

Report for the degree of PhD

**RESOURCE ALLOCATION IN ATM
NETWORKS**

Author: John Murphy, B.E., M.S.

Supervisor: Dr. Thomas Curran

Dublin City University

School of Electronic Engineering

March 1996

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD is entirely my own work and has not been taken from the work of others save to the extent that such work has been cited and acknowledged within the text of my work.

Signed: John Murphy ID No.: 91700914

John Murphy

Date: 6 March 1996

dedicated to Ed Posner

Acknowledgements

I thank Dr. Tommy Curran, my supervisor, for introducing me to this area, for his guidance and advice on both technical and other aspects of this work and for his support even through my wanderings. I also thank Prof. Charles McCorkell, for initially giving me the opportunity to return to college, for his trust in me over the last few years, and for his support and encouragement of this work and of my work in DCU. I benefited from all the students who took my classes in DCU, especially Jerry Teahan and Jenny Murphy who undertook final year projects for me, and the research students there, especially Sean Murphy.

My interest in queueing theory began when I took a class from the late Prof. E. C. Posner at Caltech, long before I started on this project. Many thanks are due to Prof. Posner for not only did he stimulate my interest in this area from his teaching and seminars, but he gave me the opportunity to return to work with him and introduce me to JPL. The tragedy of his untimely death leaves this work and my interest in this area incomplete.

I am grateful to a number of people at JPL for giving me the opportunity to work there and introducing me to new areas : Dr. Richard Markley initially made it possible to work there ; Dr. Edward Chow gave me his encouragement, enthusiasm and support while there ; Dr. Ed Upchurch and Dr. Julia George for their assistance on simulation and modelling; Dr. Michael Chelian for characterising the satellite links ; Dr. John Peterson for his support in JPL ; Dr. Timothy Hanson for the work that we have done at JPL, the work in ESIGETEL and the joint work since then.

I am grateful for the assistance of Prof. R. J. McEliece in facilitating my returns to Caltech ; Zhong Yu and Bahadir Erimli for the joint work that we did, both at Caltech and since then. Dr. Michael Mandell has given interesting and useful advice, both technically and otherwise over the last number of years, and in particular his assistance while I studied at Caltech. I have many reasons to thank Dr. Michael Lough for making my trips to Caltech pleasurable which will not be noted here.

I thank Prof. Jeffrey MacKie-Mason, University of Michigan, for all the work and insight on the pricing of networks. For the assistance and support of the simulation package I have used, I thank Peter Colaluca, SES/*europa*. I thank David Condell who has supported the computer systems in DCU, and for answering all my queries.

It would not have been possible to go to college without the support and encouragement of my parents. They have given me much advice and other assistance, and have made a lot of sacrifices, which has made studying possible over the years. The person who has given me the most confidence and encouragement to pursue my studies has been my brother. He has been my co-worker, mentor and advisor who has guided me in many directions, even from the end of a computer. From my hardest days and nights studying for my masters in Caltech, to making it possible to return to Caltech with Prof. Posner, I have many reasons to thank him.

Contents

Acknowledgements	iii
Abstract	viii
List of Symbols	ix
Glossary of Acronyms	xii
1 Introduction	1
1.1 The Need For ATM	1
1.2 ATM Standards	4
1.2.1 Virtual Paths & Virtual Channels	5
1.2.2 The Physical Layer	7
1.2.3 The ATM Layer	9
1.2.4 The ATM Adaptation Layer	10
1.2.5 ATM Services	12
1.3 Guarantees In ATM Networks	13
1.3.1 Quality Of Service Definitions	13
1.3.2 Contract Between User & Network	14
1.4 The Network & User Perspectives	15
1.4.1 The User Perspective	15
1.4.2 The Network Perspective	17
1.4.3 Future Integrated-Services Networks	19
1.5 Resource Allocation Of ATM Networks	21
1.5.1 Time Scales	22
1.5.2 Statistical Multiplexing	23
1.5.3 Effective Bandwidths	24
1.5.4 Current Resource Allocation Techniques	26
1.5.5 Will Resource Allocation Continue To Be Important ?	27
1.6 A New Model Of ATM Networking	28
1.7 Objectives of Thesis	30

Contents

1.7.1	Objectives of Chapter 2	30
1.7.2	Objectives of Chapter 3	31
1.7.3	Objectives of Chapter 4	32
1.8	Summary	32
1.8.1	Summary of Thesis	32
2	Performance Modelling & Simulation	34
2.1	Introduction	34
2.2	Overview Of Modelling	35
2.3	Source Modelling	37
2.3.1	Source Analysis	38
2.3.2	Voice Source Model	43
2.3.3	Video Source Model	45
2.3.4	Data Source Model	49
2.4	MPEG Models Of Video Sources	49
2.4.1	Overview Of MPEG	50
2.4.2	Hierarchical Models Of MPEG Video Source	51
2.4.3	Comparison Of MPEG To H.261 Models	52
2.5	Simulation Techniques	53
2.5.1	SES/ <i>workbench</i>	54
2.5.2	Source Simulation Model Design	58
2.5.2.1	Voice Source Simulation Model Design	58
2.5.2.2	One Layer Video Source Simulation Model Design	58
2.5.2.3	Two Layer Video Source Simulation Model Design	59
2.5.2.4	Data Source Simulation Model Design	60
2.5.3	Source Simulation Model Validation	60
2.6	High-Speed Simulation Problems	62
2.6.1	Decomposition Methods	64
2.6.2	Cell-Rate Methods	65
2.6.3	Step-By-Step Methods	66
2.7	Discussion	67
3	Cell Level Resource Allocation	68
3.1	Introduction	68
3.2	Motivation For ATM Over Satellite	70
3.2.1	JPL's Deep Space Network	71

Contents

3.2.2	Problems Using Satellite Links For ATM	73
3.3	Measuring & Modelling A Satellite Link	76
3.3.1	Experimental Test Set-up	76
3.3.2	Experimental Data For Error Distribution	77
3.3.3	Model Of A Satellite Link	79
3.4	ATM Adaptation Layer Over Satellite	82
3.4.1	Problems With Current ATM Adaptation Layer	82
3.4.2	Proposed ATM Adaptation Layer	83
3.4.3	Service-Based Retransmission Scheme	85
3.5	Simulation & Results For Satellite ATM	88
3.5.1	Simulation Details	88
3.5.2	Simulation Results	90
3.5.3	Result Validation & Comparison To The Theory	93
3.6	Cell Level Congestion & Traffic Control	95
3.6.1	Traffic Contract & Usage Parameter Control	95
3.6.2	Basic ATM Model	96
3.6.3	Conformance Definition – Leaky Bucket	97
3.6.4	Leaky Bucket As The UPC	99
3.7	Issues For Worst Case Traffic	101
3.7.1	Two Potential Worst Traffic Types	102
3.7.2	Finite & Infinite Buffers	103
3.7.3	Continuous & Discrete Variables	106
3.8	New Approach To Cell Loss For Worst Case Analysis	108
3.9	Counter Examples To Traditional Worst Case Traffic	110
3.9.1	Fast Serving Server	110
3.9.2	Slow Serving Server	112
3.10	Discussion	114
4	Pricing In ATM Networks	116
4.1	Introduction	116
4.2	Adaptive Connections	120
4.2.1	Defining Services In An ATM Network	120
4.2.2	Adaptive User Behaviour	122
4.2.3	Possible Types Of User	125
4.2.4	Feedback-Based Fast Reservation	129

Contents

4.3	Motivation For Pricing	130
4.3.1	Pricing Bandwidth	131
4.3.2	Price As A Feedback Signal	133
4.3.3	Price As A Component Of Charging	134
4.4	Mathematical Model	135
4.4.1	User Benefit Functions	136
4.4.2	System Problem	137
4.4.3	Distributed Pricing Algorithm	139
4.5	Simulation & Results	140
4.5.1	Simulation Model	140
4.5.2	Simulation Environment	140
4.5.3	Results	143
4.6	Implementation Of A Pricing Scheme	146
4.6.1	Generation Of The Price	146
4.6.2	Price Passing From Network To User	147
4.6.3	User's Usage Of Price	148
4.7	Different Types Of Efficiencies In Networks	149
4.7.1	Economic Framework	150
4.7.2	Users Models	151
4.7.2.1	Elastic User	152
4.7.2.2	Inelastic User	155
4.7.3	Another Distributed Pricing Algorithm	156
4.7.4	Simulation Details	156
4.7.5	Results	160
4.8	How Pricing Fits Into The ATM Controls	160
4.9	Discussion	162
5	Conclusions	165
	References	168
	Publications Arising From This Work	177

Resource Allocation In ATM Networks

Author: John Murphy

Abstract

The areas of resource allocation and congestion control in ATM networks have been investigated. ATM networks and the guarantees given to users have been reviewed and a new model of ATM networking has been proposed. To aid the analysis of ATM network issues, performance modelling and simulation methods have been reviewed. Typical sources have been designed : a two-state Markov model for voice ; a multi-state Markov one layer variable bit rate video source model ; an empirical file transfer data source model ; and some basic network elements. The models have been verified and validated on a discrete event simulator.

It was shown that there are problems when using ATM over satellite links. A model for the noise analysed from real satellite links was developed. Based on this model a new more efficient protocol for assembling ATM cells was proposed and simulated. Again at the cell level, the traffic that can pass the standardised conformance test and still produce the worst performance in the network was investigated. Counter to the traditional wisdom it was found that the on-off source does not always produce the worst case traffic.

Users have been classified with new parameters, and it has been shown that these new classes of users can still be given guarantees without giving traffic descriptors. Adaptive user classes have been modelled mathematically. A new model for efficiency has been developed, which includes both network issues and economic issues. This new model defines congestion and also describes how to allocate resources when congested. It has been shown that this economic model coupled with the adaptive user classes allow for an increase in both network and economic efficiency simultaneously for some sample cases.

List of Symbols

Notation	–	Definition
A	–	Price charged for connection set-up
B	–	Price charged for unit time
B	–	Number of place in a buffer
$B(F_j)$	–	Effective bandwidth of source j
B_q	–	Buffer size for virtual path connection q
$Bufcost_q$	–	Barrier function for buffer on virtual path connection q
C	–	Price charged per cell
C	–	Capacity of link
$CL(x)$	–	Cell loss when source x cells out of phase
F_j	–	Distribution function for source j
I	–	Increment to leaky bucket for conforming cell
J	–	Number of sources
K	–	Linear constraint on effective bandwidths
L	–	Limit on the value of the leaky bucket
M	–	Leaky bucket token capacity
N	–	Number of cells to be sent
P_t	–	Price in interval t
P^*	–	Users guess for average price
PCR_t	–	PCR assigned at time t
PCR^*	–	Users guess for average PCR

List of Symbols

Notation	–	Definition
Q	–	Buffer occupancy
R	–	Leak rate of the leaky bucket
T	–	Time interval for pricing updates
T	–	Period of the greedy on–off source
T_{on}	–	Time for the on period for peak cell emission
T_{off}	–	Time for the off period
T^*	–	Intermediate period in the three state source
V_q	–	Virtual path connection q
X	–	Number of cells in file for elastic user
X	–	Leaky bucket counter value
X^*	–	Auxiliary leaky bucket counter value
X_j	–	Load produced by source j
$X_R(t)$	–	Number of elastic user cells remaining at time t
$already-charged(t)$	–	Amount paid by user so far at time t
b_q	–	Buffer occupancy for virtual path q
b_{rq}	–	Bandwidth required for connection r on virtual path connection q
ben_{rq}	–	Benefit function for connection r on virtual path connection q
$ben\{x, v\}$	–	User benefit for file size x with delay d
$benefit$	–	Benefit user gets by sending cells
d	–	Delay of sending user file
$h(d)$	–	Penalty paid for delay of d of lost user benefit

List of Symbols

Notation	–	Definition
m	–	Mean of hyper-exponential distribution
m_1	–	Mean of lower exponential distribution
m_2	–	Mean of upper exponential distribution
p	–	Peak cell emission rate
p	–	Probability of being in lower exponential distribution
p_{ij}	–	Transition probability between state i and state j
r	–	Rate of serving cells
s	–	Standard deviation of hyper-exponential distribution
$t_a(k)$	–	Arrival time of cell k
t_{int}	–	Length of pricing interval
v	–	User benefit with no delay
x	–	Size of user file
x	–	Number of cells one source shifted from another
π	–	Congestion charge
π_q	–	Price per unit bandwidth on virtual path connection q
α	–	Forecasted number of intervals to complete transmission if nothing is sent in this interval
β	–	Forecasted number of intervals to complete transmission if send cells in this interval
γ	–	Loss probability index

Glossary of Acronyms

Acronyms	-	Explanation
AAL	-	ATM adaptation layer
ABR	-	Available bit rate
AOS	-	Advanced orbiting systems
ARQ	-	Automatic repeat request
ATM	-	Asynchronous transfer mode
BISDN	-	Broadband ISDN
CBR	-	Constant bit rate
CAC	-	Connection admission control
CCITT	-	Comite Consultatif International de Telecommunications et Telegraphy
CCSDS	-	Consultative committee for space data systems
CIF	-	Common image format
CLP	-	Cell loss priority
CLR	-	Cell loss ratio
CRC	-	Cyclic redundancy check
CS	-	Convergence sublayer
DCT	-	Discrete cosine transform
DSN	-	Deep space network
EFCI	-	Explicit forward congestion indication
EPRCA	-	Enhanced proportional rate control algorithm
FDDI	-	Fiber distributed data interface
FIFO	-	First in first out

Glossary of Acronyms

Acronyms	–	Explanation
GCF	–	Ground communications facility
GCRA	–	Generic cell rate algorithm
GFC	–	Generic flow control
GOP	–	Group of pictures
GSFC	–	Goddard space flight center
HEC	–	Header error control
ISDN	–	Integrated services digital network
ITU	–	International Telecommunications Union
LAN	–	Local area network
LCT	–	Last compliance time
LI	–	Length indicator
MAN	–	Metropolitan area network
MID	–	Multiplexing identifier
MPEG	–	Moving picture expert group
NNI	–	Network network interface
NRM	–	Network resource management
OSI	–	Open systems interconnection
PCR	–	Peak cell rate
PDF	–	Probability density function
PDU	–	Protocol data units
PM	–	Physical medium
PT	–	Payload type
QOS	–	Quality of service
RES	–	Reserved
SAR	–	Segmentation and reassembly
SDH	–	Synchronous digital hierarchy
SDU	–	Service data unit

Glossary of Acronyms

Acronyms	–	Explanation
SN	–	Sequence number
ST	–	Segment type
TC	–	Transmission convergence
TCP/IP	–	Transport control protocol / Internet protocol
UBR	–	Unspecified bit rate
UNI	–	User network interface
UPC	–	Usage parameter control
VBR	–	Variable bit rate
VC	–	Virtual channel
VCC	–	Virtual channel connection
VCI	–	Virtual channel identifier
VCR	–	Video cassette recorder
VP	–	Virtual path
VPC	–	Virtual path connection
VPI	–	Virtual path identifier
VPN	–	Virtual private network
VTC	–	Video traffic characterisation
WAN	–	Wide area network

Chapter 1

Introduction

1.1 The Need For ATM

ATM is the basis of broadband networks of the future and it will be able to carry any service, regardless of the characteristics of that service. The characteristics of a service might include the following: the bit rate needed; the time delay constraint; the cell loss constraint; the holding time. This attribute of a service independent network will future-proof ATM as the transfer mechanism for the broadband-ISDN. There is an obvious gain in designing, constructing, operating, maintaining and using a single network with all services being carried on it as compared to multiple separate networks specialised to a particular service. This efficiency is gained by having all the services using it, and so the efficiency is across the services, and no single service might be more efficient in ATM than in a specialised network. However there may be problems in trying to get all services to share a network. It may not be possible to get the same efficiency for voice over ATM as it would be on a network that just carried voice traffic. The gain made in having many services use the same network is hoped to outweigh the disadvantage of having to design the network for all services. ATM is thus a future-proof, service independent, service efficient single network.

Previous networks have been designed with mainly one service type in mind. The telephone network was designed to give guarantees on voice calls, which are a real time service which was implemented by using a constant bit rate service. While

Chapter 1 Introduction

this network is efficient at transporting voice, it is inflexible for carrying other services, for example data services. On the other hand the computer industry has a number of data networks, each of which has been designed to carry data traffic of one sort or another. This traffic may not get guarantees from the network in terms of delay or loss and may rely on the end-users to retransmit using Automatic Repeat Request (ARQ) methods. However these networks are generally non-real time networks and thus can not carry real time services like voice very well.

An ATM network will have to cater for all services, even ones that have not been planned as yet. The first criterion therefore is that it must carry the services that are known about and that exist in networks at present. These types of service range from real time voice and video, with varying quality, to telemetry data and high speed data communications. The holding times are illustrated, which give an indication of the duration of the call, along with the bit rate required for the call in Figure 1.1.

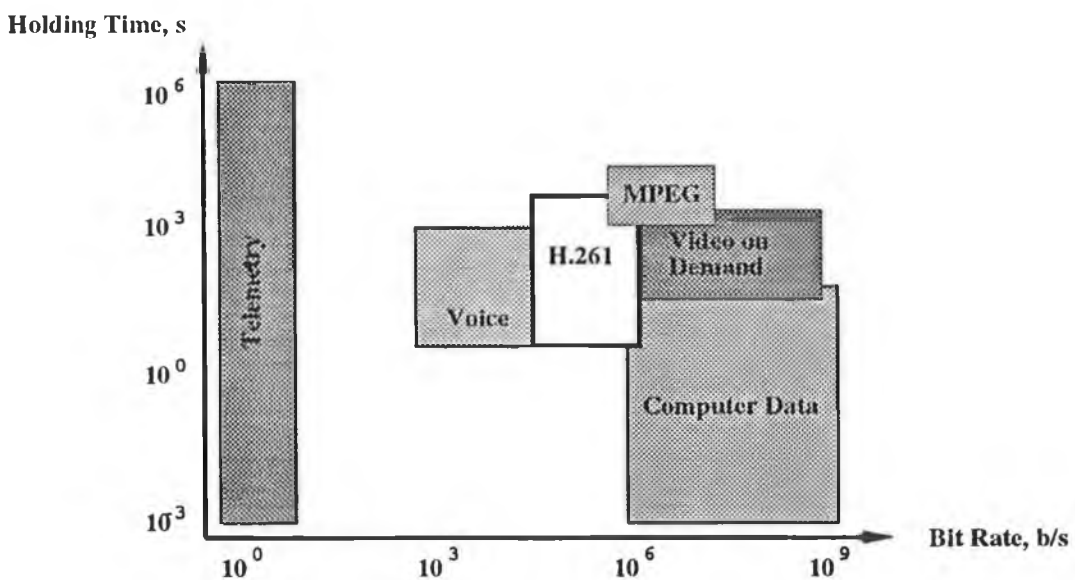


Figure 1.1: Expected Services

What is seen is that the possible services cover a huge range of possibilities, with orders of magnitude difference between them. ATM must be able to cater for all of these services and this has implications for it's design. This is accomplished by having a high speed packet switching network, that is connection oriented, with reduced functionality in the network. Therefore ATM will transport cells across

Chapter 1 Introduction

the network fast and in real time, while giving guarantees.

Because of the uncertainty in the types and characteristics of the services being carried on an ATM network, the idea of carrying the natural bit rate, or the information content, of a source has arisen. What this means is that it will no longer be necessary to change a variable bit rate source to be a constant bit rate source in order to carry it and give it guarantees of delay and loss. A good example of this is video, where at present in most video networks the video sequence is made to be constant bit rate, where the bit rate needed is the maximum bit rate needed by the source for a given quality of service. However the bit rate of the source may in fact be less than this bit rate for long periods of time, and by allowing variable bit rates the possibility of coding the source to try and get the natural bit rate or the information rate is possible, as is shown in Figure 1.2.



Figure 1.2: Natural Bit Rate Of Sources

When there are these types of sources present, then there is the possibility of multiplexing them together, and not assigning the maximum bit rate to them, in the hope that not all of them will be at their maximum at the same time. This multiplexing idea is called statistical multiplexing and it relies on the statistics of the source being available so that guarantees can be given in terms of loss or delay in a probabilistic sense. There are now two ways to lose cells in the network :

- the transmission loss, mainly due to bit error rates on the links

- the overflow loss, mainly due to the statistical nature of the multiplexing used

While it has been shown that there is a need for ATM networks, the most important fact is that they are also possible to construct. This possibility relies on a number of technologies which include high speed, low delay, low error rate, fiber optics and high speed switching techniques that can be achieved by having a small fixed sized packet.

1.2 ATM Standards

ATM networks are being specified by a number of bodies, and the two most important are the ITU, formerly the CCITT, and the ATM Forum. Both of these specifications are similar and the approach to standardising is also similar. Most standards for data networks have used the idea of divide and conquer, and this is done by layering the architecture, as is shown in Figure 1.3. Each layer in the standard is a sub-problem and is independent from the others, except by direct links at the boundaries. The basic structure for layering for data networks has been the Open Systems Interconnection, (OSI), reference model. This defines seven layers where the bottom three have to do with data getting through the network correctly and the top three layers are to do with understanding the data. The middle layer is the interface between the two sets. The bottom layer is called the physical layer and is concerned with the electrical and mechanical characteristics of the signal being carried. The next layer up is called the data link layer and is concerned with the correct transmission of the bits between two nodes directly connected. This is usually for a single link and is concerned with ARQ techniques and flow control. The next layer is called the network layer and is concerned with the correct transmission across the whole network, and so is concerned with routing and flow control. The transport layer is the fourth layer and is concerned with giving the service the required quality across the whole of the network. Each layer uses the functions of the layer directly beneath it to achieve its own goal.

ATM layering does not really follow the OSI model, but some comparisons can be made. The ATM model consists of three layers, the Physical layer, the ATM

OSI	ATM		
Transport / Network	AAL	CS	Convergence
		SAR	Segmentation & reassembly
Network / Data link	ATM		Generic Flow Control
			Cell VPI / VCI translation
			Cell multiplex and demultiplex
Physical	PHY	TC	Cell rate decoupling
			HEC header sequence generation / verification
			Cell delineation
			Transmission frame adaptation
			Transmission frame generation / recovery
	PM		Bit timing
			Physical medium

CS – Convergence Sublayer, SAR – Segmentation And Reassembly Sublayer, TC – Transmission Convergence Sublayer, PM – Physical Medium Sublayer, HEC – Header Error Control

Figure 1.3: OSI & ATM User Plane Equivalents

layer and the ATM Adaptation Layer (AAL). One possible set of relationships between the two is shown in Figure 1.3. The definition of the classes of services that is used in the ATM standards and the ITU are similar.

The basic attributes of ATM are :

- connection-oriented (Sub-Section 1.2.1)
- very small errors on links (Sub-Section 1.2.2)
- reduced header functions (Sub-Section 1.2.3)
- service independent network (Sub-Sections 1.2.4 & 1.2.5)

1.2.1 Virtual Paths & Virtual Channels

ATM networks are connection oriented networks and therefore it is possible for each connection to have a route set-up at the start of the connection. This route

Chapter 1 Introduction

will remain the same for the duration of the connection to ensure cell sequence at the receiver. The cell must contain the connection identifier within itself that uniquely identifies the connection throughout the network. Rather than have a single identifier, two are used, and a hierarchical approach is taken to the identification. A Virtual Path (VP) is the generic name for a collection of Virtual Channel (VC) links [37]. A VC is a unidirectional transport of ATM cells, and a VC Identifier (VCI) identifies a particular VC link for a given VP connection. The VC link is terminated when the VCI is changed, and a VC is originated or terminated by the assignment or removal of the VCI. VC links are concatenated to form a VC Connection (VCC) as can be seen in Figure 1.4.

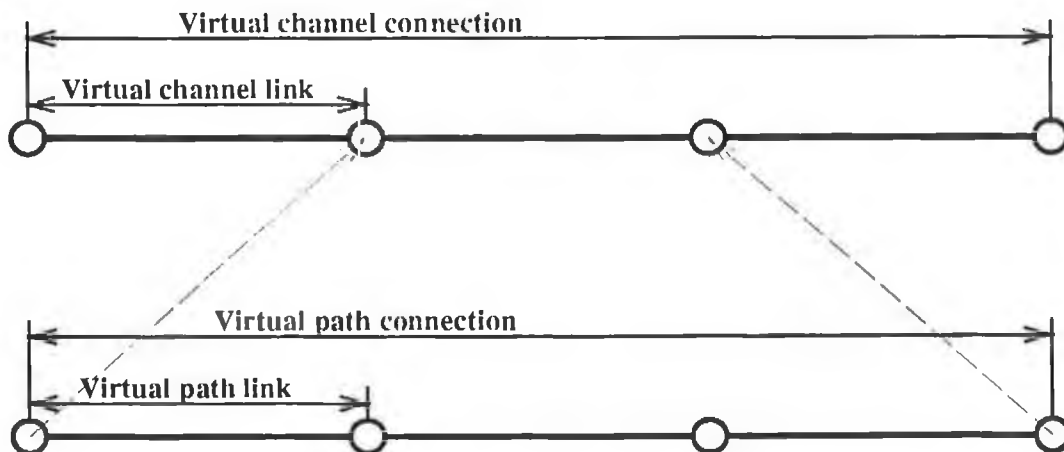


Figure 1.4: Virtual Channels and Virtual Paths

A VP is a bundle of VC links, and all the VC links in the bundle have the same endpoints, so that a VC link is equivalent to a VP connection. A VP Identifier (VPI) identifies a group of VC links that share the same VP Connection (VPC). VP links are concatenated to form a VPC, and a VPC endpoint is where the VCI changes, originates or terminates. When there is a VC switch there first must be a termination of the VPCs that support the VC links that are going to be switched. Cell sequence is preserved in a VP and also in each VC link within a VPC. In a VP switch the VC links that share a VPC must remain the same after the switch as before as is seen in Figure 1.5. In a VC switch all the VPs involved in the switching must be terminated and then originated again as can be seen in Figure 1.6.

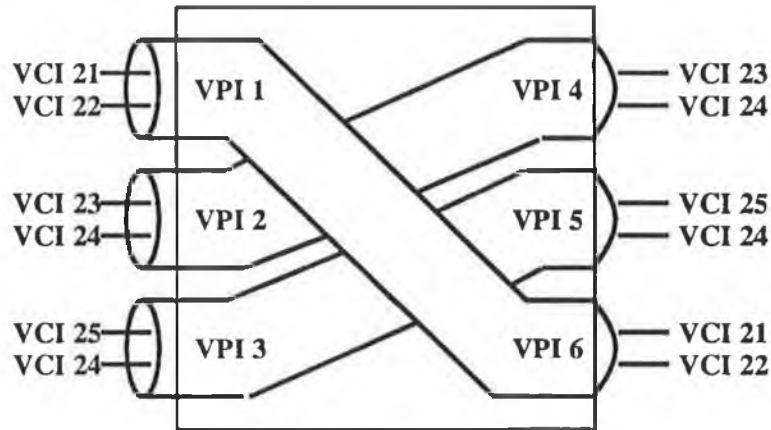


Figure 1.5: VP Switch

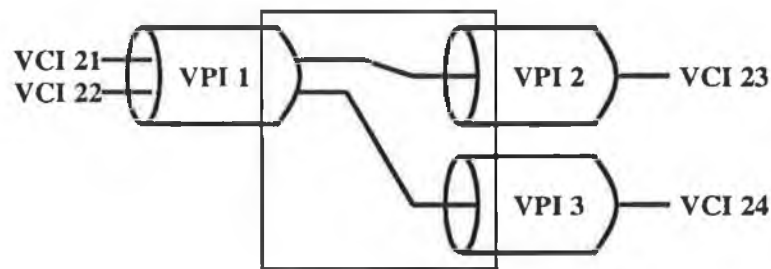


Figure 1.6: VC Switch

The VP and VC allow flexibility in the management of the resources in the network, by simplifying the routing and the resource allocation methods. It is possible for the network to lump many VCs together and then treat them as a single entity, rather than maybe hundreds.

1.2.2 The Physical Layer

The physical layer is made up of two separate functions : (1) functions that depend on the actual type of medium used, which are contained in the physical medium sublayer ; (2) functions that change the bits to ATM cells, which are contained in the transmission convergence sublayer. These sublayers are shown in Figure 1.3. ATM networks are likely to use other transmission systems to carry the cells. Typical possible transmission systems are the synchronous digital hierarchy

(SDH) or the FDDI standards. The first task for any transmission system is to get timing at the bit level, which is at the lowest physical level, and this is achieved by the physical medium sublayer. The bit rates that have been discussed are the 155.52 Mb/s and the 622.08 Mb/s systems over optical fiber. However it is also possible to use coaxial cable or radio systems. There are a number of other schemes proposed for ATM transmission, like the pilot European project, which uses 34 Mb/s transmission, and there is considerable interest in low bit rate ATM, maybe as low as 2 Mb/s. Once the bits are available to the next sublayer it is then possible to first convert the bits to the frames of the transmission systems used, and then to convert the frames to the actual cells. This can be seen in Figure 1.7 where a number of cells are put in a frame.

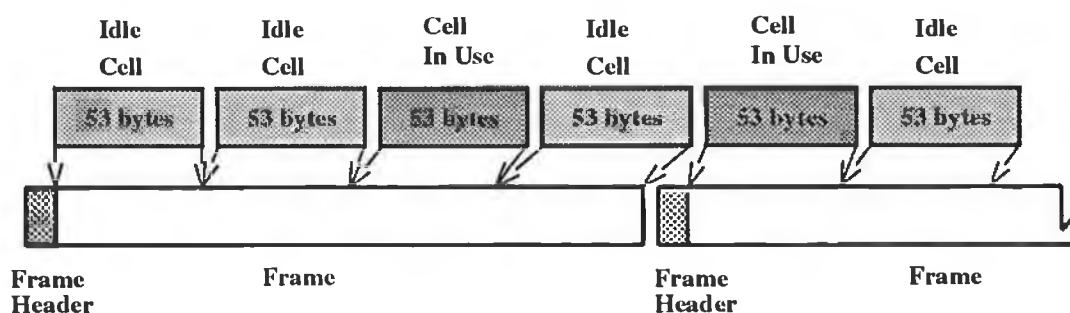


Figure 1.7: Cells To Frame Conversion At The Physical Layer

The transmission convergence sublayer has been standardised to perform the generation and extraction of the frames at the specified rates from the SDH and finding the ATM cells by looking for the HEC on the cell header and then checking the error correction code. The format for the cells within a frame of SDH is set and an overhead of 9 bytes on 270 bytes is needed for the SDH. Once the cells are found and checked the cells must be decoupled from the transmission rate of the medium and this is achieved by the insertion and deletion of idle cells in the stream, which can also be seen in Figure 1.7. When this is achieved the cells are then available to the ATM layer.

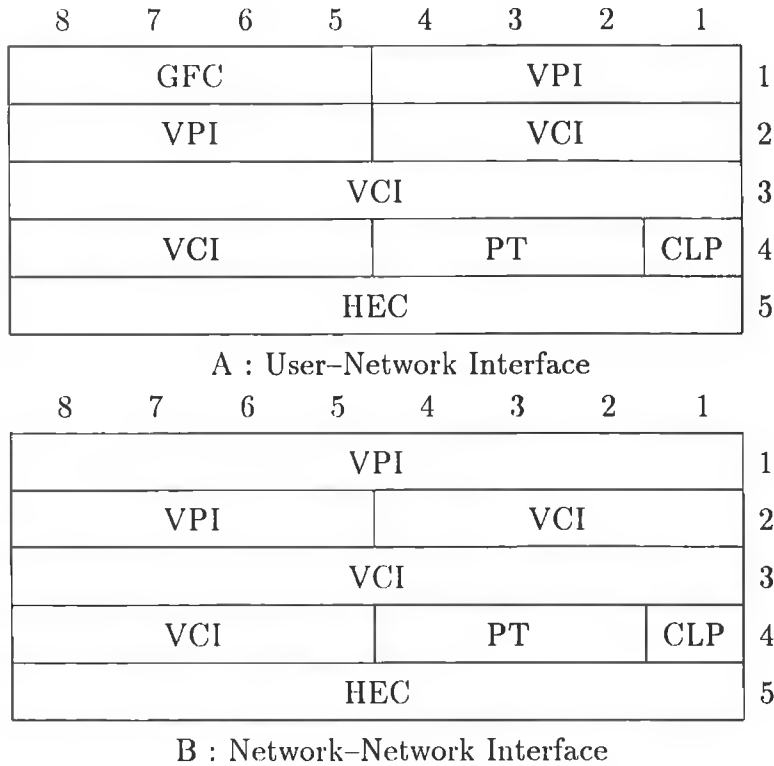
While all the physical attributes of the physical medium should be dealt with in the physical layer, this is not really possible. What remains is the assumption that the cells are being carried on a low bit error rate, low delay channel, and as

will be seen later this is not always the case. However to fix this medium dependent problem, it is not the physical layer that is analysed, but a higher layer.

1.2.3 The ATM Layer

The ATM layer is the core layer of the standard and this is the layer that routes the cells across the network and multiplexes and demultiplexes the cells together from many virtual paths on to one physical carrier. ATM is connection oriented and this is achieved by the use of a VCI and VPI, as has been seen in Sub-Section 1.2.1. There are two different ATM layer standards, the user-network interface (UNI) and the network-network interface (NNI), and these are shown in Figure 1.8. The difference between them is that in the UNI there is a field called the generic flow control (GFC) and the VPI is only one octet, while in the NNI the VPI is one and a half octets long and there is no GFC. The GFC has local significance only and the ATM switches will overwrite the value given. There are two modes of operation used, the controlled access and the uncontrolled access modes. The uncontrolled access sets all the bits to zero and the switch reads them to ensure that there are no errors. In the controlled access mode the sources are expected to modify their inputs based on the value of the GFC field. The exact nature of the method of operation is not specified in the ATM Forum UNI 3.0 [4]. Apart from routing the cells by means of the VPI and VCI, the ATM layer is also responsible for the delivery of a quality of service to the higher layers. There is a single bit in the header that is called the cell loss priority (CLP) bit and this allows two levels of priority, high and low.

There is also a payload type (PT) indicator which indicates what type of cell is being carried. This is used to distinguish between user cells and management type cells. It can also be used to show that there is congestion in the network. The last part of the header is the header error control (HEC) which is an eight bit CRC that is used to prevent errors occurring in the header itself. The ATM layer provides a service independent layer from the physical medium to the higher layers which can use the ATM cell information load of 48 bytes.



GFC – Generic Flow Control, VPI – Virtual Path Identifier,
 VCI – Virtual Channel Identifier, PT – Payload Type
 CLP – Cell Loss Priority, HEC – Header Error Control

Figure 1.8: ATM Cell Header Format

1.2.4 The ATM Adaptation Layer

While ATM is service independent, so as to be future-proof, it is also possible to cater for a number of services that are already present. The ATM adaptation layer is used to adapt the ATM layer to the services that will be using them. There are two sub-layers in the ATM Adaptation Layer (AAL) called the Convergence Sublayer (CS) and the Segmentation And Reassembly (SAR) sub-layer. The CS is service dependent and the main function of it is to adapt the service to the ATM methods. The SAR is defined to put the CS-PDU's into the cells. It will also handle the insertion of the header and trailer of the SAR-PDU if they are specified. The SAR ends up with 48 byte SAR-PDU so that when a 5 byte header is put on the unit there will be a 53 byte cell formed.

Chapter 1 Introduction

There is a standard method of classification of the possible types of service defined, based on the timing between the source and destination, the bit rate of the source and the connection mode. However not all eight possibilities are needed and there are four classes available. These are shown in Figure 1.9 and are called Classes A, B, C and D. Class A would be constant bit rate connection oriented real time services, like voice or video that is uncompressed. Class B is the same, except that the bit rate is now variable, and so would represent the compressed versions of the voice and video services. Classes C and D do not require the timing between the source and destination and are variable bit rate. The difference would be that Class C is connection oriented and Class D is connectionless. These services would be mainly data services. For each service there is an adaptation layer assigned, so that Class A uses AAL 1, and Class B AAL 2 and so on.

	Class A	Class B	Class C	Class D
Timing	Required		Not Required	
Bit Rate	Constant	Variable		
Connection Mode	Connection oriented			Connectionless
	AAL 1	AAL 2	AAL 3/4	
			AAL 5	

Figure 1.9: Service Classes In Adaptation

When the service definitions and format for the AAL's was standardised, it was seen that the AAL 3 and AAL 4 were similar and were not different in any real sense. They were amalgamated into a single AAL called AAL 3/4 as shown in Figure 1.9. What the AAL's standardise is the method of putting together the cells and what headers should be included. For the AAL 3/4 the overhead on a single cell is large and the method is complicated. Therefore a new AAL was decided upon called AAL 5, which will perform similar functions to that of the AAL 3/4, but with reduced functionality and at a higher level. A further point is that it is not necessary to use any AAL if one is not required, or it is possible to invent a specific type for a specific application as needed, although this would restrict internetworking to devices that understand the specific AAL.

1.2.5 ATM Services

ATM will provide a Quality Of Service (QOS) to connections. During connection set-up there is a need to find agreement between the user and the network on what QOS is required by that connection. The user will initiate a connection request to the network for a connection. The signalling would involve the user passing traffic parameters to the network, which in the ATM Forum 3.0 standard [4] is carried in the ATM user cell rate information element. Here the forward and backward parameters are specified, with parameters like the peak cell rate, the sustainable cell rate and the maximum burst size being typical. The network may adjust the parameters of the connection, if due to other connections QOS's, it would not be possible to connect the new connection. It will then be a matter for the user to decide if those parameters are sufficient for the connection or not.

There are a number of QOS that have been defined and they are coupled with the type of AAL that is used. It is of course also possible to have an unspecified QOS with no network guarantees. There are a number of types of services that can be supported by using these QOS types and the AAL's. The Constant Bit Rate (CBR) service is one that uses AAL 1 and requires a level of QOS from the network by specifying the mean bit rate and the cell loss acceptable as well as the allowable jitter in the cell delay. Another type of service is the Variable Bit Rate (VBR) service, which will use the AAL 2 and will also require a QOS from the network. In this service there will be specified parameters similar to the CBR but also the burst length and burst tolerance and the peak cell rate. Another type of service is the Unspecified Bit Rate (UBR) service, where there is no bit rate declared and therefore only an unspecified QOS is given to the service. To allow for some guarantees to be given a new type of service called the Available Bit Rate (ABR) service is being defined. This service will try to fill up the available bandwidth and still give guarantees to the services.

1.3 Guarantees In ATM Networks

The main reason that congestion and control are important issues in ATM, is that in ATM the network is both attempting to give a QOS to the users, as well as multiplexing the users together for efficiency. This is unlike any other data network at present. Two different types of congestion and control are possible in ATM networks, one at the cell level and the other at the connection or call level. Cell level effects are due to the possible simultaneous arrival of a number of cells from different sources. The effect of this is that if there is not a buffer large enough to store these cells until they can be served, then there will be loss of cells. At the connection level there are similar problems, except this time it could be not only the simultaneous arrival of a number of bursts from sources, but it could also be that the statistical nature of a number of sources do not interleave to allow the traffic to be smoothed. Congestion of course could also occur at either the cell or connection level due to the sources not correctly specifying or adhering to the parameters that were agreed on in the contract. Congestion occurs when not enough resources are available whereas control is the action taken to either ensure that congestion does not occur, or the action taken when congestion occurs. Connection and cell level control as well as cell level congestion are discussed in the standards [4].

