MAC and ARQ protocols need to be optimised to adapt flexibly with the service types. This is because that connection management utilises the services provided by the MAC and ARQ.

Satellite systems have limited bandwidth resource and a flexible and effective MAC protocol is needed to provide a wide range of multimedia services and to achieve high bandwidth utilisation. As discussed in Chapter 3, a combined reservation/demand multiple access has been identified as a suitable MAC scheme for the ATM-satellite system because this scheme can achieve high bandwidth utilisation for both LEO/MEO and GEO satellite system. Its access delay performance can also remain reasonably competitive. Based on this scheme, in this chapter, we present a new service-adaptive assignment-on-demand MAC protocol.

It has also been found that although the reservation-based demand assignment access has been studied (see, for example, [ZEH92][SAG95][HUN97]), the work has focused on optimising and evaluating the MAC scheme for data traffic. Very little work has been devoted to optimising and evaluating the MAC scheme to the quality of the signalling traffic. The resource allocation scheme has conventionally used a fixed resource assignment for satellite systems. The fixed resource assignment for signalling traffic is useful in view of the special requirements for regular exchange of signalling messages, stringent requirements on signalling delay and minimisation of signalling failure. However, the signalling traffic has strong on-off characteristics and short burst length in satellite systems. Thus the signalling channel stays idle during most of the signalling periods, resulting in bandwidth wastage. This situation is even worse for a GEO satellite system with a delay of 250ms.

Due to the expensive satellite bandwidth, a high bandwidth utilisation and high throughput are also strongly desired on signalling connections. Hence in this chapter we also examine
applications of the reservation/demand MAC protocol to the signalling traffic. Such an application imposes a requirement to slightly modify signalling procedures in order to incorporate the proposed MAC procedure. Thus a signalling protocol that uses the proposed MAC procedure is called “Semi-Permanent Signalling Protocol (SPSP)” in this work.

This Chapter is devoted to optimisation of the SPSP on both an error-free channel and a mobile fading channel. By evaluating performance of the SPSP and addressing its impacts on signalling delay and signalling failure probability, we aim to demonstrate improvements in channel utilization and system capacity in the presence of channel fading for the new SPSP scheme.

5.2 MAC Protocol for ATM-Satellite

5.2.1 Frame Structure

The frame structure of the MAC protocol to be considered in this work is shown in Figure 5-1. The uplink physical channel is a TDMA based. Each TDMA frame has F timeslots and is subdivided into uplink synchronisation channels (USC), resource reservation channels (RRC) and traffic channels (TCH). The RRC logical channel comprises L timeslots, each further subdivided into V minislots. Each minislot holds a reservation request packet with a size normally between 10-20 bytes [PET96]. The LV minislots in the RRC channel are slotted Aloha channels so that the packet collision may occur due to contention. The RRC channel can be divided into two groups, one is for user access and the other is for resource reservation. There are F-L data slots in the TCH channel, each holding one data packet within the extended ATM-packet-size slots for transmission of user information. The boundary between reservation slots and data slots is movable according to the traffic volume. If needed, the network controller can convert a data slot into a reservation slot with V minislots.
The downlink (DL) channel contains five types of logical channels, downlink synchronisation channel (DSC), resource assignment channel (RAC), feedback channel (FC), common signalling channel (CSC) and traffic channel (TCH). The FC channel also comprises LV minislots with a size similar to resource request message. Each FC minislot contains the feedback information regarding success or collision of each RRC minislot. If the information obtained from the FC minislots indicates that a collision occurs in the minislot in which the last request is sent, the mobile terminal will resend the last resource request packet. The RAC channel contains the resource assignment message sent to a user who requests the resource. RAC channel carries a number of successful assignments made by the Traffic Resource Manager (TRM) every frame. The size of the resource assignment message is determined by
the amount of information included but is expected to be smaller than the TCH size. The USC and DSC provide synchronisation information. We assume that all the mobile terminals are synchronised with respect to minislot, time slot and frame boundaries.

Three boundaries are defined, one between the RRC channel and the TCH channel, one between RAC channel and TCH channel and one between FC and TCH channel. These boundaries are movable so as to adapt to channel load conditions and achieve maximum channel efficiency. The suggested packet formats for RRC, RAC, FC and TCH channels are given in Figure 5-2.

<table>
<thead>
<tr>
<th>RRC</th>
<th>Traffic type</th>
<th>Request ID</th>
<th>Number of channels required</th>
<th>Reason for request</th>
<th>Other</th>
<th>CRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAC</td>
<td>Request ID</td>
<td>Number of channels assigned</td>
<td>Holding Duration</td>
<td>Assignment details</td>
<td>Other</td>
<td>CRC</td>
</tr>
<tr>
<td>FC</td>
<td>RRC minislot ID</td>
<td>Result of access contention</td>
<td>Collision resolution instructions</td>
<td>Other</td>
<td>CRC</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5-2 Major message elements for RRC, RAC, FC and TCH channels.

5.2.2 Request Methods

A resource request is initiated in three cases. It is generated at the initial access due to a call set-up or a handover request. A resource request could also be generated in-call when the terminal buffer overflow is detected in order to obtain more channel resource. For some non-real time services that do not hold a fixed assigned channel, a resource request has to be made every time when new date packets are generated.

There are two strategies for a terminal to send its resource reservation, random access (RA) and piggyback (PB). The RA strategy requests the user to send an out-of-band request in a minislot of the RRC channel. Collision information of the RRC minislot is broadcast in the downlink FC channel. Collided requests on the reservation channel will be retransmitted after a backlog randomisation delay. The PB strategy utilises a mini message embedded in the user data message. This mini message consists of only a few bits to declare how many slots it needs in the next assignment of TRM. When a mobile terminal accesses the network using the
RA strategy, it should include an ID in its request, and the network uses this ID to label the assignment message.

### 5.2.3 Scheduling Algorithm

Upon receipt of a resource request, the TRM schedules all the pending requests according to their priority and updates its request record (RCD). The TRM considers the five types of traffic, CBR, Real-time VBR (Rt-VBR), Non-rt VBR, ABR and UBR, to have a request priority given as [SAN97],

\[
\text{CBR} > \text{Rt-VBR} > \text{Non-rt VBR} > \text{ABR} > \text{UBR} \quad (5-1)
\]

The TRM first allocates the channels to requests pending in the queue. In the absence of such requests, it freely assigns remaining slots to terminals according to service priority and their QoS requirements. The channels are assigned dynamically based on a slot-by-slot basis. In the case of lack of resource, the pending request will be kept in the buffer until the next assignment.

Upon successfully completing a resource assignment, the TRM will update its resource assignment table and send an assignment message to the terminal. The number of slots assigned, the slot positions, frequencies and the resource holding time, are included in the assignment message. The resource holding time varies from one frame duration to the whole communication duration.

### 5.2.4 Service Adaptability

The proposed MAC scheme is proposed to provide three different resource reservation and obtaining methods for a radio connection. The first one is permanent reservation. The resource is reserved during the connection set-up and remains unchanged throughout the whole communication. Such a connection can only use piggyback method to slightly adjust his resource request. A connection that uses permanent reservation is called a permanent connection. The second one is semi-permanent reservation. With this type of reservation, the resource is requested only when packets are generated during the communication. The reserved channel is allocated for only a fixed time interval depending on the source characteristics and on the resource status of network. A connection that uses semi-permanent reservation is called a semi-permanent connection. The last one is mixed reservation that is a combined permanent and semi-permanent reservation. Apart from the reservation method, the
resource allocation is also designed to be adaptive based on a concept of effective bandwidth that will be discussed in chapter 7.

The MAC scheme adopts different reservation and resource allocation method for different service types. Selection of the reservation mode for CBR, VBR, UBR and ABR services needs to take into consideration the service source characteristics and QoS requirements. The detailed selection scheme will be studied and presented in Chapter 7.

5.3 Optimisation of MAC Protocol for Signalling Traffic

As has been mentioned in the introduction, satellite systems have so far adopted a fixed channel allocation scheme for signalling connections. But signalling traffic has strong on-off characteristic, short burst lengths, and channels that carry no information for most of the signalling period which results in bandwidth wastage. Therefore, there is a need to improve satellite channel efficiency by applying the proposed demand assignment MAC scheme to signalling connections.

However, we should note that the signalling traffic does have very different traffic source characteristics and service quality requirements from data traffic. Signalling traffic cannot be well represented by the traffic parameters such as Peak Cell Rate (PCR) and Minimum Cell Rate (MCR). It also has different service quality requirements, which cannot be sufficiently characterised by the QoS parameters defined to characterise performance of an end-to-end connection. The reason is that the signalling traffic has stringent requirements on both signalling delay and signalling failure probability. Thus we need to investigate the impacts of applying demand assignment MAC scheme on signalling delay and signalling failure performance and to validate for the desired application.

5.3.1 Semi-permanent Signalling Protocol

The difference between the semi-permanent signalling protocol and the conventional signalling protocol, which uses a fixed assigned channel throughout the signalling duration, mainly lies in the method of sending signalling messages on the uplink from a mobile terminal to the satellite. This difference is illustrated in Figure 5-3 and Figure 5-4. The conventional signalling protocol is referred to as "permanent signalling protocol (PSP)" in this thesis.
According to the PSP, a dedicated signalling channel with fixed bandwidth is assigned to an end user. Every time when a new signalling message is generated, the packets of this message are sent directly on the pre-assigned channel. But in the SPSP, every time that new signalling packets are generated, the mobile terminal has to go through a reservation procedure to obtain a number of timeslots, equal to the number of the newly generated signalling packets, before sending the packets to the channel.
An end-to-end signalling connection could include many different function nodes that may involve satellite switches, GTW and ATM switches. But only the mobile link between the mobile terminal and the satellite is considered to be a semi-permanent mode as shown in Figure 5-5, wherein SAT is satellite-ATM terminal and P/L is satellite payload. All other link segments are permanent modes. This is because the bandwidth limitation usually experienced at the air interface. If a permanent connection mode is implemented at the air interface, a high bandwidth utilisation and multiplexing gain would not be achieved on the mobile uplink channels.

The semi-permanent signalling procedure has two major phases. The first phase, called initial access phase, is to access the network and to be admitted by the network to have a semi-permanent connection. The network assigns a VPI/VCI or a connection identifier to a user for use in the following signalling. The second phase refers to the stage of signalling communication between end users for setting up a traffic connection. During the second phase, the network dynamically assigns the channel resource to users on a slot-by-slot basis and follows an assignment-on-demand scheme. Before transmitting any newly generated signalling packets in a terminal buffer, users have to send a resource request mini-packet on the RRC channel to the TRM to reserve timeslots. If collision information is received in the FC channel, the request will be retransmitted after a backlog randomisation delay. The received resource requests are buffered on-board and served by the TRM according to their priority.

The protocol has an additional mechanism to combat the channel fading. The mechanism is designed to send multiple copies of a resource request packet on the RRC channel to reduce a probability of packet loss due to severe fading effects. This is to avoid an unnecessary increase of the signalling delay and signalling failure.
Whilst sending signalling packets in assigned timeslots, an end user sets an end-to-end response timer. If no signalling response from the other end is received in a timer period, the packets will be retransmitted within a restricted number of retransmissions defined by a maximum signalling delay limit. If the response cannot be received after all such retransmissions, a signalling failure will be assumed and the signalling processing will be terminated. Upon receiving a response from the other end before this timer is expired, new signalling packets will be generated and transmitted following the same procedure described above.

### 5.3.2 Evaluation Approach

The proposed semi-permanent protocol is applied to a mobile originated call set-up signalling procedure shown in Figure 5-6. The first phase of this signalling is to set up an internal connection in the satellite domain. This is completed by sending an internal connection set-up message to the MCS, being accepted by the MCS and acknowledged by the GTW. End-to-end overlaying call set-up signalling is the second phase and is performed on a semi-permanent connection until the call is set up. The on-board TRM is responsible for assigning a signalling channel to perform overlay signalling.

The system used in the evaluation of the SPSP is a mobile multimedia GEO satellite system with on-broad processing capability. It uses a slotted TDMA/TDD channel at Ka and EHF band [FAN98]. The frame structure includes 10 resource request minislots, 1 random access signalling channel (RASC) and 8 TCH channels. Five types of signalling traffic, namely, mobile originated, mobile terminated, location update, handover and registration, are the system input signalling traffic sharing the random access channel.

Mobile originated and mobile terminated call arrival rates $\lambda$ are modelled by a Poisson distribution. Then the number $n$ of calls originated or terminated at a given time interval $T$ will be,

$$P[n] = e^{-\lambda T} \frac{(\lambda T)^n}{n!}$$  \hspace{1cm} (5-2)

The LU and HO rates are also assumed to be Poisson distributed with mean arrival rates derived using the mobility model presented in Chapter 3. Assuming that the LA and PA have same size, the HO and LU rates are related to $\lambda$ by the following equations,
Optimisation of MAC Protocol

\[ P_{ho} = \frac{2vR}{A_s} \left( \frac{t_h}{t_h + \frac{1}{\lambda}} \right) \]  

\[ P_{lu} = \frac{2vR}{A_s} \left[ 1 - \frac{t_h}{t_h + \frac{1}{\lambda}} \right] \]  

A low arrival rate is assumed for registration signalling. To evaluate fading channel effects, the bursty mobile channel model proposed by Zorzi is adopted. Because the mobile channel fading is normally slow fading for a low speed mobile terminal, only a slow fading channel case is taken into considerations. The related protocols and signalling procedures, such as multiple access, resource allocation, collision recovery and lost packet recovery, are also modelled and simulated.

![Diagram showing mobile originated call set-up signalling procedure](image)

**Figure 5-6** The mobile originated call set-up signalling procedure (SABM: set asynchronous balanced mode).

### 5.3.3 Performance Parameters

We use the following performance parameters to measure and characterise the performance of the SPSP protocol and compare performance with PSP.

- **Signalling service accommodating capacity**: a number of end-to-end call set-up signalling services to be accommodated per time unit.
· **Signalling resource utilisation**: ratio of occupied signalling resources to total signalling traffic resources.

· **Connection request dropping probability**: ratio of newly generated end-to-end connection request dropped by on-board buffer to newly generated calls that arrive at the buffer.

· **Semi-permanent call set-up signalling failure probability**: ratio of semi-permanent calls dropped during an end-to-end overlaying signalling state to newly generated calls that arrive at the buffer.

These parameters can reflect a level of protocol efficiency with respect to signalling delay, service accommodating capacity, resource utilisation and reliability.

### 5.3.4 Simulation Results of the Performance in an Error-Free Channel

There are three parameters that have a significant effect on performance of the semi-permanent protocol. They are input traffic load, on-board buffer size, on-board buffer arrival process and channel holding time. The on-board satellite buffer has two input arrival streams in this protocol. They are outputs of the RASC and RRC channels. The buffer arrival processes in the semi-permanent and permanent protocols are therefore different because the permanent protocol has only one output from RASC. The semi-permanent protocol has a higher buffer input rate than the permanent one, but its channel holding time is shorter. The reason is that the channel holding time is the duration of the call set-up in the permanent case and is one frame duration in the semi-permanent case.

An end-to-end call set-up signalling failure occurs only in the semi-permanent case, not in the permanent case. The reason is that the semi-permanent protocol has risks of end-to-end signalling connection dropping due to a long-period resource request blocking caused by buffer overflow. If blocking time exceeds a certain value and a call set-up response from the other end cannot be received before a timer expires, the call processing will be terminated. A signalling failure then occurs. To reduce the signalling failure, schemes such as increasing resource request retransmission times, enlarging buffering capacity, or a combination of both can be adopted. The approach of enlarging the buffer is better than that of increasing the resource request retransmission times due to the fact that the former has lower retransmission delay and results in lower channel traffic. However, on-board buffer size is a compromise value chosen to satisfy both the requirements of lower buffering delay and lower packet loss probability. Since the resource request retransmission is not preferred, if the uplink channel quality is considerably good, resending of lost requests is not implemented in order to avoid RRC channel overloading and to maintain an acceptable semi-permanent call set-up delay performance.
5.3.4.1 Simulation Results and Discussions

During the simulation, the signalling channel resource is assumed to have a separate resource pool from that of the user data traffic. The total signalling resource is a parameter varying from 2 to 8 timeslots per frame during the simulation. Optimised parameter call arrival rate $\lambda$ and on-board buffer size $Q$ are chosen to be 11.11 (arrival/second) and 5 or 2 (packets) respectively after extensive simulations over a wide range of parameters. With these parameters the system is set to work at an optimum RASC channel load of around 1 (packet/timeslot).

Simulation shows that the delay on the RRC channel is very significant in the case of the semi-permanent protocol. This is because users have to send requests on the RRC random access channel to reserve timeslots every time a newly generated signalling message arrives. Moreover the RRC channel output seriously influences the arrival distribution at the on-board buffer. Our simulations show that an inappropriate design of the semi-permanent connection signalling protocol could end up with unrecoverable system congestion.

Performances of the semi-permanent protocol are shown in Figure 5-7 to Figure 5-10. The evaluated performance parameters include connection request dropping o probability, signalling failure probability, call set-up signalling service accommodating capacity, call set-up delay, resource utilisation and mean buffering delay. We can see that when service accommodating capacity and channel utilisation respectively are as high as 13.1 (service/s) and 0.68, the call set-up delay is only 3.4 s.

![Figure 5-7 Semi-permanent signalling channel utilisation.](image-url)
Figure 5-8 Semi-permanent call set-up delay.