1.3.1 Quality Of Service Definitions

A number of parameters are used to define the QOS. These include the Cell Error Ratio, Cell Loss Ratio, Cell Misinsertion Rate, Cell Transfer Delay, Mean Cell Transfer Delay, Cell Delay Variation and the Severely Errored Cell Block Ratio. A number of classes of QOS are supported by the network and fall into either a Specified QOS or an Unspecified QOS class [4]. At present the standards only specify that the Specified QOS class 1, which is the circuit emulation service and constant bit rate video, be supported. A specified QOS may have two cell loss objectives, for the high and low priority traffic. The network gives the user no guarantees in the unspecified QOS class. However the user may give the network some traffic parameters, that the network can use for internal operation. The

network could then use these parameters internally to achieve some quality of service. These parameters can change during a connection and may not always be specified correctly. This type of traffic could be the so called best-effort traffic. This allows the network to respond to time variable resources. The unspecified QOS is optional for the network to support.

Degradation of QOS may arise for many different reasons and one of these is the ATM switch. The buffer capacity could be a complex multiple queue system with an algorithmically defined service rule that could be based on priorities. The switch may thus introduce loss under heavy load. For compliant connections the QOS will be supported for at least the number of conforming cells as specified in the conformance definition. For non-compliant connections the network does not need to support any QOS. The issue arises, when using VPs and VCs, as to what QOS does the network take note of. In other words the users are specifying the QOS on the VCs, but the network really only wants to be concerned with the VPs for ease of management. The translation between these is specified in the standards and is that the QOS of a VP will be the strictest set of QOS of any underlying VC [4]. This imposes a difficult requirement on the network, in that even if only one of many VCs has a hard QOS, the whole VP, and hence all the underlying VCs, are given that QOS.

1.3.2 Contract Between User & Network

The method used by ATM networks to provide QOS to the users is by maintaining a contract with them. When a connection requires resources there is a contract set-up between the network and the user. The user describes the connection in terms of network parameters and the network then uses a connection admission control scheme to calculate if the connection can be admitted to the network while providing that QOS to the incoming connection and also to maintain the QOS to the other connections that are already set-up. One way of achieving this is by using the concept of effective bandwidths, where the burstiness of the connection is captured in a single parameter. There is then a linear constraint on the connection admission control scheme for the connection that is about to be admitted

and it does not depend recursively on the connections that are set-up. Once the connection is in place it is important for the network to ensure that the connection is abiding by the contract. This is done by policing the connection to check on the contract traffic parameters [20]. The likely method of doing this for bursty connections is by means of the leaky bucket or generic cell rate algorithm [4, 20]. This algorithm allows cells to pass at the mean rate, but also to have some burst characteristics.

Once a contract is in place it is then likely that the user will use this information to control their own source. This type of behaviour is called traffic shaping and is similar in nature to the policer and usage parameter control that the network provides. However a difference is that the algorithm does not lose cells that do not comply, but merely stores them for further transmission and is called a regulator [16]. It is possible that the network may provide other contracts where there might be multi-level priority control on the basis of the VC that is in use and not the CLP bit in the header. This could be achieved through the signalling at the connection set-up, as is achieved in a number of early ATM switches.

1.4 The Network & User Perspectives

1.4.1 The User Perspective

The user perspective on Integrated Services Networks are discussed first to emphasise that user preferences *should* be the primary consideration. Once the network is in place only the users directly benefit from using it, by running applications which achieve higher-layer communication goals. Network owners and operators benefit indirectly, by providing services that users want or are willing to pay for.

From the users point of view, there should be as few restrictions as possible on the communication services they obtain from the network. In particular :

- the network may ask for various traffic parameters prior to accepting a connection. A user should be able to specify **any values** for these parameters or simply specify nothing about the traffic their connection will generate.

Chapter 1 Introduction

- a user should be able to demand **any values** of the various QOS parameters that the network has defined; or simply tell the network to provide the best possible service with no guarantees required.
- a user should be able to **adjust connection traffic parameters dynamically during the connection lifetime, if desired**. For example some data transfer applications are flexible regarding the delay incurred in completing the transfer and can vary the input traffic rate during the connection.
- perhaps most importantly, the user should have a **simple interface** to the network to conduct these negotiations.

The simple interface is the interface between the person and the equipment that they use, and this allows some complexity to be handled by the user's application hardware or software before the network interface. Therefore the restriction to a simple user interface to the network can be relaxed to a simple *human* interface as is shown in Figure 1.10. Any complicated processing necessary to translate user commands into actions affecting their network connection is assigned to the local processing element.

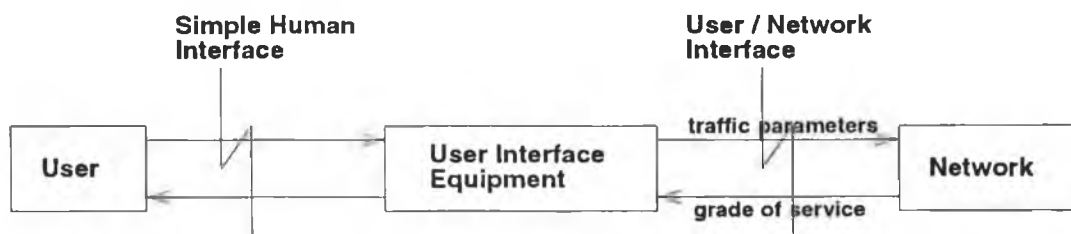


Figure 1.10: User Interface

This flexibility in user service characteristics is motivated by the observation that it is becoming more and more difficult to accurately define a *“typical” user’s requirements*. There is already a spectrum of such user traffic characteristics as mean bit-rate or peak-to-mean bit-rate ratio. In addition, technological advances may continue to change the requirements for present-day services, for instance by reducing the bandwidth needed for voice or VCR-quality-video calls. There is also a wide range of user QOS requirements even within many of the service classes proposed in the literature. For example, some “video phone” users may require a

high-quality reliable connection while others may be satisfied with poorer-quality or interruptible connections.

1.4.2 The Network Perspective

An Integrated Services Network could range from a Local Area Network (LAN) to a world-wide Wide Area Network (WAN), and could be private (all the applications controlled by one organisation) or public. The operation of a public network may be the responsibility of several organisations, within each of which the operational functions may be automated and / or distributed. Conceptually, however, the control and management functions of an Integrated Services Network can be associated with a *network operator* as if one entity was responsible for controlling and operating the network.

The network operator tries to satisfy three competing objectives (Figure 1.11) :

- **performance guarantees** should be provided and fulfilled for those users who require them. These guarantees may be deterministic or statistical.
- **all services** demanded by the users should be supported, including (ideally) future services with as-yet unknown characteristics.
- network operation should be an **efficient utilisation** of network resources, such as link capacities and node buffers. At the very least, network operation should be more efficient than a *circuit-switched* or *peak-allocation* strategy in which the maximum resources required by a connection are reserved for the duration of that connection, regardless of whether or not the resources are actually used continuously.

The difficulty with these objectives is that any two of them can be achieved relatively easily – conceptually at any rate – but simultaneously achieving all three is still an open problem. More specifically :

- *a wide range of services can be supported with efficient use of network resources, provided no guarantees have to be made to the users.* Typical values

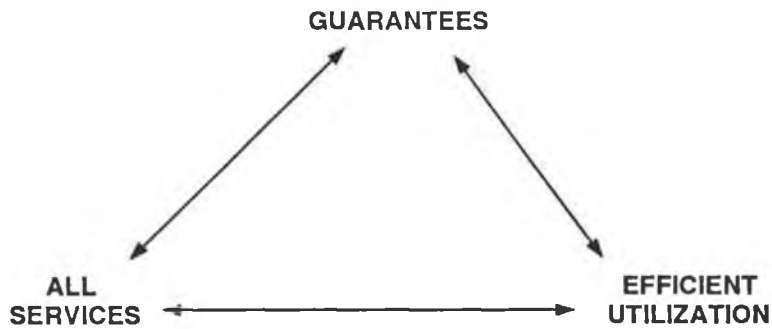


Figure 1.11: Network Objectives

of network performance measures may be a good indication of the expected QOS, aggregated over time and all users, but some user applications require more specific guarantees.

- *guarantees can be made to the users and network operation can be efficient, provided only one (or a narrow range) of service types have to be supported.* Focusing on one type of service allows the network to be optimised to efficiently deliver that service in the ways required by the users. This is essentially the traditional telephone network model.
- *a wide range of services can be supported, and guarantees demanded by the users can be offered and fulfilled, provided efficient network operation is not important.* This is usually achieved by an over provisioning of resources, such as reserving the peak bandwidth required by a connection.

The network operator should also be able to offer more customised service to individual users than is usually available today. This means that traditional network performance metrics such as average delay or packet loss may not be fine-grained enough, since typically a user cares only about the QOS their connections receive. Advances in user hardware and software, and the development of a competitive network provider industry, will require network operators to focus on satisfying the communications needs of an individual user, regardless of the size of the user's connection.

1.4.3 Future Integrated–Services Networks

The exact form of integrated–services networks of the future is also unclear, but based on current trends in communications and computing it is possible to predict some of its features with reasonable confidence. Multimedia services offering storage and transfer of voice, video, images and data will be supported, allowing remote conferencing and collaboration and replacing text processing with multimedia ‘document’ processing. Networks will be interconnected and will allow users and organisations to set up virtual private networks embedded in the physical network [2, 74]. Intelligent network interfaces will be needed at internetwork boundaries and at user access points to hide the interface details and present the image of a single network to the users. The implication of this is that the boundaries of the network become more complex while the interior of the network becomes simpler. The users will also have become more complex compared to the sort of user that would have typically used these networks. Users will demand a wide spectrum of application requirements as they become familiar with the network. Some may continue to opt for network–defined services or choose from a library of predetermined services, but others may want to customise their connections and vary connection parameters dynamically.

The view taken here is that current trends and future considerations in the communications and computing areas point to the need for a change in the traditional user–network relationship. The traditional model is centralised with passive users : the network provides well–defined services and users choose from this limited range. User feedback is long–term and inferred from the aggregate demands for services. For example the feedback may take the form of the users either not re–connecting to that service or changing to another service, in a time period that is larger than maybe hundreds or thousands of call times. Because of the small well–defined services that the network provides, the network can assume that the users are identical in characteristics, and so when getting feedback will average the response out over all the users. This is appropriate for large–scale provision of a single service, as in the phone network, but may be unsuitable for the kind of integrated multi–service network described above. The relationship

implied by many of the proposed CAC schemes is more of a ‘contract’ : users describe their traffic and make quality demands, and the network provides a stated level of service while enforcing the user commitments. Some problems with this model were outlined above, although it is a step in the right direction. Taking this process further, the view here is that :

- the network should be a provider of **basic network resources** rather than complex user services;
- the responsibility for packaging these resources into services should lie with the **users** (or in real-time their interface equipment);
- the network’s primary function should be to **co-ordinate requests** for its resources. The goal of this co-ordination could be to optimise some measure of network performance, or to maximise a suitably-defined global user satisfaction, or to ensure some degree of fairness, or some other objective.

Thus the network is viewed simply as a high-speed cell relay network in which the central issue is transporting cells rather than implementing services [46]. Furthermore ATM is supposed to be a service independent network and therefore should not be concerned with the services that are implemented at the higher layers but instead concentrate on the core issue of cell delivery. Users may actually be working with higher layer protocols such as TCP/IP running on top of an ATM network, which also argues for keeping the ATM network as simple as possible. One proposal along these lines is to move connection admission and bandwidth allocation decisions to the terminal equipment at the network access points, and somehow ensure that the combined user rates do not exceed the network capacities [81]. This gives a simpler network but it may be less efficient than previous networks.

Here the users take responsibility for requesting sufficient network resources to meet their QOS requirements. At connection set-up this involves specifying some parameters of the requested connection, either manually or in a menu-driven environment. If users want to adjust connection parameters during a connection then these parameter specifications may be automated. This kind of in-connection

negotiation is desirable for bursty connections where the actual resource usage is of interest [70]. One traffic management scheme for burst-level resource reservation is described in [80], and a form of in-connection negotiation is now commercially available in some network interface equipment. Even for static connection requests, the connection set-up process may involve a negotiation between the user and the network in which the user modifies their request to conform to the current level of resource utilisation. For example [25] outlines a real-time channel set-up procedure in which detailed information is sent to the user if a connection request is refused, allowing the user to take this feedback into account in a revised request.

Some work is underway on modifying user traffic inputs via pricing in the Internet [51] and in integrated-services networks [66, 76]. A contract-based CAC scheme is proposed in [40] in which the charge to users is related to how accurately they declared their traffic, rewarding users who provide better information on their connection characteristics.

1.5 Resource Allocation Of ATM Networks

There are many resources in ATM networks and the ones that are of interest here are the buffer spaces and bandwidths that are available in the network. The resources issue at the connection level is how to achieve reasonable network efficiency for source types, where traditional resource allocation techniques may not apply. Reactive control and feedback schemes are under investigation (e.g. [27, 38]) and may be useful in certain network environments or on longer time scales, e.g. connection level rather than burst or cell level. Experience with real users will show which schemes are feasible in an actual network as opposed to a research or test environment. The problem is complex and is likely to require a multilevel solution approach [47, 60, 85]. It is fair to say that how bandwidth will be allocated in ATM networks is still an open question. There is also some work on the cell level resource allocation in terms of the overhead and efficiency of the throughput of the cells.

1.5.1 Time Scales

There are a number of time scales that are relevant in ATM networks and there are very different approaches used to examine each of them. The smallest time interval is concerned with bit timing and whether the bits are correct or not. This bit scale level is shown in Figure 1.12, and this is relevant at the physical medium sublayer and is on the sub micro-second scale.

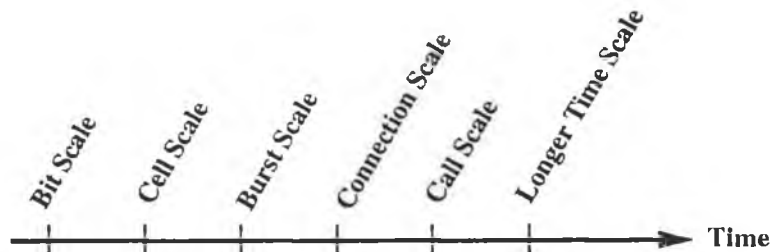


Figure 1.12: Time Levels In ATM Networks

The cell scale effects are ones that are relevant to how the cells are put together, with particular importance being placed on the overhead in the cell. This is discussed in Chapter 3, when the overhead within the cell is too large over satellite links. The burst scale is relevant when dealing with the time between cells from the same source and how cells from different sources are mixed or multiplexed together. This is also dealt with in Chapter 3 in terms of what the worst case traffic might be due to cell effects. The connection scale is on the level of milli-seconds or more and this is what the remaining chapters deal with. There are longer time scales like the call scale, but these are not covered here. All these time scales can be seen in Figure 1.12.

When dealing with a particular time scale, especially the longer ones, it is important for simulations to make a number of approximations about the lower time scale effects. When the connection level is dealt with it is convenient to make assumptions that there are no cell scale or bit scale effects that impinge on the model. If this were not the case then the time taken and complexity involved in simulation would be large and so would be unfeasible or impractical.

1.5.2 Statistical Multiplexing

One of the fundamental issues which has not yet been satisfactorily resolved is the mechanism by which user traffic will be accepted into the network, usually referred to as Connection Admission Control (CAC). When a new connection is requested by a user, the network must decide whether or not to accept the connection; and if so, how to route it through the network and what resources to reserve for its virtual channel. Packet-switched networks use higher-layer protocols to guarantee acceptable packet delivery but these are not expected to scale well to broadband speeds. In circuit-switched networks (such as most telephone networks) the CAC mechanism results in connection blocking when the bandwidth of a requested connection exceeds the available bandwidth [5]. But in an integrated-services network the traffic source may be *bursty*, so the required bandwidth of its virtual channel varies with time during the connection.

The nature of this time-varying behaviour and the mean bandwidth requirement vary widely among different sources. Therefore it is difficult to characterise the bandwidth of a requested connection. This difficulty has led to proposals to reserve the peak bandwidth of the connection (deterministic multiplexing), as required for constant bit rate (CBR) sources. However the gain in efficiency possible by taking advantage of the statistical nature of variable bit rate (VBR) sources has led to many schemes for statistical multiplexing. Such schemes assign less than the peak bandwidth required, and therefore may introduce cell loss and / or delay. The extent to which these service degradations occur is measured by the quality of service (QOS) offered to the connection.

The aim of a preventive CAC scheme is to balance the QOS offered to admitted connections against network utilisation by limiting the number of connections using the network. Many schemes described in the literature decide whether or not to accept a connection based on knowledge of the *connection behaviour*, the user's *quality requirements*, and the *current state* of the network [35, 69]. An example of this is given in [69], where the connection behaviour coupled with the QOS are mapped to specific traffic classes. The current state of the network is gained by

using predictive methods as well as reactive controls from the network. Another example is that the traffic contract that the user and network agreed on, must be available to the CAC [4] so that the CAC can base its decision on the connection behaviour and the users quality of service requirements. The network might then use a theory like “effective bandwidths”, which tells the CAC the current state of the network, to calculate if the connection can be carried without effecting the other users. Ideally a user requesting a connection would give a complete statistical description of the connection, but in practice only a limited indication of expected connection behaviour is feasible [80]. Connection behaviour is described by a set of parameters called traffic descriptors, such as mean bit rate, peak bit rate, maximum burst length, probability of cell arrival in a fixed interval, and so on. User quality requirements are usually expressed in terms of acceptable cell loss, delay and jitter. Based on these requirements, traffic sources are divided into classes and each class is provided with a different QOS guarantee, eg. [79]. The current state of the network can be determined by monitoring the utilisation of network resources and/or by characterising the behaviour of connections already admitted. For example, traffic models based on fluid flow approximations have been used in analysing network bandwidth and buffer utilisation [48]. Based on the above knowledge, CAC schemes have been developed in which each source is assigned an effective bandwidth [39] in order to meet its QOS while still permitting a statistical multiplexing gain.

1.5.3 Effective Bandwidths

Effective bandwidths are a way of summarising the statistical information of a source in a single parameter. The complex problem of resource allocation of a multi-service network can be simplified by trying to get an equivalent circuit switched model [21]. By using effective bandwidths it is possible to get a linear equation similar to the circuit switched networks and see if there is sufficient bandwidth left to admit another connection. The original idea of effective bandwidths can be attributed to Hui [36] and a summary of the uses can be found in [39]. Using that notation, here is an overview of the mathematical formulation.

Chapter 1 Introduction

The simplest model is that there are J sources that are sharing the same link which has a capacity of C . Let X_j be the load produced by source j and assume that all the X_j 's are independent random variables with possibly different distributions. The performance constraint on the systems can be given by looking at the probability of overflow of the queue and this is equivalent to the question whether it is possible to impose conditions on the distributions of the X_j 's which would ensure that :

$$\mathbf{P} \left\{ \sum_{j=1}^J X_j \geq C \right\} \leq e^{-\gamma} \quad (1.1)$$

for a given value of γ ? The answer is that there are constants α and K that depend on γ and C such that if :

$$\sum_{j=1}^J B(F_j) \leq K \quad (1.2)$$

is satisfied then Equation 1.1 is also satisfied. This is just a linear constraint. The variable $B(F_j)$ is called the **effective bandwidth** and is given by :

$$B(F_j) = \frac{1}{\alpha} \log \mathbf{E} \left[e^{\alpha X_j} \right] \quad (1.3)$$

This effective bandwidth can now be treated like in the circuit switched network and it is possible for a large range of multi-type sources that share a single queue to have a linear constraint on performance. This is true because it is possible to have an asymptotic constraint on the tail distribution of the buffers workload [21]. It is now known [40] that for a quite general models of sources and resources it is possible to associate an effective bandwidth to each source, such that if the sum of those effective bandwidths using the resource is less than a critical value then the resource can deliver the required performance [82]. For more general models of sources the simple version of the effective bandwidths given in Equation 1.3 may not be valid, however it is not known when the more general effective bandwidth is needed.

1.5.4 Current Resource Allocation Techniques

The current resource allocation methods used for services in ATM rely on many techniques. At present the CBR traffic and circuit emulation traffic is decoupled from the remaining services. This is achieved by using different VPs for each of them. This allows each sub-problem to be tackled independently. Effectively the bandwidth for the CBR traffic is split from the other traffic's bandwidth, by reserving the amount needed in advance. The other traffic can then have no effect on the CBR traffic. However not all the advantages of statistical multiplexing can be got by doing this. For VBR services like compressed voice there are a number of models using effective bandwidths that look to be useful and provide good bounds for the connection admission control. The CBR and VBR services will be policed by either a single leaky bucket or possibly a dual leaky bucket.

ABR traffic is more difficult to control due to the unknown nature of the sources and there have been a number of proposed schemes [10, 68] to control it. The factors that have influenced the decisions on the choice of schemes are numerous and vary from what types of sources are expected to what might be possible in real networks. The final decision might not even have been arrived at, and this area is still under intense research.

An early scheme proposed, is to just allocate the resources for the peak rate and resign the network to inefficiency. However this is probably only a short term solution as competitive forces will force this option out. There is also the possibility of offering different levels of priority that the user can allocate to different services that they might have multiplexed together. While this might work when the user is in control of the multiplexing, in bigger networks this might not be possible. The different levels of priority are achieved by having a number of output buffers in the switch and allocating different VCs to different buffers. As many as seven levels of priority can be offered this way. The shortcoming of this scheme is that it is not clear how to decide to distribute the levels of priority, and different users may give different priorities to the same types of service so making decisions within the network becomes complex. This solution is likely to be of use only in the local part

of the network.

One of the major questions that has been posed for ABR traffic, is whether a closed-loop scheme or an open-loop scheme is needed ? The ATM Forum have decided that closed-loop feedback is needed for ABR traffic. Therefore end-to-end feedback is needed, but what type of feedback ? There is an ongoing debate about whether a credit based flow control scheme or a rate based flow control scheme is better [77]. In a credit based flow control scheme it is claimed that there will be no loss and there is no need for policing the connections, as the credits are only created when there are sufficient resources available for the traffic. However the rate based schemes depend on the ability of the ABR traffic to fill up the rest of the bandwidth and these are able to do this at high speed, whereas credit based schemes are not. The ATM forum has specified that the rate based scheme will be the one implemented. Another question is how does the feedback actually take place, is it bit based or explicit ? With the bit based schemes the bit could be contained in the cells but then there has to a rule to decide what to do in each case of receiving this bit. However in the explicit case the new adjusted rate is specified, but it needs a resource management cell to carry the information. However this explicit feedback is the preferred method. It is also likely that rather than using queue length to measure congestion, the rate of queue growth will be used.

Two schemes that have been tested are the Explicit Forward Congestion Indication (EFCI), scheme and the Enhanced Proportional Rate Control Algorithm (EPRCA). The EFCI scheme is similar to a scheme that was used in other data networks, and is a bit based feedback scheme. The EPRCA scheme is on the other hand a new scheme and uses the resource management cells to send information from the switches to the end users, in order to control the sources.

1.5.5 Will Resource Allocation Continue To Be Important ?

Some commentators have suggested that the widespread deployment of fiber optic lines, and continuing exponential decreases in processor and memory costs, will result in these network resources becoming essentially “free” so that efficiency in

their use will not be important in the future. However three points should be kept in mind :

1. demands are continuing to increase exponentially, so that it is not clear when – if ever – network resources will be “free” ;
2. past experience suggests that application developers will have no difficulty in designing new services that use up all the available resources, perhaps after an initial adjustment period ;
3. when a significant number of users become involved in defining their service characteristics, efficient network operation will be critical in a competitive network provider environment. Put simply, if network operation is not efficient, the users will be efficient – by multiplexing their traffic before submitting it to the network, for example. Legal barriers or tariff disincentives to this kind of user behaviour may not be feasible, so network inefficiency could lead to a financial penalty for the network operator.

1.6 A New Model Of ATM Networking

Already the model for ATM service negotiation involves the users more than before in defining their own services. Taking this process further, the following basic principles of network operation are proposed :

- the network should be a provider of resources rather than services ;
- responsibility for packaging these resources into services should lie with the users, or in real-time their interface equipment ;
- the network, or a third party provider, might offer a pre-defined menu of services that the users can choose from, if there are enough users that do not want to define their own services ;
- the network's primary function should be to co-ordinate requests for its resources. The goal of this co-ordination could be to optimise some measure

Chapter 1 Introduction

of network performance, or to maximise a suitably-defined global user satisfaction, or to ensure some degree of fairness, or some other objective ;

- resources can be requested and allocated in real-time, or essentially continuously. Therefore for all but the shortest-lived connections it should be possible to adjust a connection's parameters during its lifetime.

In this model the network is simply a high-speed cell-relay network [67] in which the central issue is transporting cells rather than implementing services required at higher levels. This view is consistent with regarding a VC connection in an ATM network as a virtual wire [22], since the functionality delivered by the VC is similar to that of the physical layer in the OSI reference model. An immediate consequence of this new model is a fundamental change in user-network interaction : more flexible "contracts" can be continuously renegotiated by those users whose applications or QOS preferences permit it, while still accommodating users whose demands can be met by pre-agreed contracts.

The model proposed here agrees with the spirit of recent proposals to move decisions about network services to the edges of the network, preferably all the way to the users, and to take advantage of the flexibility of some application types :

- the proposed Available Bit-Rate (ABR) service would assign whatever bandwidth is not currently allocated for reserved-bandwidth applications to those users whose applications require little or no guarantees on delay or cell loss ;
- some proposed video schemes divide the bandwidth assigned to a connection into base and enhancement layers. The base layer bandwidth would be guaranteed so that some minimum service is obtained, while the enhancement layer would provide additional bandwidth (which presumably corresponds to improved service) whenever it was otherwise unused ;
- even for pre-agreed connection requests, the set-up phase could involve an iterative negotiation between the user and the network in which the user modifies their request to conform to the current level of resource utilisation.

For example [25] describes a real-time set-up procedure in which detailed information is sent to users whose connection requests are refused, allowing them to take this feedback into account in a revised request.

The crucial issue ignored in most of these proposals is the basis on which decisions should be made, both by the network and by users. By developing an economic framework in which incentives are provided, it is possible to make rational decisions and resource allocations in real-time. These incentives could correspond to actual money, or could simply be control signals, provided the users and the network operator react appropriately to them (for example in a private network where one entity controls both the network and the “users” i.e. applications).

1.7 Objectives of Thesis

The main objective of this thesis is to propose methods to increase the efficiency of ATM. There is the way in which the ATM cell is formed and then multiplexed together in a buffer. The way that the cell is put together can lead to a gain in efficiency as is shown when looking at ATM over satellite links, and the way the cells are multiplexed together is important as can be seen when examining the area of worst case traffic in the network. over satellite, due to the inefficiencies. Another area of interest is that of user behaviour and it is shown that when adaptive users are examined it is possible to increase the efficiency of the network. There are a number of objectives for each chapter and these are now described.

1.7.1 Objectives of Chapter 2

The objective of Chapter 2 is to develop the tools for the further chapters. These tools are the models of the sources that will be used to investigate various aspects of ATM networks and the modelling and simulation techniques needed because of the nature of ATM networks. There is first a motivation to demonstrate what sort of models are useful when examining ATM networks. It is seen that ATM networks pose some different simulation problems compared to other networks and the so-

lutions to these are attempted. There follows some justification for the simulation methods that are used in the rest of the thesis. There is some detail given about the software and simulation platform that will be used. The different models designed are models for voice, video and data sources, as well as models for the different network functions that will be needed in the rest of the chapters.

1.7.2 Objectives of Chapter 3

The two issues dealt with in Chapter 3 are firstly the method in which the cell is formed and secondly the way that cells can be multiplexed together. The first objective was to send data efficiently over a satellite link along with other services, like voice. One of the problems is that there might not be a higher layer protocol protecting the data from errors that would be encountered on the satellite link. What is then seen is that when using satellite hops, it is not possible to use the normal type of ATM protocol. However with some changes to the AAL, it would be possible to easily send data reliably across the satellite. The changes have to do with the method by which the cell is formed.

Another problem, again concerned with the cell level, is to examine if there is a type of source that adheres to the ATM policer, called the leaky bucket, that imposes the biggest constraints on the network. Put another way, is there a type of source that complies with the leaky bucket and produces the worst performance in the network. If this were to be true, then for modelling purposes, it would be possible to always use this type of source, knowing that for all other types of sources, the network performance would be better. Two sources have been suggested to be contenders for being the worst case source and what is shown here is that neither of these is the worst all the time. This of course does not rule out the possibility of there being an as yet unknown source that is indeed the worst case, but rather it tells that both types of sources should be considered when looking at performance modelling.

1.7.3 Objectives of Chapter 4

The objective of Chapter 4 is to define the problem of congestion control and propose solutions to it. Even the definition of congestion is unclear, and so a mathematical framework is proposed to deal with these issues. The mathematical model that is employed is based on economic principles, and this allows us to strictly define congestion and also efficiency in economic terms.

The user types are re-defined by using the idea of adaptive parameters, giving a new set of classification of users. By using the re-defined user types and the economic framework developed, it is possible to investigate how an economic solution to congestion, by using pricing, compares against other schemes. To define the efficiency, consideration has to be given to both economic terms and network terms. There are two types of efficiency present, network efficiency and economic efficiency. Some simple models and simulations are presented and the results show that by using both pricing and adaptive users, both the network and economic efficiency increase at the same time.

1.8 Summary

This chapter has dealt with the need for ATM in Section 1.1 and has then developed the current standards that have been applied to ATM in Section 1.2. There the basic terminology has been introduced and explained. In Section 1.3 the framework for discussing guarantees in ATM was introduced and in Section 1.4 the user and network perspectives were addressed. In Section 1.5 the current and future methods of resource allocation was presented. The new model of ATM networking was explained in Section 1.6 and the summary of the rest of this work in Section 1.8 concluded this chapter.

1.8.1 Summary of Thesis

To analyse and compare various ideas in ATM networks it is usual to try to model and simulate the system. The models that are needed for this work are the typical

Chapter 1 Introduction

types of sources that might be used in future, and the parts of the network that are relevant to the performance of those services. This is dealt with in Chapter 2 where various video sources, as well as data and voice sources, are presented. A discrete event simulator is introduced and the simulation models are verified and validated.

At the level of ATM cells there are two important issues dealt with, how to put the cell together and how to put the cells into the network, and these are examined in Chapter 3. An example of why the formation of the ATM cell is important is given by way of an example, which is a satellite link. This satellite link is modelled and then simulation carried out to examine the performance of the ATM cells with suggestions as to how the performance might be improved. The way that the cells are inputted to the network, and still abiding by the congestion control algorithms, is an important issue for network design. The different ways of achieving this are then examined and some conclusions are drawn from mathematical analysis.

The different types of users that will use the ATM networks are presented in Chapter 4. There are new types of sources presented and, even without giving a traffic descriptor, these types can still get guarantees. These adaptive sources are then modelled mathematically. A new model for efficiency is presented, where both network issues and economic issues are dealt with. By using economic principles it is possible to gain advantages compared to the traditional approach in terms of efficiency. The economic framework can also take advantage of the adaptive types of sources. Finally in Chapter 5 some conclusions are drawn from the previous work.

Chapter 2

Performance Modelling & Simulation

2.1 Introduction

The Asynchronous Transfer Mode (ATM) network is widely accepted as being the broadband network of the future [20, 46]. ATM allows the design of networks that give efficient and flexible allocation of resources by only allocating the amount of the resource that is needed. This ability of ATM networks to support variable bit rate traffic can be used to meet the dramatically changing traffic requirements of many applications such as video telephony, full motion video or high-speed data transfers. New important issues related to ATM networks, such as congestion control and cell loss probability, have to be investigated. In addition even though the ATM concept provides 'bandwidth on demand' there is a need to ascertain the network's efficiency or throughput. Performance evaluation of ATM broadband networks can be addressed by simulation studies using realistic and accurate models.

The modelling and simulation of ATM networks presents new challenges due to the small packet size and the high bit rate of the link. If reasonable run times are to be achieved on reasonable computers then ways of making the simulation efficient are essential. There are many approaches to this problem, but here two methods will be investigated, decomposition and rate control methods. By using these enhancements it is possible to achieve both accuracy and quick simulation

time.

This chapter is organised as follows. There is a discussion of modelling and in particular network modelling in Section 2.2. One of the main objectives of the modelling was to model the different source types and this is addressed in Section 2.3. In this section the video source model is concentrated on as this presents the most difficulties. To produce an accurate video model the first task is to characterise the data collected and this is addressed in Sub-Section 2.3.1, along with data for voice and data models. The models are then presented, first the voice model in Sub-Section 2.3.2, then the video models in Sub-Section 2.3.3 and finally the data model in Sub-Section 2.3.4. In Section 2.4 the video model that was presented in the previous section is compared to MPEG models. Firstly in Sub-Section 2.4.1 an overview of MPEG is given, followed in Sub-Section 2.4.2 by a hierarchical view of MPEG models and finally in Sub-Section 2.4.3 by a comparison of MPEG models to H.261 models. After the models have been designed there is a need to convert them in to simulation models. There is a discussion of simulation in Section 2.5. The package that was used to build the simulation models is called *SES/workbench* and is described in Sub-Section 2.5.1. This package provides a set of software tools for graphical animation and simulation of models. The models that were designed in Section 2.3 are converted and described in Sub-Section 2.5.2. The validation of these models is presented in Sub-Section 2.5.3 along with general discussion of validation methods. The next section deals with the problems of simulation in high-speed networks and the problems with memory and speed. This is discussed in Section 2.6. There are various techniques for overcoming these problems and these are described in Sub-Sections 2.6.1, 2.6.2 and 2.6.3. Finally there is some discussion of the overall chapter in Section 2.7.

2.2 Overview Of Modelling

A model of a system is an abstraction and approximation to the actual one. The model may be conceptual, physical, electronic or computer based. A model should simplify the analysis of the system under investigation. The purpose of modelling is

Chapter 2 Performance Modelling & Simulation

to help the analysis, design, control or understanding of a system without actually having to build the system. There are many occasions where the actual system cannot be accessible, such as future systems which are only conceptual, or expensive systems where the cost of using the system is high, or systems that are providing critical services and cannot be taken out of use. Telecommunication networks and in particular future networks are ideal candidates for modelling as they are expensive, very large and usually only conceptual at the stage of modelling. Also it is usually not feasible to build a prototype due to the inflexibility of the prototype and the expense. The types of systems that are modelled fall into these categories, for example a satellite link is an expensive item, or customer behaviour is hard to work with in practise without long trials. The models that are of interest are conceptual models and use is made of them to build computer simulation models.

There are many factors that are taken into account when modelling but the most important factor is probably the detail of the model. It is usual to have hard constraints on time, money and computer resources. It is interesting to note that for most modelling one of the main reasons is that the real system is too slow and it would take too long to wait for results from it. However in the case of high-speed networks it is the opposite and it is difficult to get close to matching the speed of the real systems by modelling and simulation. A basic principle of modelling is to try and reduce the detail in the model. This has to be done in the context of having the model accurately portraying the essential aspects that are of interest to the system. Some of the main principles used in modelling in this work are :

- use a big picture initially to find the areas of interest
- find the most important parts of the system
- start with basic small models
- decide what results and performance is required
- only include detail that effects the results of the modelling
- include detail only to the extent that the data is available
- constantly re-check the model against the aims

- it may be necessary to include more detail for credibility

Even if it were possible to build a detailed model, the excessive detail would slow down the run time and be irrelevant to the results. The detail in the model should only really represent the detail of the data that is available to us.

In terms of modelling these systems, the divide-and-conquer approach is taken. Each component of the system is modelled separately and then the resulting model is the combination of these. The main interest of the modelling is for performance and so the elements that are modelled effect the performance of the system. Individual elements are the source models and the network elements. There are also other elements modelled such as customer behaviour and performance measures. The source models are discussed in detail in Section 2.3. The network elements include the ATM cell structure, the ATM Adaptation Layers of interest, the control measures in the ATM user-network interface, the policing methods, various link rates and standards as well as buffers and multiplexors. The result that are expected from the models is to be able to build computer simulation models. These models can then be used to investigate the performance of proposed systems and compare them to each other. The performance is generally given in terms of delay and loss, but also the efficiency, both in terms of networks and economics is investigated.

2.3 Source Modelling

There is a need to model telecommunications elements for network design and planning, such as the investigation of various congestion control mechanisms. Traditionally it was sufficient to model voice traffic for the telephone network and data traffic for computer networks. With the emergence of ATM networks there will be a mixture of voice, data, video and still image traffic on one network. The sources may take advantage of the ability of ATM to provide bandwidth on demand by supporting variable bit rate traffic. In particular video coding schemes will produce bursty variable bit rate traffic. While there are a number of coding schemes proposed it is likely that the algorithms used will be similar to those in the

H.261 standard. The work here presents a contribution to the area of Video Traffic Characterisation (VTC). Video source models are presented which are based on analysis of real video phone type data and discuss the results of the simulations carried out using these models. Both one and two layer codecs are considered. The one layer codec is based on a discrete time Markov chain while the two layer model is modelled on its individual layers. Both are derived from an analysis of video conference data. The models are simulated using a commercial discrete event simulation tool. Evaluation of the characteristics of each source are presented, as are the capabilities and limitations of the models. Comparisons are drawn between actual data and the models. While the results that are produced are dependent on the model of the video codec used, the method used to design the model is general and can be adapted to any codec design. Apart from video source modelling there is also a need to model the voice and data sources and these models are also presented.

This section deals with modelling voice, video and data sources for ATM networks. The voice models are taken from the literature [12, 15]. The video data, which the video models are based on, were generated as in [87] using a H.261 one and two layer codec. This codec produced a data set consisting of bit rates for each video frame. The data model presented here is for a file transfer application and gives a bursty source. The basis for the model is typical file sizes on a computer system.

2.3.1 Source Analysis And Characterisation

Voice has been characterised by a constant bit rate source, but this has been due to the technology rather than the information transfer rate. The normal method for digitising the voice signal is to bandlimit the signal to a band from 100 Hz to 3400 Hertz and then to sample that at 8 kHz. Each sample is then given 8 bits, to produce a bit rate of 64 kb/s. A voice call between two people will be silent about half the time in each direction, unless both are going to speak together. It is also possible to remove the inter-word silence and this gives the standard model used in the literature [12, 15, 29, 64]. It is possible to achieve about 35–40 percent

compression by using both of these silence removal techniques. The resulting model can be described by a two state on-off Markov chain.

Video, like voice, is naturally bursty [12, 15, 29] but can be manipulated to give a constant output bit rate. To gain the maximum efficiency from an ATM network only the real amount of information should be sent. This will change as the information content of the picture changes and so a variable bit rate output from the codec is expected. To gain maximum compression a single one layer codec could be used. Sudden movements in the picture or scene changes will produce large information transfers and so will produce high bit rates. Small changes from frame to frame will result in low bit rates.

In ATM networks it is sometimes hard to characterise the sources. Performance evaluation of the network becomes difficult and subsequent guarantees of quality of service (QOS) are hard to achieve [65]. However a QOS for a constant bit rate source might be easier to guarantee than that of a variable source. This has led to the development of two layered codecs [87]. One layer has a constant bit rate, called the base layer, and delivers a minimum QOS to the user. The second layer contains the enhancement to the base layer and is bursty in nature. The base layer is expected to achieve very low loss rates, maybe $10E-9$ bit error rate, and this can be achieved as the layer is more or less constant bit rate. The enhancement layer can tolerate more loss and will slowly degrade the picture quality. The relevant quality of pictures of the one and two layer codecs is subject to the coding involved in the codecs, the bandwidth available and the probability of cell loss. The empirical data on which the models are based were generated using algorithms implemented as in [87].

The H.261 video codec is a method of digitising video for use in video conferencing across the telephone network, and so the bit rate that is produced is constant bit rate and is specified by a bit rate of p 64 kb/s where $p = 1, \dots, 30$. For $p = 1$ a low quality video telephony of 64 kb/s is got, but if $p = 6$ to get a bit rate of 384 kb/s, then a good quality video conferencing is achieved, and of course it is possible to use up to $p = 30$ to give a bit rate of 2 Mb/s for high quality video.

Some of the main elements of H.261 include interframe and intraframe pre-

diction, motion compensation, discrete cosine transformation (DCT), quantisation of DCT coefficients and run length coding of the DCT coefficients. Interframe prediction takes advantage of the fact that consecutive frames will be quite similar allowing just the difference between frames to be transmitted rather than the entire frame data.

Intraframe compression is when the frames are quite different and the frame is treated independently and compressed in a manner similar to an image. This is used when there is substantial new information in a frame, like at the start of a video sequence or a sudden motion/scene change.

The H.261 codec produces constant bit rate video, but the method used to code the video produced is inherently variable bit rate [87]. The way that the H.261 produces a constant bit rate is that the quantiser step is adjusted during buffering the output of the codec and a constant bit rate is produced. The picture quality then varies according to the amount of motion in the sequence shown. If the quantiser step is left fixed then a variable bit rate output is produced, and an approximate constant picture quality is achieved. This variable bit rate codec has a peak to mean of about 4 in general [87]. In this work a variable bit rate one layer codec was produced using this technique. A two layer codec was investigated as well. The two layer codec is produced by using a one layer variable codec and putting the more important information into a base layer. Such information might include the picture start codes, the coding modes, motion vectors and some of the quantised DCT coefficients. This is filled to produce a constant bit rate layer and will provide a coarse level of quality. The rest of the coding will then, if present, be put into the enhancement layer, which will be inherently bursty.

The video sequences considered here are of just a single person talking with some zooming motion. This is more typical of a video phone type picture than video conferencing because there is little changing of scenes and only moderate motion. The spatial resolution of the five minute sequence was Common Image Format, CIF, (384, 288) and the temporal resolution was 25Hz. In order to capture the data set for the encoded video sequences, on a per frame basis, the outputs of the codecs were sampled every 0.04 seconds. This produced approximately 7500

Chapter 2 Performance Modelling & Simulation

frames to be considered. It can be seen from Figure 2.1 that the bit rate of this one layer codec varies considerably over the time interval. This short sequence and the type of picture that was used are limiting factors for the reliability of the models produced.

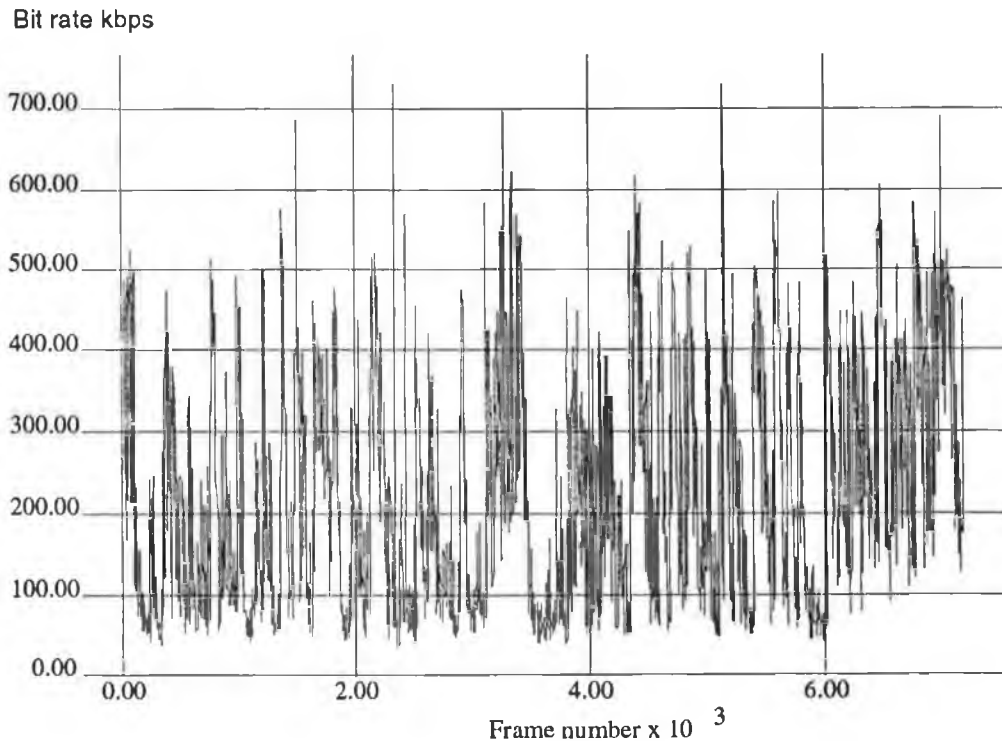


Figure 2.1: Bit Rate Of One Layer H.261 Codec

The two codecs discussed here are the one layer 256 kb/s source, whose main characteristics are shown in Table 2.1, and the two layer 384 kb/s source, whose main characteristics are shown in Table 2.2.

The 256 kb/s one layer source has a similar peak to mean ratio as is found in other studies [87]. The bit rate varies from small values of 33 kb/s to large values of 732 kb/s, depending on the information content of the sequence. The mean rate is not quite 256 kb/s due to the nature of the design of the codec, however that was the aim for the mean bit rate. It is found that the standard deviation is almost exactly the same as the mean.

The 384 kb/s two layer codec produced a combined bit rate as is shown in Figure 2.2, which is quite different to that produced by the one layer codec and shown in Figure 2.1.

Table 2.1: Characteristics Of 256 kb/s One Layer Codec

Parameter	256 kb/s One Layer Codec
Mean Bit Rate (in kb/s)	222
Lowest Bit Rate (in kb/s)	33
Peak Bit Rate (in kb/s)	732
Peak to Mean Ratio	3.3
Standard Deviation	Mean

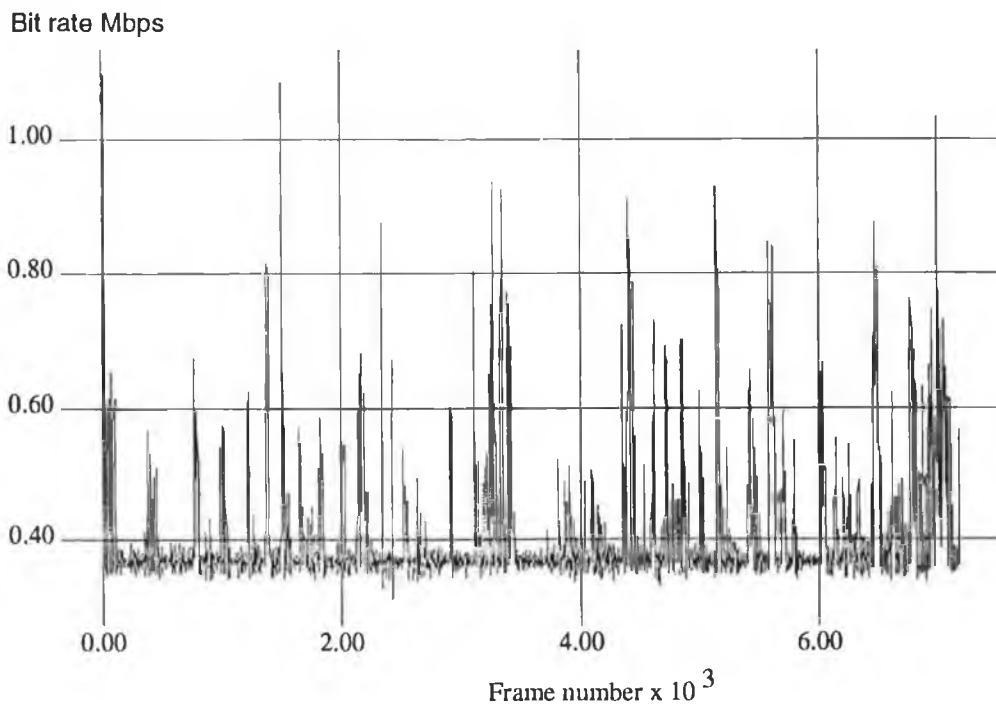


Figure 2.2: Bit Rate Of The Combined Two Layer H.261 Codec

It is seen for the 384 kb/s two layer codec that the base layer has different characteristics to the enhancement layer. The base layer is almost a constant bit rate, with a bit rate of about 384 kb/s, with the only exception is when there is not the information to send and the bit rate drops to 317 kb/s. The enhancement layer is then varying from sending no information to a high value of 725 kb/s. If the two layers are considered together then the total mean bit rate would be 408 kb/s and the peak bit rate would be 1109 kb/s and the peak to mean ratio would be 2.7, which is again close to that found in other studies [87]. There will be higher

Table 2.2: Characteristics Of 384 kb/s Two Layer Codec

Parameter	384 kb/s Two Layer Codec	
	Base Layer	Enhancement
Mean Bit Rate (in kb/s)	369	39
Lowest Bit Rate (in kb/s)	317	0
Peak Bit Rate (in kb/s)	384	725
Peak to Mean Ratio	1.04	18.6
Standard Deviation	0	High

bit rate sent with a two layer compared to a one layer due to the actual layering functions.

Data source modelling is difficult as the types of data transfer depend on the network and computers using the network. ATM networks have not been running for long enough to collect data that would represent typical data users. There are many different types of data transfer, from the almost constant bit rate large file transfer to interactive client-server applications which would be bursty in nature. Whatever type of model is used for the data source, the resulting simulations should not be specific to that model, so that if the data model were to change the results should hold true.

2.3.2 Voice Source Model

The voice model that is used here is a standard model with two states, speaking and silence [12, 29]. During speaking periods, cells are produced by the source at the bit rate of the voice. During silent periods no cells are produced. The mean duration of the speaking (or ON state) is 0.352 seconds. To allow for cells to be sent if not full the model is converted to the number of cells produced, rather than the on time in seconds. What is assumed is that even if the cell is not full, the cell still must be sent rather than waiting until it is full. This means that in the conversion from time to number of cells, the number is rounded up to the next

highest number of cells. If the number of cells produced in this period is counted, it is found that this number is geometrically distributed. In this case here this geometric distribution will have a mean of 59 cells.

The silence is made up of two parts, one from the silences when listening and the other from the silences between words or sentences. Therefore to properly represent the silence distribution there should be two distributions. However it has been found that a hyper-exponential distribution will accurately model the silences [12, 29]. A hyper-exponential distribution is a distribution that is made up of two exponential distributions, a lower exponential distribution with a mean of m_1 , and a higher exponential distribution with a mean of m_2 . There is then a probability of being in either exponential distribution, where p is the probability of being in the lower exponential distribution. The mean of the hyper-exponential distribution is given by m and can be calculated by :

$$m = p m_1 + (1 - p) m_2$$

The average silence, m , is found to be $m = 0.65$ seconds long. This is made up of the lower exponential, which models the inter-word silences, and the upper exponential, which models the listening silences. The probability of being in the lower exponential distribution is given by $p = 0.8746$ [29]. The mean of the lower and upper exponential distribution can then be calculated by using $m_1 = m / (2p)$ for the lower exponential distribution and $m_2 = m / (2(1 - p))$ for the upper exponential distribution [88]. Therefore the mean of the lower exponential distribution which models the inter-word silences is found to be $m_1 = 0.372$ seconds long and the upper exponential which models the listening silences is found to be $m_2 = 2.59$ seconds long.

From these parameters it is possible to fully characterise the hyper-exponential silence distribution. The hyper-exponential distribution relates these variables to the standard deviation, s , and this is shown in Equation 2.1.

$$p = \frac{(s/m)^2 + 1 - \sqrt{(s/m)^4 - 1}}{2((s/m)^2 + 1)} \quad (2.1)$$

From this equation it is possible to rearrange the variables to see that the second

moment of the hyper-exponential distribution is given by twice the product of the means of the exponential distributions.

$$\text{Let } Y = s^2/m^2 \quad \text{therefore } E\{X^2\} = s^2 + m^2 = (Y + 1)m^2$$

$$\text{As } p = \frac{(Y + 1) - \sqrt{Y^2 - 1}}{2(Y + 1)} \Rightarrow 2(Y + 1)p = (Y + 1) - \sqrt{Y^2 - 1}$$

$$(2p - 1)(Y + 1) = -\sqrt{Y^2 - 1} \Rightarrow (2p - 1)^2(Y + 1)^2 = (Y + 1)(Y - 1)$$

$$Y + 1 = \frac{1}{2(p - p^2)} = \frac{1}{2p} \frac{1}{1 - p}$$

$$(Y + 1)m^2 = 2 \frac{1}{2p} \frac{1}{2(1 - p)} = 2 m_1 m_2 = E\{X^2\}$$

Therefore the second moment is given by $E\{X^2\} = 1.927$. The standard deviation squared, s^2 , is equal to the second moment less the mean squared of the hyper-exponential distribution. $s^2 = E\{X^2\} - m^2 = 1.927 - 0.65^2 = 1.504$. The standard deviation is then given by $s = 1.22525$.

The resulting model assumes an input of 64 kb/s voice which if placed directly in the ATM layer would require a bit rate of 71 kb/s due to the 5 byte overhead for every 48 bytes of data. However as only the speaking portion is sent the required bit rate is 24.8 kbps. This gives a peak-to-mean ratio of 2.8. The model outlined here assumes 64 kb/s voice, but this can be dynamically changed during the simulation if desired.

2.3.3 Video Source Model

The one and two layered codecs were considered separately. For the one layer model a discrete state Markov chain was designed. This follows from work which verified that models based on multi-state Markov chains are sufficiently accurate for traffic studies [34, 71]. It was decided to take the frame as the basic unit of time for the chain and the ATM cell as the basic quantity. The objective is to capture the burstiness of the source by neither overestimating or underestimating the number of cells in each frame. The Markov chain model was produced by

Chapter 2 Performance Modelling & Simulation

carrying out the following analysis on the empirical data. The first issue was to examine the probability density function (PDF) of the size of frames produced. A comparable measure is the number of frames versus the frame size normalised to bits per second as is shown in Figure 2.3.

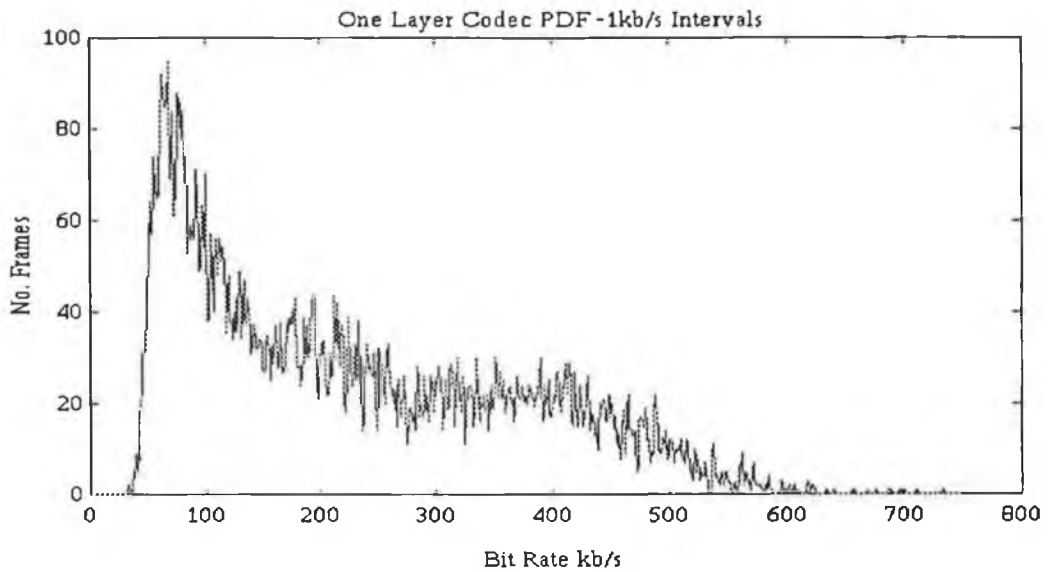


Figure 2.3: Shape Of One Layer Codec Probability Density Function (PDF)

A few criteria were used to select the states of the Markov chain. It was felt that at least five states would be needed to capture the shape of the PDF but the number could not be too large as the model would become complicated. It was felt that all states should be close to being equiprobable so that for short simulation runs they can all be reached. It was noticed that if there were a small number of states then a birth death process could be achieved. However if the number of states were larger, the assumption of a birth death process failed. The eventual number of states chosen was nine, which is a compromise between keeping a small number of states and also trying to have a birth death process. The birth death process greatly simplifies the modelling of the Markov chain as will be seen in Sub-Section 2.5.2.

Careful selection of the nine states resulted in a transition matrix where there were small entries at a distance more than one from the diagonal. The assumption made was to neglect these values. However the entries in the matrix must sum to

unity and so a method of proportioning the excess to the other entries was found. This was done by starting at either end of the Markov chain and conditioning on only the allowed transitions to the nearest neighbours. By starting at either end there would be two slightly different matrices. The eventual matrix chosen is shown in Table 2.3.

Table 2.3: Transition Probability Matrix For The One Layer Codec

State	0	1	2	3	4	5	6	7	8
0	0.81	0.19							
1	0.18	0.59	0.23						
2		0.19	0.60	0.21					
3			0.18	0.57	0.25				
4				0.22	0.57	0.21			
5					0.22	0.60	0.18		
6						0.17	0.75	0.08	
7							0.17	0.81	0.02
8								0.55	0.45

The probabilistic nature of transitions between these states is represented in matrix format by generating a count of all the transitions in and out of a particular state. This matrix is normalised in order to obtain the transition probabilities between states in the Markov-chain model p_{ij} where ij are the row column elements of the normalised transition matrix.

To ensure that the states were almost equiprobable, states were chosen so that the total frame count in that state did not exceed 15 percent of the total frame count. An attempt is made to obtain a matrix that contained transitions of length at most one state. This simplification of the Markov-chain to a ‘birth death process’ greatly simplifies the complexity involved in simulating the model. Within each state there is an upperbound and lowerbound on the bit rate of the source. This can be converted to an allowable number of ATM cells by conversion to the frame size and then fragmenting that frame size into SAR-SDU’s and by

Chapter 2 Performance Modelling & Simulation

the AAL to SAR-PDU's. This SAR-PDU joins with a five byte header to form an ATM cell. Therefore for each state there are integer numbers of upperbound and lowerbound cells allowable in that frame time. The actual number picked is uniformly distributed across that interval.

The two layer codec was modelled differently as each layer was modelled separately. The PDF of the two layers is shown in Figure 2.4. The base layer is almost a constant bit rate and was modelled by just one state. Within that state the bit rate is uniformly distributed from 350 kb/s to 388 kb/s giving the correct mean rate of 369 kb/s. The enhancement layer could have been modelled by a Markov chain similar to the one layer codec model. However here the possibility of modelling the layer by a hyper-exponential distribution is investigated. The hyper-exponential function produces a mean bit rate of 39 kb/s. To fit the distribution to the data a standard deviation of 124 kb/s was chosen.

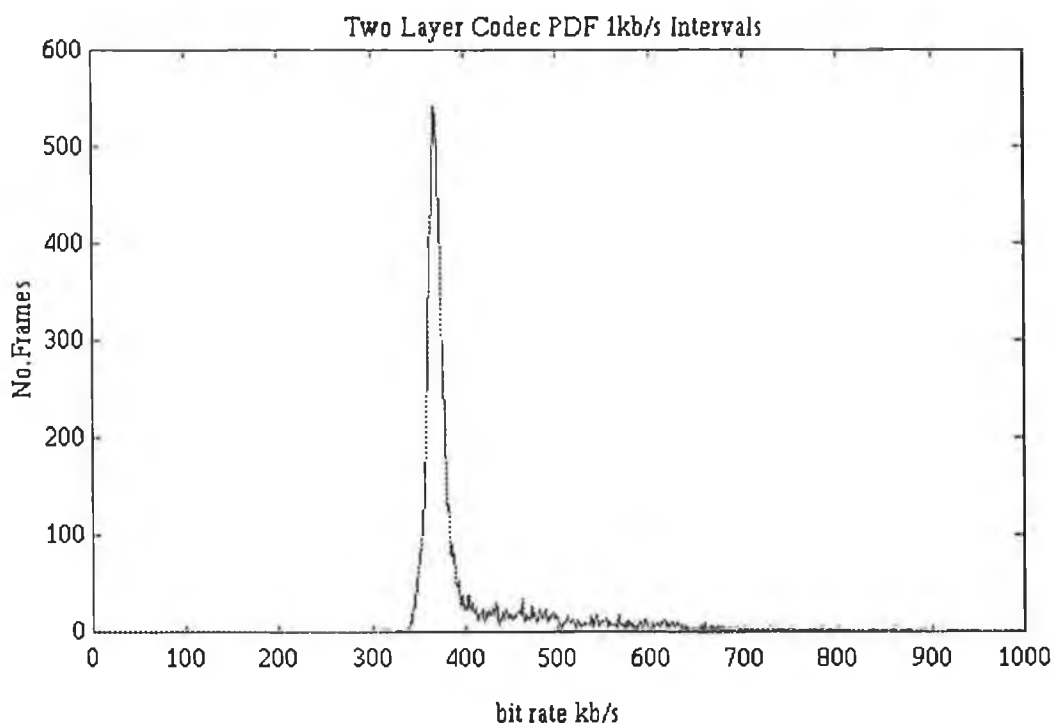


Figure 2.4: PDF For The Two Layer Codec

2.3.4 Data Source Model

A data source is one of the most difficult sources to model as the source type depends on what applications are being run and on what systems but here it is modelled by a file transfer application. This captures the bursty nature of data communications as well as its looser delay requirements relative to voice and video. A model was built based on transferring files from one computer to another. An empirical distribution for file size ranges was obtained from actual files stored on one of computers at DCU. In the simulations a range was chosen according to this empirical distribution, and then a file size was chosen from a uniform distribution within this range. The amount of data transfer can be varied depending on how many files are to be transferred. For one file per second the average bit rate is 197 kbps. The peak-to-mean ratio of this source can be high with values up around 1000. The user's quality requirement is that the file will be transferred at a rate determined by the network. A rate is given to the data source for a time interval and the rest of the file is stored awaiting transfer.

2.4 MPEG Models Of Video Sources

The current industry standard for full motion picture is the Moving Picture Expert Group, (MPEG), standard. The MPEG standard is a method of coding both the frames themselves, intra-frame coding, and between different frames, inter-frame coding. The MPEG1 algorithm is similar to and almost compatible with H.261 [87]. However one of the differences between the H.261 and MPEG is that in MPEG there are a number of different types of frames. In H.261 when there is an error it tends to propagate, due to the inter-frame coding. to prevent this in MPEG a type of frame is sent periodically with no inter-frame coding. Models for MPEG traffic sources to be used in ATM evaluation are only now beginning to emerge in the literature [73].

2.4.1 Overview Of MPEG

The frame structure for a motion picture will have considerable changes between some frames and between other frames there is little change. Therefore not as much use of inter-frame coding can be got as in the H.261 standard. There are a number of different types of frames in MPEG, to allow for both the scene changes and also to prevent error propagation caused by inter-frame coding.

The three types of frames in MPEG are : the I-frame, P-frame and B-frame. The I-frame is one where there is no reliance on previous or future frames for coding. The coding is done on only the frame itself and so there is only intra-frame coding used. This is useful when an error has been introduced to the frame sequence as it restores the correct frame again. The P-frame is coded using prediction of the previous I-frame or P-frame. This prediction is usually accomplished by means of motion compensation and motion predication. The last type of frame the B-frame is a frame that is coded by means of forward and backward prediction to I-frames and P-frames or even an estimate of the in between value of them. This type of frame is susceptible to propagating errors but has the advantage of low bit rate.

The I-frames are the frames with the most information in them and so have the largest frame size, in [73] they have a value of about 6.9 kbytes. This coding and compression is due to the discrete cosine transform similar to the H.261 coding. The P-frames have the next highest frame size with an average size of about 1.8 kbytes or about a quarter of the I-frame. This compression is achieved by the motion compensation. The lowest frame sizes are the B-frames, which have an average size of about 0.9 kbytes or about half the P-frames. It can be seen that a considerable compression can be achieved by not just sending all I-frames, but instead sending a mix of the different frame types. This mix of frames is called the group of pictures, (GOP), which is usually periodic between two I-frames, as is shown in Figure 2.5. The sequence is decided in advance and is then deterministic. A typical GOP [73] which consists of 12 frames might be 'IBBPBBPBBPBB' and then repeats again which is shown in Figure 2.5.

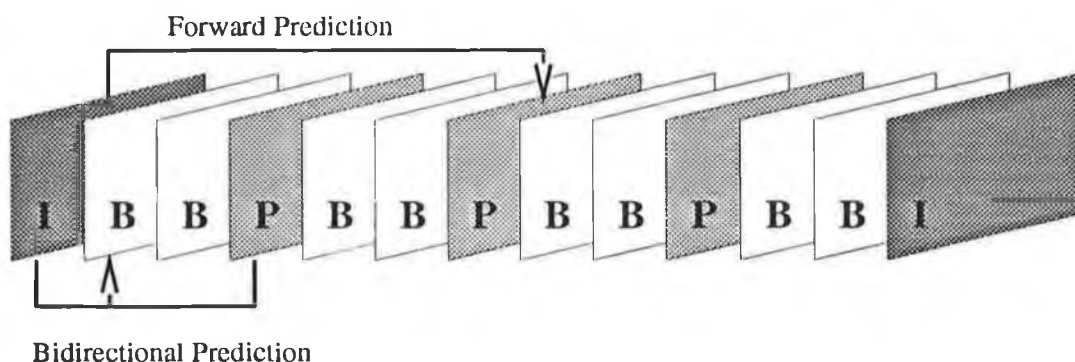


Figure 2.5: A Typical MPEG Group Of Pictures Sequence, 12 Frames / GOP

2.4.2 Hierarchical Models Of MPEG Video Source

There are many video source models used for investigating the performance of ATM with variable bit rate sources [72]. Because of the nature of MPEG frames there are a number of physical reasons for correlation's between different frames. Due to the nature of the types of frames available in MPEG there is a short term dependence on the particular GOP that is used. This short term auto-correlation is relatively fast-decaying. However there is also a long term auto-correlation due to the content of the video source which is a slow-decaying auto-correlation. It has been suggested that a hierarchical model is necessary to model the video source to a sufficient level to estimate cell loss in an ATM multiplexor [11]. It has been suggested that it is natural to classify the video traffic in terms of the time scale that they occur on. This would lead to modelling the video source in terms of the cell layer, the burst layer, the scene layer and the call layer. These have time durations of the order of microseconds, milliseconds, seconds and minutes respectively [11], as is shown in Figure 2.6.

The scene duration and the human activity of the video sequence is due to a different source compared to the effect of the actual coding and so a distinction is made between them when modelling. There is a formal description of the various models proposed in [11]. There are continuous and discrete time models based on Markov, Alternating Markov, Semi-Markov and Alternating Semi-Markov source models. If the interest is not so much in the actual cell output of one source, but rather in the combination of many source then there is also the possibility

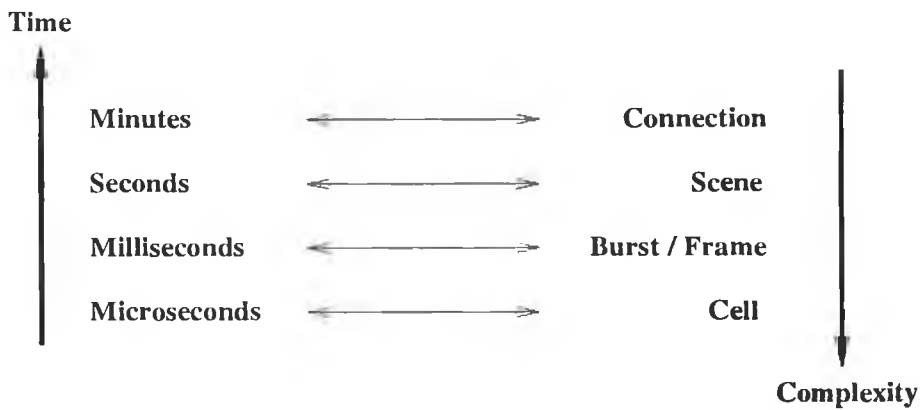


Figure 2.6: Hierarchical Model Of A Video Source

of using Fluid models which are also described by Markov, Alternating Markov, Semi-Markov and Alternating Semi-Markov in [11].

2.4.3 Comparison Of MPEG To H.261 Models

MPEG is used for mainly motion pictures while H.261 is primarily useful for teleconferencing. In teleconferencing there will in general be no scene changes and so the hierarchical model using the scene as one level is not relevant. The H.261 was meant for low quality links with a reasonably high error rate while MPEG is meant for low bit error rate networks. To overcome the possibility of propagating errors, due to inter-frame prediction, there is an MPEG-II standard which allows for layered coding and this ensures that a certain quality is maintained by using techniques similar to H.261 base layer codecs. However comparing the I-frames in MPEG with the H.261 codec they are very similar in that they use the same methods of coding. As the I-frame is the one that produces the highest bit rates it is also the one likely to be present when there is congestion due to a number of video sources reaching their peak simultaneously. Therefore when looking at congestion, as in this work, it is possible to imagine what feedback signals could be used by looking at the H.261 models alone. What is lost here is the complexity and detail in the MPEG models. However what is gained are the simple models that can be used to give an approximation of what might be possible.

What is investigated later is how a user might act in congestion and what

use would there be of a feedback signal from the network indicating congestion. It might be possible in the future for a video user to use a congestion signal from the network to lower the bit rate by increasing the quantisation step used in the coding algorithm, or possibly to reduce the number of frames transmitted per second. For MPEG it would also be possible to increase the number of P-frames and B-frames, but if the congestion is on the sub-frame level then these methods would not really be useful as they would be too slow. Therefore it is likely that similar congestion avoidance techniques are possible for both the H.261 user and the MPEG user.

2.5 Simulation Techniques

After models have been designed by normal modelling techniques there is then a choice of methods to gain insight in to the system and gain results from it through the model. It may be possible to have an analytical solution to the model but in a lot of cases this is not the case. The types of models that are dealt with here tend not to give analytical solutions in the detail that is required. This is mainly due to the fact that the system does not reduce to a product form solution, nor is it close to one. This is due to many factors but two of the main ones are that there are passive resources and there are priority schemes being used. There is also the problem that in the models there are embedded heuristic algorithms. When it is not possible to get an analytical solution computer simulation is an alternative. However there should be some crude analytical checks on the results from the simulations against the real system. The models that are designed are not static models but vary with time and so this further complicates the analytical methods.

While modelling and simulation is a slightly indirect method of studying the behaviour of systems it is sometimes a necessity when other alternatives such as observation or analysis are unavailable to us. In most cases computer modelling and simulation is the only way of evaluating new networks and control methods which are only at the design stage. The simulation is based on a stochastic model and there is the issue of how long should the model run before there is consistency

with the conceptual model. There are problems with simulation not alone in terms of verification but also due to the limits of computing power available. Most of the simulation models will have problems relating to speed and memory of the computing platform.

2.5.1 *SES/workbench*

The simulation tool used to model the sources and the network is a commercial discrete-event simulator called *SES/workbench* [88]. Although a commercial package is slower to simulate in running time than raw C code, the time to get a model working is reduced considerably. It was felt that a lot more time would be saved by using a commercial package where models can be produced quickly even though the run time would be reduced. Even with raw C code it would be necessary to speed up the simulation time and this is easier to achieve using a package where many methods can be tried out. *SES/workbench* is used for software, hardware and systems design. Its use is not limited to communication applications and it has been used extensively for non-engineering applications. *SES/workbench* consists of three modules, *SES/design*, *SES/sim* and *SES/scope*.

SES/design is a graphical interface module that is used to specify the model. It can be used under an X-windows system which is what was used in this work. This graphical interface allows the building and designing of the models in a visual way and so eases the task of creating systems. An *SES/workbench* model is a hierarchy of submodels where each submodel is a directed graph with nodes and links. There are predefined nodes available with the package and these include resource management, transaction flow control, submodel management and some other types of nodes. The nodes can be used as is or it is possible to redefine them to have another configuration. A screen dump of the 20 available nodes is shown in Figure 2.7.

The code behind the nodes is available to be altered. It is possible to write the SES simulation language called *SES/sim* or to use C. When a model is built by *SES/design* then the output of that will be a graphical model that

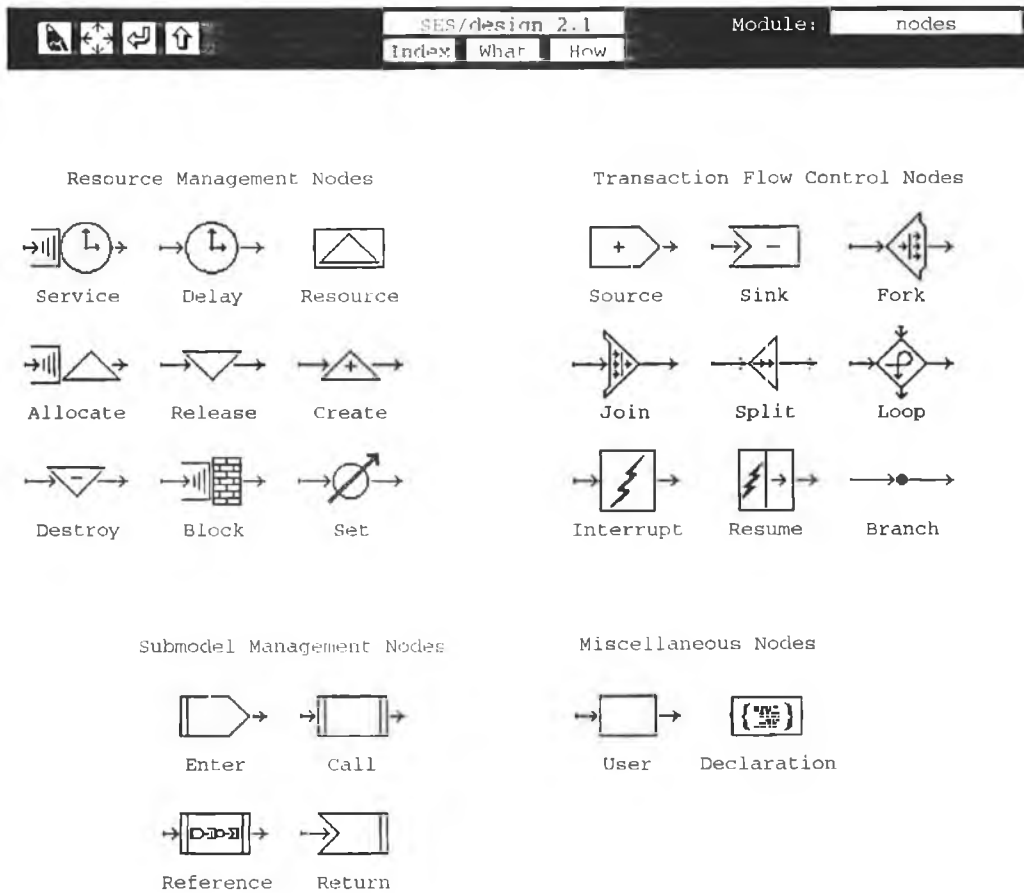


Figure 2.7: SES/design's Graphical Interface

can be translated into a SES/sim file. SES/sim is built upon the C programming language. All SES/workbench files are translated into C before being translated into executable files. SES/sim is similar to C and the expressions are the same in the two of them. Many declarations, expressions and statements are also the same, and even a function called "main" is needed to simulate a model. Because SES/sim is a superset of C it allows the user to gain access to a wealth of data underlying the model through a predefined set of variables, pointers, macros and runtime library calls.

There are predefined probability functions contained in SES/workbench, which include Poisson, geometric, binomial, Erlang, exponential, hyper-exponential, gamma, normal, uniform and triangular distributions. Apart from these it is also possible to define a probability function empirically. The probability functions either return an integer or a floating point number depending on the distribution

used. All of the probability functions and all the probabilistic modelling that is done in *SES/workbench* uses random number generators. For random number generators there are a number of predefined ones as well as the possibility of defining one.

The predefined number generators include a pre-set linear congruential algorithm, a generic linear congruential algorithm and a card-shuffle algorithm. The pre-set predefined linear congruential random number generator is the default one and the one used for the work in this thesis. This is based on the following algorithm from [41] :

$$\text{state} = \{ (a \text{ state} + b) \bmod m \} ; \quad rv = \frac{\text{state}}{m}$$

where rv is the random variable produced, 'state' is the state of the algorithm which is initialised to the seed value, $a = 1664525$, $b = 1664525$ and $m = 2^{32}$. The 'state' is initialised with the seed value of 314159, which can be altered in the code. One of the advantages in using this algorithm is that for a 32 bit machine, the algorithm is optimised and is quick to run. This is mainly because the modulus function is done automatically by having only 32 bit precision available. Another advantage of using the predefined, pre-set random number generator, is that for different runs of the program the same sequence of random numbers will appear. This means that the model will perform the same, unless the seed or random number stream is changed, and this allows for debugging of the model by going to a particular part of the run time, again and again, and getting the same results. It is possible to change the seed and the random number generator while the simulation is running, and it is possible to create many instances of the random variables. For most of the following work the random number generator used is the predefined, pre-set random number generator.

The final module in *SES/workbench* is the running of the model and the examination of the results. This is available with the use of the *SES/scope*. *SES/scope* allows the user to animate the model and so to visualise exactly what is occurring in the simulation. A typical output of the *SES/scope* is shown in Figure 2.8. It is possible to make an executable that is not animatable and this will run quicker

than an executable that is animatable. For the debugging stage it is useful to have the animation. *SES/scope* allows interaction with the model as it is running as well as controlling some parameters of the model. The operation of *SES/scope* is through a graphical interface that is similar to *SES/design*. It is possible to get detailed information from the model on the screen and to dump this to a file. Apart from statistics that are collected from the *SES/design* it is also possible to put inspectors on any variable in *SES/scope*. When a variable has an inspector attached then a graph of the performance is possible as is shown in Figure 2.8.

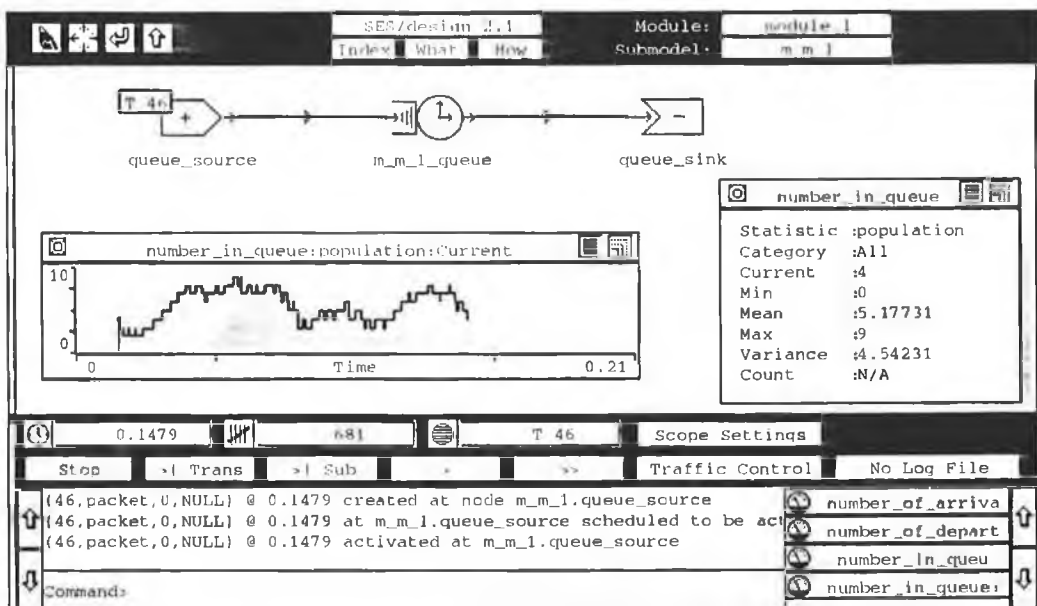


Figure 2.8: *SES/scope*'s Graphical Output

The simulation execution platform was a cluster of Sparc-10 workstations. The workstation running UNIX can support many sessions of SES at the same time. It is possible to remotely run SES from any computer that supports X-windows emulation. There are a number of methods of achieving that. The popular method is by running a UNIX operating systems and even on a PC this is possible by running "LINUX". If this option is not available then it is possible to have an X-window emulator like "X-Win". This allows SES to be controlled from any PC with a connection to the internet.