Signalling failure probability and connection request dropping probability seen in Figure 5-10 and Figure 5-11 are very dependent on the on-board buffer size. With parameter value setting at buffer size 5 (packets) and channel capacity 2 (timeslots/per frame), the connection request dropping and signalling failure probability can be reduced to 0.0083 and 0.0079 respectively. When the channel capacity is increased to 3, these values change to 0.0004 and 0.0001 respectively. Moreover, another important result is that signalling failure probability is close to zero when the channel capacity is increased to 4. From the performance of the semi-permanent protocol, an optimised value 4 for signalling channel capacity with buffer size 5 (packets) on board can be derived. Under these conditions, call set-up delay and signalling service accommodating capacity are close to optimum. Figure 5-12 shows the mean buffer delay of the resource requests in the SPSP protocol case.

Figure 5-9 Semi-permanent call set-up signalling service accommodating capacity.
Figure 5-10 Semi-permanent call set-up signalling failure probability.

Figure 5-11 Semi-permanent connection request dropping probability.

Figure 5-12 Semi-permanent call set-up onboard buffering delay.
Performance comparisons between semi-permanent and permanent call set-up under the same input traffic load and buffer size are shown in Figure 5-13 to Figure 5-14. Figure 5-13 reveals that permanent call set-up delay varies exponentially with the channel capacity but semi-permanent delay remains almost at constant. It should be noted that the advantage of the permanent connection on short call set-up delay does not always exist. During the network busy hours with buffer size over 5 (packets) on board, delay in permanent connection could be much worse than for the semi-permanent connection. Permanent connection does give a shorter delay performance of 0.342 s when the buffer size is decreased to 2 (packets). However, Figure 5-14 and Figure 5-15 show that in order to achieve this 0.342s, the service accommodating capacity and channel utilisation suffer reductions of 58% and 10%, respectively. Performance of permanent connection with buffer size 5 and 2 (packets) show an almost identical performance because the system is overloaded. The maximum capacity of the permanent connection is 0.07 channel utilisation and 5.45 (service/s) in service accommodating capacity. However, under the same traffic input condition, overload does not occur in the semi-permanent protocol. This explains why the permanent protocol has much less capacity than the semi-permanent protocol.

![Figure 5-13 Comparison of mean call set-up delay (same traffic input) between permanent and semi-permanent connections.](image-url)
5.3.4.2 Simulation Discussion

For the GEO satellite system, 0.342s may cause a slight degradation of the delay performance, but improvements of the service accommodating capacity and channel utilisation are significant enough for a multimedia satellite system to support an increasing user population. Moreover, with an increase in on-board satellite memory capacity, the semi-permanent performance can be further enhanced. From this point of view, the semi-permanent connection does offer more advantages than the permanent connection.

The above simulation results demonstrate that implementing a semi-permanent MAC protocol can considerably improve the signalling service accommodating capacity and channel utilisation, whilst retaining negligible signalling failure probability. The mean call set-up delay of the proposed scheme also demonstrates that it has a better performance during busy
hours with a buffer size larger than 5 packets on board the satellite. From the performance evaluation, we conclude that the semi-permanent call set-up protocol is capable of offering more advantages than the permanent call set-up protocol for multimedia satellite systems for a good channel condition.

<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>frame duration</td>
<td>0.0265 s</td>
</tr>
<tr>
<td>timeslot duration</td>
<td>0.00265 s</td>
</tr>
<tr>
<td>RASC channels</td>
<td>1 / frame</td>
</tr>
<tr>
<td>RRC resource request channels</td>
<td>10 / frame</td>
</tr>
<tr>
<td>minislot duration</td>
<td>0.000265 s</td>
</tr>
<tr>
<td>dedicated signalling channels</td>
<td>variable (2-8)</td>
</tr>
<tr>
<td>propagation delay from transmitter to receiver</td>
<td>0.25 s</td>
</tr>
<tr>
<td>TRM-SaT uplink speed</td>
<td>160 kbps</td>
</tr>
<tr>
<td>TRM-MCS, TRM-GTW uplink speed</td>
<td>32.768 Mbps</td>
</tr>
<tr>
<td>downlink TDM speed</td>
<td>32.768 Mbps</td>
</tr>
<tr>
<td>physical layer average processing delay</td>
<td>0.001 s</td>
</tr>
<tr>
<td>GTW-Fixed network propagation &amp; processing delay (fixed)</td>
<td>0.15 s</td>
</tr>
<tr>
<td>mean registration rate (system input)</td>
<td>0.1 arrival/s</td>
</tr>
<tr>
<td>mean location update rate (system input)</td>
<td>0.05 arrival/s</td>
</tr>
<tr>
<td>mean M-M call arrival rate (system input)</td>
<td>0.406 call/s</td>
</tr>
<tr>
<td>mean M-F call set-up arrival rate λ (system input)</td>
<td>10.56 call/s</td>
</tr>
<tr>
<td>call set-up message length</td>
<td>160 bytes</td>
</tr>
<tr>
<td>message length (proceeding, altering, connect, connect ACK message, satellite)</td>
<td>variable (2-8)</td>
</tr>
<tr>
<td>mini-slot packet length</td>
<td>8 bytes</td>
</tr>
<tr>
<td>buffer size</td>
<td>160 bytes</td>
</tr>
<tr>
<td>RASC channel retransmission timeout</td>
<td>0.34 s</td>
</tr>
<tr>
<td>RRC channel retransmission timeout</td>
<td>0.32 s</td>
</tr>
<tr>
<td>RASC backlog randomisation range</td>
<td>3 timeslots</td>
</tr>
<tr>
<td>RRC backlog randomisation range</td>
<td>10 minislots</td>
</tr>
<tr>
<td>signalling resource requirement of each connection</td>
<td>1(slot/frame)</td>
</tr>
<tr>
<td>Logic link layer ARQ scheme</td>
<td>SR ARQ</td>
</tr>
<tr>
<td>window size</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 5-1 Simulation parameters.

5.3.5 Simulation Results of the Performance with the Presence of Channel Fading

In the previous section, we have demonstrated that the semi-permanent connection offers more advantages than the permanent connection in an error-free channel. In this section, we validate the application of the semi-permanent signalling protocol in fading channel environments. To counteract the fading affect and improve the performance, the SPSP protocol is allowed to send multiple copies of the resource request on the Resource Request Channel (RRC). We now address the fading impact on the channel utilisation, signalling delay and signalling failure probability and make a comparison with the PSP scheme.

5.3.5.1 Simulation Results and Discussion

The same system parameters, e.g. on-board buffer size, system signalling traffic condition, signalling resource pool, as used in the evaluation of SPSP in an error-free channel are also
Optimisation of MAC Protocol

considered in the following evaluation. In order to analyse the performance of the semi-permanent signalling protocol in the presence of fading, the Markov channel model presented in Chapter 2 is used. The model describes the packet success/failure process on a frequency non-selective multipath mobile fading channel, in which different correlation coefficient of two successive fading channel samples are decided by the product of normalised Doppler bandwidth \( f_D T \), where \( f_D \) is the Doppler frequency and \( T \) is the packet duration. By choosing different values of \( f_D T \), we can establish fading channel models with different degrees of correlation in the fading process. In satellite systems, the channel correlation is the major factor that causes the burst error in data transmission. Therefore, in this study, a very correlated slow fading channel \((f_D T=0.01035)\) is considered.

Performance comparisons between the semi-permanent and the permanent call set-up under the same input traffic load and buffer size are shown in Figure 5-16 and Figure 5-19.

![Figure 5-16 Call accommodating capacity vs. fading margin (P - permanent protocol, S - semi-permanent protocol, Q - queue size, S - signalling channels, \( f_D T=0.01035 \)).](image_url)

The performance in terms of the accommodating capacity is shown in Figure 5-16. The semi-permanent protocol is more sensitive to the channel fading than the permanent protocol and its performance degrades quickly with a decrease of the channel condition. However, regardless of the good or bad channel condition, the semi-permanent protocol demonstrates a higher accommodating capacity than the permanent and this advantage becomes more marked with the improvement of the channel condition. When the channel is severely faded, the system fading margin cannot compensate the fading effect and an undesirable high call
blocking occurs because of the accommodating capacity decreases to 0.09 calls/second. With a heavy system load and a slow system processing due to the channel fading, accommodating capacity performance is not very sensitive to the change of the on-board buffer size.

As expected, the channel fading degrades the overall performance of the system channel utilisation. However, compared to the performance of the system using the permanent protocol to offer an utilisation less than 3.56%, the system using the semi-permanent protocol offers a much higher channel utilisation as demonstrated in Figure 5-17. Most importantly, with an increased fading margin, the channel utilisation can be further increased. This situation cannot be observed from systems that employ the permanent signalling protocol. The SPSP has the potential to offer even higher resource utilisation efficiency.
Despite good performance in channel capacity and channel utilisation, the semi-permanent call set-up delay does not show a better performance than the permanent call set-up delay with a high fading margin as shown in Figure 5-18. When an average packet error rate of 0.011 (fading margin 20dB) is encountered in the channel, the time difference between the semi-permanent and the permanent protocol can reach 7 seconds. Under this situation, even an increase in the number of the signalling channels cannot effectively minimise this delay difference. This is due to a bottleneck on the resource reservation channel, which is caused by an increased number of retransmissions of the resource request packets lost in the channel fading. However, the SPSP demonstrates an improved performance in the severe channel condition and it has the potential to offer a comparative delay performance if the bottleneck is resolved. The solution is to increase the number of resource reservation channels to resolve the channel congestion caused by an overload and to limit the number of retransmissions of each packet. It is worth mentioning that, because the channel condition is time varying, the channel condition frequently varies between good and bad states. If we increase the fading margin or the transmit power whenever a severe fading occurs, the delay performance of the semi-permanent protocol can demonstrate a competitive performance.

![Figure 5-19 Signalling failure probability vs. fading margin.](image)

The performance of signalling failure probability is shown in Figure 5-19. In contrast to the zero signalling failure in the error-free channel, the permanent protocol suffers a high signalling failure due to the call set-up delay that exceeds the call set-up timer. This failure is even greater than suffered by the semi-permanent protocol in the presence of severe channel
fading. The signalling failure of the permanent protocol is also sensitive to the buffer size. With an increase of buffer size from 2 to 5 packets, the signalling failure can decrease by around 10%. The semi-permanent signalling failure varies slowly with channel fading and buffer size and has the advantage of providing a lower signalling failure compared to the permanent protocol for some conditions. Besides, schemes such as reducing the resource request message size or increasing the number of reservation minislots can also reduce the signalling delay and failure probability.

5.3.5.2 Simulation Discussion

The simulation results show that a satellite system using the SPSP can achieve higher system capacity and channel utilization than one using the PSP in a fading channel. However, when a severe fading condition is encountered, the high bit error rate can lead to an undesirably long signalling delay and high failure ratio for both SPSP and PSP.

The higher channel utilisation and accommodating capacity offered by SPSP are significant in ensuring a multimedia satellite system to support an increasing user population. The protocol can be directly implemented to connection-oriented non-real time services. Moreover, it is possible to improve signalling delay and signalling failure performance of the SPSP by increasing the number of resource reservation channels, limiting the number of retransmissions, increasing fading margin or transmitting power. The SPSP can thus even provide an enhanced performance to support real time applications. So, the semi-permanent signalling protocol does offer significant advantages over the permanent for satellite systems.

5.3.6 Conclusion

In this Chapter, a service-adaptive MAC protocol is proposed. The protocol is based on a reservation-based demand assignment scheme. We have also validated an application of the proposed reservation/demand MAC into signalling traffic by assessing the semi-permanent signalling protocol in both error-free and channel fading conditions. The evaluation results of the performance of the semi-permanent signalling protocol (SPSP) have been presented, analysed and compared to those of the conventional permanent signalling protocol (PSP).

Our results show that in both good and bad channel conditions implementation of the SPSP can significantly improve signalling service accommodating capacity and channel utilisation. In good channel conditions, the signalling delay performance of the SPSP is comparative to that of the PSP and its signalling failure probability remains negligible. In the presence of channel fading, the signalling failure probability and signalling delay of the SPSP is still
in inferior to those of the PSP. However, the performance difference can be minimised by increasing the number of resource reservation channels, limiting the number of retransmissions and increasing fading margin. In addition, under fading conditions, if a multicopy transmission mechanism is used and system parameters such as on-board buffer size are appropriately set, the SPSP could exhibit a better performance than the conventional PSP. The SPSP is naturally suitable for non-real time services and signalling that does not have time constraints. The SPSP can be directly applied to the signalling that does not have a stringent delay constraint such as signalling establishment of non-real time services and real time services that can tolerate certain signalling delay.

From this work, we conclude that bandwidth can be saved in signalling channel resources by using a dynamic signalling channel allocation method. In this respect, the entire bandwidth required for signalling traffic can be reduced and increased numbers of signalling requests can be accommodated. We, therefore, have demonstrated the benefits of using the SPSP in a multimedia satellite system.
Chapter 6 Optimisation of ARQ Protocol

In this chapter we address two identified problems when applying the Selective Repeat ARQ (SR ARQ) protocol to the ATM-satellite. The first is the unlimited retransmission of outstanding packets in the SR ARQ that blocks the transmission of following packets. The second is that the time-varying distance of some end-to-end connections degrades the performance of SR ARQ protocol. We propose a reliability-dependent SR ARQ protocol and an adaptive timer SR ARQ protocol to solve the two problems identified. Using the optimised SR ARQ protocol, we propose a service-adaptive ARQ protocol architecture for ATM-satellite system.

6.1 Introduction

In ATM, although end-to-end error control and header error control are performed, they are unable to achieve the required QoS performance in an ATM-satellite system and may cause serious cell discarding at the ATM layer. Therefore, Logic Link Control (LLC) scheme is expected to perform error control on the radio link with minimal Cell Loss Rate (CLR) and Cell Transfer Delay (CTD). However, due to the high burst error rate on mobile channels, the poor throughput performance of conventional LLC becomes a problem for transportation of ATM traffic over satellite. It has been identified that for enhancing the performance of LLC the key component to be improved is the ARQ control procedure.

Two problems have been identified when applying the SR ARQ protocol to ATM-satellite. The first one is the inefficiency caused by the unlimited retransmission of outstanding packets. Because the ARQ provides reliable traffic delivery, packets corrupted in the channel are retransmitted unconstrained until they correctly arrive at the receiving end. However, unlimited retransmission of outstanding packets blocks the transmission of following packets. This unlimited retransmission is the factor that degrades the transmission quality. The requirement in ATM-satellite systems to support a variety of traffic types with different QoS requirements makes it necessary to study enhancements of conventional ARQ techniques for the delay constrained traffic transmission over satellite channels. The second one is the sensitivity to a distance-changing end-to-end connection. It is foreseen that future satellite systems will route the traffic through ISLs. An end-to-end connection that consists of many ISL segments of LEO/MEO constellations generally has a time-varying distance characteristics. This time-varying distance also degrades the efficiency of SR ARQ protocol.

Thus we propose a reliability-dependent SR ARQ (RDSR ARQ) protocol and an adaptive timer SR ARQ (ATSR ARQ) protocol to solve the above two identified problems. The effort is devoted to optimising the proposed ARQ protocol to the link efficiency, CLR and CTD.
Finally, a combining RDSR ARQ and ATSR ARQ protocol is proposed as a service-adaptive ARQ protocol candidate for ATM-satellite integrated systems.

### 6.2 SR ARQ Protocols

#### 6.2.1 Protocol Description

The distinguishing feature of the SR ARQ protocol from Go-Back-N (GBN) is that only the erroneous and lost packets are retransmitted before a timeout. The transmitter delivers one copy of every packet in correct order to the receiver. It keeps a copy of the packet delivered in a buffer before it is transmitted and removes it after the receiver correctly receives it. The delivered packets may be corrupted or lost in the physical channel. It is also possible that the receiver gets the packets in wrong order. For instance, when transmission errors corrupt the first packet but not the second, the second packet correctly arrives at the receiver before the first.

The receiver uses a buffer to store the packets that arrive out of order. It acknowledges the reception of the correctly received packet by sending an ACK feedback message to the sender. Once the receiver has a consecutive group of packets in its buffer, it can deliver them to the network layer. In the absence of the acknowledgement of the reception, the transmitter retransmits the packet immediately when the end-to-end timer expires. Packets received out of sequence at the receiver have to be resequenced to make sure that packets passed to the upper layer are in order. Obviously, SR ARQ requires a full-duplex link: the receiver transmits acknowledgements to the sender while the sender transmits packets to the receiver.

SR ARQ protocol requires a sufficient buffer capacity at both the transmitter and the receiver because one erroneous packet could cause all successive arriving packets to remain in the buffer. An example of the receiver buffer overflow event is shown in the Figure 6-1, assuming that transmission errors corrupt only packet 1 and its copies.
Optimisation of ARQ Protocol

Sender buffer

![Buffer Diagram]

Figure 6-1 An example of SR ARQ receiver buffer overflow.

The buffering time could be long enough to cause a buffer with a small capacity to overflow. The transmission window control is a commonly used control mechanism to avoid buffer overflow. The window size (WZ) normally is half of the buffer size in terms of packets. It controls the sequence difference of unacknowledged packets by less than the WZ so that the receiver buffer will never hold packets in excess of more than the WZ. The WZ and the finite number of packets that can be held by the buffer are design parameters of SR ARQ.