2.5.2 Source Simulation Model Design

2.5.2.1 Voice Source Simulation Model Design

The voice model produced by SES/*workbench* is on the cell level, so that each cell time a decision is taken on whether or not to produce a burst of cells. The transmission line is thus divided into cell times, and that is the smallest time unit considered. The size of the burst is the number of cells that will be transmitted and is chosen from the geometric distribution. Within the burst the cells are not back-to-back but rather at the cell rate given by the voice source while on. After the burst the silence interval is chosen and no cells are produced. When the source does not produce a cell then a time is chosen from the hyper-exponential distribution and no cells are produced in that interval but a cell will be produced in the next interval. It is possible using this model to find out how many cells will be produced in a fixed time interval and to find the number that will be produced in the interval after that as well. This allows for a speed-up of the simulation when there is no longer a need for cell-by-cell simulation.

2.5.2.2 One Layer Video Source Simulation Model Design

Simulating the one layer model is summarised in Figure 2.9¹. The basic time unit is a frame time or 0.04 seconds, which emulates a codec running at 25 frames per second. The birth death process makes the model much simpler because instead of having 64 arcs connecting all the states to each other there are only 16 arcs. The transition matrix for this birth-death process was shown in Figure reftrans.

Executing the model produces a number of ATM cells based on the range of bit rates that a particular state represents. As shown in Figure 2.9 the model can enter a new state at a rate of 25 per second. The probability of a transition between states is determined by the normalised transition matrix. On entering a state the model produces a number of ATM cells. The number produced depends on the state being entered and is uniformly distributed across the range of attainable values in

¹The bucket here corresponds to the state of the Markov-Chain and specifies the allowable bit rates, or frame sizes, for that time period.

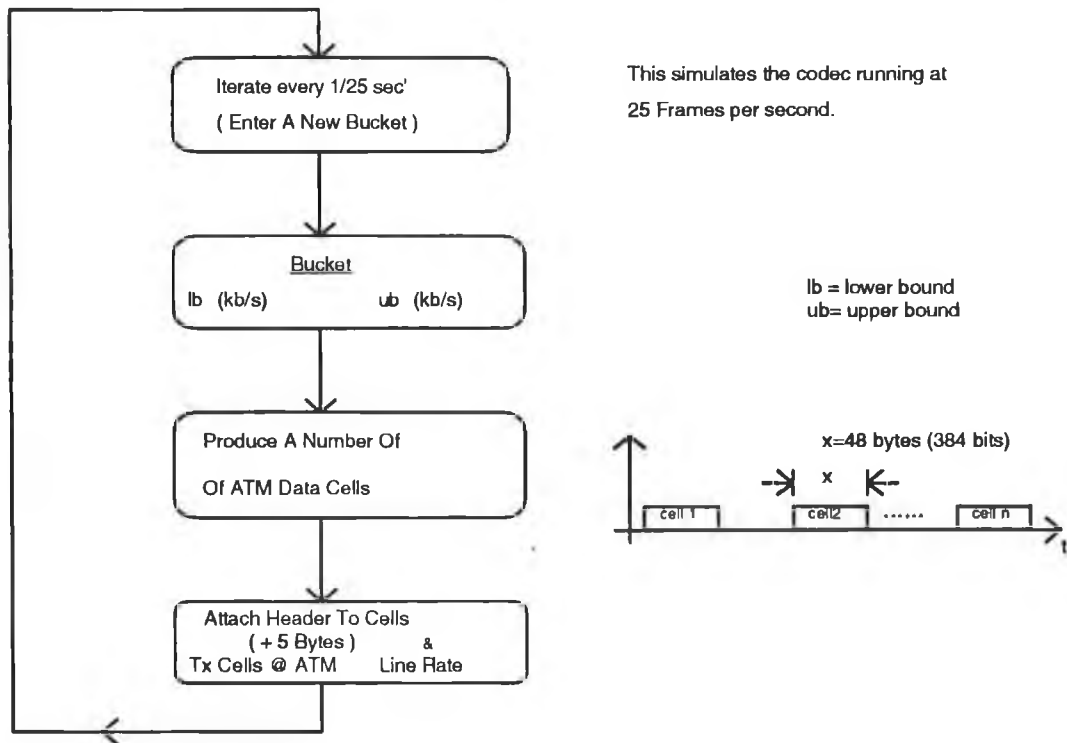


Figure 2.9: Process For Simulating The One Layer Codec

that state. Since the empirical data set for the coded video sequence presented the data on a bit rate per second format for each frame it was necessary to convert this data to ATM cells. If the cell load was not filled with video information, the cell was sent anyway as it was time critical. This is similar to rounding up the integer values of the number of cells. The AAL 3/4 was used to show a possible implementation with 44 bytes of information per cell. ATM cells are generated equally spaced in time over the frame duration, which is using source shaping.

2.5.2.3 Two Layer Video Source Simulation Model Design

Implementing the two layer model was quite similar to the one layer model. In this case however there is not a state diagram but rather a single state for the base layer and a state for the enhancement layer. The base layer produces a number of cells each frame time, where the number of cells produced is chosen uniformly between the limits of the base layer. The enhancement layer uses a hyper-exponential distribution, and so a number of cells is chosen from that each frame time. As

before the rounding of the continuous variables to the integer number of cells is up, so that cells that are not completely filled with video information are transmitted. Both the enhancement and base layers run at the same time combining to form the two layer codec output.

2.5.2.4 Data Source Simulation Model Design

The data model produced by *SES/workbench* is that of a file transfer application. There are two parameters of interest, the file size and the number of files to be transferred. The assumption is that a number of file transfers is possible simultaneously. The file size has been chosen from a probability distribution that is empirically based. This distribution was chosen from a sample of about 200 files on a PC [78], which varied in size from small text files to large executables. Both the rate of file transfer and the number of files to be transferred simultaneously are based on probability distributions. These can be changed but are currently based on uniform distributions.

2.5.3 Source Simulation Model Validation

It is generally felt that there are three steps in deciding if a simulation is an accurate representation of the actual systems considered [26, 33, 44]. The first step is to **verify** the model, that is to make sure that the computer simulation acts as it is intended to against the conceptual model that has been designed. This generally involves debugging the code and can be aided by using modules of code rather than a large program. It also includes trying the systems under a number of input parameters. The software used [88] also allows the use of traces which are a powerful method of checking on the state of the system. This was used extensively and was checked with hand calculations of the models to check the simulation. It was also possible to animate the systems using *SES/scope* and this provided a visual method of verifying the models. Another method was to check against certain conditions occurring and this was also possible to do with *SES/scope*. Once the model produced is verified then the computer simulation portrays the intended

system model. The next step is to **validate** the model so that the conceptual model can be considered to be accurately representing the real system. This validation is done by comparing the models to what is generally accepted as the real system, or in the case here the models represented in the literature. The real system for this exercise can be thought of as the previous results and characteristics of the sources. The final step in the simulation is to ensure that the models are **credible** which is to see if the simulation is accepted as being accurate and useful. For the purpose of this research this would amount to publication of the simulation details, models and methods. The relationship of these steps can be seen in Figure 2.10.

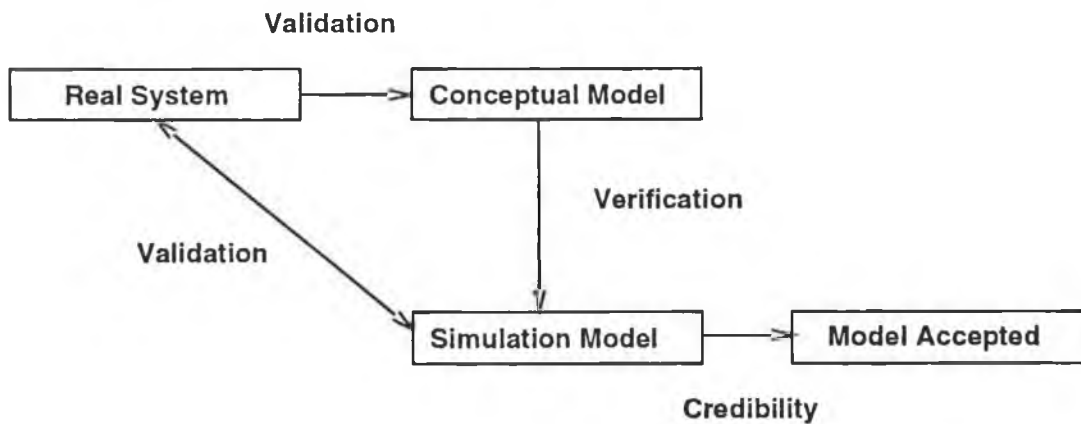


Figure 2.10: Relationship Between Verification, Validation and Credibility for Simulation Models

The voice model was verified and validated against the model of voice that is commonly used in the published literature [12, 29, 64]. The main check was to see if the mean ‘on’ time and ‘off’ time were similar to the models published. The data model was verified but validation of this model is difficult as there is not a standard model. Again the check was to see if the model was accurate with respect to the conceptual model. The validation of the model will be considered in the context of the overall modelling as the actual characteristics of the data model should not affect the results of the simulations. The video models were verified and then validated by comparison to the characteristics of the data that was captured from H.261 type algorithms. A first approach was taken by visually comparing the PDFs of the empirical data with those of the models. However due to the cell structure of the models there is a discrete effect on the PDF, and so visual

comparison was difficult. It was noted however that some edge effects were taking place at the boundaries of the states where a certain cell size was being favourably treated by being in two states. This was duly eliminated from the model. The basic measurements that were observed were the moments of the models and these are compared in Table 2.4 for the one layer model.

Table 2.4: Validation Of The One Layer Codec Model

Parameter	Empirical	Model
Mean (in b/s)	216 E+3	223 E+3
Second Moment	2.08E+10	2.22E+10
Root of Second Moment	144 E+3	149 E+3
Third Moment	1.9 E+15	2.2 E+15
Cubic Root of Third Moment	124 E+3	130 E+3

These results show that statistically the model is producing similar moments to the empirical data. To validate the model further, the correlation functions should be taken and compared for the model and the data. For the two layer model the mean bit rate produced is chosen to be identical to the empirical data and the standard deviation is similarly chosen.

2.6 High-Speed Simulation Problems

As mentioned previously there are a number of problems associated with simulation of high-speed networks. As the bit rate increases then the number of events per second also increases at the bit level. If the bit level was to be simulated then the slow-down of the simulation would be linear with the increased bit rate. However most of the work presented here is simulation at the packet or cell level. If the packet size increased according to the speed of the network then there would be no real increase in the simulation time as the cell rate would remain the same going from low-speed networks to high-speed ones. However with ATM not only has the bit rate increased but the cell size has got quite a lot smaller compared to the

Chapter 2 Performance Modelling & Simulation

previous data networks. For example Ethernet frames are about 1500 bytes while ATM cells are only 53 bytes, and the bit rate on Ethernet is 10 Mb/s compared to 155 Mb/s on ATM. This gives an overall increase in simulation time from Ethernet to ATM of $\frac{155/53}{10/1500}$, which is over 400 if the same methods and complexity of models is assumed. This decrease in cell size is counter intuitive as the higher speed networks have better error characteristics than the slower networks, and so larger packets are more efficient and also possible. However for ATM it is expected that the error control will be end-to-end and not hop-by-hop. There is also the consideration of the real time services, and especially the low bit rate ones like voice, which could not be carried in large packets due to high packetisation delays.

An indication of the efficiency of the simulation can be expressed in terms of the speed-up or slow-down factor. The speed-up factor is the number of seconds that can be simulated of the real system for one second of the processor time. The slow-down factor is the number of seconds of the CPU time needed to simulate one second of the real system. Using a Sun Sparc-10 with SES/*workbench* the speed of simulation can be estimated to be about 4,000 events per second [61]. This is comparable to other simulators that are discrete event based like Op-net. If the bit level was to be simulated and one bit were one event, an over simplistic assumption, then the slow-down factor would be nearly 40,000 for a 155 Mb/s ATM link. This would mean that one second of the real system would take about 11 hours to simulate. However if cell level simulation is done then the slow-down factor is reduced to about 100. While the aim of the simulation would be to get the simulation to run as fast as the real system there is still room for improvement. Also in any reasonable size model there will be about 20-100 events per cell, so speed-up techniques of the order of 1,000 or so are still needed.

Even if the speed of simulation were not important, there are other problems with the memory of the computing platform. Again typical models allocate about 4 kBytes of memory to each basic simulation unit, be it cell or bit. Assuming that the simulation is at the cell level and that the ATM bit rate is 155 Mb/s, or 365,566 cells per second, and if the delay in the system were to be 0.02 seconds then the computer would have to track about 7311 basic simulation units, and as each of

Chapter 2 Performance Modelling & Simulation

them needs 4 kBytes of memory, the computer needs to have about 30 Mbytes of memory. A typical Sun Sparc-10 has about 28 Mbytes and while it is possible to increase this it costs more and does not really solve the problem. While in a lot of applications delays of this order would not be encountered there are a number of areas investigated in this work where they are. One is the long delay satellite links, where round-trip delays are of the order of 0.5 seconds. A second example is when there are multiple sources, and typical numbers of sources in a 155 Mb/s single link could be of the order of 100. Therefore the maximum delay is now only allowed to be 0.2 milliseconds. This only represents a buffer size of about 75 cells.

Therefore there is a need to examine alternative methods to both speed up the simulation time and reduce the amount of memory used in the simulation. There are a number of approximation techniques used in the literature [8, 9, 26, 33, 44], and some of these are described below as they are used in the simulations.

2.6.1 Decomposition Methods

Decomposition is a method of splitting the model into smaller models and then solving each independently. When this is done the results can be aggregated to form an approximate solution. What is nice about decomposition is that for a number of communication systems the decomposition gives exact results [33]. If a model can be made to be in a product form then the solution is exact and this method is generally called the flow-equivalence method. When the model is almost product form then the decomposition produces good approximations. What is surprising is that even for models which clearly violate product form criteria the results can be good. What makes the models here non-product form is that there are passive resources which inhibit the product form solution. The most common form of passive resource that is here is the window size that is used for retransmission schemes. If this could be ignored then this problem would be removed. There is also another problem in that there are a number of heuristics that are embedded in the model, as well as a number of priority schemes and these complicate the solution. However use is still made of the decomposition method.

When using some approximation methods it is difficult to produce error bounds on the results due to the nature of the approximation. This can be particularly true for ad-hoc methods and intuition is appealed to for the justification and acceptance. While decomposition is generally applied to states of a system it can also be applied to time periods as well. In time decomposition what is achieved is to split the model into many pieces, and each one is valid for a period of time. The more detailed model takes more time to run, but this may be only needed for a small fraction of the time and many orders of magnitude can be gained by this method [55, 58, 61].

2.6.2 Cell-Rate Methods

A cell-rate method considers a burst or group of cells to be the basic simulation unit rather than the cell [8, 18]. This allows a speed-up of the simulation as the number of units to be simulated has dropped. During a time period while the burst is active the inter-arrival time of the cells is constant and what changes in the simulation is the time a burst is active and the cell rate while in a burst. Because the inter-arrival time of the cells is constant in a burst it is possible to solve the multiplexing of sources in a buffer by considering the rates rather than the cells. A typical time period of interest might be as shown in Figure 2.11 and it can be seen that the speed-up can be considerable, for example it could be many orders of magnitude.

For this method to produce good results what is needed is a large number of cells to be represented by a single event or burst. This usually means that either : 1 the burst has a long duration and the cell rate can be small or else ; 2 if the burst length is short in time then the cell rate of the source should be high. For example the gain in using this method for voice would not give good results as the bursts are short and the bit rate is low, and hence the cell rate, is low also. However for data file transfer this can improve the simulation time considerably. What can be seen is that the design of the models should be made with the approximation technique in mind to get good results.

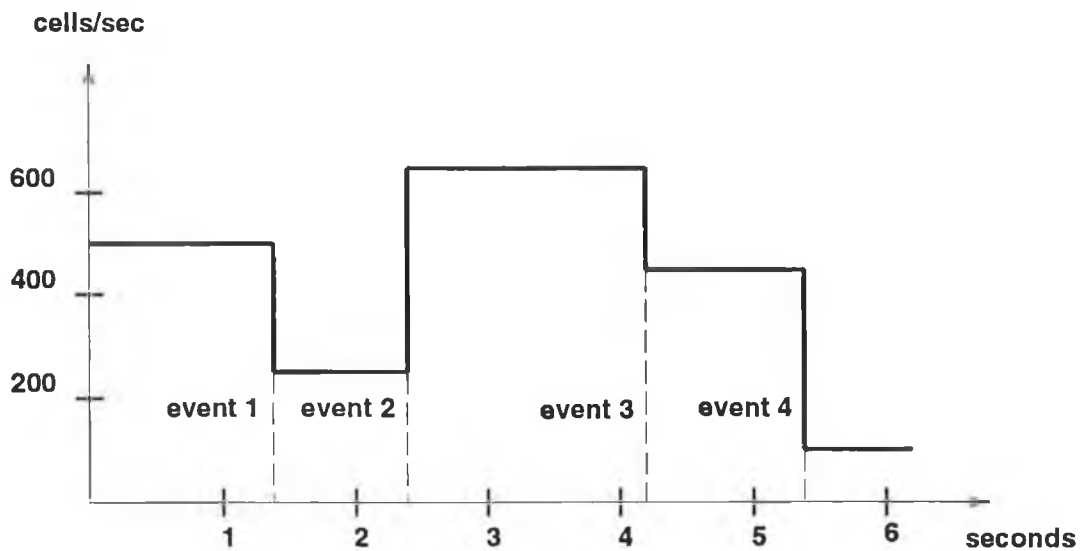


Figure 2.11: Events in a Cell Rate System

2.6.3 Step-By-Step Methods

A proposed approximation method is to break the possible times of simulation up into different parts and then to develop different models for each period. This type of simulation approximation is called a Step-by-Step method [9, 17]. While there can be many different models, it is sometimes possible to gain a considerable speed-up with just two models. The simulation would normally consist of one detailed model and one coarse model that can run quickly. The simulation runs with the coarse model until the detail in the detailed model is required and then the simulation steps into that model. Therefore it is only when the required detail is needed that the detailed model is used.

An example of this might be the loss from a finite buffer. A time interval could be chosen that is small enough so that not a lot of cells can arrive in one interval, and therefore it is possible to predict when cell loss is going to occur, but long enough to speed up the simulation. While the buffer is empty it is possible to step along in time and not to consider how the individual cells arrive. However if the buffer fills to a certain stage then the model is changed into the detailed model and maybe go at the cell time until the buffer empties and then swap back to the coarse model. Therefore at the end of each time interval the decision is made whether to stay in the coarse model or to change to the detailed model. This time

interval of course can change dynamically during the simulation. The time interval will decide on the accuracy of the simulation and the speed and there is the normal trade off between the two that the more accurate the longer the simulation time is needed and the shorter time spent in the coarse model, and the more time in the detailed model.

2.7 Discussion

The models presented in this Chapter detail the modelling design and simulation techniques that are used in the rest of this report. An introduction to modelling was given along with detail about modelling of ATM networks. Various source models were presented and validated from the design to the simulation model. The platform where the simulations were executed was discussed as well as the simulation package. Various problems were pointed out in connection with ATM networks and some approximate solution to these has been given. The whole process of modelling and simulation is circular in nature with many iterations of the design, simulation, analysis validation and approximations. All the individual stages depend considerably on each other so none can be treated in isolation.

Chapter 3

Cell Level Resource Allocation

3.1 Introduction

This chapter deals with cell level resource allocation issues. Of concern here is the structure of the ATM cell and how there is adaptation of the ATM to higher level services. There are problems encountered when using ATM in non-fiber optic systems and we investigate those limitations when using ATM over a satellite link. Again at the cell level the use of traffic controllers poses problems for using resource allocation methods due to the unclear position of the worst case traffic that can be allowed through the controllers.

It is shown here how it might be possible to adapt ATM to satellite links. ATM is a high-speed protocol designed with optic fiber as the intended transmission medium and therefore several problems arise when satellite channels are used. A solution is proposed here for the error control mechanisms to adapt to the satellite channel by moving the error recovery and detection to a higher layer of the ATM. The error recovery that is proposed here is to use the ability of the ATM to determine the service of the retransmission and to base recovery on that service. The simulation results here show that not only is there an increase in the raw data throughput for satellite channels to almost the theoretical limit, but there is an improvement in the data transfer efficiency of the ATM by 7.5%. The results also show that it is possible to guarantee data services with no loss of data under certain conditions.

ATM networks allow for the input traffic from users to vary both from one connection to another and within the connection. ATM specifies a method for controlling the traffic flow across the user network interface (UNI). This involves each user negotiating connection parameters with the network. Once these parameters have been decided then a contract is made between the user and network. The network must then enforce the contract in order to guarantee performance and quality of service to other users. What is of interest to the network is, given a particular set of users and contracts, what is the worst traffic that the users could input to the network while still abiding by their contracts? This type of input traffic would be called the worst case traffic as it would produce the lowest performance in the network. The reason why this is of importance is that for network performance it is important to have the worst case traffic inputs. Furthermore, it is important for the network traffic controller to know the possible worst case traffic so that it can assign resources accordingly.

The standards organisations have decided on an initial contract type for connection admission [4] and this is called the generic cell rate algorithm. This is not ruling out the possibility of other types of contract descriptions being implemented in the future, but allows development of systems at this stage. What is needed now is to decide what type of traffic can still pass these contracts and produce the worst or lowest network performance. This problem has been studied in the literature [23, 83, 84]. Some theoretical background is given here, to explain some of the results in the literature, and a further look is taken at some examples of types of worst case traffic sources. It is shown here for the two most common types, the greedy on-off and the three state source, that depending on the situation either can be worse [24, 62]. This would imply that neither of these types is generally the worst case traffic type as has been considered for the last number of years.

In the first part of this chapter the problem of using ATM over satellite is addressed, while in the second part the worst type traffic that can pass the generic cell rate algorithm is investigated.

1. In Section 3.2 the reasons for using ATM over satellite are motivated as well as

give a detailed description of the deep space network. The deep space network is used by JPL for communicating with un-manned space missions. Some problems with satellite ATM transmission are also examined. In Section 3.3 there is a description of the modelling of the satellite link and consideration is given to the decomposition methods used for the simulation. In Section 3.4 the issue of adapting the ATM to the services is considered. The problem with current adaptation methods is pointed out as well as a proposed solution and some notes on the ability of ATM to be service selective in terms of retransmission schemes. Section 3.5 deals with the simulation of the satellite link and the results that are gained from the simulation. There is also some detail on the validation and theoretical comparison of the results and the design of the adaptation scheme proposed.

2. Here the worst type traffic that can pass the generic cell rate algorithm is investigated. In Section 3.6 the cell level congestion and control area is explored and some detail is given about the standards approach and definition of congestion and usage parameter control. The leaky bucket is described as it is the standard usage parameter control. In Section 3.7 the issues that are dealt with here are introduced and a description of two worst case traffic types is given. There is also some discussion on the suitability of finite and infinite buffers in the analysis and the continuous and discrete variables. In Section 3.8 the new analysis used is presented for dealing with worst case cell traffic. Some supporting examples are then presented in Section 3.9 and both the case of a fast and slow server is analysed.

Finally there is discussion of the chapter in Section 3.10

3.2 Motivation For ATM Over Satellite

While ATM has been proposed for high speed services, at rates of 155 Mb/s and 622 Mb/s, there is increasing interest in ATM at lower rates [14], such as 1.544 Mb/s (T1). The interest of using ATM at low bit rates is stimulated by two main factors :

Chapter 3 Cell Level Resource Allocation

1. expensive, or lack of sufficient, bandwidth
2. lack of services requiring the larger bandwidths

It is possible to use other types of bearers like X.25 or frame relay at these lower rates, however these do not then capture the benefits of ATM like service mixing and guarantees on the service.

When dealing with satellite or mobile communication systems it is unlikely that the larger bandwidths will be available in the near future. However, there are a number of areas, civilian and military, where mobile or wireless operation is critical and there is a desire to investigate the performance and advantages of ATM in this area. There is also a current need to share the bandwidth across services in a more efficient manner. Multiplexing the services that are currently using the bandwidth is seen as an advantage of ATM.

3.2.1 JPL's Deep Space Network

The Jet Propulsion Laboratory (JPL) operates the Deep Space Network (DSN) for NASA. The aim of the DSN is to maintain communications with all the un-manned space missions. There is the possibility that the DSN may be used in future manned space missions [30]. Future space missions will require the integration of voice, video and data communications at increasing speeds and reduced costs. The DSN has two main segments as shown in Figure 3.1. The space segment will use the Consultative Committee for Space Data Systems (CCSDS) Advanced Orbiting Systems (AOS) protocols which end at the DSN ground stations [30]. The ground segment of the DSN consists of three main antenna sites, around the world. These are connected to JPL by commercial satellite links. The ground based segments may use the emerging standard for broadband communications of integrated services, ATM. ATM has been designed with the intended channel being optical fibre. The users of the space missions are connected to JPL and this is possibly going to be up-graded to ATM to support the new multimedia applications like tele-presence.

Tele-presence will play a role in future space missions. It is the use of

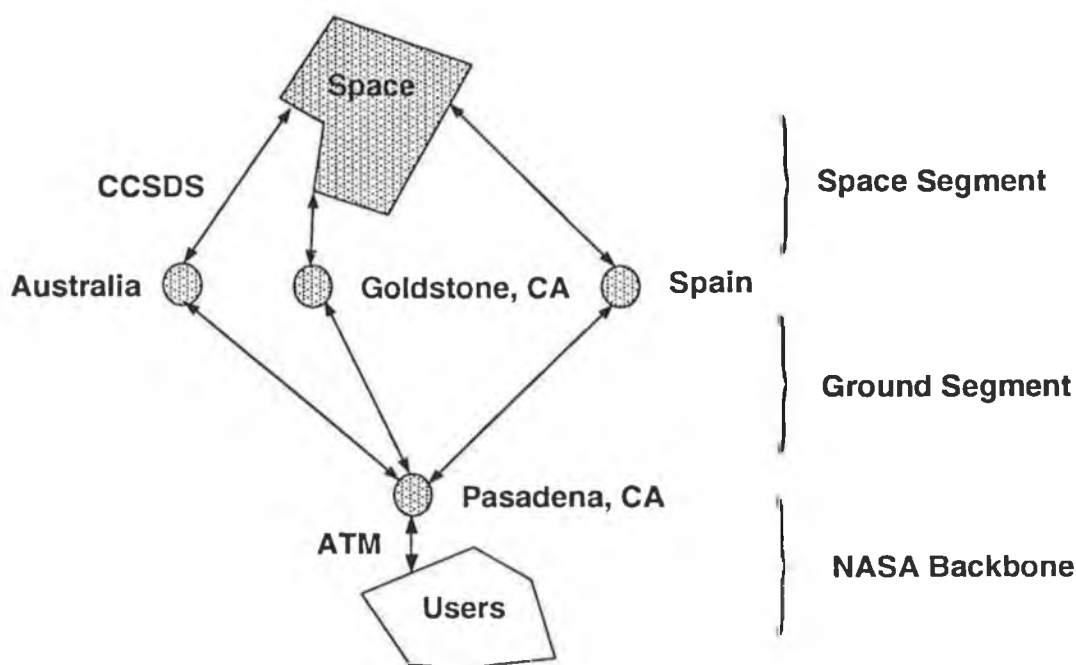


Figure 3.1: Deep Space Network

communications to provide a remote projection of human presence, for instance this can be done with robotic rovers remotely piloted by earth based scientists. Robotic rovers require stereo vision using low rate compressed images. Therefore tele-presence methods require transmission of integrated video and data. Future manned missions will require integrated voice, data, and video communications systems [31, 32]. Cost minimisation requires the establishment of standardised communications techniques that may support these services.

The bit rates being used on the commercial satellite links in the ground segment are generally T1 or less. The services that use these links vary from command files for spacecraft and ground distribution of telemetry, to voice communications among operators and test data. There is also the future possibility of using these links for digital video. The accuracy of the data from and to spacecraft is critical while other data and voice to the other earth sites may not be as important. At present there is no automatic method of differentiating between the different services that use the network.

3.2.2 Problems Using Satellite Links For ATM

Because ATM has been designed with optic fiber as the expected transmission medium this means that the expected errors will be produced by Gaussian noise and will introduce random geometrically distributed bit errors. The bit error probability is denoted by p_b and can be expected to be approximately $p_b = 10^{-10}$ for fiber optic cable. ATM cells are 53 bytes long, or 424 bits, and so it is possible to calculate the probability for cell error. It is possible to further calculate how many of these errored cells have just a single bit in error, and then work out what the probability of more than a single bit in error is.

$$\begin{aligned} \text{Pr}_{\text{cell in error}} &= 1 - (1 - p_b)^{424} \\ &= 4.2410^{-8} \end{aligned}$$

$$\begin{aligned} \text{Pr}_{\text{single bit in error}} &= 424 p_b (1 - p_b)^{423} \\ &= 4.2399 10^{-8} \end{aligned}$$

$$\begin{aligned} \text{Pr}_{\text{more than single bit in error}} &= \{ 1 - (1 - p_b)^{424} \} \\ &\quad - \{ 424 p_b (1 - p_b)^{423} \} \end{aligned}$$

Using an expansion for the $(1 - p_b)^n$ term the following equations are got, and a bound on the error on the result from the series expansion because of the alternating series can be found :

$$\begin{aligned} \text{Pr}_{\text{more than single bit in error}} &= \{ 1 - (1 - 424 p_b + \frac{424 \cdot 423}{2} p_b^2 - O(p_b^3)) \} \\ &\quad - \{ 424 p_b (1 - 423 p_b + O(p_b^2)) \} \\ &= \{ 424 p_b - \frac{424 \cdot 423}{2} p_b^2 + O(p_b^3) \} \\ &\quad - \{ 424 p_b - 424 \cdot 423 p_b^2 + O(p_b^3) \} \\ &= 89676 p_b^2 - O(p_b^3) = 8.967 10^{-16} \approx 10^{-15} \end{aligned}$$

Chapter 3 Cell Level Resource Allocation

What can be seen is that nearly all the errors will be single bit errors or less than 1 in 10^7 errors will be more than a single bit error. In fact for a 155 Mbps system, more than a single bit error would only occur every 80 years or so. Therefore error detection and single bit error correction will almost fully protect the data.

In satellite links bursty errors are common as will be seen in Sub-Section 3.3.2. These bursts may be a result of Gaussian noise, but rather than producing randomly distributed bit errors they produce a large group of bit errors due possibly to the coding of the channel. It is also possible that degraded performance may occur on the link causing increased bit errors which may destroy the link for a fraction of a second or more. On top of these bursty errors there is also background Gaussian noise which might produce error of the order of 10^{-6} . This introduces a different type of error distribution to what can be expected from optic fiber cables. The probability of having more than a single bit error, can also be calculated :

$$\begin{aligned} \text{Pr}_{\text{more than single bit in error}} &= \{ 1 - (1 - p_b)^{424} \} - \{ 424 p_b (1 - p_b)^{423} \} \\ &= 89676 p_b^2 - O(p_b^3) = 8.967 \cdot 10^{-8} \approx 10^{-7} \end{aligned}$$

This would mean that the same 155 Mbps system would have more than a single bit errored cell every 30 seconds or so and 1 in about 4000 errors will be more than single bit errors. Therefore single bit error correction is not likely to protect the data. The probability for loss of the data and the header of an ATM cell is shown in Figure 3.2 versus the bit error probability for Gaussian noise.

It is seen that for optic fiber media there is about one cell lost every century for a 155 Mb/s link whereas for satellite links the loss will be about one cell every minute. Therefore different error-detection methods must be implemented when using ATM over satellites. If the errors were likely to be all bursty then it is possible to use the cell header to protect the data and the incorrect cells could be discarded. For example if the bit error probability was 10^{-7} but all the errors occurred in blocks of 400 bits, then when an error is got either one or two cells will be in error. The cell error probability is then less than $2 \cdot 10^{-7}$ which is better

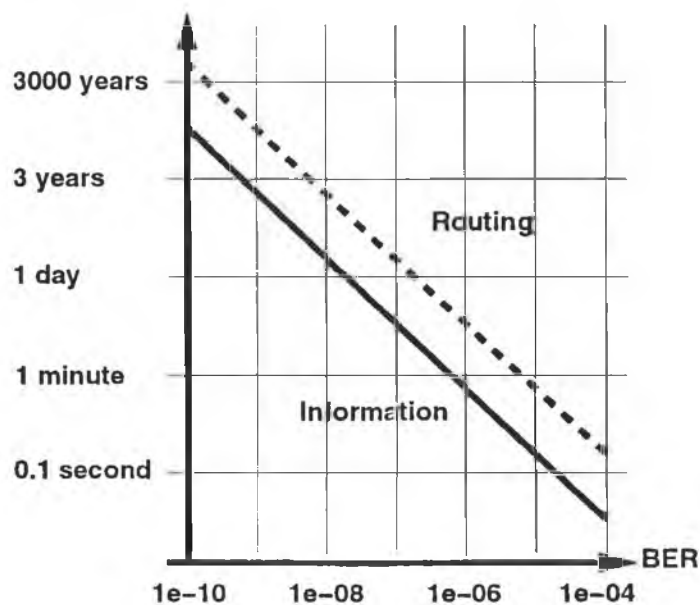


Figure 3.2: Loss For 155 Mb/s Link For Various Bit Error Rates (BER)

than optic fiber with $p_b = 10^{-9}$. However with both bursty and random errors there have to be other schemes implemented to combat the errors and to protect both the ATM cells and the data that is contained within them. There are many schemes proposed to adapt ATM to non optic fiber environments [7] however these do not take into account the type of error found on the satellite links used by JPL. With any of the schemes proposed so far there is always the probability that the cell header may be correct and the data contained within the cell is corrupted. It is assumed that there are some higher layer protocols, like TCP/IP, that look after these errors. However in the DSN this may not be the case.

Another problem in using ATM over satellite links is that the delay of the link is large and therefore there can be a large number of cells, or indeed higher layer data units, outstanding on the link at any time. When designing the numbering scheme in ATM, and the adaptation layers, for fiber optic cables the delays are relatively small. This means that the numbering schemes in ATM may not be sufficiently large to allow the satellite link to operate in continuous transmission mode. This problem is called sequence number starvation [6] and can be overcome by either : 1 allocating more bits, and therefore more numbers, for the same cells or data units; or 2 numbering larger data units than before where the cells were

numbered.

3.3 Measuring & Modelling A Satellite Link

The satellite link is characterised by the delay across the link and the bit rate of the link. In this model the delay across the link is taken to be 270 milliseconds, and the bit rate of the link is T1, or 1.544 Mb/s. There is also the error due to incorrect decisions on the bits at the receiver. There are many models proposed to model this bit error and most are based on reasoning why the errors occur. Rather than take one of the predefined models or be involved in the actual coding and transmission schemes being used on the link, an empirical model based on actual test data taken from the DSN is used here. There is an amount of data that was collected from the DSN and the method in which this was collected is described in Sub-Section 3.3.1 and then the actual data is presented in Sub-Section 3.3.2. While looking at the data it is important to keep in mind that the end result will be to perform simulations. There are problems in simulating long delay systems and so the modelling of the noise is critical.

3.3.1 Experimental Test Set-up

There was a study done on the DSN satellite links between the start of January 1992 and the end of October 1992, a 10 month period [50]. This report examined the satellite links to produce results for both the circuit availability and also the circuit error characteristics. The only part that is of concern here is the circuit error characteristics. The Ground Communications Facility (GCF) at JPL consists of satellite links and the digital connections that connect the satellite links to the central site in Pasadena, CA. The unit of data that is being used on these links is a NASCOM standard where the blocks of data are 4800 bits in length [50]. The links that are of concern to the report are the prime and backup circuits which are :

Chapter 3 Cell Level Resource Allocation

- Prime Circuits
 - Goldstone System 2, 1.544 Mbps, direct satellite
 - Canberra, 512 kbps, direct satellite
 - Madrid, 512 kbps, direct satellite
- Backup Circuits
 - Goldstone System 1, 1.544 Mbps, direct satellite
 - Canberra, 64 kbps, undersea cable
 - Madrid, 56 kbps, terrestrial & satellite via GSFC

It is only the satellite links that are directly connected that are of concern, which are all the prime circuits and the Goldstone backup circuit. 12 error rate tests were conducted as part of the study, and these results were reported on in the report. It was also possible to examine the receiving test set that was stored on electronic tape. Naturally the error rate tests were not carried out on the whole of the link capacity, as the links had to remain operational at all times. The error rate tests are shown in Table 3.1.

The tests numbered 2, 3 and 8, in Table 3.1, are not all satellite connections and so are not analysed here. In the tests numbered 5 and 6, in Table 3.1, it was found out, after the tests were completed, that there was not enough transmitter power being used and so these will not be analysed here either. So the tests that are further examined are 1, 4, 7, 9, 10, 11, 12. The error rate tests that were carried out were one way tests where the data being sent was generated and received at the far end by a “Firebird” test set. These test sets count the bit errors, block errors, blocks transmitted and other statistics. Every 10 minutes the contents of the test set is available at the output.

3.3.2 Experimental Data For Error Distribution

The experimental data used to model the error distribution are from fractional T1 links, which vary in bit rate from 56 kb/s to 224 kb/s and the tests were taken over a period from less than a day to about two days. The results from the tests, that were described in Sub-Section 3.3.1, are summarised in Table 3.2.

Table 3.1: Error Rate Tests On The GCF Of The DSN

Test	Link	Bit rate in kbps	Duration in hours
1	Canberra to JPL prime	56	41
2	Canberra to JPL backup	64	45
3	Madrid to JPL backup (via GSFC)	56	51
4	JPL to Goldstone prime	56	44
5	Madrid to JPL prime	112	7
6	Madrid to JPL prime	56	52
7	Goldstone to JPL prime	224	17
8	JPL to Madrid backup (via GSFC)	56	51
9	JPL to Goldstone prime	224	10
10	JPL to Goldstone prime	224	24.5
11	Canberra to JPL prime	224	35
12	JPL to Madrid prime	56	51

Table 3.2: Detailed Selected Error Rate Tests

Test number	Bit rate in kbps	Duration in hours	Blocks transmitted	Block errors	Bit errors	Bit Error Rate
1	56	41.28	1,728,990	0	0	0
4	56	43.7	1,835,213	16	3,151	3.58E-7
7	224	16.67	2,799,954	11	191	1.42E-8
9	224	10	1,679,379	17	28,813	3.57E-6
10	224	24.5	4,115,943	36	58,420	2.96E-6
11	224	34.78	5,843,957	575	1,376,946	4.91E-5
12	56	50.59	2,549,630	0	0	0

A total of 10^{11} bits of data were used from the tests that were done to produce the models here. The overall bit error rate from these tests exceeds the

NASCOM standard of 10^{-7} on 4 of the 7 tests done. More detailed examination of the results will show how many bits are in error for each error event. An error event is when a number of bits are in error contiguously. From these results it is possible to examine the duration of each error event and from this a model of the error distribution can be found.

3.3.3 Model Of A Satellite Link

While the data was being analysed possible decomposition methods were in mind. With time decomposition it would be possible to speed up the simulation by many orders of magnitude. The term burst event is defined to mean a time duration when there are a number of bits in error either back to back or with smaller bursts of errors periodically occurring over a short time, as can be seen in Figure 3.3. It may be possible to have single bit errors occurring as well within the burst event. A non-burst event is defined to mean a time when there are only single bit errors which occur in a random manner. Therefore the result is a combination of burst event errors and single bit errors due to Gaussian noise as is shown in Figure 3.3.