6.2.2 Packet Numbering

The SR ARQ numbers packets using a concept of Numbering Modulo [WAD91]. In this scheme, the \( n \)th packet is numbered as,

\[
\text{Modulo (n)} = \frac{(n - 1)}{2^w}
\]

(6-1)

where the \( w \) is the window size. The ACK of the \( n \)th packet has the same modulo number as the \( n \)th packet. Using this modulo numbering modulo, the receiver will not be confused during the transmission. Otherwise, confusion may occur. For instance, when using numbering modulo 7 instead of 8 for window size 4, the sender numbers the packets using the successive numbers \( \{0, 1, \ldots, 6\} \). In Figure 6-2 (a), for no transmission errors, the second packet numbered 0 that arrives at the receiver is the eighth packet. In Figure 6-2 (b), transmission errors corrupt the first ACK0. The second packet 0 received by the receiver is a retransmitted copy of the first packet. The receiver cannot distinguish between these two
cases. Consequently, the receiver does not know whether to store or to discard the second packet numbered 0. We can conclude from this example that the numbering modulo 7 is not suitable. The receiver is confused because it may receive different packets with the same number.

Packet sequence  
1 2 3 4 5 6 7 8  
Packet modulo  
0 1 2 3 4 5 6 0  

Received packet sequence 1 2 3 4 8

(a)

Packet sequence  
1 2 3 4 Timeout 1 (copy)  
Packet modulo  
0 1 2 3 0

Received packet sequence 1 2 3 4 1 (copy)

(b)

Figure 6-2 SR ARQ with numbering modulo 7.

6.2.3 Protocol Efficiency

ARQ protocol efficiency is characterised by channel throughput and average packet transfer delay. Throughput efficiency $\eta$ is defined as the average fraction of time that the protocol transmits new packets [WAD91]. During the rest of the time, the protocol waits for ACKs or for timeouts, or it retransmits copies of packets. The efficiency of the ARQ protocol is mainly dominated by the packet error rate in the channel, window size, receiver buffer size and the retransmission timeout. An exact analysis of SR ARQ with finite buffer is complicated. Usually the analysis is carried out based on many assumptions such as a negligible round-trip delay, an error-free feedback channel or an infinite buffer size. Different assumptions in the analysis result in a difficulty to compare their results. However, with the assumption of an infinite window size, an upper bound on the efficiency of SR ARQ with finite-size buffers is given as,
Optimisation of ARQ Protocol

\[ \eta = 1 - \varepsilon \]  

(6-2)

where the \( \varepsilon \) is the average packet error rate in the transmission channel. This upper bound is usually taken as a reference in the evaluation of SR based ARQ protocol. In the satellite systems that our research is based on, the SR ARQ protocol offers unsatisfactory efficiency. The protocol itself is not robust enough and is prone to many factors.

6.3 Reliability-Dependent SR ARQ Protocol

The impact of satellite channel impairments, user mobility and constellation dynamics on SR ARQ protocol result in the network having to incorporate a complex error protection and recovery procedure. Extra bandwidth and processing resource will be consumed and user QoS will be degraded because of long packet transfer delay and delay jitters. This system performance degradation prevents ATM from keeping up with high-speed traffic transmission requirements. The Reliability-Dependent SR ARQ (RDSR ARQ) protocol is proposed to overcome this problem by discarding some cells during the transmissions.

6.3.1 Scheme Description

The RDSR ARQ protocol is shown in the Figure 6-3 and Figure 6-4. The RDSR ARQ can perform two modes of traffic transmission, conventional reliable transmission (CRT) and relaxed reliable transmission (RRT). CRT can be applied to real-time services that require a low CLR and non-real-time services with a CLR commitment obtained from the network. Provided that the required CLR can be maintained, RRT can be implemented for some real-time services, which tolerate certain cell loss, to meet the required CTD by discarding some cells. These cells are delayed in the network more than the specified CTD and are considered to be of less value, or no value at all to the application [ATF96]. If a CRT cannot meet required CLR or worsens system congestion, discarding these cells before the next retransmission will be necessary. RRT is most suitable for unspecified bit rate (UBR) services because the network offers no traffic related service commitment as to the cell loss ratio experienced by a connection. With this mode, the network can be assisted to relieve its serious congestion condition.
Optimisation of ARQ Protocol

Receiver Architecture.

Figure 6-3 RDSR ARQ protocol diagram – receiver architecture.
6.3.1.1 Relaxed Reliable Transmission

The Relaxed Reliable Transmission (RRT) can perform relaxed reliable transmission rather than the Conventional Reliable Transmission (CRT). The differences in protocol between RRT and CRT ARQ lie in the operation of the packet receiving process. No difference is made to the packet transmission process. The differences have the following features. Firstly, RDSR operates CRT or RRT according to the transmission mode that is set during the
Optimisation of ARO Protocol

connection set-up phase. Second is that RDSR ARQ utilizes two types of acknowledgments (ACK). The two types of ACKs are the ACK with Receiving (ACKR) and ACK with Discarding (ACKD). The transmitter will proceed with the transmission upon receipt of any of the two types of ACKs. Thirdly, a Discarding Timer (DT) is used to restrict the waiting time (WT) for a current outstanding packet at the receiver. The operation is as follows;

1. Upon receiving a packet correctly, the receiver sends an ACKR to the sender and sets the next expected packet A as the outstanding packet and meanwhile sets the DT.

2. If a packet is received before the DT is expired, the receiver starts from operation step 1.

3. If a packet has not been correctly received when the DT has expired, the receiver sends an ACKD or ACKR to the transmitter. Meanwhile, the receiver empties the receiving buffer and delivers the correctly received, but inconsecutive packets, to the upper layer.

4. Upon receiving the ACKD, the transmitter discards the packet as if it has been correctly received by the receiver. The transmitter continues to transmit the following packets waiting in the transmission buffer. To avoid a discarding of many consecutive packets that can result in high CLR, the receiver and transmitter are not allowed to discard any two consecutive packets, or n consecutive packets, provided that the CLR is in the line with QoS requirements.

5. If there is no ACK of packet A received when the retransmission timer expires, the packet A will be retransmitted and marked with the Nth number of the retransmitted packet.

The timeout value of DT varies from packet to packet. It is derived based on parameters such as Maximum Cell Transfer Delay (MaxCTD), retransmission timeout (RT), window size W and single propagation delay Ds. Let the $t_{interval}$ denote the time interval between times when packet (n-1) and packet (n) are transmitted at first time, n denote the number of retransmissions of packet (n-1) before it is correctly received, $t_{n-1,r}$ denote the storage time of packet (n-1) at the receiver after it is correctly received, $t_{now}$ denote the instant in time when the packet (n-1) is to be delivered to the network layer and the discarding timer for packet (n) to be set up. Then the transmission events of two consecutive packets can be understood easily from Figure 6-5. Clearly, the DT timeout value, $DT_n$ for packet n will be

108
\[ DT_n = \text{MaxCTD} + t_{\text{interval}} - nRT - D_s - t_{n-1,s} \]  

(6-3)

In the above equation, all the parameters except \( t_{\text{interval}} \) are known by the receiver. The \( t_{\text{interval}} \) can be estimated. Assuming that

\[ WT \geq 2D_s \]  

(6-4)

where \( T \) is the transmission interval between two transmission timeslots. Then the \( t_{\text{interval}} \) can be estimated as

\[ t_{\text{interval}} = e_{\text{avg}} RT \]  

(6-5)

where \( e_{\text{avg}} \) is the average packet error rate in the fading channel.

---

**Figure 6-5** Two consecutive packets transmission events on the time axis.

### 6.4 Adaptive Timer SR ARQ

Although SR ARQ is a widely studied technique, it faces new challenges in the new advanced LEO/MEO satellite constellations such as Iridium and Teledesic that are equipped with intersatellite links (ISL). Due to the inherent dynamics of ISLs and short satellite visibility period of LEO/MEO constellations, an end-to-end connection routed in the sky suffers from path length variations and frequent connection handovers. The major reason is the continuous
distance variations on the connection caused by interplane ISLs connecting satellites in adjacent co-rotating orbit planes and discontinuous distance variations on the connection caused by the deactivation of interplane ISLs in the polar regions [WER97C]. When one or more serving satellites on the route passes over the polar or moves out of visibility range, connection handover has to be performed to keep the continuity of the communication link [UZU97B]. The continuity and discontinuity of distance change have been found to have considerable impact on the conventional SR performance.

In this section, we address the impacts of distance-varying end-to-end (DVE) connections on the performance of the SR ARQ and propose an enhanced SR ARQ, named Adaptive Timer SR (ATSR) protocol, to work efficiently and robustly on such connections. We focus on the study of the average packet transfer delay and link efficiency of the proposed protocol with respect to data generation rate, buffer size and retransmission timer. We will compare the results of the SR protocol in both changing-distance and fixed-distance connections.

6.4.1 Problem Identification

For a SR protocol to work effectively, the retransmission timer should be set slightly above the end-to-end round-trip delay. In a fixed-link distance (FLD) connection case, a fixed retransmission timer can be set up during the call set-up time. However, with continuous and discontinuous distance-changing connection, the round-trip delay varies with time throughout the call holding duration in an unpredictable manner. Therefore, SR cannot perform efficiently due to the following:

1. End-to-end propagation delay varies with time and thus an appropriate retransmission timer value is difficult to determine.

2. An inappropriate retransmission timer causes low channel throughput. When the retransmission timer is smaller than the end-to-end round-trip delay, redundant retransmission will occur. When the timer is larger, the channel may carry no traffic during some time intervals resulting in bandwidth wastage.

3. Continuous or discontinuous changes in the connection length can cause the packets to arrive at the destination out of order.

4. Buffer capacity requirements for a given packet loss ratio are difficult to dimension on DVE connections.

These problems can degrade the mean packet transfer delay and link efficiency of the SR protocol. This degradation will be addressed in the following section of this chapter.
6.4.2 ATSR ARQ Protocol Description

A key cause of the problems mentioned above is the changing end-to-end round-trip delay differing from the retransmission timer value during the communication. If the timer can be maintained constantly at a value corresponding to the instantaneous round-trip delay, the above problems would be overcome. Following this approach, we propose an enhanced version of the SR protocol, Adaptive Timer SR (ATSR) protocol, to improve the protocol efficiency.

Similar to the operation of the SR protocol, only the erroneous and lost packets are retransmitted. The transmitter keeps the newly generated packets in a buffer until they are correctly received by the receiver. Before a packet is transmitted, a timer is set up in order to retransmit the packet in the absence of reception acknowledgement from the receiver. The receiver acknowledges the reception of the packet by an ACK feedback message if the packet is correctly received or by a Non-Acknowledgement (NACK) feedback message if the packet is received in error. Received packets are re-sequenced by the receiver and delivered to the upper layers in order.

Unlike the conventional SR ARQ protocol using a fixed retransmission timer, ATSR uses a varying retransmission timer. The value of the timer is adjusted dynamically according to the information derived from periodically tracking the end-to-end round-trip propagation delay. The principle of ATSR is illustrated in Figure 6-6.
The tracking of round-trip delay is performed by detecting the time difference between the retransmission time of a packet and the timer’s remaining time when the acknowledgement of this packet is correctly received at the transmitter. Before a packet is transmitted, a buffer storing the predicted current round-trip delay is read and the retransmission timer is set accordingly. The retransmission timer value has to be stored and indexed by sequence number of the packet for use in deriving the next retransmission timer.

Upon correct reception of a packet, the transmitter can derive the feedback time, $f_d$, defined as the difference between the packet’s retransmission (RT) timer and RT timer’s remaining time. To let

$$P = \frac{f_d}{\text{RT timer}}$$  \hspace{1cm} (6-6)

and denote $G_t$ as the newly derived round-trip delay, $T_R$ as the new RT timer, we have

$$G_t = f_d \quad \text{if} \quad 0.5 \leq P \leq 1$$  \hspace{1cm} (6-7)

or

$$G_t = f_d + \text{RT timer} \quad \text{if} \quad P \ll 0.5$$  \hspace{1cm} (6-8)

Then

$$T_R = G_t + \Delta t$$  \hspace{1cm} (6-9)

where $\Delta t$ is an empirical value to ensure that the RT timer value remains slightly larger than the round-trip delay so as to compensate extra delays on the transmission path through the network caused by queuing, packet processing etc. (In some systems, $\Delta t$ is set to be the mean variance of the end-to-end distance). The newly derived $T_R$ will replace the old $T_R$ and is stored in the memory for setting the RT timer of next transmitted packet.

### 6.4.3 Evaluation Method

To carry out the performance investigation, we firstly need to study path length variation in an end-to-end connection routed through intersatellite links. This variation depends on many factors such as satellite constellation, end users' geographical position, system routing
algorithm, satellite visibility period and call holding time. These factors make the formulation of a general model of DVE connection very difficult. We therefore base our investigation on a typical DVE connection using the IRIDIUM constellation. It will be shown that the results can be generalised to any DVE connection.

It is well known that the IRIDIUM constellation supports two intraplane connections to satellites forward and backward and two interplane connections to satellites in each neighbouring plane. An ISL topology for IRIDIUM is assumed as that shown in Figure 6-8 [WER97A]. We choose a DVE connection with 4 intermediate satellites positioned (t=0) as given in Table 6-1. The illustration of such an end-to-end connection is shown in Figure 6-7.

![An end-to-end connection with three ISL segments.](image)

The distance of the selected connection versus time calculated by software tool [SUR96] is given in Figure 6-9. The corresponding distribution of end-to-end propagation delay can therefore be derived. In our investigation, we also limit the simulation to 10 minutes, which are the typical LEO satellite visibility period, in order to avoid connection handovers.

<table>
<thead>
<tr>
<th>Satellite No.</th>
<th>Latitude°</th>
<th>Longitude°</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.06</td>
<td>-69.053</td>
</tr>
<tr>
<td>2</td>
<td>-16.263</td>
<td>-38.626</td>
</tr>
<tr>
<td>3</td>
<td>-32.577</td>
<td>-8.418</td>
</tr>
<tr>
<td>4</td>
<td>-48.871</td>
<td>21.150</td>
</tr>
</tbody>
</table>

Table 6-1 Four intermediate satellites' position at t = 0.
Figure 6-8 Example Iridium ISL topologies (dashed lines denote the interplane ISLs, solid lines denote the intraplane ISLs).

Figure 6-9 Distance of end-to-end distance-changing connection versus time.

An error free channel is assumed in this investigation in order to isolate the packet retransmission coexisting effects.

6.4.4 Simulation Results and Discussion

We now address the impact of distance-varying end-to-end (DVE) connections on performance of SR ARQ and discuss performance of the proposed ATSR protocol by way of making a comparison between ATSR and SR ARQ for both the changing-link distance (CLD) case and the fixed-link distance (FLD) case. With FLD, three fixed distances, equal
respectively to the maximum, minimum and mean of evaluated changing-distance connection shown in Figure 6-9, have been used. Their retransmission timer is set to be slightly larger than the end-to-end propagation delay.

<table>
<thead>
<tr>
<th>Packet generating rate (packets/s)</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission rate (bits/s)</td>
<td>24k</td>
</tr>
<tr>
<td>Packet length</td>
<td>80 bytes</td>
</tr>
<tr>
<td>ACK packet length</td>
<td>16 bytes</td>
</tr>
<tr>
<td>Buffer size (packets)</td>
<td>400</td>
</tr>
<tr>
<td>Window size</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 6-2 Simulation parameters

A. Impacts of DVE connections on SR protocol performances

We first evaluate the SR ARQ performance varying with different retransmission timer on a DVE connection, and then examine sensitivity of the protocol performances to data generation rate and buffer size.

In a DVE connection, the variation of round-trip delay is difficult to predict throughout the call holding time. The performance degradation of packet transfer delay and link efficiency due to inappropriate timer values on the DVE connection are show in the Figure 6-10 and Figure 6-11.

When the RT timer is smaller than the round-trip delay, packet retransmission occurs leading to bandwidth wastage and increases of the packet transfer delay. As seen from Figure 6-10, degradation of the delay performance varies almost linearly with the negative offset of the retransmission timer from maximum propagation delay (the transmission delay on the selected connection that has maximum distance) at a ratio of 0.4564 during the communication. Compared to the performance on the fixed distance connection, we can see that even the delay on the longest distance, which is 0.04625s, is much better than the performance on the DVE connection when the RT timer is smaller than the round-trip delay.

Using a large RT timer sufficient to cover the maximum round-trip delay is one option to cope with CLD. But it wastes channel bandwidth because the channel carries no information during some time intervals. When the RT timer is larger than the maximum propagation delay, the packet transfer delay is improved. However, as seen from Figure 6-11, the degradation of link efficiency remains about 11.55% lower than that of fixed distance. This
clearly indicates that increasing the retransmission timer can only improve the packet transfer delay but not improve the serious bandwidth wastage.

Due to inappropriate retransmission timer values, the network moves towards instability. This is demonstrated by an increased sensitivity of protocol performance to the key system parameters such as source packet generation rate, buffer size and window size.

Sensitivity of the SR protocol performance to packet arrival rate can be observed from Figure 6-12 and Figure 6-13. When the RT timer value is below the maximum round-trip delay, a slight variation of source packet arrival rate can cause significant performance degradation. However, this sensitivity can be reduced by increasing the timer value to a level beyond the maximum propagation delay of CLD. But again it can not improve the link efficiency.

Figure 6-10 Mean packet transfer delay versus offset of RT timer value from maximum round-trip delay on a DVE connection.
Figure 6-11 Link efficiency versus offset of RT timer value from the maximum round-trip on a DVE connection.

It is worth noting that the protocol performances are also sensitive to the buffer size. In order to satisfy the user required QoS, terminal buffer size has to be carefully selected to avoid high packet loss. On a DVE connection, the performance of packet loss ratio is more sensitive to the size of terminal buffer. With a comparison of packet loss ratios given in Table 6-3, a high packet loss ratio up to 33.6% can be observed on a DVE connection. Such a high ratio is not acceptable. A larger buffer size is therefore needed to keep the packet loss ratio at a required level.

B. Performance of ATSR protocol

The performances of the proposed ATSR protocol are shown in Figure 6-14 and Figure 6-15. The results are compared to those of the conventional SR protocol.

It is seen from these results that, by using the ATSR protocol, the mean packet transfer delay is reduced from 6 to 0.4178 seconds and link efficiency is also improved by around 8.8%. The sensitivity of ATSR protocol performances to the retransmission (RT) timer, buffer size and packet arrival rate are also greatly reduced. ATSR can maintain this performance level, which is the best performance achievable by the conventional SR on a CLD connection, throughout the whole communication even with an improperly selected RT timer value.
The way that ATSR maintains its performance is by automatically adjusting its timer value to be slightly above the current round-trip transmission delay. In this way, the impacts of changing-distance connection are almost eliminated. So the determination of the initial RT timer value during the call set-up time only needs to take the current round-trip delay as a reference.

As demonstrated in Figure 6-14 and Figure 6-15, ATSR does not outperform the SR protocol when the timer value is larger than the maximum round-trip delay. This is due to the error free channel assumption.
Besides the improvements made on the system performance, The ATSR protocol can also contribute to flow control on the connection. A congestion that occurs in the network can be detected by a terminal by observing the delayed packet transfer. In this case, ATSR can behave in the same way as it does for an increasing-distance connection. This will lead to a corresponding increase of the ATSR retransmission timer. Therefore, the total traffic injected into the congested switch will be reduced accordingly.

Table 6-3 Comparison of packet loss ratio.

<table>
<thead>
<tr>
<th>Buffer Size (packets)</th>
<th>CLR on CLD Connection RT = 0.086s</th>
<th>CLR on CLD Connection RT = 0.0752s</th>
<th>CLR on Fixed Distance Connection</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.1388</td>
<td>0.3362</td>
<td>0</td>
</tr>
<tr>
<td>127</td>
<td>0.0979</td>
<td>0.2953</td>
<td>0</td>
</tr>
<tr>
<td>155</td>
<td>0.0555</td>
<td>0.2529</td>
<td>0</td>
</tr>
<tr>
<td>182</td>
<td>0.0147</td>
<td>0.2120</td>
<td>0</td>
</tr>
<tr>
<td>209</td>
<td>0.0001</td>
<td>0.1711</td>
<td>0</td>
</tr>
<tr>
<td>236</td>
<td>0</td>
<td>0.1303</td>
<td>0</td>
</tr>
<tr>
<td>264</td>
<td>0</td>
<td>0.0879</td>
<td>0</td>
</tr>
<tr>
<td>291</td>
<td>0</td>
<td>0.0467</td>
<td>0</td>
</tr>
<tr>
<td>318</td>
<td>0</td>
<td>0.0061</td>
<td>0</td>
</tr>
<tr>
<td>345</td>
<td>0</td>
<td>0.0010</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 6-14 Mean packet transfer delay of ATSR protocol in comparison with the conventional SR protocol.
C. Extension of ATSR to other connection patterns

Our evaluations have been conducted in a 4-hop connection with distance changing characteristics as shown in Figure 6-9. However, the nature of ATSR is its ability to track the changing end-to-end transmission delay and adjust its RT timer accordingly. Its effectiveness does not depend on any particular connection pattern and variation characteristics of connection distance. Hence, the results and conclusions that we have obtained in this work do not lose their generality.

Simulation results show that, under a good channel condition, the ATSR protocol can offer higher link efficiency and shorter average packet transfer delay than the conventional SR protocol on a distance-changing connection. The protocol is robust and efficient without significant modification being needed to the SR protocol. It could be used to replace the conventional SR protocol on a connection formed by ISLs in LEO/MEO satellite systems to achieve a higher bandwidth utilisation.

6.5 The Architecture of ARQ Protocol for ATM-satellite Systems

In the above two sections, we have demonstrated that the proposed RD SR and ATSR ARQ protocols can overcome the problems caused by an unlimited retransmission and changing
distance of an end-to-end connection. They are more efficient than the conventional SR ARQ protocol. These two ARQ protocols can be merged together to build up a new ARQ protocol. The merge is simply to equip the RDSR ARQ protocol with some additional memory and the adaptive timer algorithm of ATSR, which monitors the end-to-end packet transfer delay and changes the time value accordingly. The merged protocol operation is shown in Figure 6-16 to Figure 6-18. We thus propose this protocol for the use in ATM-satellite systems.

The proposed ARQ protocol can therefore have three different operation modes as shown in Figure 6-19. The adaptive timer can only be operated when the reliable transmission mode is selected and the end-to-end connection is not affected by server channel fading. The proposed protocol provides different operation modes for different service types, e.g. CBR, Rt-VBR, Nrt-VBR, UBR and ABR. The selection of the operation mode will be presented in Chapter 7. Negotiation in determining the suitable operation mode for a specific connection should be completed at the connection set-up phase. However, the operation mode can be renegotiated and changed upon request from the user or the network during the communication.
Optimisation of ARQ Protocol

Receiver architecture.

Figure 6-16 Receiver architecture of the ARQ protocol proposed for ATM-satellite system.
Optimisation of ARQ Protocol

operation at sender

Packet from higher layer

Buffer overflow?

Yes

Overflow report to RCL layer

No

Read ARQ operation mode

No

Buffering

Yes

UT mode

Generate unnumbered packet

Insert modulo

Record last sequence number released

Release packet

Release packet

Outstanding packet ACKed

Queuing

Window size reached?

No

Packet storage

Set acknowledgement state

Yes

Link control process

Retransmission timeout

Packet retransmission

Set retransmission timer

Process A

Record outstanding 1-frame sequence

Ready for transmission

To MAC layer

Transmission control from MAC

S-frame

I-frame Piggyback ACK

ACK sequence

Piggyback successful after T2

Yes

Transmission control from MAC

I-frame

Prioritise packet transmission sequence

Figure 6-17 Transmitter architecture of the ARQ protocol proposed for ATM-satellite system.
Figure 6-18 Additional processing architecture for the proposed ARQ protocol.
6.6 Conclusions

We have proposed a service-adaptive SR ARQ protocol for ATM-satellite systems based on reliability-dependent SR (RDSR) and adaptive-timer SR (ATSR) ARQ protocols. RDSR ARQ protocol can successfully discard insignificant cells and speed up the cell transfer between two end stations. It is more suitable than the conventional SR ARQ for cell transfer delay constrained real-time services in satellite systems. The adaptive timer algorithm can overcome the performance degradation of the SR ARQ caused by a changing connection distance under a reliable channel condition. The ATSR offers a higher efficiency than the conventional SR on distance-varying end-to-end connections without much added complexity to it. The protocol that combines the RDSR and ATSR has the advantages of both protocols and thus we have proposed such a combined protocol to be the ARQ protocol for mobile ATM-satellite systems. The proposed ARQ protocol has three different operation modes and provides different operation modes for different service types.
Chapter 7 A QoS-Provisioning Connection Management Scheme

Using the services provided by MAC and ARQ, a novel radio connection management scheme (CMS) for provisioning the transport of ATM traffic over satellite links is proposed. The proposed CMS effectively manages the air interface connections for ATM services through diversifying connection types, establishing QoS-based connections and deploying an efficient connection mapping and control scheme. The scheme tailors its bandwidth into different types of connections according to service types and QoS requirements of active users. The performance of CMS is evaluated by characterising the performance in terms of cell loss rate and cell transfer delay of voice traffic in the presence of channel fading.

7.1 Introduction

ATM-satellite services are connection oriented. This means that a connection must be established between the parties involved in the communication, before any traffic exchange can commence. At the end of the communication, the connection must be released and the resource must be freed. In practice, part of the resource held by the user can be released during the communication. The connection management (CM) scheme is devoted to the radio connection set-up, maintenance and release. It is a function of the radio connection control layer.

The CM is required in order to set up the radio connections for different ATM services with different QoS requirements. It should also be able to manage the resources used to implement connections and the information needed to allocate the resources. The established connection should provide a satisfactory QoS on transportation of the traffic of the individual ATM service over the satellite network.

There are four major factors that have considerable influence on the ATM traffic transmission in satellite channels, namely, the physical channel, the MAC protocol, and ARQ and resource allocation scheme. The above four elements have usually been studied separately in previous work. However, the requirement of effective ATM traffic transportation over satellites can only be truly achieved by studying them concurrently.

In this chapter, we present a novel radio connection management scheme (CMS) for a QoS-provisioning the transport of ATM traffic over satellite links. The proposed CMS effectively manages the air interface connections for ATM services through diversifying connection types, establishing QoS-based connections and deploying an efficient connection mapping and control scheme. The scheme aims to provide the specific performance objectives of CLR and CTD with minimal bandwidth provisioning for individual services. The scheme enables the network to minimise the unused bandwidth on a connection by dynamically tailoring its bandwidth into different types of
connections according to service types and QoS requirements of active users. The performance of CMS is evaluated by characterising the performance of CLR and CTD of voice traffic in the presence of channel fading. Related simulation results are presented and analysed.

7.2 System Model

To set up a connection across the satellite network to provide traffic transportation with high bandwidth utilisation and agreed QoS, three important components have been identified. These are connection definition and diversity, connection mapping algorithm and signalling control protocols. The model shown in

Figure 7-1, which includes the three components, is proposed as a reference model for the design and implementation of the CM scheme in ATM-satellite.

The concept of an ATM-satellite connection supported at the air interface firstly needs to be defined firstly to cope with various external traffic types. A mapping algorithm on the network side is then required to map the external ATM services onto a suitable air interface connection type, based on the traffic parameters and performance objective parameters submitted by the individual service. Connection control signalling procedures have to be designed to assist set-up, release and maintenance of such a connection. The study of connection control signalling will be presented in the next chapter. So the focus of this chapter will be on the connection definition and mapping.

Figure 7-1 System model.
7.3 The Proposed Connection Management Scheme

In the previous chapters, we have optimised the MAC and LLC layers for different service types. The proposed novel radio CMS utilises the services provided by these optimised MAC and ARQ protocols. The proposed CMS aims to provide a QoS-provisioning transport of ATM traffic over satellite links. As mentioned, the CMS consists of three parts: a novel connection definition, a QoS-based connection mapping algorithm and connection control signalling.

7.3.1 Connection Definition and Diversification

The novel concept of an ATM-satellite connection at the air interface is defined via three parameters: connection configuration, connection mode and connection reliability.

\[ \text{ATM-satellite Connection} = (\text{Configuration}, \text{Connection Mode}, \text{Reliability}) \]

This definition diversifies connection types to accommodate, with minimal provision of satellite channel resources, various external ATM traffic with different QoS requirements. Connection diversification can minimise the unused bandwidth on connections by shaping connections into different categories according to their service types and quality requirements. Each set of three parameters has been defined to include the following elements,

- **Connection Configuration** = \{point-to-point bi-directional, point-to-point uni-directional, point-to-multipoint uni-directional\}

The connection configurations are distinguished by the number of parties involved in the communications. Three types of connection configurations are provided to accommodate a wide range of services, i.e. Internet service, broadcast service and video-conferencing.

- **Connection Mode** = \{permanent, semi-permanent, mixed-mode\}

According to different resource-request methods and resource-holding times, a connection is categorised as permanent, semi-permanent or mixed mode. A permanent connection, provided with In-Band Resource Request (IBRR) to request extra bandwidth, holds the resources assigned during the connection set-up throughout the duration of communication. A semi-permanent connection relies on Out of Band Resource Requests (OBRR) or IBRR to reserve resources. It is allocated the required resources for only a fixed time interval depending on the source characteristics and on the resource status of network [FAN99]. In mixed mode connections, a user can utilise both IBRR and OBRR to adjust his temporary resource demand.
Provided that the required CLR is met, a connection can be set to a conventional reliable transmission (CRT) connection, a relaxed reliable transmission (RRT) connection or an unreliable transmission (UT) connection in order to meet the required CLR/CTD and to speed up the end-to-end packet transfer. CRT connections use traditional ARQ protocols and RRT uses modified ARQ protocols which can perform a relaxed reliable transportation of ATM traffic by discarding cells that have lost significance for the application. With RRT, a packet transfer can be accelerated to meet specific cell transfer delay and cell delay variation requirements. Unreliable transmission connections provide a non-guaranteed transmission mode. The lost packets are not recovered using ARQ.

7.3.2 Connection Mapping Scheme

Based on users' traffic characteristics and the connection definition at the air interface, a connection mapping algorithm (CMA) is proposed. The concept of connection mapping is shown in Figure 7-2. The CMA maps an external ATM service onto a suitable ATM-satellite connection that is associated with performance objective parameters, which are related to the original QoS requirement of the ATM traffic. The algorithm, taking into consideration source characteristics, QoS requirements, channel impairments and bandwidth allocation, aims to achieve the performance objectives of CLR and CTD with minimal bandwidth allocation. However, it is worth mentioning that the higher bandwidth efficiency that the algorithm can offer, the more complex the algorithm will be and the more processing resource will be required. This trade-off will always exist.

The task of the CMA is to determine the configuration, mode and reliability of the ATM-satellite connection. The performance parameters of CLR and maximum CTD at the radio interface are used to determine the suitable radio connection type for a particular ATM service. The connection configuration is decided directly by the number of involved calling parties. In the following, we focus on the mapping of the connection mode and the reliability of the connection.

![Figure 7-2 Connection mapping concept (RCL: radio control layer).](image-url)
7.3.2.1 Mapping the Connection Mode

Mapping of the connection mode contains two major tasks: to determine the resource request method and the amount of resource to be allocated. The mapping needs to take into consideration the QoS requirements and the channel impairments. The concept of effective bandwidth [GUE91], which is the bandwidth needed for an ATM service to ensure its required QoS on the ATM-satellite link, is adopted in this study. In the following, we discuss the mapping method for different ATM service categories. The symbols $B_{\text{req}}$, $B_{\text{eff}}$, $B_{\text{tp}}$ are used to denote respectively the required user bandwidth at the ATM layer, effective bandwidth and temporary bandwidth reserved through out-of-band resource request.

- **CBR services**

For Constant Bit rate (CBR) services that generate traffic with a constant bit rate, a permanent connection is the suitable mode. To counteract the fading influence of the satellite channel and ensure the real-time QoS, the effective bandwidth rather than the CBR peak cell rate needs to be allocated by the network. Assuming a large window size and an accurate prediction of the channel packet error rate (PER), the $B_{\text{eff}}$ of a single CBR source can be computed using a simple queuing model as shown in Figure 7-3.

![Figure 7-3 CBR service effective bandwidth $B_{\text{eff}}$ calculation model.](image)

In this model, a buffer with an average input cell arrival rate $\lambda_c$ and average output cell rate $\lambda_i$ is considered. If we assume the cell loss probability due to the buffer overflow is $P_b$ and the cell flow into the channel with average rate of $\lambda_i$, we then have [SCH96],

$$\lambda_i = \lambda_c - \lambda_c P_b$$  \hspace{1cm} (7-1)

Assuming that the PER in the mobile channel is $\varepsilon$ and the link capacity is $C$, the correctly received packet departure rate from the channel $\lambda_f$ can be written as

$$\lambda_f = (1 - \varepsilon) \cdot C$$  \hspace{1cm} (7-2)

When the considered packet transmission reaches an equilibrium point, the following equation holds,
\[ \lambda f + \lambda d + \lambda r \]

where \( \lambda_d \) is the packet loss rate due to the channel impairments and \( \lambda_r \) is the packet retransmission rate. In a reliable link transmission, a packet lost in the channel is retransmitted until it is correctly received. Therefore, \( \lambda_d \) is considered to be zero. From equation (7-1) to (7-3), we can derive \( C \) as,

\[ C = \frac{\lambda_c \cdot (1 - P_b)}{1 - \varepsilon} \]  

If \( CLR_{\text{req}} \) denotes the required cell loss ratio by the service, then \( P_b \) should satisfy the following condition in order to meet the required CLR,

\[ P_b \leq CLR_{\text{req}} \]

The effective bandwidth \( B_{\text{eff}} \) can be regarded as the assigned link capacity \( C \) that allows a cell loss ratio of \( CLR_{\text{req}} \), so the \( B_{\text{eff}} \) is given by,

\[ B_{\text{eff}} = \frac{\lambda_c \cdot (1 - CLR_{\text{req}})}{1 - \varepsilon} \]  

- **Rt-VBR services**

The mixed connection mode applies to Real-time Variable Bit rate (Rt-VBR). Because Rt-VBR services are real time and the cells have to be transmitted within the required maximum CTD time, effective bandwidth is required to be allocated on a permanent basis to guarantee the QoS. Meanwhile the out-of-band resource request method of the semi-permanent connection mode is also provided to solve the serious cell loss problem due to the peak rate of a traffic burst or a long error burst of a seriously faded channel.

Due to the unpredictable traffic characteristics of the VBR services, computing the effective bandwidth is difficult. An approximate approach, which is presented in [MOH97] to calculate the effective bandwidth for a single on-off VBR source, could be a reference. As derived in [MOH97], the following equation is used to calculate the effective bandwidth for the traffic transmission on a channel with low bit error rate.
Where

\[ F(\zeta) = \frac{F(\xi) - \sqrt{F^2(\xi) - 4(\lambda_1 \lambda_2 \zeta + \lambda_1 \mu_1 \zeta + \lambda_2 \mu_2 \zeta)}}{2\zeta} \]

(7-6)

\( \Pi, \lambda_1(\lambda_2), \mu_1(\mu_2), CLR_0 \) respectively denote the terminal buffer size, packet rate in off (on) state of the VBR model, transition rate from state off to on (on to off) of the VBR model, cell loss ratio due to buffer overflow. However, further study of the effective bandwidth for VBR services is needed in order to achieve a more accurate approximation. This work is out of the scope of this thesis and it is suggested for the future research.