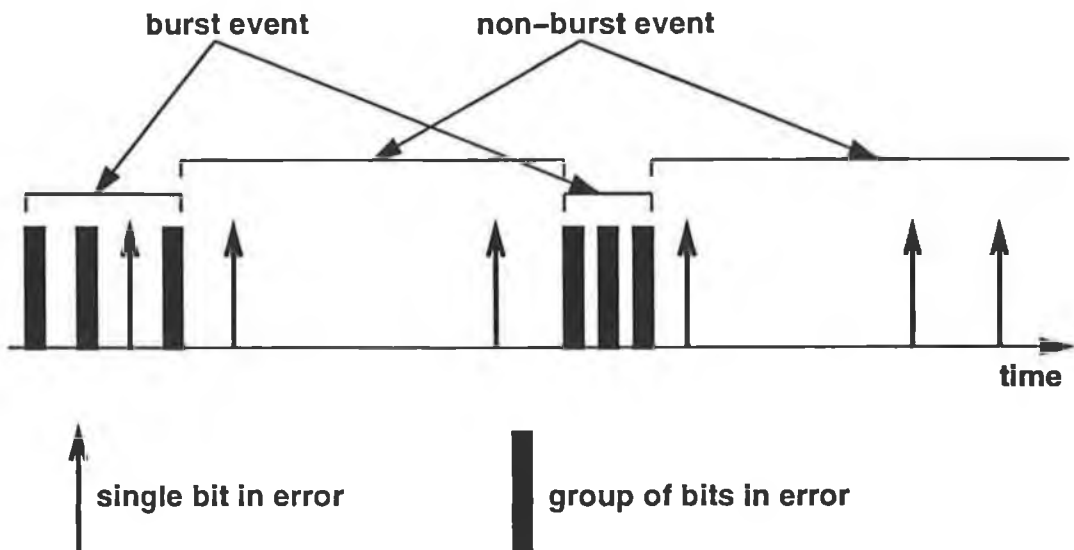


Figure 3.3: Burst Event And Non-Burst Events

Using this distinction between non-burst events and burst events, it is found that the links operate in the non-burst event mode for about 99.75% of the time, and so operate in the burst event mode for about 1/400 of the time. Because

Chapter 3 Cell Level Resource Allocation

the single bit error mode of operation can be analysed separately and easily, it is now possible to use time decomposition between the burst events and the non-burst events. The reason why it is possible to analyse the single bit error mode differently is that in most cells there will be only a single bit in error, for example when $p_b = 10^{-7}$ the probability of cell error is $4 \cdot 10^{-5}$ of which all but 10^{-9} are single bit errors. In the header of the cell the probability of error is $4 \cdot 10^{-6}$ of which all but $8 \cdot 10^{-12}$ are single bit errors. The single bit errors will be corrected by most CRC algorithms, including the ATM header one. The probability of undetected cell header error is sufficiently low to be neglected at $8 \cdot 10^{-12}$, which is only one cell in error every 4 days on a 155 Mbps link. If only the burst events are simulated then there will be a speed-up factor of about 400.

What is needed now is to look at the burst event in detail. During this time there will be periods of complete loss of the bits and there will be single bit errors occurring as well. To model the groups of continuous bit errors both the duration of the group and the interarrival of these groups have to be determined. Both of these distributions are modelled by empirical distributions. Thus there are two distributions :

1. distribution of the time between the groups of errors, called the inter arrival of burst errors which is shown in Figure 3.4;
2. distribution for the duration of the burst of errors, called the length of burst errors which is shown in Figure 3.5.

In the burst event the probability distribution for the inter arrival of burst errors is shown in Figure 3.4 and the mean time between bursts is approximately 36 seconds, and the variance is 24 seconds. The duration of the burst itself is given by the burst length and the probability distribution for this is shown in Figure 3.5 and the mean burst duration is 0.15 seconds with a variance of 0.26 seconds.

Thus a burst may have tens of thousands of bits in error continuously. Both the burst length and the burst inter arrival are modelled in terms of time rather than in terms of bits so that this can be applied to different bit rate links. The possible causes of this error distribution is unclear and may be a combination of the

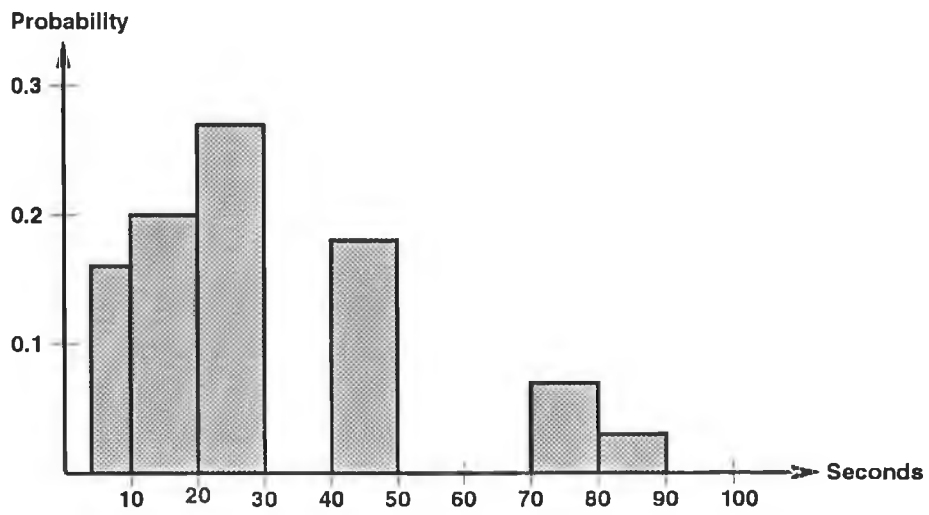


Figure 3.4: Probability Distribution For Burst Inter Arrival

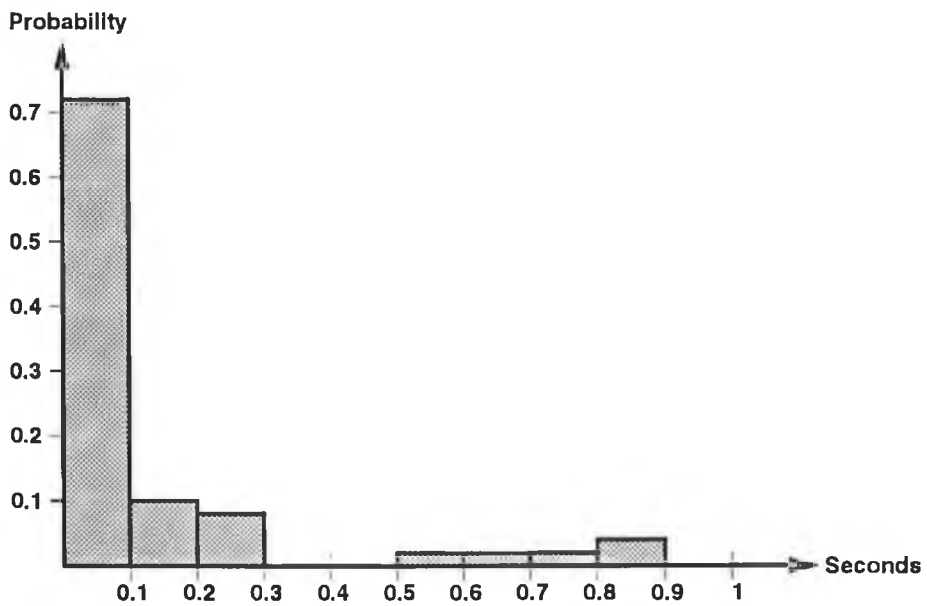


Figure 3.5: Probability Distribution For Burst Length

coding schemes used in the satellite systems and the interference on the physical link. This alone is a difficult area of study and will not be treated here. The causes of the errors can be due to the coding, the weather, airplanes, or indeed faulty equipment.

3.4 ATM Adaptation Layer Over Satellite

3.4.1 Problems With Current ATM Adaptation Layer

The ATM protocol has three layers corresponding approximately to the bottom two layers of the OSI model [22]. The bottom layer, which is called the physical layer, is concerned with the physical transmission of the bits. The second layer, the ATM layer, takes a payload of 48-bytes and puts a 5-byte header on to it to form a 53-byte cell. This header has the routing and addressing information in it, as well as an 8-bit cyclic redundancy code (CRC) that detects header errors [43] and corrects single bit errors. The 48-byte payload comes from the third layer, the ATM Adaptation Layer (AAL), which adapts the cells to the different services. The lower part of the AAL is called the Segmentation And Reassembling (SAR) sublayer which breaks a message up into cells. There are a number of proposed AAL's specified for different applications [20, 43]. The proposed AAL for data transfer is AAL 3/4, which uses a 44-byte information load and a 2-byte header and trailer as shown in Figure 3.6.

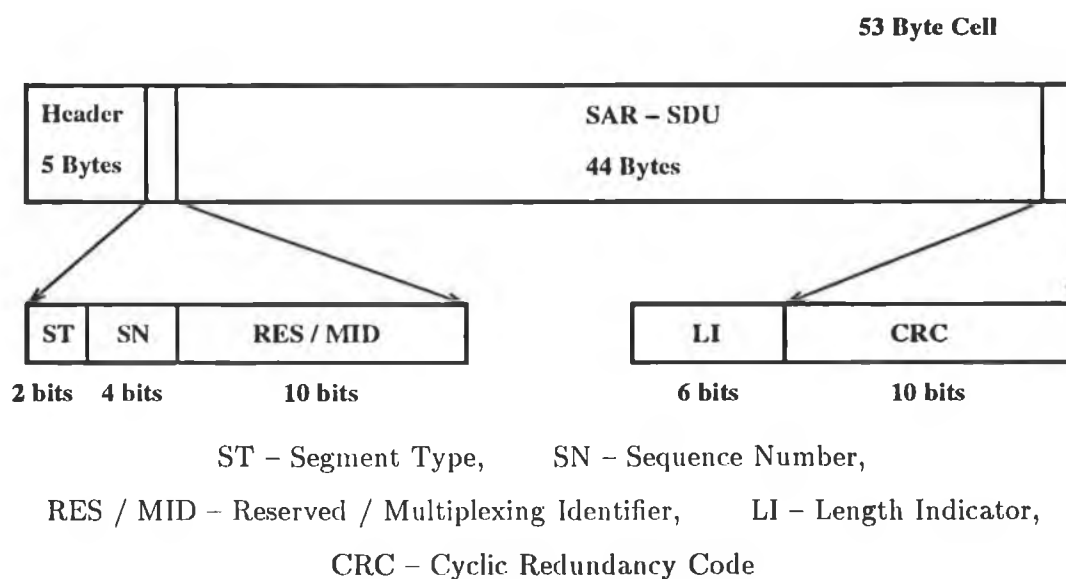


Figure 3.6: The AAL 3/4 Cell Format

In the trailer of AAL 3/4 there is a 10 bit CRC that protects the information

Chapter 3 Cell Level Resource Allocation

load from errors. However due to the small CRC, the potential undetected error rate is high. For the 10 bit CRC there are 2^{10} different polynomials that can be used, no more. If the SAR-SDU suffers an error there is then a $1/2^{10}$ probability that it will match the polynomial, or the undetected rate could be as high as 10^{-3} . There is also no correction for double bit errors.

There is also another problem in that there are only four bits allowed for the sequence number. This means that there can be at most 2^4 SAR-SDU's outstanding on the link, if there is to be a retransmission scheme working correctly. If the round trip propagation delay is 540 milliseconds then the maximum throughput will be given by :

$$\text{Maximum Bit rate throughput} = \frac{53.8 \times 2^4}{0.54} = 12,563 \text{ kbps}$$

This limits the maximum transmission rate for a single link. Even though there are methods that can be used to overcome this problem, it is a non-ideal solution, and the maximum throughput should be larger. To get to a stage where you could have 155 Mbps throughput there would need to be 18 bits for the sequence number. This would impose 4.7% overhead, or an extra 3.6% overhead from the previous case, and this is probably unacceptably high.

3.4.2 Proposed ATM Adaptation Layer

It is proposed here that it is more efficient to consider moving the CRC and the sequence numbers to the higher Convergence Sublayer (CS) [6]. By moving the sequence number to this layer larger blocks are got to put error detection and correction data on. For the AAL 3/4 the overhead is about 8% but to modify that for a T1 satellite link an 11 bit sequence number would be needed. Also by moving the CRC to the CS level, and making it a 32 bit CRC, the undetected error is reduced to less than $1E-9$. The ATM standards specify many AAL's that can be used, but it is permissible to specify a different AAL for a particular application. It is suggested here that high-speed, long-delay satellite links need a unique AAL. This can be achieved by inserting the sequence number and the CRC at the CS sublayer, requiring a type of framing. The AAL that is proposed here is capable of

supporting 155 Mb/s throughput with a 270 milliseconds delay on a single channel. The proposed framing structure is shown in Figure 3.7.

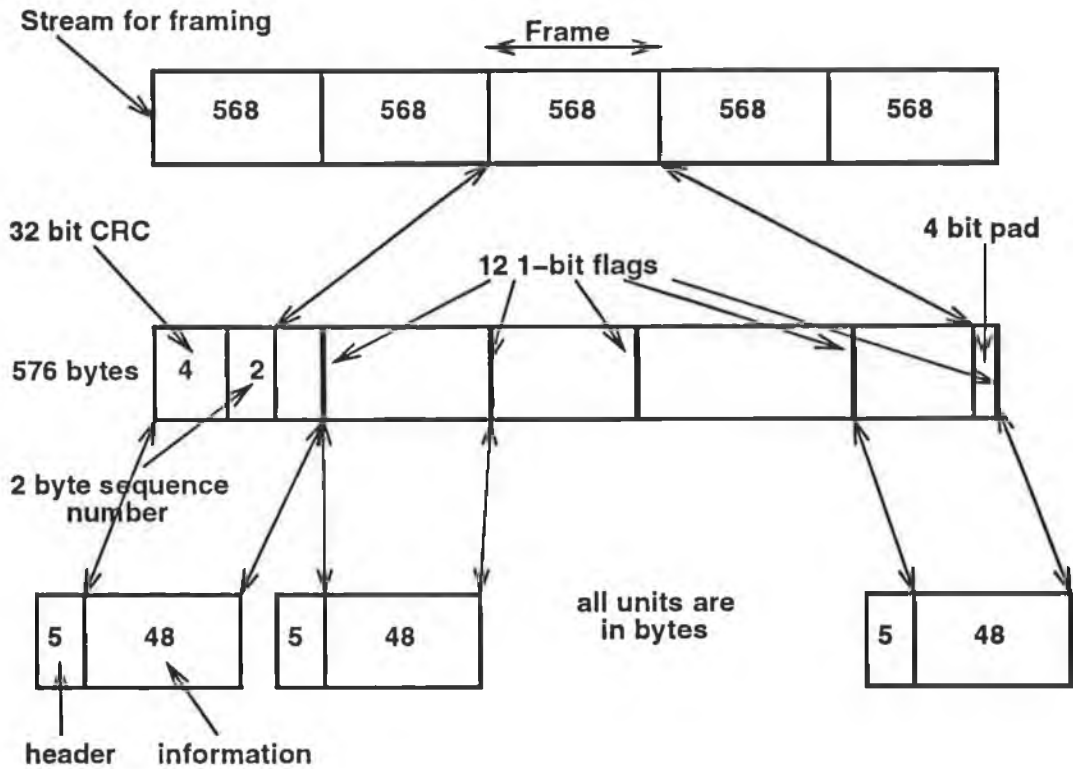


Figure 3.7: Proposed Convergence Sublayer AAL

An important consideration of the framing structure is that it should be possible to put a number of cells into the frame without stuffing it. Therefore an integer number of cells is chosen to make the frame. The performance of the link will depend on this size of frame and a compromise between the overhead associated with the frame and the length of the frame is needed. It will be seen that a frame size of about 12 cells is optimum. This means there will be 636 bytes in the frame of which 568 bytes are information bytes. This gives an upper bound on the efficiency of about 89%. Compared to the AAL 3/4 efficiency of about 83% this is a 7.5% improvement in throughput in terms of overhead alone. The 636 bytes in the frame are made up of the data, the ATM headers and the frame overhead. To protect the data from errors there is a CRC-32 which takes 4 bytes and this provides a lot more protection to the data against errors than before with the AAL 3/4 which is a major factor for the critical data.

3.4.3 Service-Based Retransmission Scheme

It is also proposed that there should be service selection to improve efficiency for reliable delivery. This is needed because real-time data and voice usually cannot tolerate delays of one and a half times the round trip propagation delay. For example, voice will tolerate the 270 milliseconds propagation delay, but if it is in error, it will take another 540 milliseconds for retransmission, even assuming no congestion delay. The total delay is 810 milliseconds, which is unacceptable. Therefore there is no use in trying to detect errors or retransmit voice cells. In fact, any service that is delay sensitive in the sense of less than a few seconds is considered here to be real time and so is not framed for error detection and retransmission. For real-time services no framing is required. Even if the link is fully occupied the room for the retransmission of reliable data is got by discarding some of the real-time cells, e.g., the voice cells. The way in which the real-time cells mix with the framed cells is shown in Figure 3.8. It can be seen that the real-time cells only enter the buffer if they expect to have less than a 50 millisecond queuing delay. Also the framed cells are kept in a time out buffer in case they are needed for retransmission.

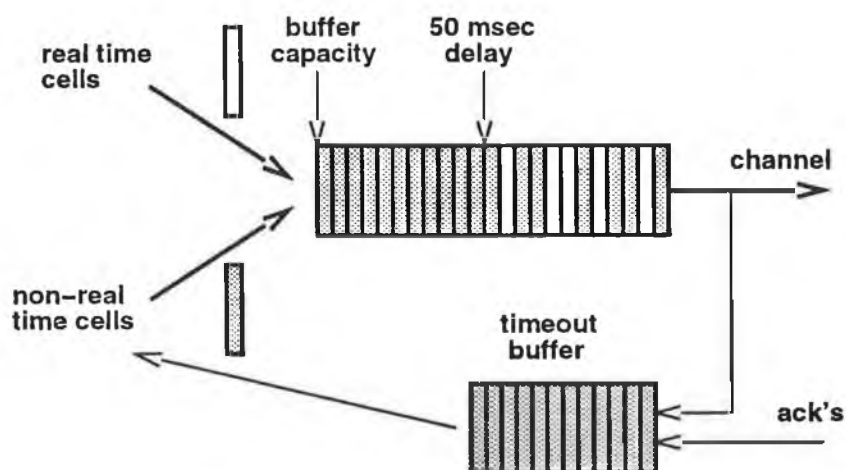


Figure 3.8: Proposed Retransmission Scheme

It can be seen that the frame is at the CS sublayer and takes 568 bytes of user data and adds 8 bytes of overhead to form a 12 by 48-byte frame to be split up by the SAR into 12 cells. The 8 bytes of overhead is made up of a 32-bit

Chapter 3 Cell Level Resource Allocation

CRC, which could be the standard IEEE 32-CRC [6]. There is a sequence number of 16 bits which allows a full 155 Mb/s of data on a link with the delay as given above and the size of frame chosen above. For the optimal size frame the error characteristics need to be known in advance, however the probability of bit error will be in the order of 10^{-7} for satellite links, and possibly some degradation to around 10^{-6} should be allowed for. Knowing this and that the overhead on the frame is 8 bytes then the optimal frame size is 992 bytes which is in the same region as what has been chosen. It is shown in the results that a frame size of about 12 cells is a fair estimate for the conditions of the channels investigated here. Also 12 cells in a frame corresponds reasonably well to the currently used NASCOM standard block size of 4800 bits [50]. The algorithm that the transmitter uses to send framed and real-time cells is shown in Figure 3.9.

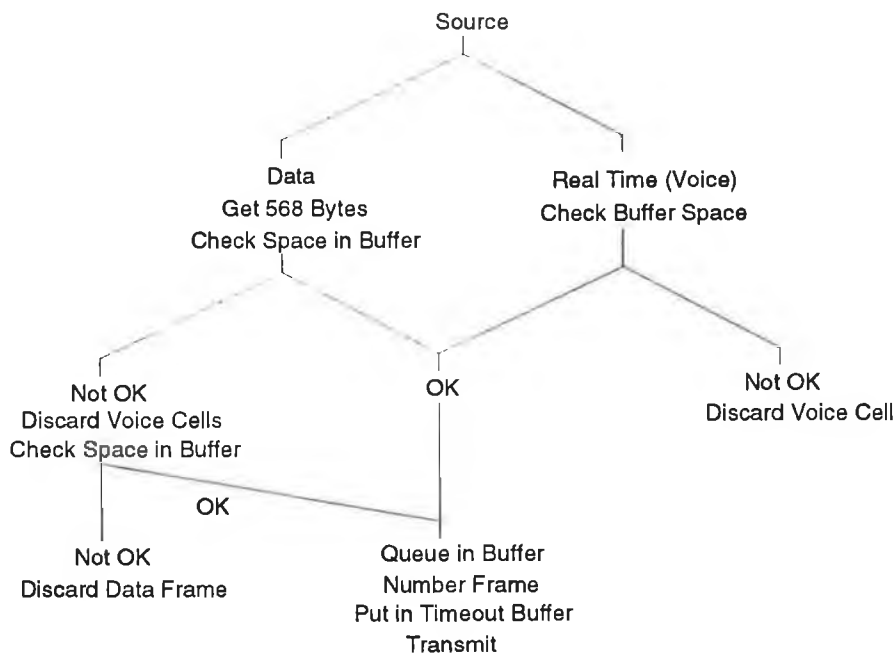


Figure 3.9: Transmitter Algorithm For Selective Retranmission

At the receiver there is a problem when a cell is lost in that it is not known whether it is a voice cell or is one of the framed cells. If the cells from one frame are sent back to back on the link then these problems would not occur, but that takes away from a lot of the advantages of ATM. Instead the VCI gets the cells in order for the framing. However when the header is lost there are problems. If the header could be decoded correctly then it would be possible to conclude from

Chapter 3 Cell Level Resource Allocation

the VCI what type of cell it was. However with no header, all that is known is that a cell has been lost. What could happen if the cell were framed is that there would not be the required number of cells in the frame and the framing would lose synchronisation. To overcome this and to maintain the framing synchronising, a one-bit flag is proposed, and this coupled with the VCI will show the start of the frame. This one-bit flag is the last bit of the cell, for framed cells, and tells whether the cell is the first in the frame or not depending on whether it's a 1 or a 0. The algorithm that the receiver uses to receive framed and real-time cells is shown in Figure 3.10.

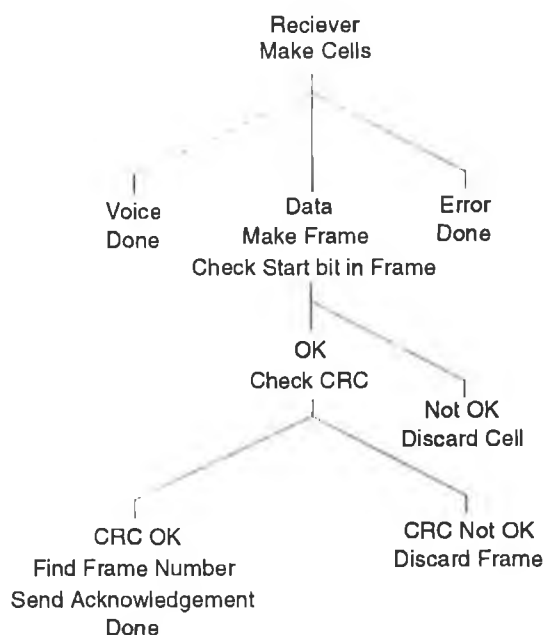


Figure 3.10: Receiver Algorithm For Selective Retranmission

The efficiency that is gained from this proposed ATM adaptation layer is twofold :

1. there are only 8 bytes of overhead on a frame of 568 user information bytes and
2. there are the sequence numbers to transmit over long delay links which allows large throughput and possible retransmission schemes to be implemented.

The gain in efficiency compared to AAL 3/4 is that instead of carrying 528 bytes

of user information bytes in 12 cells 568 bytes are carried, which is a 7.5% increase in efficiency over the whole frame.

3.5 Simulation & Results For Satellite ATM

3.5.1 Simulation Details

The proposed models were simulated using *SES/workbench*. *SES/workbench* is a discrete-event package that allows hardware and software simulation. The model of the retransmission scheme and the error probabilities here are created by use of the graphical interface. The *SES/workbench* compiles the code to C and runs on a four-processor Sparc-Server 629. Each simulated data point is the equivalent of 17 days on a T1 satellite link, but rather than modelling the $2E+12$ bits, only the burst events are simulated. This reduces the simulation to $5E+9$ bits and further optimisation of the model compresses this to $5E+7$ events. In real time this can take between 3 to 36 hours of run time as this configuration of computing can execute between 400 to 5000 events per second depending on the complexity of the event and the detail of the model. In total 1300 hours of CPU time, or nearly 8 weeks continuously computing, was expended in the simulation tests. Without the speed up in the modelling and simulation it would have taken at least 1000 years to get the equivalent results.

The model consists of the generation data cells, which are framed, and real-time cells, as is shown in Figure 3.11. There is a buffer at the transmitter, which varies in size, that can be used to influence the throughput. The real-time source cells check the delay in the transmit buffer before joining the queue. If the delay is more than a specified value, in this case 50 milliseconds, then the channel is assumed to be in burst error and the cells are discarded.

The data source generates cells, and these are framed together first. Then when a frame is ready this is attempted to be placed in the transmitter buffer, if there is space available. If the transmitter buffer is full, it checks to see if there are any real-time cells there, and if there are they are removed to make room for the

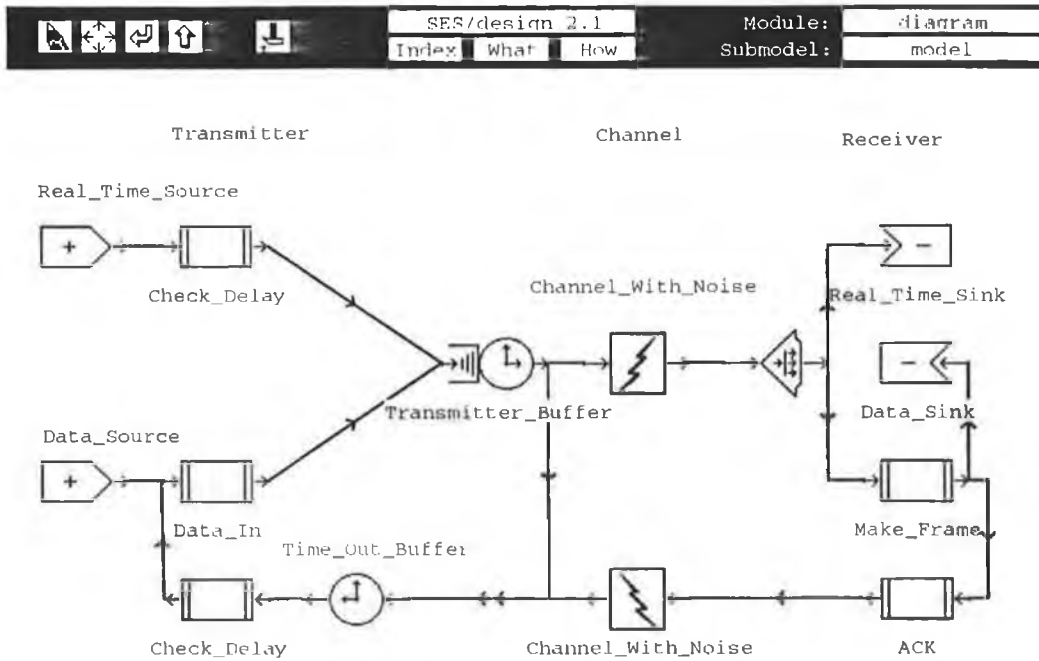


Figure 3.11: Simulation Model

framed data cells. If there is still not enough room then the frame is lost. However if there is room in the transmitter buffer, then the frame is added to the buffer and given a sequence number with modulo 65,536. When the frame is transmitted the frame is also kept in the time-out buffer in case it is needed for retransmission. This buffer is 111.3 kbytes, which accommodates the two-way propagation delay plus a small processing delay.

The channel has two sources of noise, burst errors and Gaussian noise. At the receiver the real-time cells and all the errored cells are released, while the correct framed cells are kept to attempt to make a frame. If any bits in the frame are in error, or if a cell is missing, which is seen by a missing one bit flag, then the whole frame is discarded. If the 32 bit CRC of the frame is okay then an acknowledgement is sent back to the transmitter, through a channel in which it also may encounter errors.

When an acknowledgement arrives from the receiver, the frame which corresponds to that sequence number, is discarded from the time-out buffer. Otherwise after the time-out delay, the frame in the time-out buffer, skips to the head of the transmitter buffer, after checking that there is sufficient room in the buffer,

and is retransmitted. If there is not sufficient space then real-time cells are discarded from the buffer, and if this does not release enough space then the frame is discarded.

3.5.2 Simulation Results

It is possible to guarantee delivery of the framed cells, with the given error probabilities by means of retransmission providing there is either excess bandwidth to allow for the retransmission or else the percentage of real time services is reasonably high. However if the mix of framed cells to real-time cells is high then some framed data will be lost due to the burst errors. Tests were carried out for three types of errors, first only bursty errors, then only Gaussian errors, and finally for both Gaussian and bursty errors. For each of these types of errors the effect of the buffer size is investigated. Also, the effect of changing the framed data mix on the cell loss is investigated.

For burst errors, the cell loss for framed data varies depending on the mix of framed data and on the transmitter buffer as is shown in Figure 3.12. The cell-loss rate for framed data can be decreased to whatever value is required by either backing off the framed data mix or increasing the transmitter buffer size. However as the percentage of framed data gets toward 100% the effect of the buffer is small and the graph becomes almost vertical meaning that either all the data gets through or gets lost.

As the loss of framed data cells is dependent on the mix of traffic types, it becomes important to find the mix of framed data to unframed data that allows all the framed data to get through without error. For both bursty error and Gaussian noise the buffer required for no framed data loss is shown in Figure 3.13.

As expected, as the percentage of framed data increases towards 100% the buffer size required for no loss increases fast and the buffer size becomes unreasonably large. It is therefore apparent that the best scheme is to have a reasonable mix of framed data and real-time traffic.

Gaussian noise introduces mainly single frame errors for bit error rates lower

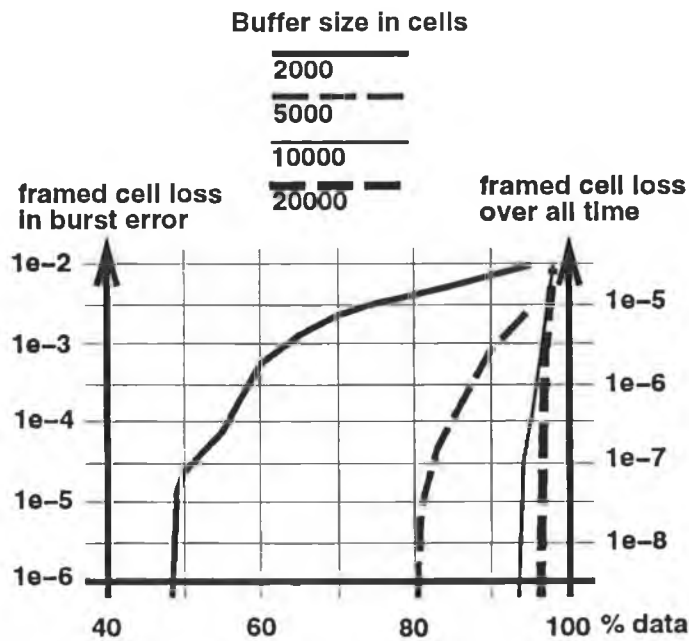


Figure 3.12: Framed Data Loss For Burst Error, Varying Buffer Size

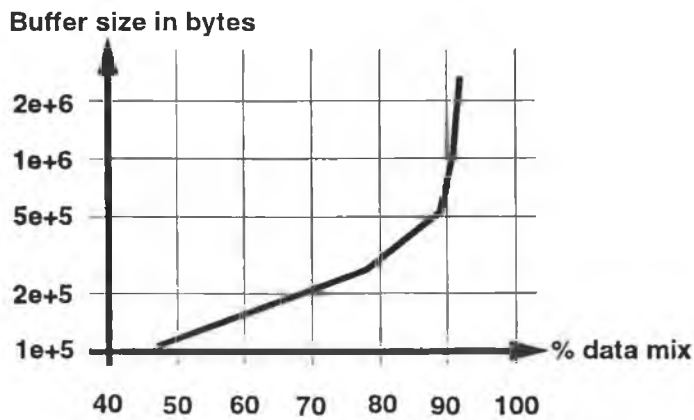


Figure 3.13: Transmitter Buffer Required For No Framed Cell Loss

than 10^{-5} . Hence the Gaussian error model on its own will not be sensitive to transmitter buffer size. This can be seen by comparing how the bursty error, with a traffic mix of 90% data, and the Gaussian error, with a traffic mix of 95%, vary the framed data loss probability with transmitter buffer size, as shown in Figure 3.14. The variation on loss by varying the transmitter buffer size is almost negligible in the Gaussian error case compared to the large variation in the burst error case.

The gain of getting no loss for the framed data cells has a penalty side by having to retransmit the ones in error. If there is no excess capacity, as is assumed

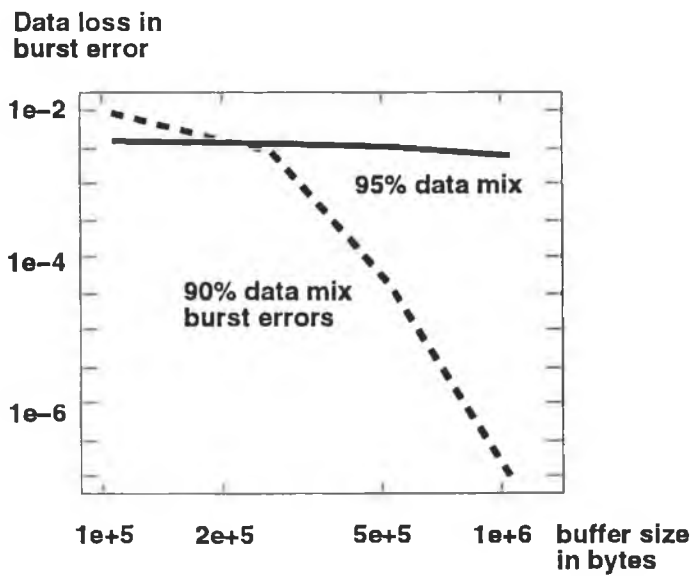


Figure 3.14: Effect Of Varying The Buffer Size

in this model, then there is going to be loss in the real-time cells. To see the effect of this loss, a plot of real-time cells, or voice cells, lost due to burst events for various transmitter buffer sizes and against the framed data mix is shown in Figure 3.15.

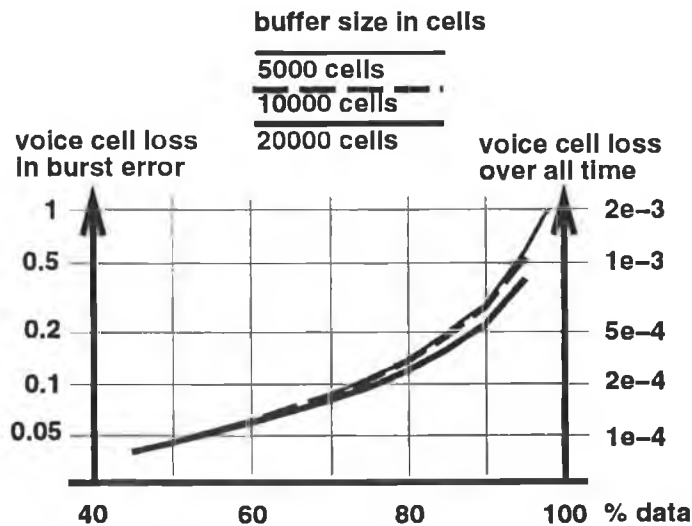


Figure 3.15: Real Time Cell Loss, Varying Buffer Size

In the burst event it is possible to lose all the voice cells if the mix of framed data is high enough. This means that there is no more bandwidth available for retransmission. However this loss when averaged out over the day is of the order of

1E-3. Also the loss is concentrated into intervals of seconds. There may be schemes to lessen the effect of the loss of this voice and spread it out more evenly. However if the critical objective is not to lose framed data this is the price that may have to be paid. Also the size of the transmitter buffer has almost no effect. This is because when the delay expected in the transmitter buffer is more than 50 milliseconds, the real-time cells are discarded. The loss of the voice cells is dependent almost entirely on the error statistics of the model and not on the retransmission model.

3.5.3 Result Validation & Comparison To The Theory

The frame size is an important decision and the optimum value depends on the error rate of the link as well as the distribution of errors. Small frames have high protocol overhead but have the advantage that when an error occurs only a small number of bits are lost. Large frames have lower protocol overhead but when one is lost a large number of bits are lost. It is really the Gaussian error that determines the size of the frame that should be used. The burst errors affect the large and small frames almost the same, because a number of continuous frames will be in error. Throughput when the Gaussian error rate is varied between 1E-5 and 1E-6 and the percentage of framed data varied between 50% and 80% is shown in Figure 3.16.

What is noticed is that for the lower mix of framed cells, for example 50%, the lower frame size is preferred, but for higher percentage of framed data, for example 80%, larger frames would achieve higher throughput. There is a need to compromise the throughput by picking a fixed frame size. The frame size was chosen to be 12 cells long, which is a good compromise for the types of errors that are investigated here. If there is only one link then there could be a optimum-sized frame, or even an adaptive frame size as is already implemented on other applications.

The theoretical throughput limit of a pure selective repeat retransmission scheme plotted against the proposed scheme is presented in Figure 3.17. What is noticed is that the limit is approached almost over the whole range of bit error rates.

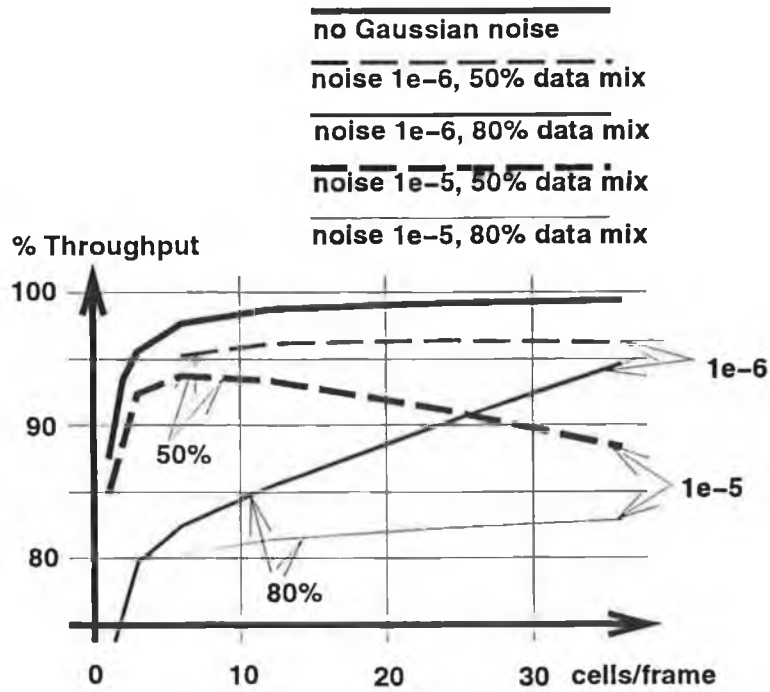


Figure 3.16: Variation Of The Number Of Cells In A Frame

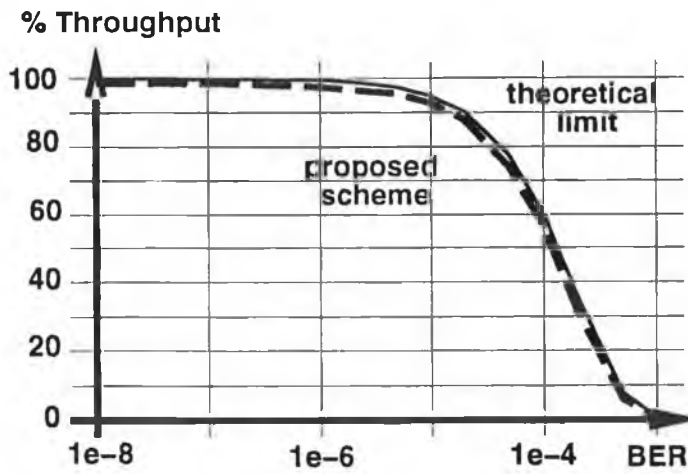


Figure 3.17: Comparison Of Proposed Scheme With Theoretical Limit

Therefore there is no need to use more complicated retransmission schemes [75] for these satellite links as there is only a small amount of efficiency left to be captured with an increasing cost of transmitters and receiver protocols.