- **Nrt-VBR, ABR, UBR services**

Nrt-VBR, Available Bit Rate (ABR), Unspecified Bit Rate (UBR) services are all non-real time services. Because they are CTD and CDV tolerant service types, allocation of effective bandwidth is not necessary for these services and the semi-permanent connection mode can apply to ABR and UBR services to improve the channel utilization. The resource assigned to Nrt-VBR services consists of two parts. One is the ATM layer requested cell rate and the other part is the bandwidth dynamically reserved by out-of-band resource requests. So the non-real time VBR services are mapped onto a mixed connection mode.

### 7.3.2.2 Mapping the Connection Reliability

For some real-time services, cells that are delayed in the network by more than the specified CTD are considered to be of less value, or no value at all to the application [ATF96]. If a reliable traffic transmission cannot meet the required CTD when the effective bandwidth is provided, discarding these cells at the logic link layer will improve the CTD performance. This type of transmission is called as relaxed reliable transmission in this study.

The relaxed reliable transmission has to ensure that the total cell loss ratio \( CLR_T \) resulting from the buffer overflow and the non-ideal physical channel meets the required \( CLR_{req} \). Let \( CLR_0 \) denote the cell loss ratio due to the buffer overflow and \( CLR_C \) denote the cell loss ratio due to the non-ideal physical channel, \( CLR_T \) can be given as per [MOH97],

\[ CLR_T = CLR_0 + (1 - CLR_0) CLR_C \]
Let the $T_{\text{timer}}$ denote the retransmission timer of the ARQ protocol and MaxCTD denote the maximum cell transfer delay required, then $N$, the maximum transmission number of each packet at logic link layer before the maximum cell transfer delay is reached, is given as

$$N = \left\lfloor \frac{\text{MaxCTD}}{T_{\text{timer}}} + 1 \right\rfloor \quad (7-7)$$

Since each packet can only be transmitted $N$ times, the packet loss ratio in the channel CLR$_c$ is,

$$\text{CLR}_c = e^N \quad (7-8)$$

To ensure that the required CLR and CLR$_T$ satisfy the inequality,

$$\text{CLR}_T \leq \text{CLR}_\text{req} \quad (7-9)$$

this requires the following;

$$\text{CLR}_O + (1 - \text{CLR}_O)e^N \leq \text{CLR}_\text{req} \quad (7-10)$$

As indicated in Eq.7-10, if we can choose a buffer size for which CLR$_O$ satisfies Eq.7-10, then the relaxed reliable mode can be used. In the case that a large buffer size can be provided, then for most of the $e$ values, CLR$_O=0$, we have

$$e^N \leq \text{CLR}_\text{req} \quad (7-11)$$

Given that Eq.7-10 can be satisfied, CBR services and Rt-VBR services, which have a certain level of the cell loss tolerance, can be considered to use the relaxed reliable transmission to improve the CTD performance.

The Nrt-VBR and ABR services require a low cell loss ratio and have no strict requirements on CTD and CDV, so the reliable transmission modes are adopted. For UBR services, because the network does not commit guaranteed service quality to this type of service, relaxed reliable transmission or unreliable transmission could be utilized when the system is congested. The connection-mapping scheme is summarized in Table 7-1.
7.4 Simulation Results

The channel model proposed by Zorzi [ZOR95] was used to analyse the performance of the proposed scheme. Since the voice traffic can tolerate certain levels of cell loss, the relaxed reliable finite buffer ARQ protocol presented in the previous chapter is modelled and used to investigate the performance of voice traffic in which some voice cells that are delayed in the network by more than the specified CTD are discarded.

In this study, we focus on investigation on the performance of mapping the voice traffic onto an ATM-satellite connection to which the $B_{eff}$ is allocated. The $B_{eff}$ is computed by using Eq.7-5. The performance of the transmission using both reliable and relaxed reliable transmission mode was evaluated and compared in order to verify that the relaxed reliable transmission could provide a better CTD performance as well as a satisfactory CLR performance for individual users in the presence of channel fading conditions.

The round-trip propagation delay between the SAT and GTW is considered to be 500ms. The voice traffic generates voice packets at 32kbps. Since voice packets are required to be delivered in real time, the maximum transmission delay, MaxCTD, is considered to be 1450ms in a relaxed transmission mode. Partial voice packets that have not been transmitted within the required MaxCTD are discarded. The MaxCTD is regarded to be infinite in a reliable transmission. The Retransmission timeout at the logic link layer is set to be 750ms for a GEO satellite system.

Figure 7-4 to Figure 7-7 show the performances of CLR and CTD for the voice traffic transmitted with assigned $B_{eff}$ in a flat fading channel using reliable transmission mode and relaxed transmission mode. At each simulation point in these figures, the effective bandwidth, which is calculated according to the corresponding CLR$_{req}$ and fading margin related to the average channel PER by Zorzi’s channel model, is utilised.
A QoS-Provisioning Connection Management Scheme

Figure 7-4 CLR of the relaxed reliable transmission mode vs. fading margin.

Figure 7-5 CTD vs. fading margin (CLR_{req}=0.0001).
The performance of the CLR in a relaxed reliable transmission mode under different CLR_{req} requirements of 0.0001, 0.005 and 0.01 are shown in Figure 7-3. The buffer size is 172.8kbits. The CLR performance of a reliable transmission under the same simulation parameters are all zero, therefore their results are not displayed in Figure 7-4. The results show that the cell ratio discarded in the relaxed transmission mode has a polynomial relationship with the fading margin. When the fading
margin is less than 16dB, the number of discarded cells becomes high and sensitive to the fading margin. However, these performances demonstrate that the relaxed reliable transmission can adequately maintain the CLR to the required level as the reliable transmission mode. Partial discarding of the packets delivered over time does not damage the transmission stability.

The performance of the mean end-to-end CTD under different CLR\textsubscript{req} requirements of 0.0001, 0.005 and 0.01 are shown in Figure 7-5, Figure 7-6 and Figure 7-7 respectively. From the CTD delay performance revealed in Figure 7-5 and Figure 7-6, we observe that the effective bandwidth computed approximately by Eq.7-5 can provide a guaranteed CTD delay performances for both transmission modes whilst the required CLR\textsubscript{req} is met, given that the fading margin is higher than 14dB. The performance is degraded when the fading margin is less than 14dB. However, it offers a performance that is close to the required CTD of 1450ms. In Figure 7-7, where the CLR\textsubscript{req} is 0.01, the approximated $B_{\text{eff}}$ cannot meet the required CTD even with a large fading margin of 31dB. The reason for this is that the approximation of the effective bandwidth given in Eq.7-5 is optimized to the performance objective of the CLR and not to the maximum CTD. To guarantee the performance of both CTD and CLR under any channel fading condition, a more accurate $B_{\text{eff}}$ approximation, which may be a function of channel fading, buffer size, MaxCTD, round-trip delay and retransmission timeout, should be considered for further study.

The poor CTD performance at low fading margin can be improved by increasing the allocated bandwidth. However, this improvement is limited. When the limit is reached, a more powerful forward error control is needed to counteract the channel fading rather than an ARQ protocol alone. This is due to two reasons. Firstly, in order to provide a sequence-in-order service to the higher layer, some correctly received packets are kept in the receiver buffer waiting for the lost packets that have smaller sequence numbers. The waiting delay in the receiver buffer is difficult to trace and control. This is the major reason that the mean packet transfer delay is longer than 1450ms. A second reason is that, in the relaxed reliable transmission mode, only partial packets, which have a packet transfer delay longer than MaxCTD, are discarded. This is to protect the packet transmission from suffering an uncontrollably high packet discarding when a severe fading is encountered. Some correctly received packets having a long CTD are delivered to the higher layer.

It is also observed from Figure 7-7 that the relaxed reliable transmission mode can provide a much better CTD performance than the reliable transmission mode at a high packet error rate and can maintain the delay performance closer to the required delay level than the reliable transmission mode.

Because the computation of the effective bandwidth is based on an accurate prediction of the average packet error rate in the channel, we also investigate the impacts of an inaccurate packet error rate prediction. The effects of an inaccurately predicted channel packet error rate on the performance of
the CLR and CTD are shown in Figure 7-8 and Figure 7-9. The CLRreq, the buffer size and the average channel PER used in the simulation are 0.005, 90kbits and 0.070, respectively. The approximated $B_{eff}$ according to Eq.7-5, is 51.72kbps. Figure 7-8 and Figure 7-9 show that the performances of CLR and CTD are critically dependent on the predicted packet error rate (PER) value. However, increasing the buffer size at the transmitter can solve this sensitivity problem. It is found that the CTD performance benefits more from an increased bandwidth than an increased buffer size because the CTD performance is mainly decided by the allocated bandwidth, channel fading condition and the round-trip delay.

The above simulation results show that the effective bandwidth approximation obtained from Eq.7-5 can ensure the required CLR for the voice transmission over satellite given the buffer size larger than...
90kbits. It is also able to provide a CTD performance that is close to the required MaxCTD. When a high PER occurs in the mobile channel, the CTD performance may be further improved by using the relaxed reliable transmission mode. The effective bandwidth approximation given in Eq.7-5 offers the advantage of a close approximation to the accurate effective bandwidth, which can guarantee both CLR and CTD for any channel condition.

7.5 Conclusions

A novel radio connection management scheme for QoS-provisioning transport of ATM traffic over satellite links has been presented in this chapter. The proposed connection management scheme effectively manages the air interface connections for ATM services through diversifying connection types, establishing QoS-based connections and deploying an efficient connection mapping and control scheme. The scheme is optimised to target the specific performance objectives of cell loss rate for individual services, and enables the network to keep unused bandwidth on a connection to a minimum. Computation of the effective bandwidth for constant bit rate services has been detailed.

The performance of the connection management scheme is evaluated by characterising the performance of Cell Loss Rate (CLR) and Cell Transfer Delay (CTD) of voice traffic transmission in the presence of a channel fading. The simulation results show that the effective bandwidth approximation obtained from the computation model can ensure the required CLR for voice transmission over the satellite. It can also provide a CTD performance that is close to the required MaxCTD. When a high packet error rate occurs in the mobile channel, the CTD performance can be further improved by using the relaxed reliable transmission mode. The approximation model proposed in this work offers a reasonably good approximation to the accurate effective bandwidth.

However, accurate computation algorithms for the effective bandwidth of constant bit rate services and variable bit rate services under different channel conditions, which can satisfactorily provide both required CLR and CTD with a minimal bandwidth provisioning, needs to be extensively investigated. This is suggested for the future work.

The proposed radio connection management scheme together with the optimised MAC and ARQ provides a framework of interworking protocols for ATM over satellites. It can also find applications in other systems that involve the integration of terrestrial protocols and mobile satellites.
Chapter 8 ATM-Satellite Air Interface Call Handling

Based on the system architecture, multiple access scheme, error control scheme and connection management scheme discussed in the previous chapters, we propose a connection control and call handling scheme for an ATM-Satellite integrated system in this chapter. A reservation meta-signalling is proposed for setting up signalling connections at the user-network radio interface. Call processing techniques such as authentication, location update, paging, handover and call routing, are discussed and call set-up procedures for both mobile originated and terminated calls are presented. The objective of this chapter is to demonstrate the call control and routing implementation for the ATM-satellite system.

8.1 Introduction

The new ATM-satellite signalling procedures must be flexible enough to support terminal mobility and satellite access functions whilst utilising sufficiently the fixed ATM signalling protocols. For this reason, the ATM-satellite call control function should be based upon the Q.2931 signalling scheme. However, signalling integration always faces the difficulty of compromising the trade-offs between the signalling simplicity, bandwidth efficiency and integration cost. In chapter 4, we have suggested an ATM Q.2931 compliant approach of the ATM-satellite system protocol architecture.

Based on this suggested approach, in this chapter, we address the call control related issues at the control and mobility plane and propose a call handling solution for ATM-satellite integrated systems. At the control plane, we address call routing strategies and call handling management that are specific to establishing, maintaining and releasing a link over satellite. At the mobility plane, we address terminal authentication, location update, paging and handover procedures.

The discussion in this chapter is centred on mobility specific connection set-up and call routing procedures. A reservation meta-signalling is proposed to set up a dedicated signalling at the air interface. Signalling implementations such as incoming call routing, outgoing call delivery, handover routing and paging techniques, are presented. The call set-up procedures for point-to-point and point-to-multipoint signalling protocol procedures are described at the end of chapter.
8.2 Service-Based Network Access Control

As we can see from the MAC scheme that has been proposed for the mobile ATM-satellite network, mobile terminals initially access the network using the random access method. The performance of random access has been discussed in the previous chapter. Here we address the mechanisms for controlling the user's access. A network usually needs mechanisms to perform system load control, security control and collision control. In advanced telecommunication systems, the access control mechanism is expected to provide more access options such as user-specified access control and group-specified access control. Consequently, supplementary services and possibly signalling procedures are required to set up the user defined or the group-defined access rules.

The security control service can adopt password-based access control. The collision control requests the network to perform a collision resolution algorithm that executes the collision access rules to resolve the collisions. The simplest and most commonly used collision resolution scheme is the random backlog delay, as used in our proposed MAC protocol.

The access rules that control the access load is important considering a heavy traffic to be carried in the future communication systems. A good access load control method can also reduce the collision rate in a random access channel. As far as the signalling and protocol performance is concerned, the system access load control is investigated here.

Because the access method in the proposed MAC scheme adopts a random manner, the throughput can be very low. The throughput of the random access (RA) channel decreases quickly when the channel load rises beyond 1. The load control has to ensure that the access channel does not become a bottleneck for the network operation. It is expected that a fast overload detection and fast overload control can be deployed in the network. Once an overload is detected, immediate overload control can quickly relieve the overload condition. An ideal throughput performance of the RA channel with effective control scheme is illustrated in Figure 8-1.
The load control can adopt a priority-based access control mechanism. In an ATM-satellite system, since the priority is mainly decided by the service type, we propose a service type based priority control scheme. Meantime, rules should be used to prioritise the calls with special needs such as emergency access and handover access.

This service type based control scheme aims to ensure a fast access and high throughput of delay sensitive services. The control procedure is as follows:

1. The network first prioritises all the service requests according to the following priority order and assigns a priority code to each of them according to their services type.

   \[\text{CBR} > \text{Rt-VBR} > \text{Non-rt VBR} > \text{ABR} > \text{UBR}\]

2. When a terminal accesses the network, its priority code should be included in the initial access message.

3. The network keeps a record of access and collision statistics and periodically dimensions the traffic condition on the channel. The network decides whether the channel is in overload, normal or underload condition, according to the statistics record.

4. If the network detects an overload, it broadcasts a "stop access" message in the downlink feedback channel with the lowest priority codes included. The active users whose code is included in the message will temporarily stop accessing and wait for a further notice.
5. If the overload condition is not changed after a time period, the second lowest priority group will be stopped to access the network by network broadcasting their priority code in the downlink feedback channel (FC). The process will continue until the channel enters a stable normal traffic condition.

6. When the system is recovered from the overload, the network will broadcast a “start access” message in the feedback channel. Upon receiving this message, those terminals will start to access the network as normal.

7. The network can gradually block the access of different priority groups one after another. It can also block the access of several priority groups at the same time in order to achieve a fast load reduction.

8.3 Reservation Meta-signalling for Signalling Connection Set-up

8.3.1 Meta-signalling in ATM

As we know the user-to-network signalling connection set-up adopts two methods in ATM. For a point-to-point configuration, a dedicated signalling channel with default value VCI=5 is used. In a point-to-multipoint user-network interface (UNI) case, signalling connection is set up on demand by a meta-signalling protocol. This is needed because of the large bandwidth and the wide variety of services supported in ATM requiring a number of different signalling procedures. Thus B-ISDN signalling channels are not permanently defined.

The objective of the signalling connection set-up is to request both the VPI/VCI and channel resource for setting up a dedicated signalling channel. Meta-signalling is a simple protocol with 5 messages. All the commands are given space in one ATM cell, each via a reserved ATM channel with VCI=1 and default VPI=0. Meta-signalling is invoked by the user following a call establishment request from the application layer. It could also be invoked by the network to maintain the on-going signalling connections. A major task of meta-signalling is to obtain both the requested bandwidth and the VPI/VCI identifier for the traffic connection. From this point, the VPI/VCI also become a type of resource similar to the bandwidth resource that needs to be dynamically assigned on demand.

The satellite system is a typical point-to-multipoint UNI configuration system. Meta-signalling therefore needs to be considered to undertake the task. However, the efficiency of
adapting this scheme directly to a satellite system must be addressed before implementation. In the following, we will show that a direct implementation is not efficient. Two different approaches to encapsulate the ATM meta-signalling into an ATM-satellite system have been identified. These two solutions imply two different protocol reference models of ATM-Satellite integration, as shown in Figure 8-2.

8.3.2 Solution A—Random Access Meta-signalling

As there is no dedicated signalling channel for meta-signalling at the air interface, the only possible method for meta-signalling to undertake signalling connection set-up is to implement the random access protocol on a random access channel. The same “ASSIGN REQUEST”, “ASSIGNED” and “DENIED” messages in the ATM meta-signalling protocol can be adopted to complete the signalling connection set-up. All users share this common random access channel. Whenever a user wants to set up a connection, the user sends its meta-signalling message to the network asking for channel and VPI/VCI resource.

A disadvantage of this protocol lies in the inefficiency of channel utilisation. In this solution, the meta-signalling message is sent directly on the random access channel. Because each meta-signalling message is 53 bytes which a big random access burst size, the access throughput is relatively low and collision is relatively high. This causes bandwidth wastage. Besides, some information provided in the “ASSIGN REQUEST” message is not necessarily sent at the initial access stage in the satellite case. However, an advantage of this protocol is its short access delay without reservation requirement.

8.3.3 Solution B—Reservation Meta-signalling

In this solution, connection set-up can be split into two steps. The first step is to use a small reservation packet to reserve an uplink timeslot to transmit the meta-signalling message. The
second step is to perform meta-signalling on a reserved timeslot to request VPI/VCI. Unlike the ATM meta-signalling that requests channel resource and VPI/VCI resource at the same time, reservation meta-signalling separates channel reservation from VPI/VCI reservation. The flexibility of this scheme is that small packets can be used for reservation so that random access can be limited to a small area. All other channels will be reserved and assigned without contention. Throughput and channel utilisation are therefore higher than in random access. The disadvantage of this scheme is that an extra round-trip delay may be introduced. This solution requires a radio control layer below the ATM layer in order to create and process reservation packets. The protocol stack is shown in Figure 8-2 (a).

8.3.4 Performance Comparison

We should note that the major task of the above two connection set-ups is to complete both VPI/VCI and channel resource assignment during a connection set-up. Solution A and B represent different ways to achieve this aim.

Comparisons between these two protocols are shown below. First, solution A can have better delay performance than solution B because there is no reservation required in solution A. Packets are sent to the channel directly. However, this method is not flexible in terms of organising the channel frame. In some cases such as paging and channel measurement reporting, smaller packets will be more appropriate than satellite-ATM cells and can save on the limited satellite channel resources. Solution B overcomes the aforementioned drawbacks, but can lead to a longer delay since the reservation needs to be made before meta-signalling.

To decide upon which protocol is optimal for an ATM-satellite system, we have studied their performance in bandwidth utilisation and connection set-up delay. Considering that the satellite channel resource is so demanding that it has become the major concern for a system designer, we set our criterion for protocol selection as utilisation of bandwidth.

It is seen that solution B possesses the potential to meet this criterion. A careful design of the proportion of mini-reservation slots to packet slots in the satellite channel can further improve channel throughput. In addition, solution B can also outperform solution A on delay performance in some cases given that they both have the same call arrival rate and the reservation slots are significantly small in solution B.

If we let $T_f$ be the channel frame duration, $T_g$ the guard time between timeslots, $T_s$ the timeslot duration and $N$ the number of timeslots per frame. Then in the case of solution A, we have:
\[ T_r = N(T_s + T_g) \]

In the case of solution B, we assume that \( M \) out of \( N \) timeslots are used for reservation. Each one of \( M \) frames can split into \( L \) mini-reservation slots. The guard time between mini-reservation slots is the same as that between data timeslots. Let \( T_m \) be the duration of the mini-reservation slot, then we have

\[ T_i = (N - M)(T_s + T_g) + M \times L(T_m + T_g) \]

Let us assume that both channels are under the channel load \( G \) that is equal to 1. Then the throughputs of solution A and B on the random access channel are both \( e^{-1} \). The channel utilisation of solution A is \( e^{-1} \). In solution B, if we can satisfy the following conditions

\[ \frac{L \times M}{N} > 1 \quad (8-1) \]

and

\[ L \times M \times e^{-1} \leq (N - M) \quad (8-2) \]

Then the channel utilisation in solution B is \( \frac{L \times M \times e^{-1}}{N} \) which is higher than in solution A.

It is not difficult for B to fulfil requirements of Eq.8-1 and Eq.8-2. Reservation packets only contain reservation cause and reference number. They are very small packets compared to the ATM-satellite cell which is in excess of 53 bytes, and normally can accommodate more than 3 minislots. Once the \( L \) is chosen, we can carefully select \( M \) and the expected throughput will be achieved.

Based on the above considerations, solution B is preferred to undertake the signalling connection set-up in an ATM-satellite system. Performance of this protocol is dominated by the reservation channel performance, e.g. the slotted aloha protocol. If we assume that the mean throughput of the slotted aloha channel is \( \eta_{sa} \), then the channel throughput of solution B, \( \eta_B \) are given as:

\[ \eta_B = \min \left\{ \frac{N - M}{M \cdot \eta_{sa}} \right\} \]
8.3.5 Simulation Results of RMS Connection Set-up

The performance of the reservation meta-signalling (RMS) connection set-up is measured by means of set-up delay and link efficiency. Link efficiency is defined as the utilised data traffic bandwidth to total bandwidth held. Two parameters, reservation packet size (RPS) and the numbers of reservation channels, are taken as variables in the simulation. Here we assume that there are \( k \) timeslots per frame and \( n \) out of the \( k \) timeslots are used for reservation. Each of the \( n \) slots can be divided into \( m \) mini-reservation slots.

The distributions of the delay and link efficiency variations with the RPS are shown in Figure 8-3 and Figure 8-4, given that the reservation request arrival rate is 20.76 calls/s, \( k \) is 10 and \( n \) is 1. It can be seen that reducing RPS leads to a fast decrease of set-up delay and increase of link efficiency. Link efficiency can reach 0.53 and is much longer than the 0.368 of slotted aloha. However, the set-up delay has a minimum bound and the link efficiency a maximum bound. Further improvement of delay and link efficiency cannot be achieved by reduction of RPS.

Figure 8-5 and Figure 8-6 show performance of set-up delay and link efficiency for different \( n \) that determines the amount of bandwidth used for resource reservation. Clearly, increasing \( n \) can enhance both the set-up delay performance and link efficiency. However, increasing \( n \) implies a decrease of the available bandwidth for data traffic. If this increase causes the total available bandwidth for data traffic to go below the total traffic bandwidth required by existing active users, then the traffic bandwidth will be fully utilised. This is the reason that the link efficiency shown in Figure 13 drops to 0.5 and 0.4 when \( n \) is increased to 5 and 6 respectively.

The achievable minimum set-up delay and maximum link efficiency are dominated by the reservation request load, reservation packet size and available bandwidth for reservation. Under conditions of known system load and RPS, an optimal bandwidth for the reservation channel exists, for which the maximum system link efficiency can be achieved.

Simulation results show that RMS exhibits good performance on connection set-up delay and link efficiency. Its throughput performance is better than that of random access. Moreover, the achievable able maximum link efficiency improves with the reduction of RPS whilst the connection set-up delay can also be improved to a satisfactory level.
Figure 8-3 Mean set-up delay vs. reservation packet size.

Figure 8-4 Link efficiency vs. reservation packet size.
Figure 8-5 Mean set-up delay vs. Ratio of reservation slots to data slots.

Figure 8-6 Link efficiency vs. Ratio of reservation slots to data slots.
8.4 Logical Channels

The logical signalling channels required at the air interface of the integrated ATM-satellite system have been identified. We focus on those logical channels devoted to connection control and call control protocols. The definition and description of these signalling channels are given in Appendix B and their applications and multiplexing mode are summarised in Figure 8-7. These channels are defined based on the operational requirements of the proposed ATM-satellite signalling system.

<table>
<thead>
<tr>
<th>Channel name</th>
<th>Description</th>
<th>Application</th>
<th>From</th>
<th>To</th>
<th>Multiplexing</th>
</tr>
</thead>
<tbody>
<tr>
<td>USC</td>
<td>Uplink Synchronisation Channel</td>
<td>Synchronisation</td>
<td>SAT</td>
<td>P/L</td>
<td>Slotted Aloha</td>
</tr>
<tr>
<td>DSC</td>
<td>Downlink Synchronisation Channel</td>
<td>Synchronisation</td>
<td>P/L</td>
<td>SAT</td>
<td>TDM</td>
</tr>
<tr>
<td>RRC</td>
<td>Resources Request Channel</td>
<td>Resource Reservation</td>
<td>SAT</td>
<td>TRM</td>
<td>Slotted Aloha</td>
</tr>
<tr>
<td>RAC</td>
<td>Resources Assignment Channel</td>
<td>Resource Assignment</td>
<td>TRM</td>
<td>SAT</td>
<td>TDM</td>
</tr>
<tr>
<td>FC</td>
<td>Feedback Channel</td>
<td>Feedback Information</td>
<td>TRM</td>
<td>SATs</td>
<td>TDM</td>
</tr>
<tr>
<td>BC</td>
<td>Broadcast Channel</td>
<td>Broadcast Channel</td>
<td>MCS</td>
<td>SATs</td>
<td>TDM</td>
</tr>
<tr>
<td>CSC (SAT)</td>
<td>Common Signalling Channel</td>
<td>Common Signalling</td>
<td>MCS</td>
<td>SATs</td>
<td>TDM</td>
</tr>
<tr>
<td>DMSC (SAT)</td>
<td>Dedicated Meta-Signalling Channel</td>
<td>Meta-signalling</td>
<td>SAT</td>
<td>MCS</td>
<td>DAMA</td>
</tr>
<tr>
<td>DSCC (SAT)</td>
<td>Dedicated Signalling Control Channel</td>
<td>Dedicated Signalling</td>
<td>SAT</td>
<td>MCS</td>
<td>TDM</td>
</tr>
<tr>
<td>PGC</td>
<td>Paging Channel</td>
<td>Paging</td>
<td>MCS</td>
<td>SAT</td>
<td>TDM</td>
</tr>
<tr>
<td>SASC (GTW)</td>
<td>Slow Associated Signalling Channel</td>
<td>Associated Signalling</td>
<td>SAT</td>
<td>GTW</td>
<td>TDM</td>
</tr>
<tr>
<td>TCH</td>
<td>Traffic Channel</td>
<td>Traffic Channel</td>
<td>SAT</td>
<td>GTW</td>
<td>TDM</td>
</tr>
<tr>
<td>FSCC</td>
<td>Forward Satellite Control Channel</td>
<td>Resource Management</td>
<td>MCS</td>
<td>TRM</td>
<td>TDM</td>
</tr>
<tr>
<td>RSACC</td>
<td>Return Satellite Control Channel</td>
<td>Resource Management</td>
<td>TRM</td>
<td>MCS</td>
<td>TDM</td>
</tr>
<tr>
<td>DSCC (GTW)</td>
<td>Dedicated Signalling Control Channel</td>
<td>Dedicated Signalling</td>
<td>MCS</td>
<td>GTW</td>
<td>TDM</td>
</tr>
<tr>
<td>CSC (GTW)</td>
<td>Common Signalling Channel</td>
<td>Common Signalling</td>
<td>MCS</td>
<td>GTW</td>
<td>TDM</td>
</tr>
<tr>
<td>GRRRC</td>
<td>GTW Resource Request Channel</td>
<td>Resource Reservation</td>
<td>GTW</td>
<td>TRM</td>
<td>Slotted Aloha</td>
</tr>
<tr>
<td>GRAC</td>
<td>GTW Resource Assignment Channel</td>
<td>Resource Assignment</td>
<td>TRM</td>
<td>GTW</td>
<td>TDM</td>
</tr>
</tbody>
</table>

Figure 8-7 Logic signalling channel summary.
8.5 Call Processing Techniques

In this work, we assume that the ATM-satellite system maintains the relationship between the ATM-satellite network address space and the fixed ATM network address space, i.e. at each ATM-satellite network element, it will always be possible to acquire the ATM-satellite network address given its ATM network domain identity or vice-versa.

With respect to system mobility signalling, which is used to route an incoming, handover or outgoing call, the approach of centralised location management architecture is considered. We assume that the location area (LA) and paging area (PA) are both defined as same as per the spotbeam coverage, the home location register (HLR) attached to the MCS centrally controls the location update and paging for the coverage of one satellite.

8.5.1 Authentication

In this work, we assume that the user authentication centre (AC) resides with the HLR, but the visitor location register (VLR) will be temporarily granted parts of the authentication functions when the user resides in the visited location area [MOU92]. Two types of authentication are identified to be integrated into the control plan of the ATM-satellite specific protocol: SAT-MCS Authentication and the MCS-MCS Authentication. The SAT-MCS authentication mechanism will be used to authenticate the user terminal at the location update and call set-up time. The authentication procedure will be either user or network initiated. Dependent on the system security requirements, the authentication can be a mutual authentication between two MCSes. For MCS-MCS authentication, a similar procedure to SAT-MCS authentication can be used.

8.5.2 Location Update and Paging

As the location area is defined as the area of a spotbeam, location update is expected to be performed when a mobile terminal crosses the location area boundary. Three different crossing cases are identified; in-call crossing, standby-mode crossing and idle-mode crossing.

8.5.2.1 In-call crossing

In-call crossing refers to the boundary crossing when a mobile terminal is in an active call. The terminal will firstly carry out the handover followed by a location update procedure. The location update can be initiated by either the mobile terminal or the network.
8.5.2.2 Standby-mode crossing

A terminal may cross the boundary in a standby mode. The location update procedure is initiated by a mobile terminal once the change of the residing spotbeam is detected. In this case, the terminal has to access the network and obtain a signalling channel before the location update procedure can be executed.

8.5.2.3 Idle-mode crossing

In the idle mode, a terminal is in detach mode and is power-off. It is considered to be temporarily unreachable by the network. The location update procedure will not be executed until the terminal is switched on again. The location update procedure is performed immediately after the terminal synchronises itself to the network.

Because the location area and the paging area are the same, the HLR has the accurate information of the terminal's residing spotbeam unless the terminal is currently in a state of detachment. Thus the terminal is only necessarily paged in the spotbeam it resides in. Paging is a single message sent on the downlink paging channel. Upon receiving the paging message, the terminal accesses the channel and sets up the connection.

8.5.3 Incoming Call Delivery

The gateway is responsible for fetching the location information and routing the call towards the MCS through which the subscriber can obtain services at this instant (the visited MCS). When an incoming call arrives at a gateway, the gateway looks up its database to determine the HLR of the called mobile terminal. The gateway visits the HLR and requests the location and routing information of the HLR. The HLR verifies the identification and accessibility of the called mobile terminal and advises the gateway of the address of the correct MCS for which the called mobile terminal is currently registered, and therefore to which the call should be routed. The address obtained is then sent back to the gateway. Gateway delivers the incoming call to the MCS visited by the called user. The MCS looks up the database and finds the spotbeam in which the mobile terminal resides at that instant. The MCS pages the mobile terminal in the spotbeam until the terminal accesses the channel to respond to the paging. After both channel and VPI/VCI are assigned to the user, the authentication will be performed. Finally, the call set-up request from the calling user is delivered to the called user.
8.5.4 Outgoing Call Routing

Each MCS is assumed to be connected to an ATM switch node via a gateway. An outgoing call is firstly routed to the MCS and further routed to the gateway that can perform the interworking functions. In this case, an outgoing call is firstly authenticated by the MCS or HLR and is then routed by the MCS to a traffic resource manager in order to perform a resource reservation. If the resource is available, the call is further routed to the gateway. The MCS does not have an interworking functions in this case, so that all calls are routed to their destinations in the fixed ATM network via the gateway.

In the case that the call is raised from a mobile user for another mobile user, the MCS first asks the VLR to authenticate the calling mobile user and then obtains the routing information of the called mobile user from the HLR. The HLR sends call-routing information back to the calling MCS. The calling MCS sends a call set-up request to the called MCS requesting for the incoming call delivery information. The called MCS/VLR decides if the call can be delivered. Once the decision made is positive, a paging command will be sent by the called MCS in the location area to page the called mobile user. The mobile will response to the paging and starts the connection set-up procedure.

8.5.5 Handover Routing

The handover in the ATM-satellite system is a hard, backward, network initiated handover. A mobile terminal submits the results of link measurements to the network regularly or upon requests from the network. In the ATM-satellite system that we have considered in this work, we assume two types of handovers, the inter-MCS handover and intra-MCS handover. Intra-MCS handover is handled locally by the MCS to transfer the call from one spotbeam to another. Inter-MCS handover will be performed under the cooperation of two neighbouring MCSes.

Intra-MCS handover is performed simply via updating the routing table and re-allocating resources. The same VPI/VCI can be used but the actual channel allocation will change. Inter-MCS handover can take one of two methods, path extension and path re-routing [ATF97]. In the ATM-satellite system, the path extension is not preferable because of the longer propagation delay. Path re-routing is thus considered in this work.

The handover is determined by the current MCS which first searches for a target MCS and then sends a handover request to the targeted MCS. After the handover request has been accepted, the targeted MCS sends the handover acceptance to the source MCS and begins to
transfer the necessary information. After completing the process, the source MCS serves the old connection.

### 8.5.6 Radio Connection Type Negotiation

As proposed in the previous chapter, a suitable ATM-satellite air interface connection type has to be determined for each ATM service requesting a connection during the call set-up phase. The mapping algorithm that maps the individual ATM service to a suitable type of ATM-satellite connection has been presented in the previous chapter. Special signalling procedures are needed to perform and establish the connection type. The proposed procedure will be performed when the call set-up message is submitted to the MCS during the call set-up phase. The network decides what connection type to be used and sends the connection type information back to the mobile terminal via a signalling message. However, a negotiation of the connection type is permitted if the connection type that the network has proposed is not accepted.

### 8.6 Call Level Service Quality Requirements

Being more difficult than the terrestrial ATM network, the ATM-satellite network has to ensure a call level QoS in addition to the cell level QoS. The call level QoS is usually decided by the system traffic load, total resource available and handover quality. The network must closely monitor the system operation condition and efficiently control the traffic to ensure low call blocking, signalling set-up delay and signalling failure probability.