3.6 Cell Level Congestion & Traffic Control

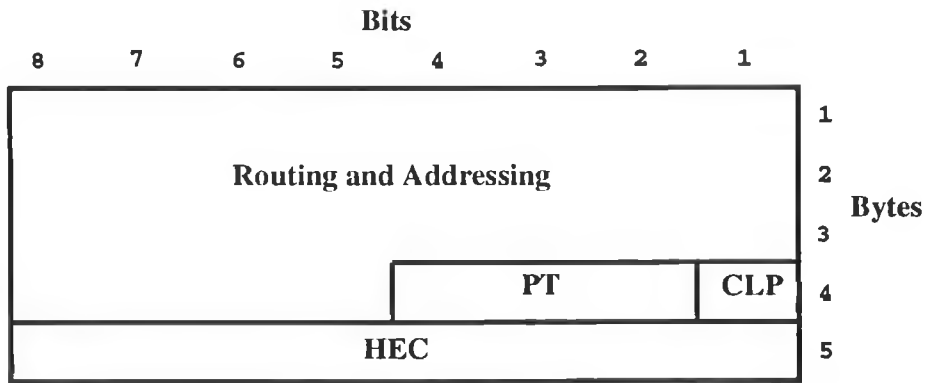
The primary role of traffic control and congestion control is to protect the network and the user in order to achieve network performance objectives. An additional role is to optimise the use of network resources. The traffic control and congestion control mechanisms should not rely on other higher layer protocols which are either application or service specific to achieve the performance objectives required. However protocols may make use of the information in the ATM layer to increase their efficiency. The level of congestion control that is investigated here is concerned with the cell level. Given that there is congestion control used, what is of interest is the cell patterns that will cause the worst performance in the network.

3.6.1 Traffic Contract & Usage Parameter Control

ATM connections specify unidirectional Quality of Service (QOS) parameters. These parameters are specified at connection set-up and are guaranteed by the network. To guarantee the QOS, the network must be able to obtain enough information from the user about the connection and be able to ensure that no other connections that share the resources degrade the QOS. A user must enter into a contract with the network about the parameters of the connection. Then the network will implement traffic control to avoid problems with degraded QOS before they occur. This includes Network Resource Management (NRM), Connection Admission Control (CAC) and Usage Parameter Control (UPC). The network will also control the case where congestion does occur by implementing Explicit Forward Congestion Indication (EFCI), selective cell discard and reaction to UPC failure.

Within the ATM cell there are a number of bits available for congestion and priority setting. These include the Payload Type (PT) indicator and the Cell Loss Priority bit (CLP) both of which are contained in the header of the ATM cell and can be seen in Figure 3.18.

The first bit of the PT indicator tells that the cell is a user cell and the next



PT – Payload Type, CLP – Cell Loss Priority, HEC – Header Error Control

Figure 3.18: A Generic ATM Cell Header

bit of it tells if congestion has been experienced by the cell in the network. The last bit differentiates between two different types of ATM-SDU's. The CLP bit is for high and low priority setting of the cells and can be done by the user and / or by the network. A cell entering the network with low priority is subject to being discarded by the network in times of congestion.

Traffic control is necessary to protect the network so that it can achieve the required performance objectives. UPC enforces a contract between the user and the network about the nature of the connection. This prevents any one user from causing excessive congestion that would degrade the quality of service provided to the other users. It is necessary to determine what is the worst traffic a user can inflict on the network while still abiding by UPC. The Leaky Bucket Algorithm is commonly used to implement the UPC function.

3.6.2 Basic ATM Model

A simplified ATM switch model consists of N users feeding a finite First In First Out (FIFO) buffer with B places and this is shown in Figure 3.19.

The arrival process from the users is random, but the UPC algorithm for worst case analysis makes the arrival process to the FIFO buffer deterministic. There are two types of arrival patterns that are considered after the UPC and both

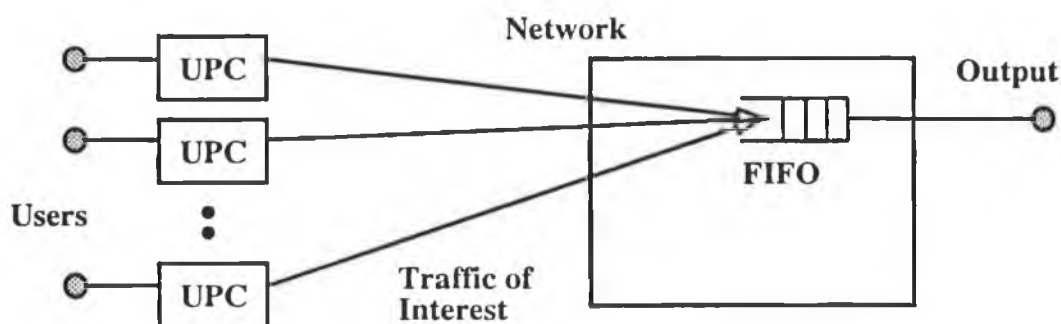


Figure 3.19: Simplified Model

of them are periodic. The service process is also deterministic (at rate r). The worst case traffic is that which creates the highest cell loss for a certain type of UPC, like the leaky bucket. A cell is lost when a cell arrives and gets through the UPC functions and finds that the buffer is full.

3.6.3 Conformance Definition – Leaky Bucket

The traffic contract specifies the negotiated characteristics of the connection. A connection traffic descriptor consists of 3 parts :

- the set of traffic parameters in the source traffic descriptor
- the cell delay variation tolerance
- the conformance definition

The conformance definition is used to decide which cells are conforming in an ATM connection at the UNI [4]. A typical conformance definition is the leaky bucket [1] or Generic Cell Rate Algorithm (GCRA). The conformance definition should not be interpreted as the UPC algorithm, as the network provider may use any UPC as long as the operation of the UPC does not violate the QOS objectives of compliant connections. It is possible to use many such algorithms together in tandem, and a typical implementation is the dual leaky bucket where there are two GCRA algorithms in series. The CAC will use the connection traffic descriptor to allocate resources and to derive parameters for the UPC. Any connection traffic descriptor must be enforceable by the UPC. Even though a cell is found to be non-conforming

Chapter 3 Cell Level Resource Allocation

that does not mean that the connection is not compliant. The precise definition of a compliant connection is left to the network provider. However a connection where all the cells are conforming must be specified by the network provider to be compliant. The traffic contract consists of the connection traffic descriptor and a requested QOS for each direction of the connection. The private UNI may support a different traffic contract to the public UNI.

The contract will contain the Peak Cell Rate (PCR) and the cell delay variation as part of the source traffic descriptor, and will also contain the cell delay variation tolerance. Sustainable cell rate and burst tolerance are optional parameters. For best effort traffic the only parameter specified is the PCR and the network may not reject the connection because that bandwidth is not available but it may impose a different PCR. CAC is used to decide if a connection should be accepted or continue to be accepted in the network. It is required that the traffic contract be accessible to the CAC so as to efficiently allocate resources. This access can be achieved by signalling for switched VCs or via a network management system for permanent VCs. The prime concern is to achieve the required QOS for the new or re-negotiated connection as well as to ensure that the connection will not degrade the QOS of all the other connections in the network. As well as deciding to accept the connection the CAC must determine the parameters needed by the UPC and route and allocate the resources to the connection. Even if high and low priority are not set the network may set them for non-conforming cells.

The leaky bucket algorithm or GCRA has been standardised by the ATM Forum [4] for the conformance definition and it is formally defined as shown in Figure 3.20. There are a number of variables needed to define the algorithm. The Last Compliance Time (LCT) is defined to be the last time that a conforming cell passed through the algorithm. How full the bucket is, is given by X the leaky bucket counter, and there is an auxiliary variable X^* for predicting the value of the leaky bucket counter. The leaky bucket counter can pass cells conforming up to a limit given by L , and for each cell that is passed conforming the bucket is incremented by I . The examination of the leaky bucket occurs when cell k arrives at time $t_a(k)$. The variables L and I specify the operation of the algorithm.

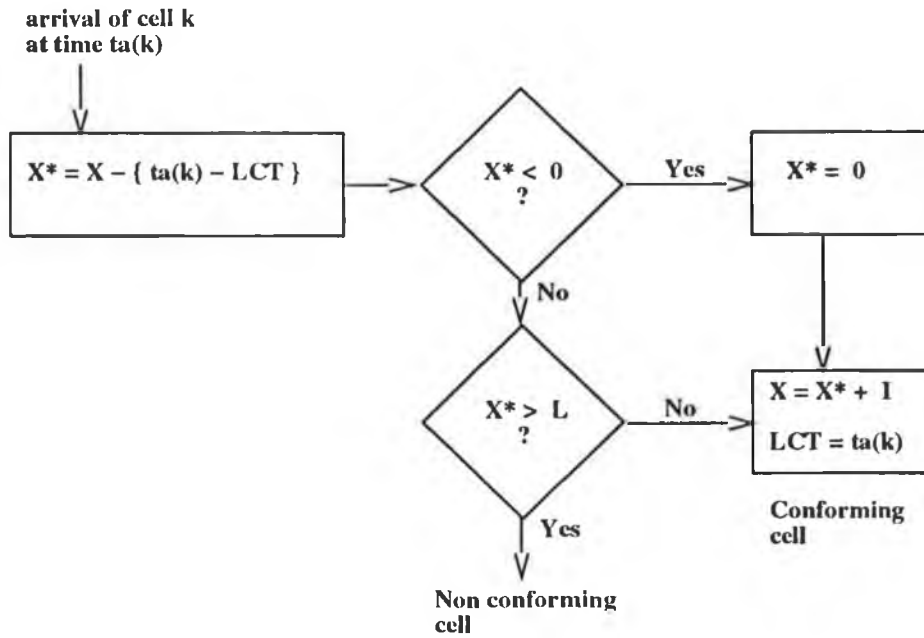


Figure 3.20: Conformance Definition (Leaky Bucket) Algorithm

When a cell k arrives at time $t_a(k)$ the variable X^* is set to the value that the counter will have at this time. This is given by the previous counter value less the amount that the bucket will have leaked away since the last compliant cell, or $X^* = X - (t_a(k) - LCT)$. If this is less than zero then the cell is compliant and the counter X^* is set to zero and passes the cell. When a conforming cell is passed the counter X has to be updated by the increment I by $X = X^* + I$ and the last compliance time is the time of this cell, or $LCT = t_a(k)$. If the bucket has not completely leaked away then the bucket has to be checked to see if the limit is going to be exceeded or $X^* \geq L$. If the limit is not exceeded then the cell is conforming and the cell can be passed as before. Otherwise the cell is non compliant and no updates are done on the variables.

3.6.4 Leaky Bucket As The UPC

UPC is the set of actions the network take to monitor and control traffic. Even though the UPC definition is left to the network provider, it is likely that many network providers will use the leaky bucket to implement it. This is because the same functions have to be carried out by the conformance definition and the UPC.

The UPC also has other functions like the validation of the VPI and VCI values for the active VPCs and VCCs.

The operation of the leaky bucket as the UPC is shown in Figure 3.21. The operation of the leaky bucket is that a splash is added to the bucket (counter increment) for each incoming cell when the bucket is not full. When the bucket is full cells cannot pass through to the network un-marked but the bucket leaks away at a constant rate. The important parameters to be defined in this system are the leak rate of the bucket (R), the bucket capacity (M) and the peak cell emission (p). It is assumed here that if the cells cannot pass the leaky bucket without being marked then they are lost because they are non conforming. These lost cells do not count in the cell loss rate as the cell loss rate is only specified for conforming cells. This is because the network will only give guarantees to the conforming or un-marked cells.

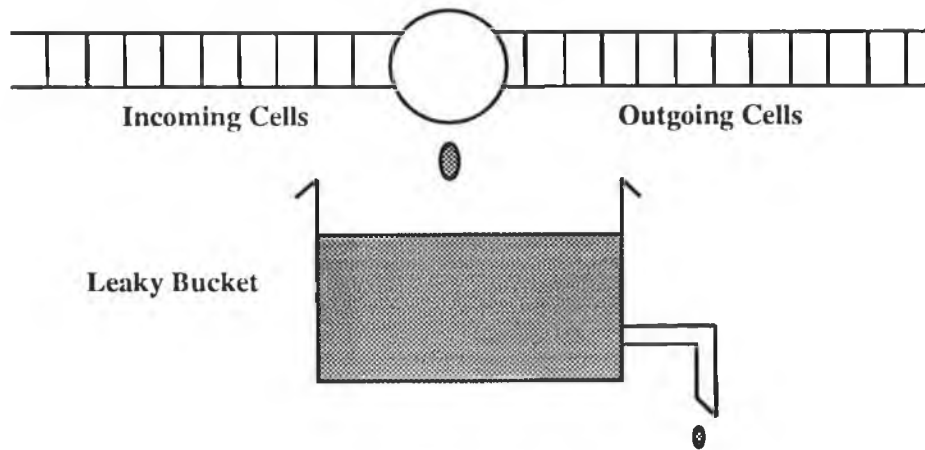


Figure 3.21: Leaky Bucket Algorithm As The UPC

The operation of the UPC shall not violate the QOS objectives of a compliant connection. However the excessive policing actions on a compliant connection are part of the overall network performance degradation and so safety margins should be engineered to limit the effect of the UPC. The UPC can also fail to take action on a non compliant connection where it should have done so. Policing actions on the non conforming cells are not to be allocated to the network performance degradation of the UPC. At the cell level the UPC may pass a cell, change the priority of the cell or discard the cell. A low priority cell is discarded

by the UPC if it is non conforming. Following the UPC, traffic shaping may be implemented on the conforming cells to reduce cell clumping. It is optional for the network operator to allow the UPC to initiate the release of a non compliant connection. Care must be taken when two levels of priority are used as the UPC may discard high priority cells. This can be the case even when if the UPC were placed only on the high priority cells they would be conforming. This is because the total of the two priorities are non conforming and there is the possibility that a previous low priority cell has been passed and so can not be taken back again.

The UPC and CAC are operator specific and should take into account the traffic contract to operate efficiently. It is specified that the signalling should take into account experimental traffic parameters that could be proprietary to either the manufacturer or network operator. It is optional to allow the operation of these parameters across the UNI by mutual agreement. This could be achieved as before by either signalling for switched VCs or via a network management system for permanent VCs. It is optional for the user to be allowed to mark cells as low and high priority. It is also optional for the network to mark cells as low priority if they are not adhering to the traffic contract. The cell loss ratio for low priority cells must be higher than for high priority.

A commonly used UPC is a double leaky-bucket that controls the two different cell rates R and p , each with a separate bucket. The peak rate, p , is controlled by a bucket of size 1 and leak rate, p , and the second bucket operates as explained above. For a cell to be a conforming cell, it needs to be conforming with both buckets at the time it's transmitted.

3.7 Issues For Worst Case Traffic

It is of interest here to find the types of sources that can pass the conformance algorithm and produce the highest cell loss in the network when multiplexed together. This type of source will be called the worst case traffic. There are a number of issues arising when looking at the worst case traffic.

- The first item to consider is the potential types of sources that might produce high cell loss.
- The model of the buffer is important as will be seen and there are a number of questions raised due to the use of infinite buffers. It is possible to use infinite buffers to approximately analyse random finite buffer systems, however this is not the case when the system is deterministic.
- There are further complications due to the discrete nature of the problem. It is possible to simplify the problem by using continuous variable analysis and continuous variables but this can produce misleading results as can be seen by examining [23].

3.7.1 Two Potential Worst Traffic Types

When considering which type of sources will produce the worst performance in the network while still maintaining the contract the first point to note is that the type of source will not allow the leaky bucket to ever overflow. In other words the total available number of cells that are allowed to enter the network will enter to produce the lowest performance. Two types of sources that have been proposed that could be the worst types are a two state source and a three state source as shown in Figure 3.22.

The two state source (greedy on-off) emits a burst of cells at the peak cell emission until the bucket is about to overflow and then falls silent waiting for the bucket to empty. This occurs periodically depending on the parameters of the system. The three state source is similar to the greedy on-off source except that it keeps emitting cells at the leak rate after a burst. Therefore its operation is to emit a burst of cells at the peak cell emission rate until the bucket is about to overflow and then emit cells at the leak rate of the leaky bucket for some time and then fall silent to allow the bucket to empty. This would then be repeated again. What can be seen is that the three state source would have a longer period than the greedy on-off for the same system parameters.

The general belief is that the greedy on-off source gives rise to the worst

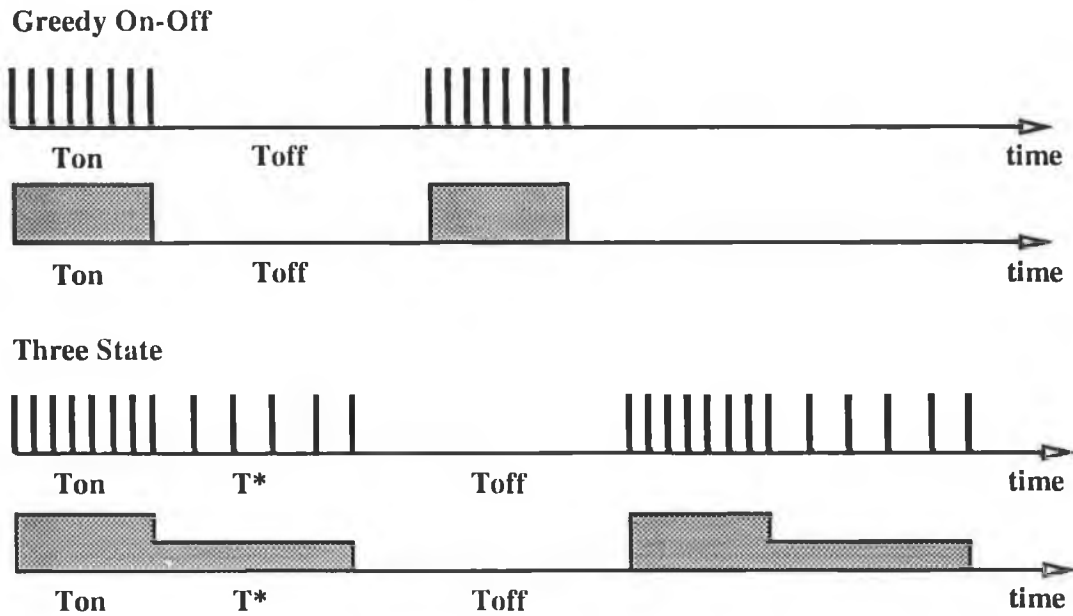


Figure 3.22: Source Types

case traffic [83] as it would have the largest variance possible. However the three state source has been proposed because it produces longer queues in an infinite buffer and therefore potentially larger loss in a finite buffer [84].

3.7.2 Finite & Infinite Buffers

A study of the three state source [84] compared its performance to that of the greedy on-off source, for a number of sources into an infinite buffer. What was found in those simulations is that the three state sources produced higher buffer occupancy than the greedy on-off sources. The survivor function, $P[Q > q]$, was found where Q was the buffer occupancy. This function was then assumed to approximate to cell loss in a finite buffer. However this would only be true for input traffic that would be statistically independent of the queue lengths, which is not the case here as the input traffic is periodic and deterministic. It is possible to demonstrate this fairly simply with a small example as is shown in Figure 3.23. Here the buffer occupancy for the infinite case of both the greedy on-off and the three state is plotted on the same axis. Note that only a single source is being considered here on its own.

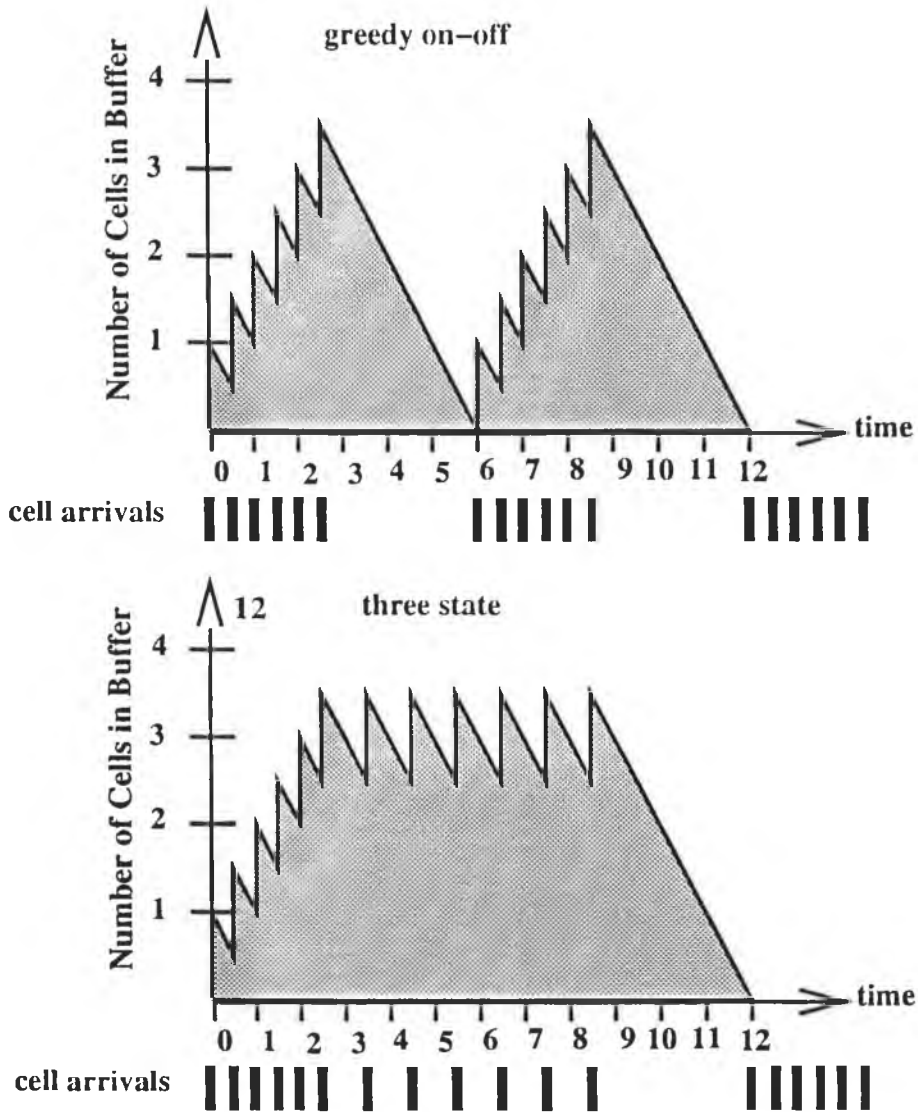


Figure 3.23: Buffer Occupancy In An Infinite Buffer

The period T of the greedy on-off is chosen to be half of the period of the three state source so that it is possible to examine a single period of the three state and compare it to two periods of the greedy on-off for cell loss. There are 2 cells arriving per cell time while the source is emitting at the peak rate p and this continues until the bucket is full to capacity M , which in this example is 4. Then the bucket leaks away at a rate of 1 cell per cell time R for the greedy on-off but for the three state source it emits at this rate R for 6 time units, which is denoted by T^* and here is equal to the period of the greedy on-off source. Look over 12 time units as both repeat exactly the same after this. Over this period there are 12 cells inputted to the buffer which is the same as the leak rate R multiplied by

Chapter 3 Cell Level Resource Allocation

the time period. The three state source gives rise to higher buffer occupancy than the two state source as can be seen in Figure 3.23.

This would lead us to believe that the three state source has higher loss compared to the greedy on-off source because of higher queue occupancy for the three state source. When considering a finite buffer of capacity 2, as shown in Figure 3.24, it is found that the three state source loses 2 cells, one at time period 1 and one at time period 2. However the greedy on-off source loses one at time period 1 and one at time period 2 and again one at time period 7 and one at time period 8. So the cell loss rate for the three state source is 0.166 while the cell loss rate for the greedy on-off is 0.333.

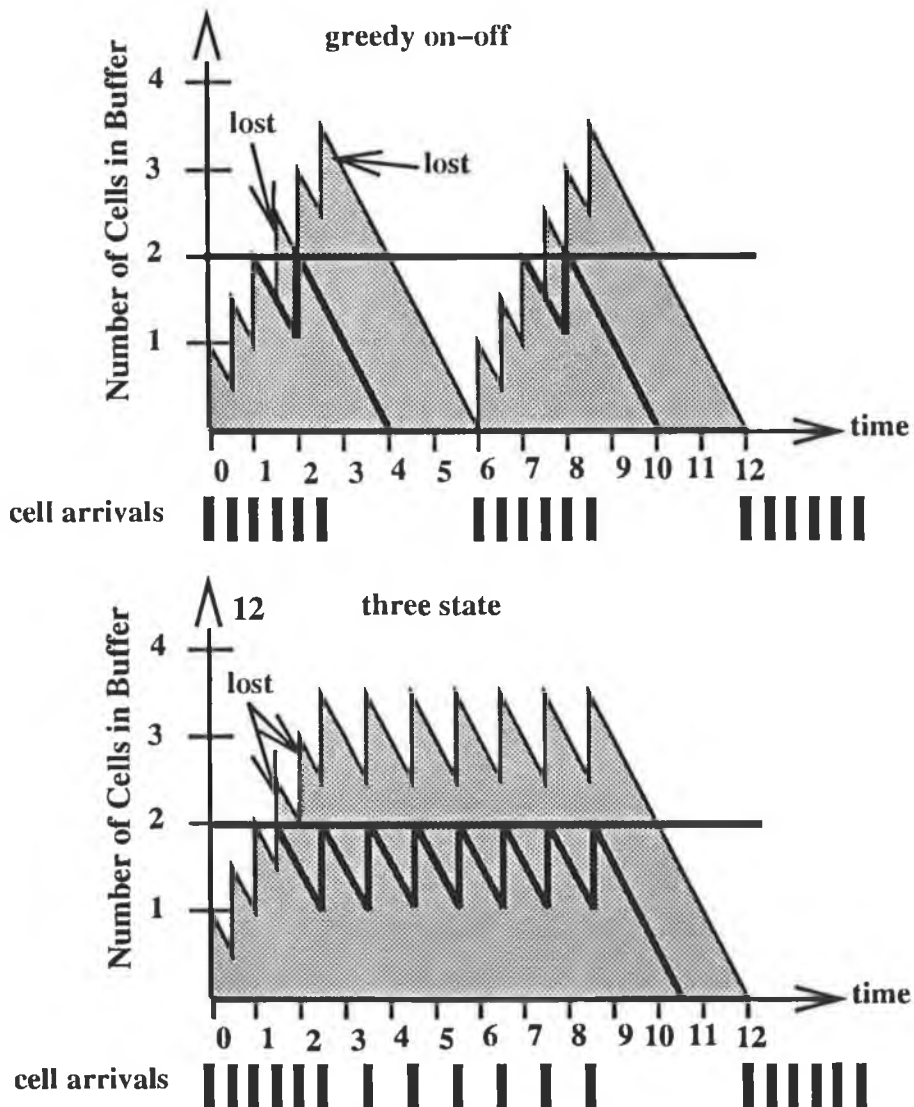


Figure 3.24: Buffer Occupancy In An Finite Buffer

The fact of higher buffer occupancy in the infinite case does not necessarily mean higher cell loss rates in the finite buffer when the arrival process is deterministic. For deterministic arrivals the loss in a finite queue is not necessarily related to the buffer occupancy in the infinite case but more on the method of arrivals, or the process of arrivals, past the finite places in the queue. This problem was simulated [83] although the results were not explained.

3.7.3 Continuous & Discrete Variables

The problem contains both continuous and discrete variables and therefore it is important to distinguish between the two types. A simplified model is considered here which is similar to that published elsewhere in the literature [23] where there are two identical, independent users feeding a finite buffer. If the users were not independent then the worst case would be the greedy on-off source with all sources in phase emitting together. However when the users or sources are considered to be independent then the phase between the sources is random as shown in Figure 3.25. This then gives rise to the probability of cell loss because there is a probability of phase difference between the sources. The assumption is made that the two sources are randomly phased, which means it is possible to consider one source as a reference source and the other one is then out of phase by some amount. The amount out of phase will determine the number of cells lost. The more in phase the more loss is expected. The probability of each of the possible combinations of out of phase is just the reciprocal of the number of the possible combinations.

Two possibilities are considered here, either 2 greedy on-off sources or 2 three state sources. The cell losses are calculated for every distinct combination of the two traffic patterns in Sub-Section 3.8. To get the average cell loss for either case the losses are averaged. The cell loss rate for each scenario is also compared in Sub-Section 3.8. There are a number of constraints on the problem due to the discrete nature of the variables which are described here :

1. The sources may be out of phase only in cell time units which means that it

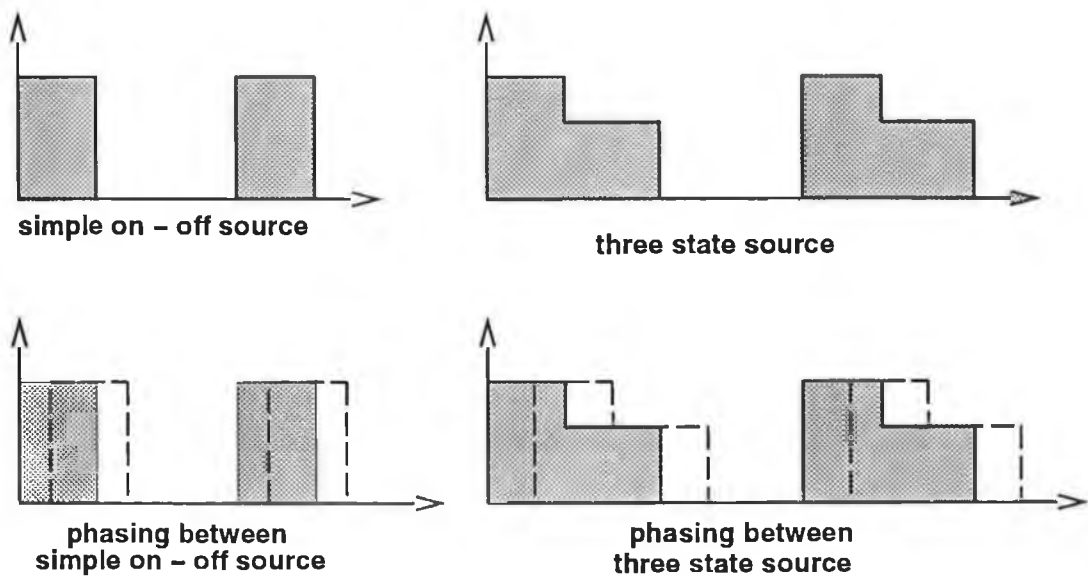


Figure 3.25: Sources Out Of Phase With Each Other

is not possible to use integration as was used in previous work [23].

2. The on period of the two state source is the time taken for the bucket to just about overflow whilst emitting at the peak rate and also to emit any cells which arrive during this time. In the discrete case this is equal to $\lceil M/(p - R) \rceil - 1$.
3. Cells are only lost as units – no fractional cell loss. Therefore it is necessary to ensure that in using general formulae the cell loss is truncated to an integer. Therefore $M \geq 1$.
4. The discrete nature of the cell also implies that the buffer size must be an integer.
5. The service process can be assumed to be continuous. This means a cell can be served as it arrives. Alternatively it is possible to think of the server waiting until the cell has fully arrived before starting to serve it.

3.8 New Approach To Cell Loss For Worst Case Analysis

The comparison of cell loss from both source types in the two identical user system is revisited assuming the cell arrival and service processes to be discrete. The cell loss rate is computed as the total number of cells lost in one period divided by the total number of cells emitted in one period. The total number of cells lost in any one period is the average of the cells lost by each combination of the two traffic patterns. It is assumed that the minimal phase difference between two sources is one cell time and there are T (T is the period) distinct combinations of the two traffic patterns. The total number of cells emitted in one period T equals $2RT$. A diagram showing the variables and how they relate to each other is shown in Figure 3.26.

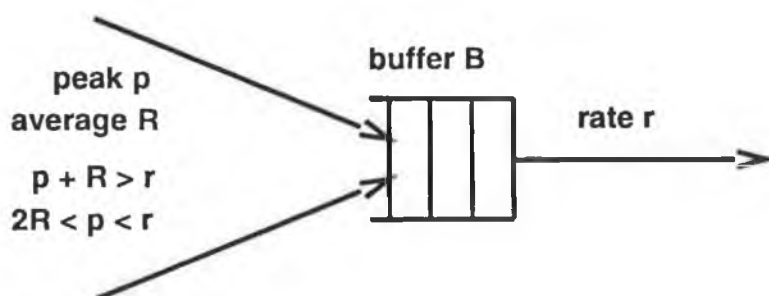


Figure 3.26: Set-up For Calculating Cell Loss For Worst Case Traffic

Furthermore the following stipulations are placed on the parameters :

- The service rate is at least equal to the peak cell emission rate, $p \leq r$
- The leaky bucket rate is less than or equal to half the peak cell rate (and hence the service rate), $R \leq p/2$
- Together the peak cell emission rate and the leaky bucket rate exceed the service rate, $p + R \geq r$
- The buffer size must be small enough so that cell loss is guaranteed when both sources are in phase, $(2p - r)T_{on} \geq B$

These stipulations ensure that cell loss only occurs when at least one of the sources

Chapter 3 Cell Level Resource Allocation

is emitting at the peak rate. If the service process is assumed to begin as the cell is arriving this is called a fast server. Then the cell loss for two greedy on-off sources, x places out of phase with each other, is denoted by $CL(x)$ and is calculated in Equation 3.1

$$CL(x) = [(2p - r)(T_{on} - x) - B] \quad (3.1)$$

where the symbols have the usual meanings. When the two sources are on at their maximum rate they exceed the service rate by $(2p - r)$ and so the queue is building up. This lasts for the amount of time they are on at this rate $(T_{on} - x)$ and so the queue builds to $(2p - r)(T_{on} - x)$ and so the loss is what this is in excess to the buffer size B . Hence Equation 3.1 is explained. Similarly the cell loss for the three state source can be represented by Equation 3.2.

$$CL(x) = [(2p - r)(T_{on} - x) + (p + R - r)(x) - B] \quad (3.2)$$

The reason for the difference in the upper and lower bounds is due to the way in which the cells arrive at the end of the three state source. The upper bound, for the on-off source, is due to the fact that a fractional cell loss means that we have in fact lost that cell. A fractional cell loss in the three state source comes from the overlap of the end of the peak rate and the start of the mean rate, but here they will not generate loss as the cells arrive as quick as the server can serve them. Therefore the lower bound is chosen.

Alternatively if it is assumed that the server waits until the cell has arrived in the buffer before starting to serve the cell, which is called the slow server, then the expressions for cell loss are modified for the greedy on-off source as follows in Equation 3.3

$$CL(x) = [(2p)(T_{on} - x) - (T_{on} - x - 1/p)r - B] \quad (3.3)$$

What is happening is that the time to serve has increased by $1/p$ and so this must be included in the equations. For the three state slow server the cell loss will be given in Equation 3.4.

$$CL(x) = [(2p)(T_{on} - x) - (T_{on} - x - 1/p)r + (p + R - r)(x) - B] \quad (3.4)$$

To calculate the cell loss ratio allow x to vary over all phase possibilities and then divide by the number of combinations and also divide by the number of cells transmitted by both sources. The number of cells transmitted by both sources will be $2RT$ and the number of phase combinations will be T , so the cell loss ratio, CLR , is given by Equation 3.5.

$$CLR = \frac{\sum_{x=0}^{T-1} CL(x)}{T 2RT} \quad (3.5)$$

It is difficult to plot these equations due to the number of variable used and also because there are constraints on the variables that depend on the other variables. The constraints could be like $T_{on} \geq x$ for cell loss to occur. Instead of attempting to plot or simulate these equations, some examples will be provided.

3.9 Counter Examples To Traditional Worst Case Traffic

To show that there is no single worst type traffic for two identical sources a counter example to the traditional theory of the greedy on-off being the worst case is presented. It is shown here that the three state source can produce higher loss for integer values of variables chosen. As mentioned previously two different types of servers, the fast serving server and the slow serving server are considered.

3.9.1 Fast Serving Server

Assuming that the cell is served as it arrives here is an example showing the three state source to give rise to greater cell loss in a finite buffer than the two state source. The following system parameters are used :

- p , peak cell emission rate = 1
- r , service rate = 1
- R , leaky bucket rate = .5

Chapter 3 Cell Level Resource Allocation

- M , leaky bucket capacity = 8
- B , buffer size = 12
- T^* (for three state source) $\geq ((2p - r)T_{on} - B)/(p - R) = 6$
(This is chosen so that there will be no loss when the two sources are out of phase by more than T^* but there is a probability of cell loss if they are out of phase by less than T^* .)

It takes 8 cell times to fill the bucket up but in that time 4 more cells are allowed through because of the constant leak rate of the bucket and while they try to fill the bucket 2 more arrive and then finally one arrives and the bucket is full. Therefore the amount of time that the source is on and emitting cells at the peak rate is given by $T_{on} = \lceil M/(p - R) \rceil - 1 = 15$. When the bucket is full it takes M/R seconds to empty normally, however here one time period has already elapsed so $T_{off} = 15$.

There are 30 phase combinations for the greedy on-off source patterns, however sources that are too far out of phase do not give rise to cell loss. For the greedy on-off sources by examining Equation 3.1 sources that are 3 or more time units out of phase do not produce any cell loss. Remember that this loss can occur when the second source is a little advanced from the reference and also when it's so advanced that it is almost back in phase with the reference. This can be seen in Figure 3.27. The cell loss for the two state sources is in total 9 cells, which is calculated from 3 cells lost when in phase, 2 cells lost when either one out of phase and also 29 out of phase, and 1 cell lost when either two out of phase or 28 out of phase. For all other phases there is no cell loss. Get 9 cells lost in total out of a possible 900 cells over the 30 phase and 2 sources. Therefore using Equation 3.5 the cell loss ratio can be calculated to be $CLR = 0.01$.

For the three state sources cell there are 36 possible phase combinations and by examining Equation 3.2 loss can occur up to 4 units out of phase. The total number of cells lost over all 36 possible phases can be seen in Figure 3.27 and is 15 cells. By using Equation 3.5 the loss can be calculated to be $CLR = 0.01157$. Therefore the three state source produces higher loss than the greedy on-off source for the fast server.