Handover must maintain the continuity and integrity of calls at a satisfactory quality level. The handover delay should be as short as possible to minimise the interruption to the communication and ensure that the handover decision still remains valid for the new mobile terminal position after the handover process is complete. The handover procedure should aim to preserve the requested QoS of the original connection. When this cannot be achieved, a procedure of QoS re-negotiation may be needed to minimise the communication degradation.

An acceptable call level QoS relies on an efficient management of traffic and handover release. It is expected that the desired QoS can be maintained with minimal trade-off of the channel bandwidth.
8.7 Call Establishment Procedures

Based on the above call processing and routing discussion, we discuss and present the examples of call set-up procedures in this section. The main objective is to demonstrate the implementation of the designed ATM-satellite call handling protocols.

8.7.1 Point-to-Point Call Establishment

Given the above discussion of call routing, point-to-point (PTP) call control procedures for mobile originated and terminated call are shown in Figure 8-8 and Figure 8-9. The major signalling procedures performed include authentication, resource reservation, connection type set-up, routing and paging.

Mobile originated PTP call set-up procedure

![Figure 8-8 Mobile originated call set-up.](image-url)
Mobile terminated PTP call set-up procedure

8.7.2 Point-to-Multipoint Call Establishment

A point-to-multipoint (PTM) ATM-satellite call set-up procedure adopts the "add party" and "leaf join" procedures from fixed ATM. However, with an ATM-satellite PTM call, the leaf terminals could include both fixed ATM end users and mobile ATM terminals. Similar to the ATM "add party" procedure, a root connection is set up first, using point-to-point signalling between calling user and the first party user. The rest of the leaf terminals join the group upon the "add party" request sent by the root or use the "leaf join" procedure which allows the leaf to initiate a join to an active connection without the intervention from the root. However, the difference in the PTM call set-up procedure between ATM-satellite and fixed ATM is that all the mobile terminals in the former case will be paged and interrogated before the join.

There are two potential problems associated with mobile PTM call set-up procedures. Firstly, the call set-up delay could be undesirably long due to the long propagation delay and many leaf parties involved. Secondly, because each leaf requires a separate connection session, the
total bandwidth taken from the uplink/downlink could be large and inefficient. In order to save the uplink bandwidth, it is proposed to share the bandwidth by all the connection sessions on the uplink between the root SAT/GTW and the satellite payloads. So if applicable, the same uplink traffic channels are expected to be shared by all the leaf parties in the same spotbeam. A further study concerning the implementation of this bandwidth-sharing scheme remains for future investigation.

### 8.8 Conclusion

This chapter has presented a reservation meta-signalling for setting up signalling connections at the user-network radio interface, and a mobility-enhanced call handling protocol promoted from Q.2931. The advantage of the proposed ATM-satellite call control signalling is that it supports terminal mobility and satellite access functions whilst keeping to a minimum modifications to the fixed ATM signalling protocol. Call processing techniques such as authentication, location update, handover, incoming call delivery and outgoing call routing have been discussed. Call set-up procedures for both mobile originated and terminated calls were presented.

We have demonstrated the incorporation of mobility specific signalling into the ATM Q.2931 standard and possible call handling procedures that can be used in the ATM-satellite network. The proposed signalling protocol provides a protocol reference model for ATM-satellite integrated systems.
Chapter 9 Conclusions and Future Work

This thesis has addressed and presented a radio connection management scheme and a mobility-enhanced signalling protocol system for ATM-satellite integrated mobile communication systems. This chapter summarises the major contributions, presents a brief discussion of the results and discusses directions for the future work.

9.1 Conclusions

This thesis has addressed and presented a radio connection management scheme and a mobility-enhanced signalling protocol system for ATM-satellite integrated mobile communication systems.

The issues involved in integrating the ATM and mobile satellite technologies were identified at the beginning of the work. As described early, the use of satellite to carry ATM traffic and to provide user mobility presents a number of problems due to the nature of the satellite environment. Amongst other factors the high bit error rate, long propagation delay and LEO/MEO constellation dynamics, personal mobility differ substantially from those around which ATM was conceived. These factors impact on the transportation of ATM traffic over satellite and on connection management related functionalities at the air interface. It also imposes a requirement to incorporate the mobility-related functions into the ATM protocol. These identified issues/problems are also common to other research areas that involve the integration of terrestrial and satellite protocols.

This work has thus been directed to addressing these problems and researching a viable solution to an effective radio connection management and mobility-enhanced call handling procedures. We target at minimising the difference in performance between terrestrial ATM and ATM over satellite and providing mobility extension into the ATM protocols whilst maintaining a high satellite channel efficiency and keeping as little as possible signalling modifications.

One of the contributions of this work is that a new and fully integrated ATM-satellite signalling protocol architecture, which accounts for mobility of ATM users, has been proposed. The proposed architecture is developed from the ATM protocol. The mobility functions and MAC, LLC functions are incorporated into the existing ATM layers. The proposed integration scenario, network elements and connectivity, protocol reference model and layered protocol functions have been presented. This architecture is a general structure for the overall system design.
Conclusions and Future Work

In order to achieve an efficient connection management scheme to support different ATM services with a minimal provisioning of satellite bandwidth, the MAC and ARQ protocols need to be optimised to adapt flexibly with the service types. This is because the connection management utilises the services provided by the MAC and ARQ. It relies on the MAC protocol to access and obtain the channel resource and utilises the ARQ to provide a reasonably reliable transmission that can satisfy the ATM layer QoS.

A service-adaptive MAC protocol is proposed. The protocol is based on a reservation-based demand assignment scheme. The proposed MAC scheme provides three different resource reservation and obtaining methods and adopts different resource reservation and allocation method flexibly with service types as shown in Table 7-1. The major advantage of this scheme is that it offers high channel efficiency.

Observing that current satellite systems use fixed assignment for signalling and that no work has been done to validate the application of the demand assignment MAC scheme to the signalling traffic, we further optimised the MAC scheme for its application into signalling to minimise the bandwidth wastage. This is a significant contribution of this work.

A signalling procedure is also proposed in order to apply the proposed demand assignment MAC scheme onto the signalling connections. Signalling protocols that use the demand assignment MAC scheme are called semi-permanent signalling protocols (SPSP). In contrast to the conventional permanent signalling protocol (PSP) where a dedicated signalling channel with a fixed bandwidth is assigned to an end user, the proposed SPSP requires the mobile terminal to go through a reservation procedure before sending the packets to the channel every time new signalling packets are generated. The number of timeslots that the mobile requests each time is equal to the number of the newly generated signalling packets.

The semi-permanent signalling protocol is a novel idea proposed in this work. The performance evaluation, in both error-free and fading channel conditions, shows that the proposed SPSP considerably improves the channel utilisation and signalling-request accommodation capacity. It also shows that the SPSP demonstrates comparative signalling delay performance to that of the PSP in good channel conditions whilst its signalling failure probability remains negligible. In spite of slightly inferior signalling failure probability and signalling delay of the SPSP to those of the PSP in the presence of channel fading, the performance difference can be minimised by increasing the number of resource reservation channels, limiting the number of retransmissions and increasing fading margin.
Conclusions and Future Work

The SPSP can be directly applied to the signalling that does not have a stringent delay constraint such as signalling establishment of non-real time services and real time services that can tolerate certain signalling delay. The SPSP is applicable to both satellite and terrestrial mobile systems.

A service-adaptive ARQ is also proposed. The proposed ARQ scheme can provide a reliability-dependent ARQ transmission to handle services that request different levels of cell loss ratios and cell transfer delay, and an adaptive timer retransmission to overcome the low throughput of the connections that are established via ISL and have changing link distance. The proposed ARQ adopts different transmission reliability according to different ATM services types as shown in Table 7-1.

Another significant contribution of this work is the proposal of the reliability-dependent SR ARQ (RDSR ARQ) protocol. This protocol, first proposed in this thesis, is the first such proposal that suggests discarding of some of the outstanding non-received packets, which have exceeded the required maximum CTD at the receiver, to speed up the transmission. The proposal is based on the observation of the inefficiency caused by the unlimited retransmission of outstanding packets in the SR ARQ that ensures a totally reliable packet delivery. However, unlimited retransmission of outstanding packets blocks the transmission of following packets and degrades the transmission speed. The requirement in ATM-satellite systems to support a variety of traffic types with different QoS requirements makes it necessary to enhance the conventional ARQ techniques for the delay constrained traffic transmission over satellite channels.

This scheme is capable of offering a better CTD performance and controlling the CLR compliant with the maximum allowable CLR. The scheme is suitable for mobile systems with long round-trip delays such as GEO satellite systems and other mobile systems that intend to provide high-speed transmissions to mobile terminals.

The Adaptive Timer SR ARQ (ATSR ARQ) that we have proposed is to overcome the low throughput of SR ARQ on a connection that consists of many ISL segments of LEO/MEO constellations. These connections generally have a time-varying distance characteristic. The study of the performance of SR ARQ protocol on end-to-end connections established via ISL is the very first research work in this area. The evaluation results show that the connections with time-varying distance degrade the efficiency of the SR ARQ protocol. The results also demonstrate that the adaptive timer SR ARQ (ATSR ARQ) is a viable and effective solution to improve the conventional SR ARQ performance on such a connection. It offers a good link efficiency and good cell transfer delay performance under a good channel condition. It can be
applied directly to the reliable connections such as connections between two satellites or between two fixed earth stations licked by ISLs.

However, the performance of the adaptive time SR ARQ on a changing-length connection in the presence of channel fading needs further comprehensive and in-depth study. It is suggested to be carried out in the future work.

The combined RDSR/ATSR ARQ is proposed as the ARQ protocol most applicable to the mobile ATM-satellite networks. In fact, it is a strong candidate for future mobile satellite networks that intend to offer multimedia services.

Based on the proposed MAC and ARQ protocols, we have put forward an innovative proposal for a QoS-provisioning connection management scheme to transport the ATM traffic over satellite links. The proposed QoS-provisioning connection management scheme is the most important contribution of this work. The proposed connection management scheme effectively manages the air interface connections for ATM services through diversifying connection types, establishing QoS-based connections and deploying an efficient connection mapping and control scheme.

The advantage of the scheme is that it can provide the specific performance objectives of CLR and CTD with minimal bandwidth provisioning for individual services. It enables the network to tailor the available bandwidth efficiently for different services with the different connection types associated with different transmission characteristics. It can also provide compensation for the QoS degradation caused by the mobile radio environment.

The performance of the connection management scheme is evaluated by characterising the performance of CLR and CTD of voice traffic transmission in the presence of channel fading. The simulation results show that the scheme can ensure the required CLR and CTD for the voice transmission over the satellite using both reliable transmission and relaxed reliable transmission. When a high packet error rate occurs in the mobile channel, using the relaxed reliable transmission mode instead of the reliable transmission mode can improve the CTD performance whilst maintaining the required CLR.

The proposed radio connection management scheme requires an accurate approximation of the required effective bandwidth of the individual service. The accurate computation algorithms of the effective bandwidth for constant bit rate services and variable bit rate services under different channel conditions, which can satisfactorily provide both required
Conclusions and Future Work

CLR and CTD with a minimal bandwidth provisioning, need to be further investigated. This is suggested for the future work.

Finally, based on the proposed protocol architecture, multiple access scheme and radio connection management scheme, a reservation meta-signalling for setting up signalling connections at the user-network radio interface, and a mobility-enhanced call handling protocol derived from Q.2931 are proposed.

The advantage of the proposed ATM-satellite call control signalling is that it supports terminal mobility and satellite access functions whilst keeping a minimal modification to the fixed ATM signalling protocol. Call processing techniques such as authentication, location update, handover, incoming call delivery and outgoing call routing have been discussed. Call set-up procedures for both mobile originated and terminated calls are presented. We have demonstrated the incorporation of mobility specific signalling into the ATM Q.2931 standard and the possible call handling procedures that can be used in the ATM-satellite network.

9.2 Major Contributions

To highlight the work accomplished, the major contributions are summarised as following:

- Proposal of a new and fully integrated ATM-satellite signalling protocol architecture that provides mobility for ATM users.

- Proposal of a service-adaptive MAC protocol. Validation of the application of demand assignment MAC scheme into signalling traffic and proposal of a semi-permanent signalling protocol.

- Proposal of a service-adaptive ARQ protocol. Proposal of a reliability-dependent SR ARQ protocol to support a variety of traffic types with different QoS objective, especially those delay-constrained services.

- Study for the first time the impact of distance changing end-to-end connections on performance of the Selective Repeat (SR) ARQ protocol, with particular reference to the connection routed via ISL of an LEO/MEO satellite constellation.

- Proposal of an Adaptive Timer SR ARQ (ATSR ARQ) protocol to be used on a distance-changing connection for improved throughput performance.
Conclusions and Future Work

- Proposal of a QoS-provisioning radio connection management scheme for ATM via satellite, taking into accounts the influence of all four factors: the fading channel, MAC protocol, ARQ protocol and resource allocation.

- Proposal and verification of a Reservation Meta-signalling (RMS) for setting up signalling connections at the air interface in an ATM-satellite system.

The proposed signalling protocol architecture provides a protocol reference model for ATM-satellite integrated systems. The verification and demonstration of the advantages of the semi-permanent signalling protocol, which offers a new method to improve the system channel efficiency on signalling connections, have been achieved. The proposed reliability-dependent SR ARQ protocol, which provides an enhanced and service-adaptive traffic transmission over satellite channels, provides a novel approach to optimise the transmission throughput to support a variety of traffic types with different QoS requirements in ATM-satellite systems.

The optimised MAC and ARQ protocol contribute to the efficient handling of the transportation of ATM traffic over satellites with user required QoS, and calls between the ATM backbone system and the mobile satellite system.

The proposed radio connection management scheme together with the optimised MAC and ARQ provides a framework of interworking protocols for ATM over satellites. It can also find applications in other systems that involve the integration of terrestrial protocols and mobile satellites.

The research work that has been accomplished herein provides a solution and guidance to the design of signalling protocols for mobile satellite systems to implement ATM technology or indeed other future protocols. It is perceived to be equally applicable to IP over integrated ATM-satellite networks.

9.3 Future Work

This research work has so far focused on the optimisation of the signalling and protocols for the point-to-point (PTP) connections. Further research work is expected to contribute to the optimisation of the point-to-multipoint (PTM) related protocols, such as PTM ARQ and PTM call control. In addition, other areas such as resource allocation and call admission control, which have not been comprehensively investigated in this work, require further research work as well. Drawing heavily upon the invaluable research conducted so far, including the work
Conclusions and Future Work

done in the ACTS projects, SECOMS and ASSET, we herein provide some suggestions for future work with respect to the above-mentioned aspects.

9.3.1 Performance of ATSR in the presence of channel fading

The performance of ATSR ARQ protocol on end-to-end connections established via ISL is studied assuming a reliable channel. However, the performance of the adaptive time SR ARQ on a changing-length connection in the presence of channel fading needs further comprehensive and in-depth study. It is suggested to be carried out as future work.

9.3.2 Point-to-multipoint ARQ scheme

Issues in PTM ARQ for ATM-satellite systems are different from those of PTP ARQ. They are more complex than those of PTP ARQ. This is mainly due to the following reasons:

1. PTM includes non-equal distances between senders and receivers. The delay between receivers and the root may be very different. This is different from a simple multicast problem. Most research work on PTM ARQ has assumed the simultaneous reception of packets from the root, however other non-simultaneous aspects should be appropriately addressed.

2. Retransmitted packets are sent over the multi-destination connection. If one receiving station suffers a relatively higher error rate than other stations, then the throughput of all the stations is essentially limited to the throughput achievable by that station. This may seriously degrade the real-time services.

Using the conventional GBN and SR ARQ in PTM will result in low throughput and long packet transport delays. In the fading channel, time and bandwidth are wasted on transmitting the same packets to all the receivers even though the retransmission is caused by only one receiver that fails to receive the packets correctly.

Unfortunately, a limited amount of work has been accomplished on the point-to-multipoint error control scheme [DEN93]. A review of relevant research work is given in [SAM96]. Little work has been done to apply or adapt the GBN or SR scheme to a PTM connection and evaluate the corresponding performance. However, most of this work assumes the same distance between transmitters and receivers, which may be invalid in the case of ATM-satellite PTM connections. Performance in fading conditions has not been investigated and schemes to overcome low throughputs need to be addressed.
Therefore, more work needs to be carried out to investigate the impacts of large numbers of terminals on the ARQ performance in a fading condition, and to enhance and optimise the ARQ scheme to improve the protocol throughput efficiency.

9.3.3 **Point-to-Multipoint call control signalling design**

As mentioned in chapter 8, there are two potential problems associated with mobile PTM call set-up procedures in satellite systems. Firstly, the call set-up delay could be unreasonably long due to the long propagation delay and many leaf parties involved. Therefore, it is necessary to investigate the “add party and leaf join” call set-up performance within the satellite environment. The call set-up delay and the channel efficiency should be evaluated in a fading channel and the leaf dropping probability should be addressed. For a non-real time PTM connection, the semi-permanent signalling protocol can also be used and its performance needs to be evaluated.

It is foreseen that PTM will suffer from performance degradation for the channel fading and that enhanced ARQ protocols will be required to combat such effects.

Secondly, because each leaf requires a separate connection session, the total bandwidth taken from the uplink/downlink could be large and inefficient. In order to save on the uplink bandwidth, it is proposed to share the bandwidth between all the connection sessions on the uplink between the root SAT/GTW and the satellite payload. Thus, further research work is required to design a signalling procedure that allows such bandwidth sharing on the uplink or downlink.