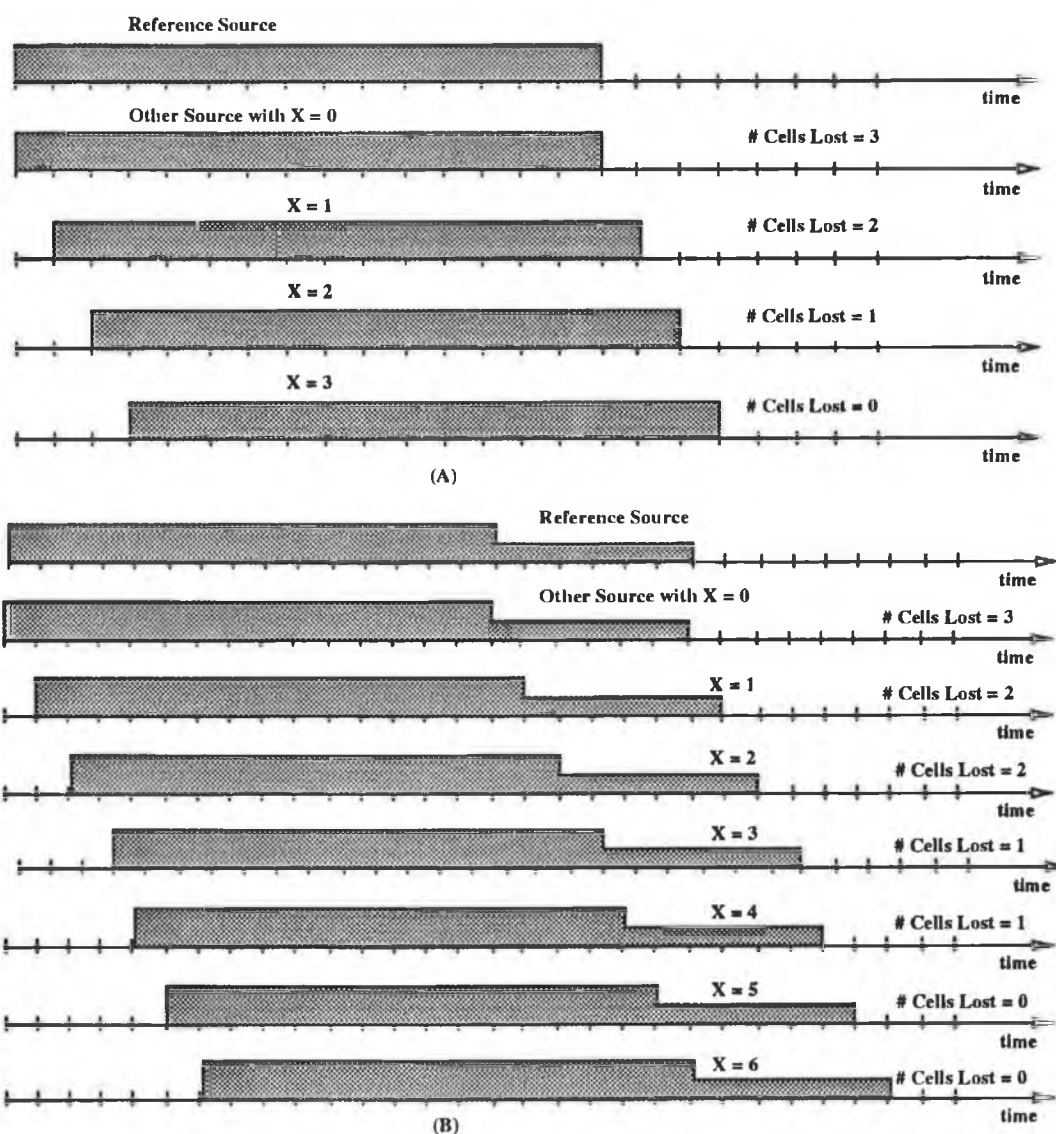


Figure 3.27: Fast Server: A. Greedy On-Off Source, B. Three State Source

3.9.2 Slow Serving Server

If the service process is assumed to begin serving a cell only after it has fully arrived into the buffer here is a similar counter example. The following system parameters are used :

- p , peak cell emission rate = 1
- r , service rate = 1
- R , leaky bucket rate = .5

Chapter 3 Cell Level Resource Allocation

- M , leaky bucket capacity = 8
- B , buffer size = 14
- T^* (for three state source) $\geq ((2p - r)T_{on} - B + r/p)/(p - R) = 4$

Similar to the fast serving server $T_{on} = T_{off} = 15$. For the greedy on-off sources by examining Equation 3.3 sources that are 2 or more time units out of phase do not produce any cell loss. This can be seen in Figure 3.28. The cell loss for the two state sources is over all possible phase combinations equal to 4 cells and by using Equation 3.5 the cell loss can be calculated to be $CLR = 0.00444$.

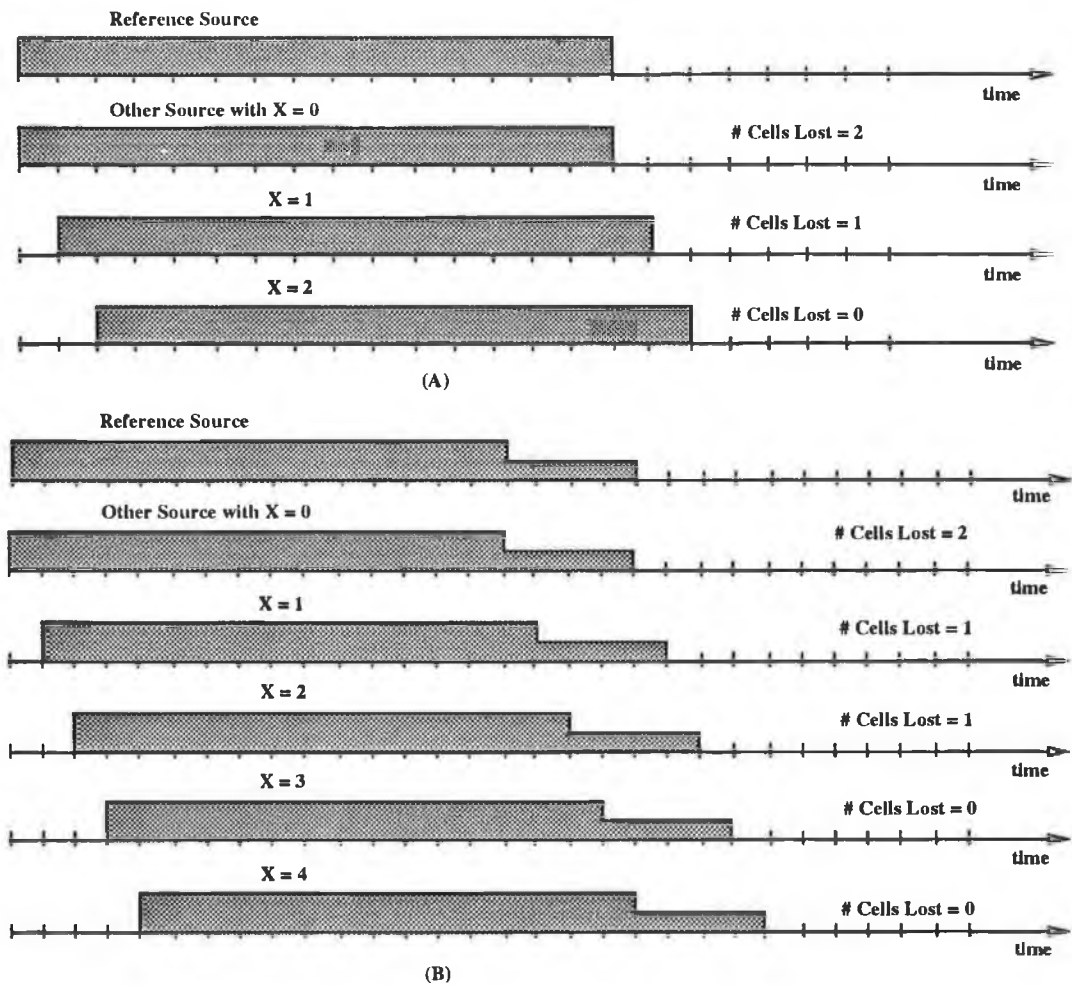


Figure 3.28: Slow Server: A. Greedy On-Off Source, B. Three State Source

For the three state sources cell loss can occur up to 2 time units out of phase. This is concluded by examining Equation 3.4 for the cell loss for a three

state source and this is also seen in Figure 3.28. Here in total 6 cells are lost over all 34 possible phase combinations and so again by using Equation 3.5 the cell loss ratio can be calculated to be $CLR = 0.00519$. Therefore the three state source produces higher loss than the greedy on-off source for the slow serving server.

Therefore regardless of how the service is achieved in the buffer there is an example of the three state source producing more loss than the greedy on-off source. It is therefore shown that there is no single type of worst case traffic source for the two identical source problem, considering these types of sources.

3.10 Discussion

A number of issues have been dealt with here that arise when considering resource allocation at the cell level of the ATM. The structure of the adaptation schemes for satellites have been examined and the worst case traffic for the standard leaky bucket traffic controller has also been investigated.

It has been shown here that it is possible to use ATM over satellite links. It is proposed here to place the error control in the convergence sublayer of the ATM rather than the segmentation and reassemble sublayer. This improves both the efficiency of the protocol and the efficiency of the error detection. The moving of the error control functions is compliant with the ATM standards. It is also proposed here to differentiate the recovery mechanism, or re-transmissions, on the service. This allows a guarantee to be given to data services of no loss while not affecting the real time services. While the simulation was for a T1 link, this approach can be incorporated into a data stream that is part of a larger link, such as a 45 Mb/s link.

From the analysis here it may be concluded that the greedy on-off source does not always give rise to the worst case traffic. The first point that is made is regarding the infinite versus finite buffer and it is shown that the occupancy in an infinite buffer will not relate to the loss in a finite buffer when the traffic is deterministic. The second point that is made here is that some of the variables are discrete and therefore attention must be paid when selecting examples and formula.

Chapter 3 Cell Level Resource Allocation

This is then concluded by showing two counter examples against the greedy on-off source and were able to say that there is no single worst case traffic type for the two identical source problem. The problem is far from totally solved as the result for this two source case is not binding on an n source situation, indicating one area of future work. On the other hand, the network needs to know what the worst case is in order to give the QOS to the users. This brings up the question of under what conditions is one case worse than others or if there is a single case that is not known about that performs universally badly for a given traffic controller. The progression could take either an analytical or simulation track, however care would be needed for either approach.

Chapter 4

Pricing In ATM Networks

4.1 Introduction

Admission control and bandwidth allocation are important issues in telecommunications networks, especially when there are random fluctuating demands for service and variations in the service rates. In the emerging broadband communications environment these services are likely to be offered via an ATM network. In order to make ATM future safe, methods for controlling the network should not be based on the characteristics of present services.

There are many problems with traditional approaches to the connection admission control area. Users may not know enough about the statistics of their connections to provide accurate traffic descriptors to the network. This is particularly true for certain data transfers with long holding times. For example, in browsing a remote image database, it may not be possible to give a reasonable estimate of the mean bit rate except over a long period of time. Even if users know their traffic descriptors they may be unwilling to reveal them, or may try to misrepresent them. This implies the need for some form of traffic policing and enforcement at the network access points, in order to ensure that each source is complying with what it declared at connection set-up. Assuming accurate traffic descriptors, a complicated analysis is required to determine whether sufficient resources are available to support the requested connection's QOS while not degrading the QOS of admitted connections. Current computational methods cannot

do this analysis on the short time scale in which CAC decisions must be made [80]. Consequently either simplifying assumptions about connection and resource interactions are introduced, or connections are grouped into classes whose members are assumed to have similar characteristics. Making simplifying assumptions can lead to an over-controlled network, or produce unreliable results. For example, recent studies have shown that LAN and VBR video traffic is self-similar [45], which is very different to the behaviour predicted by Poisson-based or fluid flow models.

There is a trade-off between simplicity, which implies a small number of traffic classes, and the large number of classes needed to cover a wide range of different traffics and services. There is likely to be a wide variation in user requirements even within the same traffic class, due to individual preferences and internetworking considerations. Even if accurate traffic and network models were available, technological advances may quickly render their assumptions invalid. For example, the bandwidth required for video or image transfer applications may continue to decrease due to improvements in compression / decompression algorithms and technology [3]. In addition, it is impossible to predict the characteristics of all future applications, or the application mix when users become familiar with new network capabilities. It may be unwise to base CAC and bandwidth allocation schemes on the requirements and properties of current services.

ATM networks are expected to accommodate a wide range of users including some who can tolerate a certain amount of cell loss and / or delay. There are also likely to be some users who can modify their traffic inputs in response to feedback signals from the network. A feedback scheme is proposed to increase network efficiency by taking advantage of this flexibility.

The proposed approach here is based on pricing bandwidth to reflect network utilisation, with users competing for resources according to their individual bandwidth valuations. The prices may be components of an actual tariff or they may be used as control signals, as in a private network. Simulation results show that there is an improvement possible in terms of cell loss probability with the scheme here versus a more traditional approach, like the leaky bucket. This is partly due to the fact that a closed-loop feedback scheme is being implemented while the traditional

methods are open-loop schemes. In general closed-loop schemes can provide better utilisation than open-loop schemes. Simulation results here will show that a small queue with pricing can be efficient to multiplex heterogeneous sources. The reason that only a small queue is needed is that once the number of cells that can be in excess in a feedback interval can be buffered, then the feedback can control the flows reasonably well. There is also consideration given to the economic efficiency of the network which addresses whether the traffic that is being carried is the most important or not. There are other simulation results that show that it is possible to increase both the network utilisation and the economic efficiency together [53, 63].

Two fundamental points about the approach here and pricing need to be spelled out :

- price is *one possible* feedback signal which has some attractive properties (compactness, quantifiable, etc). Economists have developed a large body of theory on pricing mechanisms, and there is a lot of experience with the use of prices in real-world markets. However it is possible that there are other feedback mechanisms that, for one reason or another, may be preferable in communication networks. These other feedback mechanisms may be pricing schemes, like bandwidth auctions [51], or they might be traditional feedback schemes like Explicit Forward Congestion Indication, (EFCI), which eventually gets fed back to the user.
- when most people think of prices, they think in monetary terms, e.g. dollars and cents. However, *there is nothing “inherently” monetary in applying pricing principles to communication networks.* Pricing can be thought of as merely a mathematical framework or tool to be used to find the optimum operating point. Other non-pricing feedback schemes can be seen as special cases of the pricing scheme. For example the main reason that feedback will give the desired response is if the users change their traffic flows, and algorithms that change traffic flows according to signals from the network can be modelled as an economic system. As long as the appropriate cost and valuation functions can be defined, a pricing mechanism can be applied even if money is not directly involved. An example is a private network where

Chapter 4 Pricing In ATM Networks

one organisation controls all the users. Here the valuation functions could be based purely on seniority. In this case the users (or their applications) are co-operative and can be programmed to obtain a desirable traffic mix. Of course some incentives are needed to ensure that the users respond as if the prices were going to be applied to their usage, but these incentives could be based on aggregated usage or some other indirect measure.

This chapter is organised as follows. In Section 4.2 the users are examined in detail in this light and they are re-classified using a new parameter called adaptive. There are a number of new types of users investigated which can use a feedback signal from the network to control themselves and these are used in the rest of this chapter. In Section 4.3 there is the motivation for pricing in networks and in particular how to use price as a feedback signal and how price might be part of the overall charging scheme. In Section 4.4 the mathematical model is built for the pricing and the benefit functions. First the user benefit function is addressed, and then the system or network problem. Then a proposed distributed pricing algorithm is discussed. The following section, Section 4.5 deals with the network efficiency simulation model, the simulation environment and the results of the simulation. How to actually implement the pricing scheme suggested is shown in Section 4.6. There the generation of the price, the passing of the price from the network to the user and how the user uses the price to control the source is discussed. The idea of having both network and economic efficiency is discussed in Section 4.7 which starts off with an economic framework and the user models that might exist. In particular two of the five possible types of users are modelled, the elastic and inelastic users. Then another algorithm for pricing is proposed and the simulation details and the results are presented. In Section 4.8 it is shown where pricing fits into the overall picture of control of resources at the connection level. Finally in Section 4.9 there is some discussion of the chapter.

4.2 Adaptive Connections

4.2.1 Defining Services In An ATM Network

An ATM-based network is expected to allow users to negotiate their own service definitions. In this way a user can get a service appropriate to their individual needs. This assumes that some users will take a more active role in defining their network services than in present-day networks : these users are assumed to enter into negotiations with the network at the **connection level**, whenever a new connection is requested.

Some applications require guarantees from the network, on loss or delay or both. The user-network relationship for these applications in many of the proposed Connection Admission Control (CAC) schemes is a **contract** : users describe their traffic and make QOS demands, and the network provides a stated level of service while enforcing the user commitments. When a new connection is requested, the network must decide whether or not to accept the connection; and if so, how to route it through the network and what resources to reserve for it. The connection request is refused by the network if accepting it would lead to QOS degradation for one or more currently-active connections which are complying with their contracts. Some CAC schemes provide feedback on the reason for the refusal [25], which guides the user in submitting a revised request.

Early CAC schemes reserved the peak bandwidth needed by an accepted connection – *deterministic multiplexing* – as required for constant-bit-rate (CBR) sources. However the gain in network utilisation possible by taking advantage of the statistical nature of variable-bit-rate (VBR) sources has led to the idea of *statistical multiplexing*. Statistical multiplexing is where it is possible to assign less than the peak bandwidth required but cell loss and/or delay may be introduced.

The aim of a preventive CAC scheme is to balance the QOS provided to admitted connections against network utilisation by limiting the number of connections using the network. Most proposed CAC schemes decide whether or not to accept a connection request based on knowledge of the *traffic description*, the

Chapter 4 Pricing In ATM Networks

user's *QOS requirements*, and the *current state* of the network :

- ideally a user requesting a connection would give a complete statistical description of their traffic, but in practice only a limited indication of expected behaviour is feasible. Connection behaviour is described by a set of parameters called traffic descriptors, such as mean bit-rate, peak bit-rate, maximum burst length, probability of cell arrival in a fixed interval, and so on.
- user QOS requirements are usually expressed by networks researchers in terms of acceptable cell loss, delay and jitter. How to relate these quantities to parameters the user is concerned with (such as picture quality in a video call) is an active area of study.
- the current state of the network can be determined by monitoring the utilisation of network resources and/or by characterising the behaviour of connections already admitted.

Based on this knowledge, CAC schemes have been proposed in which an effective bandwidth [39] is associated with each source in order to meet its QOS requirements while still permitting a statistical multiplexing gain.

Accepted connections are provided with a traffic contract and as long as they comply with that contract, their cells are treated as high priority by the network. Cells submitted in excess of their contract may be discarded at the network access point, or may be marked as low priority and discarded if the network is congested. Therefore the network provides no guarantees for low-priority traffic.

Not all applications require guarantees from the network. For example, some types of non-real-time data transfers can tolerate a certain amount of cell loss; or the application may recover from cell loss or delay at a higher layer of the protocol stack. These **best-effort** applications request an Unspecified QOS service [4], in which the network always accepts the connection request but may assign a different Peak Cell Rate (PCR) than what was requested (for example, a connection may be assigned a PCR of zero). Cells which are submitted in excess of the PCR are marked as low priority.

4.2.2 Adaptive User Behaviour

In the previous Sub-Section, network services were divided into “guarantees required” and “best-effort”, and an enforceable traffic description was needed from users who wanted guarantees from the network. This breakdown of services does not capture all possible user desires now, let alone in the future, because it does not account for **adaptive** users. A user is adaptive if they are able *and* willing to respond to feedback signals from the network during their connection lifetime by changing their offered traffic. The addition of willingness-to-respond reflects the fact that some users may run applications which are capable of adaption, but do not want to be concerned with network feedback during the connection. A typical adaptive scheme is where there is a limited set of traffic parameters given to the network and the network gives a limited set of QOS to the user. However these parameters may not affect the user but would affect the user’s terminal equipment and would be instructed in advance by the user on how to adapt to the network feedback signals. Not all users will be adaptive and therefore the network will contain a mixture of user types. Two types of users, an adaptive and non-adaptive, are shown in Figure 4.1. A user who is willing to wait (based on network feedback) to set up their connection but who wishes to be insulated from such feedback once their transmission starts is adaptive only at the connection level, and such users are not regarded as adaptive.

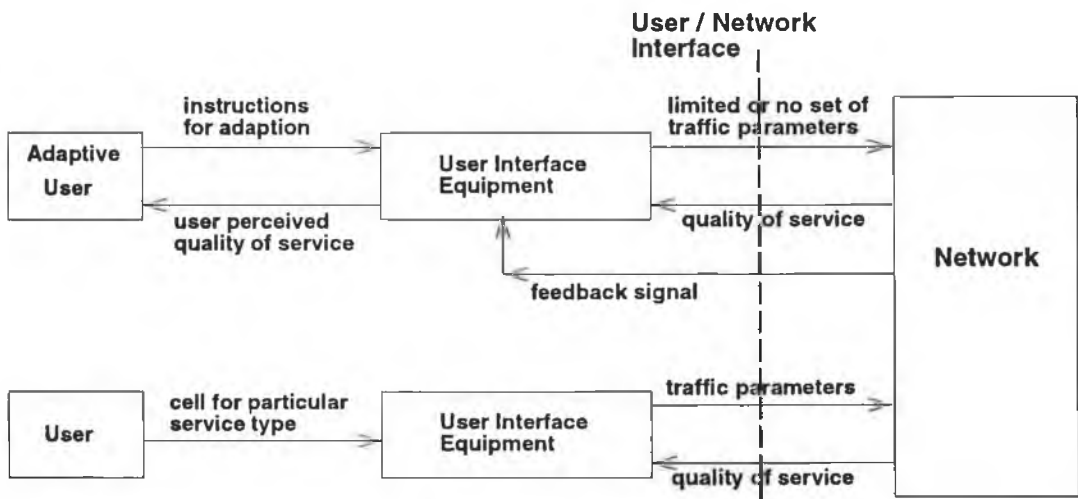


Figure 4.1: Proposed Adaptive Feedback Scheme

Adaptive users can help to increase network efficiency if they are given appropriate feedback signals. When the network load is high, the feedback should discourage adaptive users from injecting cells; when the load is low, the feedback should encourage these users to send any cells they have ready to transmit. Up to now, adaptive users have been regarded as lower priority than non-adaptive users who want guarantees from the network. Typically the latter are served first if possible, and any remaining bandwidth is shared out among the adaptive users on a best-effort basis. There is an implicit assumption in such schemes that the more demanding users pay premium prices in order to get the service they require, and the adaptive users pay lower prices as a reward for their flexibility. The bandwidth allocation of such schemes places the non-adaptive users, like voice, video and some data applications, on the network first and then the remaining bandwidth is filled with the adaptive users, as is shown in Figure 4.2. Here assume that the whole bandwidth is used, which is what would expect from a good adaptive scheme, as long as the traffic is available to the network.

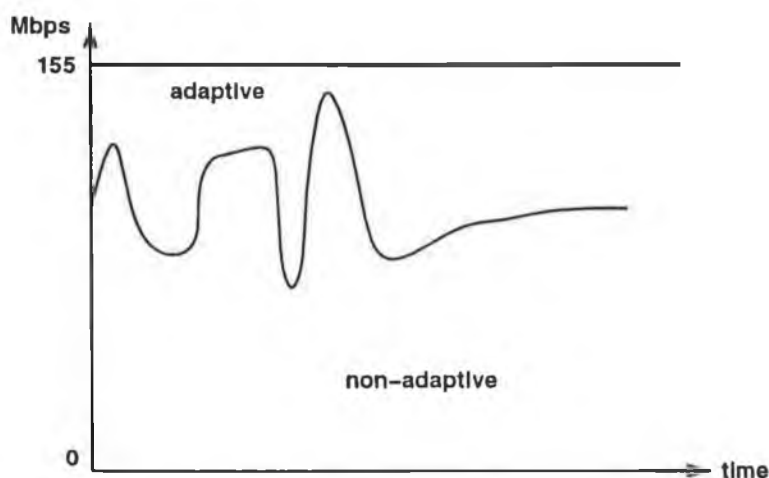


Figure 4.2: Traditional Bandwidth Allocation For Adaptive Users

One of the difficulties of having these two different types of users sharing the same network, is that the adaptive users will have to accommodate the non-adaptive users in times of congestion. However this does not mean that the network has to give higher priority to the non-adaptive users in all cases. An example might be where the network has split the bandwidth resource into two parts, one for the adaptive users and the other for the non-adaptive users. While each traffic type

is within its own part of the bandwidth, it has high priority, but when it uses the other's part of the bandwidth then it has low priority. In Figure 4.3 it can be seen that the non-adaptive users have been allocated 100 Mbps of bandwidth and the adaptive users are allocated 55 Mbps. When the adaptive users fall into the 100 Mbps of the non-adaptive users bandwidth, then their priority is changed from high to low. However when the non-adaptive users use part of the 55 Mbps of the adaptive users bandwidth then their priority is changed from high to low. However the adaptive scheme will lower the adaptive users input traffic to accommodate the variations in this non-adaptive traffic.

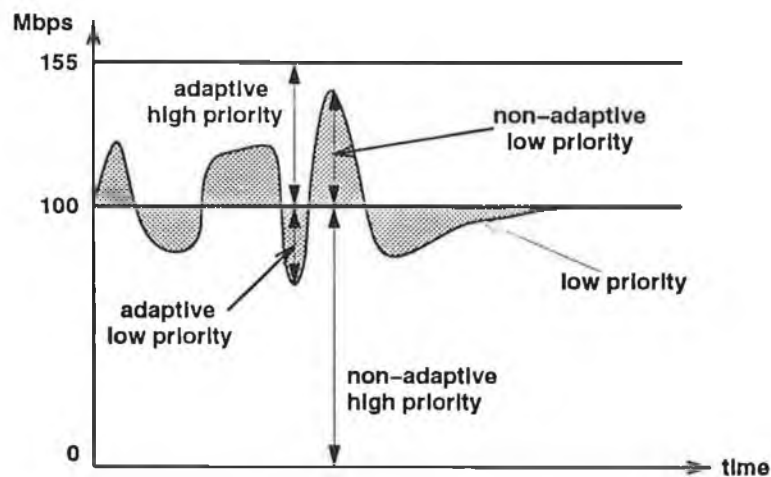


Figure 4.3: Example Of Priorities For Two Traffic Classes

Whether or not a network operator gives higher priority to more demanding users is a policy or business decision, not a technical one, and it is not the intention here to argue for one side or the other. From a technical point of view it is as easy to give either higher priority. Naturally there are technical implications in giving one or the other higher priority. For example network dimensioning changes depending on the priority allocation. However, technical problems do arise when some adaptive users also require guarantees on cell loss. For example, data transfers using compression may be very flexible with respect to delivery time but require zero cell loss to ensure data integrity, if there is little or no redundancy in the transmitted data. The classical approach to such users is to retransmit when errors occur, but in congested conditions this approach tends to increase the network load and worsen the congestion. This is especially true when the data is protected by

some higher layer protocol and one cell loss may mean tens or hundreds of cells have to be retransmitted as the block of data that has to be retransmitted is much larger than the cell. This would mean that in times of congestion rather than run the risk of loss it would be better to delay the transmission of the data.

A novel idea is to provide a limited set of guarantees to adaptive users who require them. This could be giving loss guarantees to an adaptive user that can wait to send the cells. The network would give the guarantees by informing the user when there is a low load and therefore no expected cell loss in the network. However this is a difficult problem, since adaptive users cannot give an accurate traffic description at connection set-up, the traffic they send depends in some way on network conditions at the time their application generates it. Before the proposed solution here is described to this problem, a description of general classification of user types which includes adaptive users is done.

4.2.3 Possible Types Of User

Four criteria for classifying user types are considered here (other criteria may also be useful and it is not claimed here that all possible user behaviour are described) :

- a traffic description – in terms of network traffic descriptors – is available at connection set-up;
- guarantees are required on cell delay. Since the user knows (or can determine) their own processing delays, the application's delay tolerance can be translated into a maximum acceptable delay within the network;
- guarantees are required on cell loss. In practice this means that the network guarantees to treat the user's cells as high priority;
- the user is adaptive, in the above sense

This classification is summarised in Table 4.1.

Some of the entries in Table 4.1 can be ruled out as follows :

Table 4.1: Classification Of User Types

Type	Adaptive	Traffic Descriptor	Loss Guarantee	Delay Guarantee
1	Yes	Yes	Yes	Yes
2				No
3			Yes	
4			No	
5		No	Yes	Yes
6				No
7			Yes	
8			No	
9	No	Yes	Yes	Yes
10				No
11			Yes	
12			No	
13		No	Yes	Yes
14				No
15			Yes	
16			No	

- for non-adaptive users, a traffic description is required in order to get guarantees on loss and/or delay. Therefore rows **13**, **14** and **15** can be eliminated;
- for non-adaptive users who provide suitable traffic descriptors to the network, combine the cases where they require guarantees. Therefore rows **9**, **10** and **11** can be combined into a single user type;
- for users who do not require guarantees, no traffic descriptor is needed (it is possible to think of these users as marking all their cells to be low priority, as if they had been given Unspecified QOS service with a zero PCR). Therefore rows **4** and **12** can be eliminated;

- by definition, adaptive users do not have a traffic descriptor at the start of a connection. Therefore rows 1 – 4 can be eliminated;
- adaptive users who require guarantees on loss and delay have to decide a minimum acceptable bandwidth. This is because it is not possible for the network to give guarantees to a user who cannot describe its characteristics and has to send the cells immediately and requires a certain loss probability. So this type of user essentially converts their traffic into a non-adaptive component which has a traffic descriptor and requires loss and delay guarantees (row 9), and an adaptive component which requires no loss guarantees (row 7). Therefore row 5 can be eliminated.

Table 4.2: Expected User Types

Type	Old Type	Adaptive	Traffic Descriptor	Loss Guarantee	Delay Guarantee
1	9,10,11	No	Yes		
2	16		No		
3	8	Yes	No	No	No
4	7			No	Yes
5	6			Yes	No

When these entries are eliminated, the user types shown in Table 4.2 are left. A particular user can divide their traffic input into one or more of these types :

1. *non-adaptive, requires guarantees, has a traffic description at connection set-up.* This is one of the user types addressed by various proposed CAC schemes, so it will not be discussed further in this work.
2. *non-adaptive, no guarantees required.* This is non-adaptive best-effort traffic, such as the traffic a user injects in excess of their contract with current CAC schemes.

3. *adaptive, not delay-sensitive, no loss guarantees required.* This type of user waits until feedback from the network indicates that the load is low, then sends cells on a best-effort basis. This could be implemented by negotiating a traffic contract specifying zero mean and zero peak rates. Then there is no reservation of resources and all cells will be marked as low priority. These users are called **adaptive best-effort** users here.

4. *adaptive, delay-sensitive, no loss guarantees required.* This type of user responds to network feedback by modifying how many cells they transmit, but sends them immediately on a best-effort basis. Otherwise they are similar to adaptive best-effort (type 3) users, and they are called **inelastic** users.

5. *adaptive, not delay-sensitive, requires loss guarantees.* This type of user waits until feedback indicates the network is lightly loaded, then transmits and requires that their cells are not lost in the network. These users are called **elastic** users here.

The names that are given to the types that are listed in Table 4.2 are shown in Table 4.3 with a description.

Table 4.3: Description Of Expected User Types

Type	Name	Description
1	Normal	Non-Adaptive Guarantees
2	Best-Effort	Non-Adaptive No Guarantees
3	Adaptive Best-Effort	Adaptive No Guarantees
4	Inelastic	Adaptive Delay Guaranteed
5	Elastic	Adaptive Loss Guaranteed

Traditionally the view has been that without an enforceable traffic description, no guarantees can be expected of the network. One possibility is that bursty sources can be well-described during a burst. The need for burst-level CAC has

been widely discussed, e.g. [70, 80]; however burst start time and duration can be unpredictable.

The view is taken here that adaptive users can be accommodated with a **fast reservation** scheme. A form of in-connection negotiation is proposed here, on a time scale shorter than the dynamics of connection set-up and tear down but longer than cell-scale effects. Elastic users who comply with these dynamic contracts transmit high-priority cells in the interval between successive feedback signals from the network.

4.2.4 Feedback-Based Fast Reservation

Ideally, feedback would be provided continuously to adaptive users. In practice the signalling and computation required means that time must be divided into *feedback intervals*. Feedback intervals on the order of a millisecond are expected here, so the adaptive user responses would have to be automated.

At the start of each feedback interval, the network sends a signal to each adaptive user. The responses of inelastic and adaptive best-effort users were outlined in the previous Section. Elastic users presumably have an application-dependent deadline by which transmission must be completed or the application will be aborted; assume this deadline is long-term compared to the feedback dynamics and therefore is not a factor in their response to the feedback signals.

Based on the feedback signal and their application requirements, each elastic user signals the network with the number of cells they wish to transmit in this feedback interval. The network uses buffers to take care of possible cell-scale congestion, and decides if it can accommodate each elastic users request with zero cell loss. If not, the request is denied and the user waits until the next feedback interval to try again (they can also transmit the cells immediately, at low priority, but that essentially converts them into adaptive best-effort users). If an elastic user's request is accepted, then they transmit their cells with high priority during the feedback interval and are guaranteed that the network will not lose those cells. A mechanism is probably needed to enforce the user requests, since otherwise users

who are accepted but decide not to transmit unnecessarily block other users.

The main problem with this kind of dynamic contract negotiation is that the network has to do the corresponding computations every feedback interval. If the users contending for shared network resources are geographically separated, the time to do the required signalling may be excessive even if the necessary computational power was available, and the feedback scheme would be potentially unstable.

One possibility is to localise the contention, for example by deterministically multiplexing the VP's within the network. This reduces the network to a set of "virtual channel connections" between source-destination pairs : the capacity of each such virtual channel connection is fixed for the feedback interval, and interactions within the network between traffic streams corresponding to different source-destination pairs are eliminated. Therefore the only contention between users is at the network access points. While this model is a first cut, two points should be noted :

- the loss of efficiency by not statistically multiplexing the VP's within the network may be outweighed by the increased efficiency obtained by using feedback to multiplex the users at the network edges;
- the VP capacities could also be updated, on a longer time scale than the feedback intervals, to match overall dynamics of network traffic.

4.3 Motivation For Pricing

Most researchers divide users into *classes* according to their application requirements and traffic characteristics; for example, real-time video, real-time audio, stored (off-line) video, and off-line data transfer. Each class is then regarded as having a generic user for analytical purposes. Users are usually assumed to be capable of negotiating a "contract" at connection set-up, which specifies their traffic parameters and service level. The network then polices the offered traffic and enforces the contract if users misbehave.

Two obvious user characteristics are usually ignored :

- even within a class, user preferences can be expected to be heterogeneous. If this “internal” variation within a class is substantial, then any assumptions about generic user behaviour will be inaccurate for many users in this class. One way to capture this heterogeneity is to say that users *value* their network service differently, according to their subjective requirements for the quality of their application.
- some users may be capable of responding to control signals from the network *during* their connections, not just at connection set-up. Users are part of the network and in many cases can be thought of as intelligent peripheral devices at the network edges. These intelligent peripherals can regulate how much and when traffic is input, given appropriate indicators of the network status.

4.3.1 Pricing Bandwidth

The system made up of the users together with the network has various resources that can be used to meet service demands. However in all realistic systems these resources are limited and some method of allocating them is needed when total demand is greater than the resource limit. The resources are the capacities of the ATM connections. The bandwidth allocated to a user is considered to be a commodity which is sold by the network to the user [19, 52, 54]. The view here is that the users place a benefit, or willingness-to-pay, on the bandwidth they are allocated. Given a price per unit of bandwidth, a user’s benefit function completely determines that user’s traffic input. Users are assumed to act in their own best interests and to be capable of responding to changes in the price for bandwidth.

In the formulation here, network constraints such as virtual path capacities [13] are translated into cost functions on the bandwidth demands. This is because one user’s consumption of bandwidth gives rise to fairness issues involving all the other users. This can be represented as a social ‘cost’ (in terms of longer delays, less available bandwidth, etc) which is borne by all users. The network operator sets the prices so that the marginal benefit the users place on their band-

widths is equal to the marginal cost of handling that traffic in the network¹. The network operator dynamically adjusts the prices based on monitored network conditions. The assumption here is that when the price of bandwidth increases, users reduce their demands. There is no need for enforcement of connection parameters in the scheme here because users do not declare anything about their traffic.

A typical configuration that is envisaged is shown in Figure 4.4. The user programs the terminal equipment with their benefit function. The bandwidth in cells per second allocated to the user by the network depends on this benefit function and the price that is sent from the network to the terminal equipment.

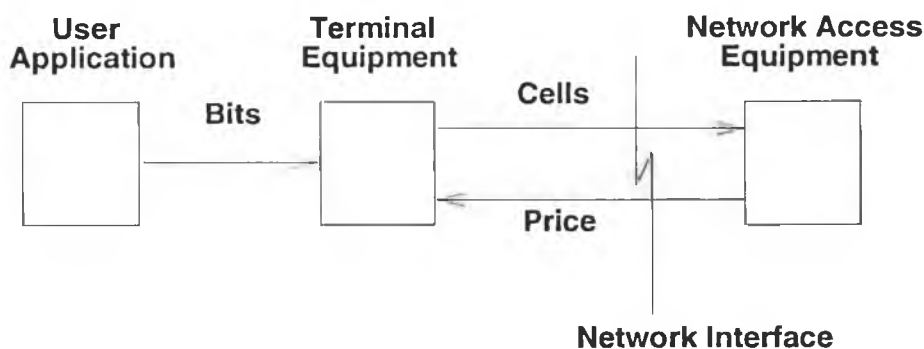


Figure 4.4: Proposed Configuration For Access

The approach here is intended to be applicable to general ATM networks and to the full range of user and service types, such as voice, video, real-time data and non-time-sensitive data. However a special case that is particularly suitable is that of a Virtual Private Network (VPN) in which one operator controls the network and all applications running on it. Why this is special is because when there is a single operator it is possible to implement pricing algorithms on all user nodes even without having any charging taking place. In a public network it would not be possible to ensure that all users would install the algorithms in the first place and then run them without tampering with them. The relative importance of various applications in the VPN could then be varied over time and from node to node. For example, the operator may place a high priority on operational data transfer and low priority on other data transfers; or interactive video and voice

¹Only address the variable costs corresponding to network constraints. The recovery of fixed network costs could be done by an access charge or some other method.

might get higher priority than other traffic types.

Dynamic priority assignment is where the priority can change from one application to the next, or even during an application, for the same type of application. This can be used to differentiate between the importance of even the same types of applications. However assigning dynamic priorities is difficult. What is also needed is adaptive priority assignment so that it could be possible to give real time services priority if there is no chance of loss but change to give data applications priority if there is likely to be loss. For example if the real-time applications such as voice and video are given priority to ensure timely delivery, then data traffic may suffer higher loss though it may not be able to tolerate cell loss as well as voice. On the other hand, if priority is given to data and large buffering is employed, then real-time applications may suffer large variable delays.

What is needed is a dynamic adaptive inter-temporal priority scheme. The priorities should change to track changes in the network state or in the application requirements over multiple time periods. Rather than have a complicated priority scheme, a pricing scheme like the one proposed could be used. The operator would set the benefit functions for the different applications, and could also set different benefit functions for applications of the same type. Each application would then input traffic according to its assigned benefit function and the current state of the network, as reflected in the prices.

4.3.2 Price As A Feedback Signal

There are two distinct scenarios that need to be considered when deciding on an appropriate feedback signal : private and public networks. In a private network (such as a LAN / MAN or a company-wide network) the “users” are applications owned and controlled by one organisation. Therefore the users are co-operative, since their responses to feedback can be programmed to obtain a desirable traffic mix. In a public network the users must be considered as separate entities, with their own private rules for deciding their traffic inputs. The network cannot assume the users will be co-operative without being given the right incentives.