In any cases where the ATM Q.2931 protocols are found to be inefficient for the satellite system, we recommend to find an improved efficient solution for the set-up in point-to-multipoint connections rather than simply to adopt the ATM standard.

9.3.4 **Resource Management**

Resource management has two major aspects in the ATM-satellite system. The effective resource allocation for signalling and data, and the resource management for handovers.

As the concept of effective bandwidth is used in packet transmission over satellite, an accurate effective bandwidth allocation scheme is required to allocate a suitable bandwidth to users. The allocated bandwidth needs to be the minimum bandwidth that can just ensure the
users' QoS throughout the communications. Such a scheme should take into consideration fading channels.

Handover resource allocation is another challenging area in which further work needs to be done. When the request to implement a handover is received by an MCS, it will initialise the handover procedure in the satellite network. A handover resource allocation scheme should be able to ensure sufficient resources to support the connections available along the new route. These resources include radio interface bandwidth, buffers needed in the MCS and the types of connection identifiers to be used. If these resources are not available for a connection, the connection cannot be supported with the quality required. In this case a new route with more available resources should be identified. The resources for handovers can be allocated using a procedure similar to the resource allocation needed at the connection set-up. However, a suitable resource allocation scheme for mobile ATM-satellite needs more thorough investigation.
Conclusions and Future Work
1. A Reliability-Dependent ARQ Protocol for ATM Traffic Over Satellite Links
   Fan B, Tafazolli R, Evans BG
   To be submitted to IEE Electronics Letters.

2. Connection Management for Broadband Mobile Satellite Systems
   Fan B, Tafazolli R, Evans BG
   To be submitted to International Journal of Satellite Communications.

   Fan B, Tafazolli R, Evans BG

   Fan B, Tafazolli R, Evans BG
   Proceedings of 18th AIAA International Communication Satellite Systems Conference and Exhibit, April 10-14, 2000, Oakland, CA, U.S.A.

5. An On-board Resource Allocation Scheme for EuroSkyWay System in ASSET
   Fan B, Vahid S, Tafazolli R, Evans BG

6. A Connection Management Scheme for Broadband Satellite Systems
   Fan B, Vahid S, Tafazolli R, Evans BG

7. Application of IN Framework to the Design of Mobility Management in ASSET
   Vahid S, Fan B, Tafazolli R, Evans BG

8. Adaptive Timer Selective Repeat ARQ Protocol for Dynamic Satellite Constellation with Intersatellite Links
   Fan B, Tafazolli R, Evans BG

   Fan B, Tafazolli R, Evans BG
   Proceedings of the 3rd European Workshop on Mobile/Personal Satcoms (EMPS'98), Venice, Italy, pp. 121-130, November 4-5, 1998.

10. A Fully Integrated Air Interface Signalling and Protocol for ATM via Satellite
    Fan B, Tafazolli R, Evans BG
11. SECOMS Interworking Scenario and Interconnection With B-ISDN
   Mertzanis I, Fan B, Deli Priscolli F, Marziale V, Podda V, Tafazolli R.
   Proceedings of the 3rd ACTS Mobile Communications Summit’98, Rhodes, Greece, Vol.2,

12. Signalling and Protocol Schemes for SECOMS the ACTS European Project
   Mertzanis I, Fan B, Tafazolli R, Evans BG.
   Proceeding of the 4th European Conference on Satellite Communications (ECSC-4), Rome,
   Italy, pp 216-221, November 18-20, 1997.

TECHNICAL REPORTS

1. Modelling and Performance Evaluation Scheme for SECOMS Signalling and Protocols

2. SECOMS Signalling and Protocol System Performance Optimisation

3. Connection Management for ASSET Multimedia Satellite Communication System

4. EuroSkyWay Connection Management Scheme Consolidation
## System design parameters of some proposed mobile satellite systems

The recent satellite systems, including Iridium, GlobalStar, ICO, Odyssey and Teledesic, are intended to offer a variety of services covering voice, data, paging and fax. The major system design parameters including the constellation parameters are given in the following table.

<table>
<thead>
<tr>
<th>Uplink parameters (Earth-to-space)</th>
<th>LEO</th>
<th>MEO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iridium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Globalstar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Odyssey</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICO</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Orbit

<table>
<thead>
<tr>
<th></th>
<th>LEO</th>
<th>MEO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude (km)</td>
<td>780</td>
<td>1414</td>
</tr>
<tr>
<td></td>
<td>10354</td>
<td>10355</td>
</tr>
<tr>
<td>Satellite separation (degrees)</td>
<td>32.7</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>72</td>
</tr>
<tr>
<td>Number of satellites</td>
<td>66</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Orbital planes</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Inclination angle (degrees)</td>
<td>86</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>45</td>
</tr>
<tr>
<td>Orbital period (min)</td>
<td>100</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>359.5</td>
<td>358.9</td>
</tr>
<tr>
<td>Satellite max. Visibility time (min)</td>
<td>11.1</td>
<td>16.4</td>
</tr>
<tr>
<td></td>
<td>94.5</td>
<td>115.6</td>
</tr>
</tbody>
</table>

### Frequency bands, GHz

<table>
<thead>
<tr>
<th>Mobile-to-satellite, service</th>
<th>GHz</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Uplink frequency</td>
<td>1.616 - 1.6265</td>
<td>1.610 - 1.6265</td>
<td>1.610 - 1.6265</td>
<td>1.98 - 2.01</td>
</tr>
<tr>
<td>Downlink frequency</td>
<td>1.616 - 1.6265</td>
<td>2.4835 - 2.50</td>
<td>2.4835 - 2.50</td>
<td>2.17 - 2.20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Earth station-to-satellite, service</th>
<th>GHz</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Uplink frequency</td>
<td>27.5 - 30</td>
<td>5.09 - 5.25</td>
<td>29.1 - 29.4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Downlink frequency, GHz</td>
<td>6.875 - 7.055</td>
<td>19.3 - 19.6</td>
<td>7</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------------------</td>
<td>---------------</td>
<td>-------------</td>
<td>---</td>
</tr>
<tr>
<td>Intersatellite crosslink, GHz</td>
<td>23.18 - 23.38</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

**Satellite characteristics**

<table>
<thead>
<tr>
<th></th>
<th>48</th>
<th>16</th>
<th>61</th>
<th>163</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of spot beams per satellite, for the service links</td>
<td>12</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Frequency reuse, cells per cluster</td>
<td>1.8x10^5 to 7x10^5</td>
<td>6.3x10^5 to 2.3x10^6</td>
<td>9.7x10^5 (6.3°)</td>
<td>5x10^5 to 2x10^6</td>
</tr>
<tr>
<td>Beam size, km^2</td>
<td>-20</td>
<td>-15</td>
<td>-20</td>
<td>-20 (peak)</td>
</tr>
<tr>
<td>Average beam sidelobes, dB</td>
<td>31.5</td>
<td>30</td>
<td>37.9</td>
<td>34</td>
</tr>
<tr>
<td>Total satellite output power, dBW</td>
<td>17 to 25</td>
<td>N/a</td>
<td>24 to 28</td>
<td>30</td>
</tr>
<tr>
<td>Average gain / beam, dBi</td>
<td>700</td>
<td>450</td>
<td>2207</td>
<td>1925</td>
</tr>
<tr>
<td>Minimum elevation angle, (degrees)</td>
<td>8.3</td>
<td>10</td>
<td>20</td>
<td>10</td>
</tr>
</tbody>
</table>

**Transmission parameters**

<table>
<thead>
<tr>
<th></th>
<th>2.4 / 4.8</th>
<th>2.4/4.8/9.6</th>
<th>4.8</th>
<th>4.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice telephony, kb/s</td>
<td>2.4</td>
<td>7.2</td>
<td>2.4 - 9.6</td>
<td>2.4</td>
</tr>
<tr>
<td>Data for BER &lt; 10^-3, kb/s</td>
<td>QPSK</td>
<td>QPSK</td>
<td>QPSK</td>
<td>QPSK</td>
</tr>
<tr>
<td>Modulation</td>
<td>FEC</td>
<td>FEC</td>
<td>FEC</td>
<td>FEC</td>
</tr>
<tr>
<td>Coding</td>
<td>FDMA/TDMA</td>
<td>FDMA/CDMA</td>
<td>FDMA/CDMA</td>
<td>FDMA/TDMA</td>
</tr>
<tr>
<td>Access scheme</td>
<td>TDD</td>
<td>FDD</td>
<td>FDD</td>
<td>FDD</td>
</tr>
<tr>
<td>Duplex scheme</td>
<td>90</td>
<td>N/a</td>
<td>N/a</td>
<td>40</td>
</tr>
<tr>
<td>Frame length, ms</td>
<td>50</td>
<td>N/a</td>
<td>N/a</td>
<td>36</td>
</tr>
<tr>
<td>Burst rate, kb/s</td>
<td>N/a</td>
<td>N/a</td>
<td>N/a</td>
<td>36</td>
</tr>
</tbody>
</table>
Appendix B

Logic Channels

The definition and description of the logic signalling channels required at the air interface of the integrated ATM-satellite system are as follows and their applications and multiplexing mode are summarised in the following table. These channels are defined based on the operation requirements of the proposed ATM-satellite signalling system and logic signalling channels of GSM.

<table>
<thead>
<tr>
<th>Channel name</th>
<th>Description</th>
<th>Application</th>
<th>From</th>
<th>To</th>
<th>Multiplexing</th>
</tr>
</thead>
<tbody>
<tr>
<td>USC</td>
<td>Uplink Synchronisation Channel</td>
<td>Synchronisation</td>
<td>SAT</td>
<td>P/L</td>
<td>Slotted Aloha</td>
</tr>
<tr>
<td>DSC</td>
<td>Downlink Synchronisation Channel</td>
<td>Synchronisation</td>
<td>P/L</td>
<td>SAT</td>
<td>TDM</td>
</tr>
<tr>
<td>RRC</td>
<td>Resources Request Channel</td>
<td>Resource Reservation</td>
<td>SAT</td>
<td>TRM</td>
<td>TDM</td>
</tr>
<tr>
<td>RAC</td>
<td>Resources Assignment Channel</td>
<td>Resource Assignment</td>
<td>TRM</td>
<td>SAT</td>
<td>TDM</td>
</tr>
<tr>
<td>FC</td>
<td>Feedback Channel</td>
<td>Feedback Information</td>
<td>TRM</td>
<td>SATs</td>
<td>TDM</td>
</tr>
<tr>
<td>BC</td>
<td>Broadcast Channel</td>
<td>Broadcast Channel</td>
<td>MCS</td>
<td>SATs</td>
<td>TDM</td>
</tr>
<tr>
<td>CSC (SAT)</td>
<td>Common Signalling Channel</td>
<td>Common Signalling</td>
<td>MCS</td>
<td>SATs</td>
<td>TDM</td>
</tr>
<tr>
<td>DMSC</td>
<td>Dedicated Meta-Signalling Channel</td>
<td>Meta-signalling</td>
<td>SAT</td>
<td>MCS</td>
<td>DAMA</td>
</tr>
<tr>
<td>DSCC (SAT)</td>
<td>Dedicated Signalling Control Channel</td>
<td>Dedicated Signalling</td>
<td>SAT</td>
<td>MCS</td>
<td>SAT</td>
</tr>
<tr>
<td>PGC</td>
<td>Paging Channel</td>
<td>Paging</td>
<td>MCS</td>
<td>SAT</td>
<td>TDM</td>
</tr>
<tr>
<td>SASC</td>
<td>Slow Associated Signalling Channel</td>
<td>Associated Signalling</td>
<td>SAT</td>
<td>GTW</td>
<td>SAT</td>
</tr>
<tr>
<td>TCH</td>
<td>Traffic Channel</td>
<td>Traffic Channel</td>
<td>SAT</td>
<td>GTW</td>
<td>SAT</td>
</tr>
<tr>
<td>FSCC</td>
<td>Forward Satellite Control Channel</td>
<td>Resource Management</td>
<td>MCS</td>
<td>TRM</td>
<td>TDMA</td>
</tr>
<tr>
<td>RSCC</td>
<td>Return Satellite Control Channel</td>
<td>Resource Management</td>
<td>TRM</td>
<td>MCS</td>
<td>TDM</td>
</tr>
<tr>
<td>DSCC (GTW)</td>
<td>Dedicated Signalling Control Channel</td>
<td>Dedicated Signalling</td>
<td>MCS</td>
<td>GTW</td>
<td>TDM</td>
</tr>
<tr>
<td>CSC (GTW)</td>
<td>Common Signalling Channel</td>
<td>Common Signalling</td>
<td>MCS</td>
<td>GTW</td>
<td>TDM</td>
</tr>
<tr>
<td>GRRC</td>
<td>GTW Resource Request Channel</td>
<td>Resource Reservation</td>
<td>GTW</td>
<td>TRM</td>
<td>Slotted Aloha</td>
</tr>
<tr>
<td>GRAC</td>
<td>GTW Resource Assignment Channel</td>
<td>Resource Assignment</td>
<td>TRM</td>
<td>GTW</td>
<td>TDM</td>
</tr>
</tbody>
</table>

Logic signalling channel summary.
1. Synchronisation Channels

**USC (Uplink Synchronisation Channel)** - This channel enables the user terminals to perform an initial acquisition to get synchronised to the network. The adopted access scheme is slotted ALOHA for the user terminals.

**DSC (Downlink Synchronisation Channel)** - This channel contains the information used by each terminal to synchronise itself to the downlink reception by recognising the start of each frame. It contains the DLUW, the frame counter, the satellite position, the spot beam ID, the carrier identifier and carrier configuration. The adopted access scheme is fixed.

2. Multiple Access

**RRC (Resources Request Channel)** - This channel enables the user terminals to request capacity by a means of an out-of-band reservation with the TRM. The adopted access scheme is slotted ALOHA for the user terminals.

**RAC (Resources Assignment Channel)** - This channel contains the resource assignment to the terminal. The information includes the assigned carrier frequency error, number of timeslots, holding time and channel positions. The adopted access scheme is fixed.

**FC (Feedback Channel)** - This channel contains the access results of the corresponding uplink RRC request and network access control information. The adopted access scheme is fixed.

**BC (Broadcast Channel)** - This channel contains the network announcement message which should be listened to by all the terminals. The adopted access scheme is fixed. The adopted multiplexing scheme is TDM (Time Division Multiplexing).

**CSC (Common Signalling Channel)** - This channel supports general network control related signalling procedures. It enables the MCS to send the signalling messages to the terminals. The adopted multiplexing scheme is TDM (Time Division Multiplexing).

3. DLC

**DMSC (Dedicated Meta-Signalling Channel)** - This channel supports metasignalling. It enables the dialogue between the SAT and the MCS or SAT to the GTW (from the GTW to the MCS). The adopted access scheme is assigned TDMA (Time Division Multiple Access).
DSCC (Dedicated Signalling Control Channel) - This channel supports authentication and connection control procedures; it enables the dialogue between the SAT and the MCS or SAT to the GTW (from the GTW to the MCS). The adopted access scheme is assigned TDMA (Time Division Multiple Access).

4. Radio Connection Control

PGC (Paging Channel) - This channel contains the paging message to call for a terminal to access network. The information includes the terminal ID, the paging reason and the calling user ID. The channel is from MCS to SAT and it is a TDM multiplexing channel.

SASC (Slow Associated Signalling Channel) - This channel is a low rate signalling channel associated with a dedicated TCH channel. This channel is used to convey some information regarding the channel measurement, handover reports (from SAT to MCS or from SAT to GTW). This channel adopts TDMA multiplexing.

5. Call Control

TCH (Traffic Channel) - This channel supports traffic transportation between SAT and GTW or between SAT and MCS. It enables the dialogue between the SAT and the MCS or SAT to the GTW (from the GTW to the MCS). The adopted access scheme is assigned TDMA (Time Division Multiple Access).

6. Channels between MCS - TRM

FSCC (Forward Satellite Control Channel) - This channel supports traffic resource management and traffic control related signalling. It enables the MCS to send the signalling messages to TRM (from the MCS to the TRM). The adopted access scheme is TDMA (Time Division Multiple Access).

RSCE (Return Satellite Control Channel) - This channel supports traffic resource management and traffic control related signalling. It enables the TRM to send the signalling messages to MCS (from the TRM to the MCS). The adopted multiplexing scheme is TDM (Time Division Multiplexing).

7. Channels between MCS - GTW

DSCC (Dedicated Signalling Control Channel) - This channel supports connection control procedures for each connection; it enables the dialogue between the GTW and the MCS (from
the GTW to the MCS) in order to set up a traffic connection. The adopted access scheme is TDMA (Time Division Multiple Access).

*CSC (Common Signalling Channel)* - This channel supports the generally signalling communication between MCS and GTW. It is a fixed channel and its multiplexing scheme is TDM (Time Division Multiplexing).

8 Channels between GTW – TRM

*GRRC (GTW Resource Request Channel)* - This channel supports traffic resources request from GTW to TRM. It enables the GTW to request extra bandwidth in the case that a buffer overflows on a connection (from the TRM to the GTW). The adopted multiplexing scheme is Slotted Aloha (Time Division Multiplexing).

*GRAC (GTW Resource Assignment Channel)* - This channel supports traffic resources assignment procedures. It enables the TRM to send the detailed resource assignment message to GTW (from the TRM to the GTW). The adopted multiplexing scheme is TDM (Time Division Multiplexing).
References


References


References


References


[SUR96] Software Package, “SPOC+ (Simulation Package of Orbital Constellation + Network Analysis)”, Developed by University of Surrey, 1996.