A number of proposals have been made concerning the use of feedback in ATM networks, e.g. [28, 38, 86]. A crucial issue ignored in most feedback proposals is the basis on which decisions should be made, both by the network and by users. The scheme that is proposed addresses this issue by developing an economic framework in which incentives are provided to enable rational decisions and resource allocations. These incentives would correspond to actual money in a public network as this would be a good method of getting the users to adapt their behaviour. However in a private network the prices could correspond either to actual money or to control giving incentives to the users signals. These control signals would summarise the state of network resources such as bandwidth or buffer space.

4.3.3 Price As A Component Of Charging

The price that is being discussed can be thought of as a component of the total charge that the user faces from the connection. The total connection charge may be made up of many parts, like a connection charge, a time charge, a cell charge and then maybe a congestion charge as shown in Equation 4.1.

$$\text{total charge} = A + Bf\{t\} + Cf\{N\} + \pi \quad (4.1)$$

A is the amount charged for each connection set-up, B is the amount charged for each time unit, C is the charge per cell and N is the total number of cells sent in the connection and finally π is the congestion charge that the connection has encountered.

Therefore there might be a need to have a connection last a long time if the connection charge, A , were large and the time charge, B , were small. If the opposite were true and the connection charge, A , was small and the time charge, B , large then there would be many small duration connections for a single call, which is the reverse of what was there before. What can be seen is that the actual types of connections that may be seen in practice would depend on the charging strategy that is employed, that is the relative values given to A , B , C and π . The choice of these values might not be decided by purely technical considerations but may be decided by business and marketing decisions. No matter which charging

scheme is chosen the cost of its implementation is likely to be considerable. This is a point that is true for most pricing schemes and in ATM it might be one of the most important factors in deciding which scheme to implement. In this work only the congestion charge is considered and how this is related to the other charges would be a different matter.

4.4 Mathematical Model

The ATM network is modelled here as consisting of virtual paths and virtual channels, as shown in Figure 4.5. A virtual channel carries one call from its source to its destination on one physical connection, so the cells arrive in order at the destination. A connection is set up when the call arrives, and a virtual channel is established through the network. A virtual path is established by the network to simplify routing by grouping virtual channels [13].

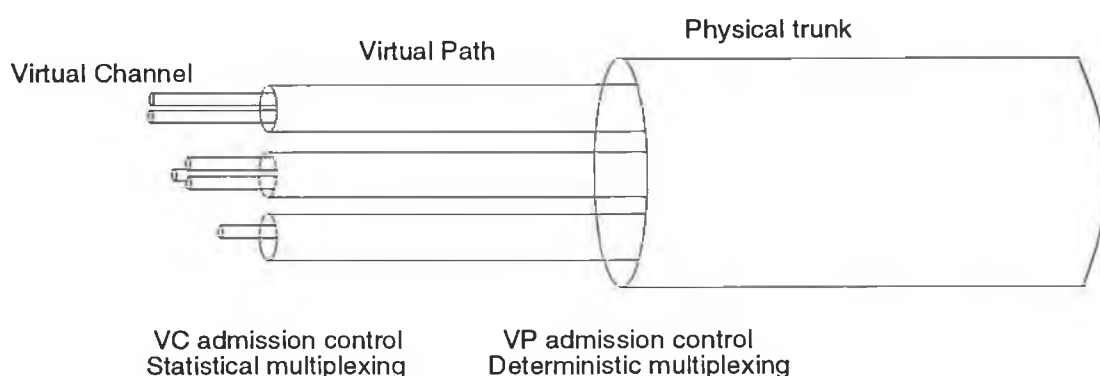


Figure 4.5: Virtual Path / Virtual Channel Connections

Assume that a virtual path is set up for every source–destination pair, and all virtual channels using a virtual path are for the same source–destination pair. Thus when a call arrives at the access to the network it is assigned a virtual channel, which is amalgamated on one virtual path with the virtual channels for other calls from that source to destination.

The virtual paths are carried by the physical network trunks, whose capacities are assumed to be fixed. Assume that statistical multiplexing of the sources is done in the sense of allocating a capacity to a virtual path which is less than

the sum of the virtual channel peak bandwidths using this virtual path. The network decides the capacities of the virtual paths and these satisfy the physical trunk capacity constraints. For example, the network could simply do deterministic multiplexing at the virtual path level, as in Figure 4.5. Deterministic multiplexing means that the sum of the virtual paths that use a physical trunk would be less than or equal to the capacity of the physical trunk. This is unlike statistical multiplexing where it is possible to divide more than the capacity of the physical trunks among the virtual paths. By doing deterministic multiplexing different virtual path capacities are not allowed to interact, at least not in real time. The virtual path capacities are assumed to be fixed. This assumption can be relaxed by having the virtual paths interact with each other at a longer time scale than the time scale of the pricing [57] without affecting the model proposed.

4.4.1 User Benefit Functions

The users connected to an ATM switch wish to send traffic to various destinations. Their benefit versus bandwidth curve for a particular connection is assumed to be concave increasing in general, as shown in Figure 4.6. This follows the usual economic assumption that users value the first part of a commodity the most, with diminishing incremental valuation as they get more of the commodity [19, 89].

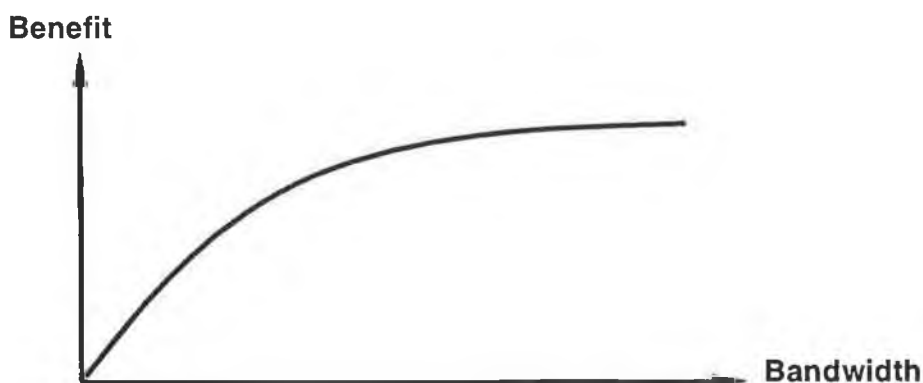


Figure 4.6: User Benefit As A Function Of Traffic Rate

Consider the problem faced by a user attached to ATM node i in deciding their bandwidth demand for a particular connection to destination ATM node j . Let the virtual path corresponding to this source–destination pair be denoted V_q .

Chapter 4 Pricing In ATM Networks

The network quotes all users a price per unit of bandwidth on V_q, π_q . Each user then uses their private benefit function ben_{rq} for connection r to decide their bandwidth at that price, b_{rq} . This decision process can be formulated mathematically as follows :

$$\text{maximise } ben_{rq}(b_{rq}) - \pi_q \cdot b_{rq} \quad (4.2)$$

The user solves a maximisation problem of this form for each connection to destination ATM node j , for each j . The optimality condition for Equation 4.2 can be solved [49] and is given by Equation 4.3 :

$$\frac{\partial ben_{rq}}{\partial b_{rq}} - \pi_q = 0, \quad \forall V_q \quad (4.3)$$

4.4.2 System Problem

Consider now the system made up of all the users and the ATM network, and suppose for the moment that there is a single 'system manager' who knows all the problem data. It will be seen that no such 'system manager' is needed.

The problem facing this (hypothetical) system manager is to choose the bandwidth demands to maximise aggregate benefit, subject to the virtual path capacity constraints. Here the virtual path capacity constraint is replaced by a corresponding buffer capacity constraint, which ensures that the ATM switch buffer occupancy, denoted as b_q , is less than the buffer size B_q but allows the input rates to temporarily exceed the output rate of the virtual path :

$$b_q \leq B_q, \quad \forall V_q \quad (4.4)$$

This constraint is dropped and a barrier function term is added to the system objective function, $Buffcost_q$. This function will be a convex function and the shape of this cost term is shown in Figure 4.7, and ensures that as the buffer capacity is approached, the 'cost' of assigning bandwidths increases to infinity.

The system problem can be formulated as follows :

$$\text{maximise } \sum_{\text{all } V_q} \left\{ \left[\sum_r ben_{rq}(b_{rq}) \right] - Buffcost_q \right\} \quad (4.5)$$

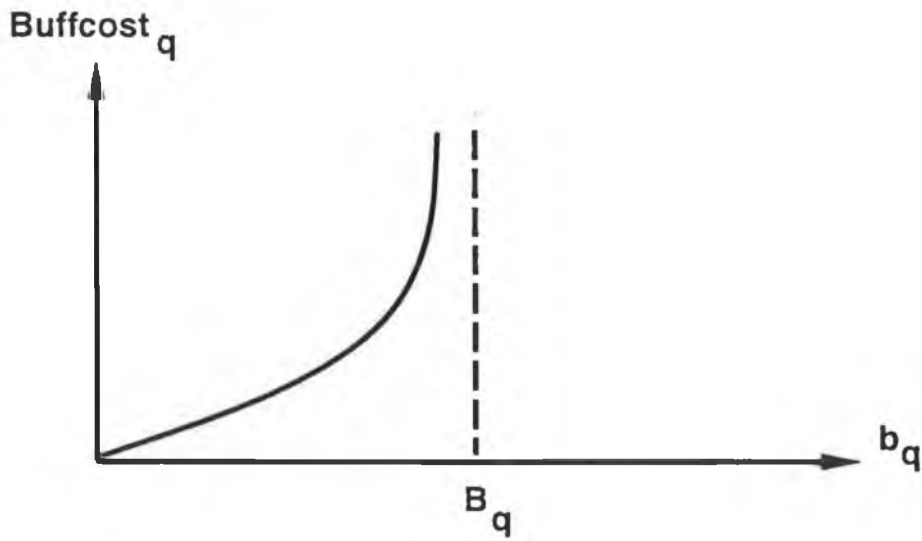


Figure 4.7: Barrier Function For q^{th} Buffer Capacity Constraint

Because of the concave nature of the functions defined, there is a guarantee that there will be an optimal solution. These optimality conditions for the system problem are [49] :

$$\frac{\partial ben_{rq}}{\partial b_{rq}} - \frac{\partial Buffcost_q}{\partial b_q} \cdot \frac{\partial b_q}{\partial (\sum_r b_{rq})} = 0, \quad \forall r, V_q \quad (4.6)$$

Comparison of Equations 4.3 and 4.6 shows that by setting :

$$\pi_q = \frac{\partial Buffcost_q}{\partial b_q} \cdot \frac{\partial b_q}{\partial (\sum_r b_{rq})} \quad (4.7)$$

individual user benefit maximisation (Equation 4.3) coincides with system optimality (Equation 4.6). Therefore if the network operator sets the prices according to Equation 4.7, the users will self-select their optimal bandwidth demands while competing for the virtual path capacity.

One of the major advantages of this scheme is that there is no need to have a 'system manager' who knows all the user benefit functions as well as the network data. This is because the prices can be set as in Equation 4.7 by the network without any knowledge of any user benefit function. The users will then use the price to maximise their private benefit functions and this in turn will give the network information to have the optimal conditions for the network. Therefore the network operator or provider can induce system-optimal user behaviour by setting the prices as in Equation 4.7 without any knowledge of the user benefit functions.

4.4.3 Distributed Pricing Algorithm

Let time be divided into successive pricing intervals of length T . Within each interval the benefit functions of all the users are assumed to be fixed. However there is nothing stopping the users changing these benefit functions from one interval to the next if they wish to do so. The price per unit of bandwidth on virtual path V_q , π_q , is announced by the network at the start of each pricing interval, and remains fixed for the duration of the interval. Hence each user solves a maximisation problem to decide their optimal average bandwidth b_{rq} , and transmits exactly $b_{rq} \cdot T$ cells during the interval. At each ATM node, the resulting buffer occupancies are measured, enabling the network operator to calculate the marginal cost of each buffer with respect to its total traffic input rate, as given by the right-hand side of Equation 4.7.

If these marginal costs are equal to the prices that were announced, then the system optimality conditions in Equation 4.6 are satisfied. If not, then the prices must be updated to correct for the difference. A summary of the negotiation between the network and the users is given by the following Bandwidth Allocation Algorithm :

- Step 0.** network operator chooses initial values for the prices $\{\pi_q\}$
 - Step 1.** network operator announces these prices to the users at the start of the current pricing interval
 - Step 2.** users respond by choosing their $\{b_{rq}\}$ according to Equation 4.3
- Pricing Interval
- Step 3.** network operator calculates the marginal costs of the buffers with respect to the bandwidths, as on the right-hand side of Equation 4.7
 - Step 4.** network operator adjusts the prices $\{\pi_q\}$ to reduce the difference between the terms in Equation 4.7 \Rightarrow go to Step 1

Note that the Bandwidth Allocation Algorithm is purely local to an ATM

node and so can be done at the ATM node to which the users are attached. This is because there is the assumption that the virtual path capacities are fixed.

4.5 Simulation & Results

4.5.1 Simulation Model

To investigate the proposed algorithm here and its efficiency, a model of a single link which represents a single virtual path was built. Typically there could be diverse services using this path and so a number of models of sources were needed. Three basic source models are used here, representing some of the services that will be demanded in an ATM network : voice, video and data. For simulation purposes a link with a bit rate of 2.048 Mbps is investigated, although the results here are expected to scale to higher link speeds.

The voice model that is used is a standard model with two states, speaking and silence [29, 12] and was described in Sub-Section 2.3.2. One of the major intended uses of ATM networks is for video applications. The video source model that is used here is a standard one for video conferencing [87] and was described in Sub-Section 2.3.3. A data source is one of the most difficult sources to model as the source type depends on what applications are being run and on what systems. For this analysis the model that is used was described in Sub-Section 2.3.4.

4.5.2 Simulation Environment

The network and source models were simulated using *SES/workbench* [88], a discrete-event simulator that allows hardware and software simulation. The models in this work were mainly created by use of its graphical user interface. *SES/workbench* compiles the graphical code to C and creates an executable. The simulation execution platform was a cluster of Sparc-10 workstations. The run time for the simulations was 1000 seconds, or 5 million cells. Sharing the 2.048 Mbps link were 20 voice sources, 2 video sources and a data source. A schematic of the simulation

set-up is shown in Figure 4.8.

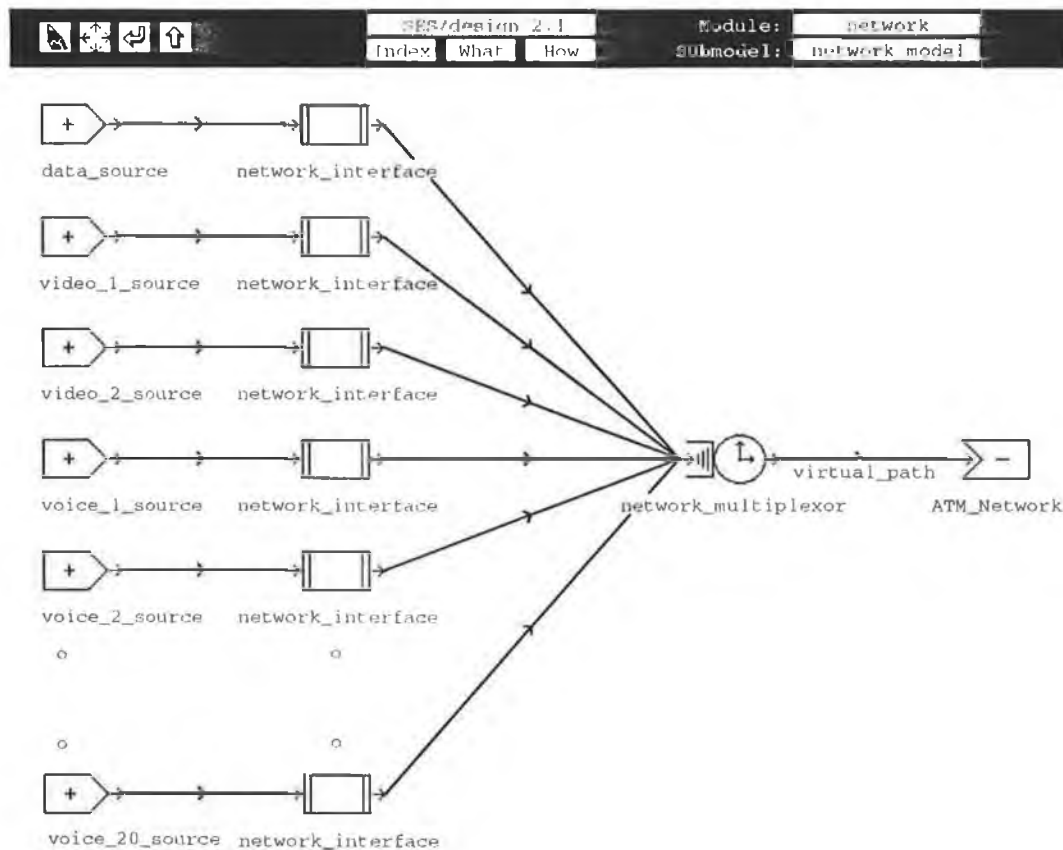


Figure 4.8: Simulation Schematic

The simulation model is made up of submodules, each of which performs a well defined function. The sources generate cells which are input to a network interface submodule. The network interface takes the source bit stream and forms ATM cells. For voice cells 5 bytes of header are added to the 48 byte information load. For data and video cells AAL 3/4 was implemented, which uses 44 bytes of information with a 4 byte AAL header and trailer in addition to the 5 byte ATM header. The cell stream from an interface is then input to the ATM switch buffer submodule. This submodule smoothes the arrival of cells to the ATM network and so takes care of cell scale congestion. This buffer is the limited resource critical to the operation of the model. In a typical implementation the buffer is managed by an input control mechanism, such as a leaky bucket scheme.

The leaky bucket is a method for policing the mean rate of a source and

still allowing bursts to occur. The mean rate is the rate at which tokens are given to allow cells past the network interface and into the network. If there were no storage of tokens then this would be a pure mean rate policing function. If the source can store tokens, it could use up all the tokens in one go by bursting cells into the network. If a source has no remaining token allowance then rather than dropping that cell, it is usually marked. If there is congestion in the network and cells have to be discarded the marked ones are discarded first. The leaky bucket does not lower the cell loss in a buffer; it merely selects which ones to lose. The pricing scheme here actually lowers the probability of cell loss by smoothing the traffic adaptively based on the buffer occupancy to avoid congestion.

In the proposed scheme here a price is generated by the network based on the present state of the network buffer, and the sources adapt their demands based on this. The leaky bucket can also be implemented on top of this scheme so that if there is cell loss the marked cell will be discarded first. The model takes in cells over an interval of about 50 milli-seconds and gives a price to all the sources sharing the link. The price reflects the congestion in the buffer (if any) and hence on the virtual path. The higher the buffer occupancy, the higher the price sent back to the sources. For these simulations the following functions are used :

$$Bufcost_q = \frac{1}{B_q - b_q} - \frac{1}{B_q} \tag{4.8}$$

The corresponding price that is generated is then the solution to Equation 4.7 which is given by Equation 4.9.

$$\pi_q = \frac{1}{(B_q - b_q)^2} \cdot \frac{\partial b_q}{\partial (\sum_r b_{rq})} \tag{4.9}$$

This price is received by the network interfaces which use it to calculate how many cells they *should* input in the next interval. For example, a data source may calculate that at the given price it should input 100 cells to the network, but if no cells are actually generated by the data source in that time interval then obviously there is no input. The outline of the distributed pricing algorithm that is used is contained in Section 4.4.3.

In an interval of about 50 milli-seconds the link can serve about 250 cells and the buffer size that is chosen is 250 cells. This means that when the buffer is full there should be no more than 250 input cells over the next interval. The input sources are shaped as much as possible by the time the cells are formed at the network interfaces, so the cells arrive approximately equally spaced in time. It is standard to give guarantees to real-time services, that if their cells get through, then the maximum access delay they experience is the buffer queueing delay, which in this case is small (50 milli-seconds).

4.5.3 Results

In practice voice and video calls may be inflexible with respect to bandwidth, and even if they are flexible they do not permit a huge statistical multiplexing gain (except for the enhancement layer of a two layer codec) compared to data. In these simulations here the assumption is that the voice and video have inflexible bandwidth demands. There are 20 voice sources which have a combined mean of 492 kbps and a potential maximum of 1413 kbps. The video has two channels, both one layer codecs, with a combined mean of 507 kbps and a potential maximum of 1773 kbps. Even without a data source there is clearly a non zero probability of cell loss. The data source has a mean of 237 kbps and a potential maximum equal to the bit rate of the link, 2048 kbps.

The results of the simulations are summarised in Table 4.4 for the two cases of no pricing used and the pricing scheme proposed here used.

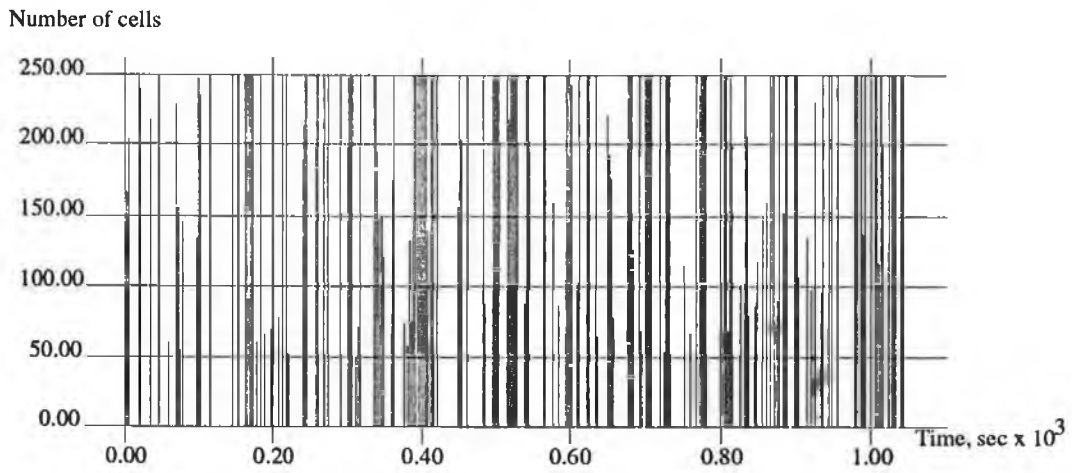
When a leaky bucket scheme alone was implemented the results in Figure 4.9.A were obtained for the sampled buffer occupancy. In Figure 4.9.B the plot of buffer occupancy is shown over the interval when the pricing scheme here was used. The buffer is not as occupied as in Figure 4.9.A . This is apparent when the loss probabilities are compared.

The pricing scheme produced no cell loss, either of marked or unmarked cells². There is a small decrease in the quality of service provided to the data

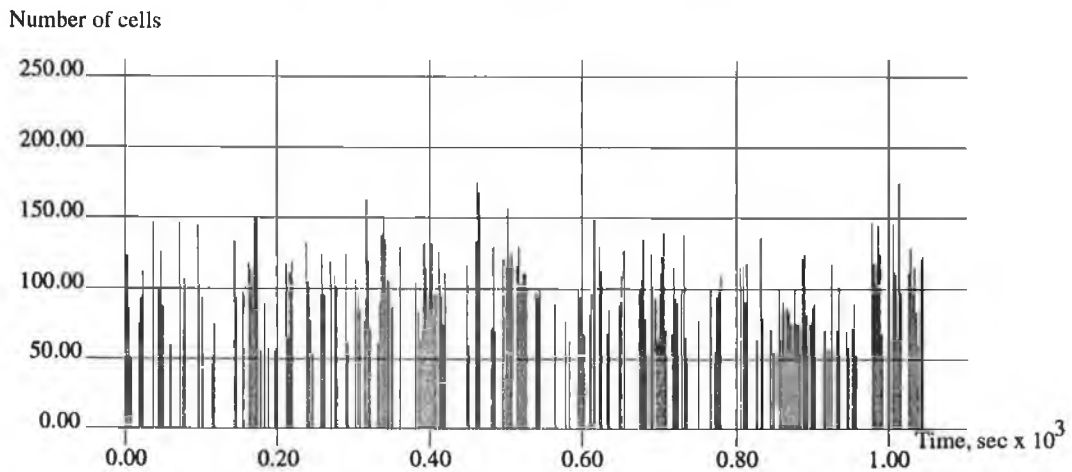
²It should be noted that there could be cell loss if a longer interval was chosen.

Table 4.4: Results For Both No Pricing & Pricing

No pricing used	
Total number of cells generated	2,914,191
Number of marked cells lost	32,268
Number of unmarked cells lost	191,044
Percentage of cells lost	7.7%
Pricing used	No cell loss



A. Without Pricing



B. With Pricing

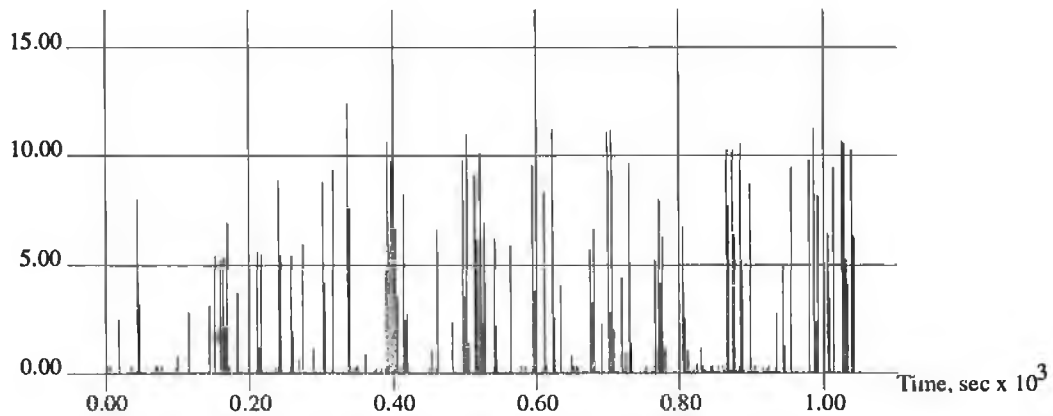
Figure 4.9: Buffer Occupancy With And Without Pricing

source to gain this improvement. Figure 4.10 A shows the sampled number of

Chapter 4 Pricing In ATM Networks

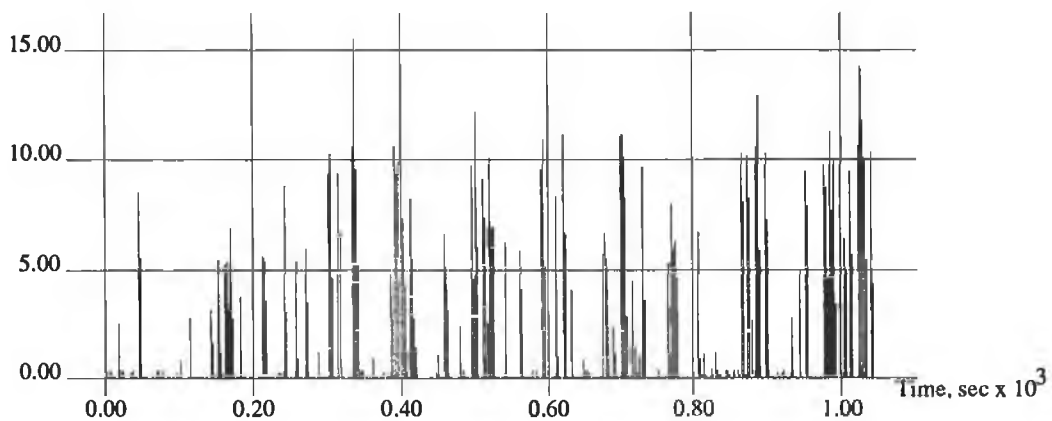
waiting data cells of data over the interval when the leaky bucket is used. By comparing Figure 4.10 A to Figure 4.10 B for the pricing scheme, it can be seen that the mean waiting time for the data source has increased.

Number of cells $\times 10^3$



A. Without Pricing

Number of cells $\times 10^3$



B. With Pricing

Figure 4.10: Delay, In Cell Times, For Data Cells With And Without Pricing

4.6 Implementation Of A Pricing Scheme

The possible implementation of any pricing scheme involves many practical difficulties, such as the volume of data to store, where to store the data and security of the data. With usage based pricing, as presented here, there are additional problems associated with the measuring and recording of the usage and hence charge to be made to a user. A simplification is made here that it will be possible to implement the pricing at the edge of the network, instead of say within the network as in a hop-by-hop or segment-by-segment fashion. The proposed set-up could look like that shown in Figure 4.11.

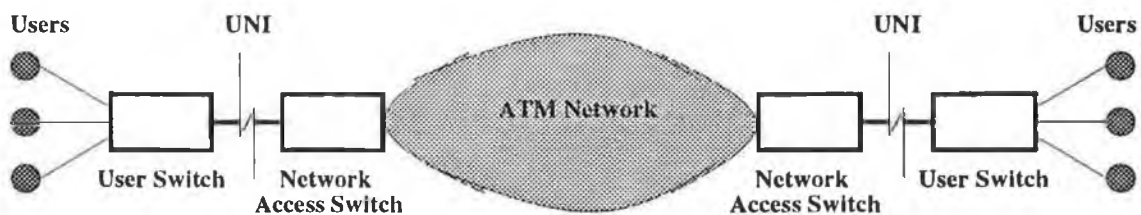


Figure 4.11: End-To-End Implementation Of A Pricing Scheme

4.6.1 Generation Of The Price

The measure of congestion that is used to determine the price from the network is the buffer occupancy. The buffer in question is the buffer for the virtual path that the connection is carried on from the access switch. Therefore all connections that are using the same virtual path at the access switch will get the same price from the network. It is implicit that with the above assumptions that the congestion of the connections using the same virtual path will be similar. However it is possible that this may not be the case and so other virtual paths that are concatenated to the first virtual path may have to signal back across the network to the access switch what their state of congestion is.

The buffer occupancy could be measured and averaged over a number of

cell times, or it could be a moving average, again over a number of cell times. However what it imagined here is that the instantaneous cell occupancy at every pricing interval is used to calculate the price. Within a switch this would mean that each virtual path would need to calculate the price from its buffer occupancy every pricing interval. However it would be possible to stagger the measurements and calculations so that they do not occur together and impose a load on the switch processor.

4.6.2 Price Passing From Network To User

To pass the price information from the switch to the users is one of the most difficult issues. One possible suggestion is to pass the price back to the users by means of a specially dedicated operations and maintenance, (OAM), cell. If the pricing interval is small, for example 1.7 milli-seconds, then there are 600 OAM cells generated per second, which equals a bit rate of 254.4 kbps. This may not seem excessive, but this would be for every connection that is connected to the switch. So if you had relatively small bit rate connections, for example sub-Mbps connections, then the overhead would be excessive. It might instead be possible to pass the price back through an OAM cell to each device connected to the switch and then let each device use that price to control all of it's connections. This would be helpful especially for data communications, as it is possible that a single processor could control multiple connections, some of which might be dormant on the second time scale.

Another possibility is to try not to generate an OAM cell for each price but rather to pass the price back to the users within other cells that are flowing in the reverse direction. The price that is generated could more than likely be encoded with 2 bytes of code, i.e., there would be at most a choice of 65,536 different prices in total throughout the network. This seems a reasonable number when consideration is given to the fact that the price is based on the buffer occupancy and so there are unlikely to be more than 65,536 different levels of buffer occupancy, or 65,536 cells in a buffer. Therefore the 2 byte price would be embedded in an reverse cell to the user from the access switch. Rather than try and use a complicated method

of interfering with U-plane cells it would be preferable to use the M-plane cells that will be generated by the network anyway to inform the customer premises equipment about the state of the connections and the network as specified by the ATM forum [4].

It is possible to imagine more complicated methods of reducing the load of transferring the price to the users. For example to stop the switch passing the price back to users who are dormant, a scheme might exist where the first cell generated by a user gets a returned price from the network, but then no other price is given until either a number of cells have been inputted within a time interval, or if no cell is inputted then no price is given back within a time interval. This would stop the generation of pricing cells to users who really have no intention of sending cells into the network.

4.6.3 User's Usage Of Price

From the viewpoint of the user, the reaction to the price will mainly be an automatic response. This could be accomplished by a computer program that would be running on the workstation where the connections are terminated. The price would be used to automatically decide on what to do about the various connections that might be in use by the same user. The programming of the response to the price could be done in advance by the user and this is what would be expected. However it is also possible that the user could alter the response of the workstation to the price if the service that was being delivered was not sufficient. This is of course feedback but at a different time scale and is shown in Figure 4.12.

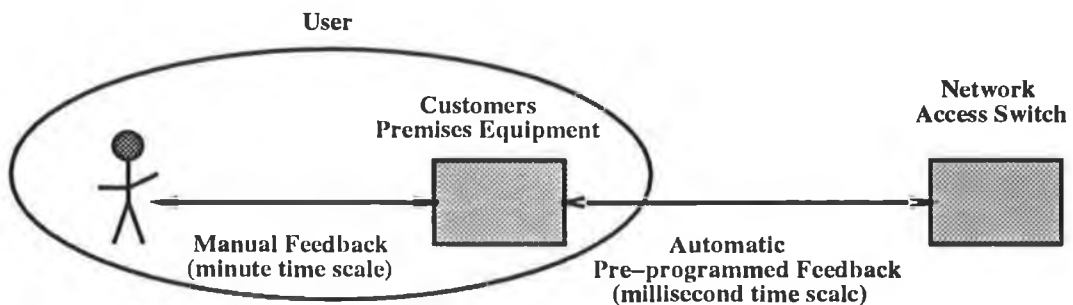


Figure 4.12: Users Response To Price

The programming of the intelligent interface to handle the price can be also customised to allow for the user to change the overall performance, without going into the details of pricing scheme, by application programs that could be written for interfacing with it.

4.7 Different Types Of Efficiencies In Networks

In focusing on user preferences, two very different notions of efficiency need to be distinguished :

- **Network efficiency** refers to the utilisation of network resources such as bandwidth and buffer space.
- **Economic efficiency** refers to the relative valuations the users attach to their network service.

If a network can maintain an acceptable level of service over time while minimising the resources necessary to provide this service, its operation is (network) efficient. If users who value network service more are served ahead of users who value it less, then that operation is (economically) efficient.

Feedback signals from the network can help to increase both types of efficiency. When the network is heavily loaded, users who are flexible about when they input traffic can be signalled to wait; when the network is lightly loaded again these flexible users can be encouraged to transmit. In this way many of the congestion problems that can occur if the offered load is regarded as fixed can be *avoided*. The flexible users absorb the longer delays and do not tie up network resources, which can therefore be used to serve less flexible users.

A cost measure can be defined for any level of network operation. Typically this cost is made up of several components corresponding to fixed (sunk) costs, variable operational costs, and so on. One crucial component is the *congestion cost*: the cost imposed on other users when a particular user adds traffic that increases network congestion (degrades service quality). If the prices quoted to

Chapter 4 Pricing In ATM Networks

the users reflect the current costs of operating the network, then those users who value network service more will choose to transmit, and users who value it less will wait for lower prices. If the network is lightly loaded then the prices will be close to zero and all users can be accommodated.

These increases in efficiency (of either type) require flexible or price-sensitive users to be more closely involved in controlling their connections than is currently the case. Not all users are capable of, or want to be involved in, dynamically adjusting their traffic inputs. How the network owner wishes to differentiate between these user preferences is not of concern here (for example, by offering a “discount” to flexible users, or charging “premium” prices to users who want a predetermined price at connection set-up).

4.7.1 Economic Framework

In this Sub-Section one possible method for computing adaptive prices is shown. The users discussed below are thus adaptive users (either adaptive best-effort, inelastic, or elastic). A more detailed description can be found in [56, 57]. Other pricing schemes for communication networks have been suggested, e.g. [48, 51].

The network and its users are considered to form an *economy* or economic system. The system has various resources such as link bandwidths and buffer spaces that can be used to meet user demands for service. Network constraints such as buffer sizes or link capacities are translated into cost functions on the demands for resources. This reflects the fact that one user’s consumption of bandwidth or buffer space gives rise to a “cost” (in terms of longer delays, less available bandwidth, longer buffer occupancies, etc) which is borne by all users.

Each adaptive user is viewed as placing a benefit, or willingness-to-pay, on the resources they are allocated. Given a price per unit of bandwidth or buffer space, a user’s benefit function completely determines that user’s traffic input. A benefit function could follow the usual economic assumption of diminishing incremental benefit as more of the resource is consumed (Figure 4.13(a)). Or it could be a simple threshold rule, or series of threshold rules, for deciding how much of

the resource to request based on the current price (Figure 4.13(b),(c)). Users are allowed to change their benefit functions every feedback interval so the examples in Figure 4.13 are for a particular interval.

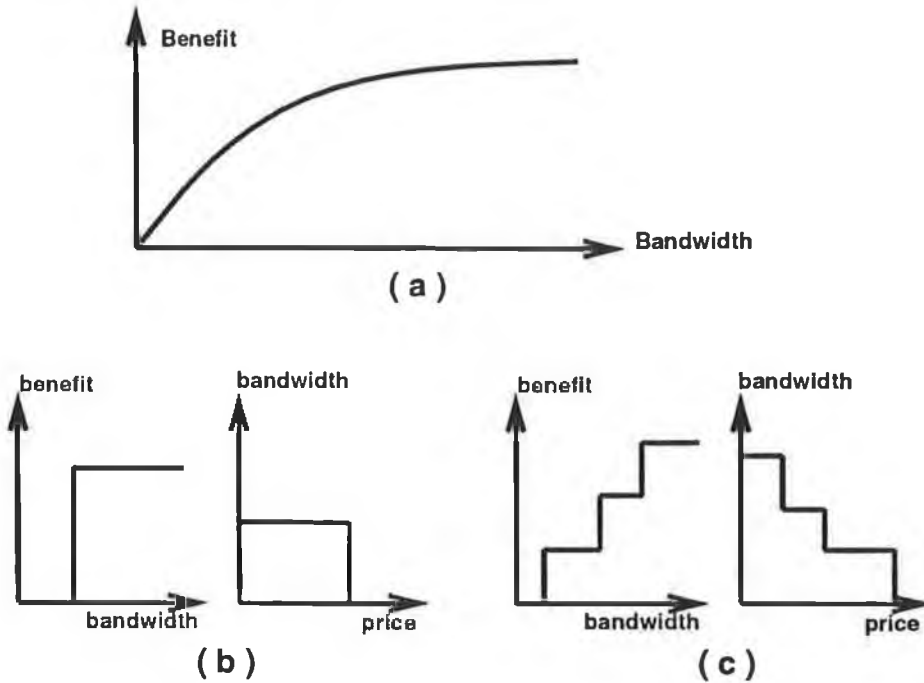


Figure 4.13: Possible User Benefit Functions

The network operator sets the prices so that the marginal benefit the users place on their resource allocation is equal to the marginal cost of handling the resulting traffic in the network³. The basic requirement is that price should go to infinity as usage of the resource approaches capacity. The network operator dynamically adjusts the prices based on current network conditions. It turns out that it is not necessary for the network operator to know the user benefit functions; therefore this pricing scheme is suitable for both public and private networks.

4.7.2 Users Models

In Sub-Section 4.2.3 the possible types of users were discussed. In this Sub-Section the **elastic** and **inelastic** users are modelled. The users have a certain file size to send over a number of pricing intervals and what has to be decided is how much to

³These prices only address the variable costs corresponding to network constraints.

send in each interval if any at all, based on the price that is given from the network.

4.7.2.1 Elastic User

There is one essential element to preferences for elastic traffic which is that the traffic has value even if it experiences delay. As shown before, however, the value should be declining the longer is the delay (if it isn't, then the decision rule to send is simple : never send unless the price is zero, or the lower bound, if not zero). Another characteristic that is not necessarily intrinsic to all elastic traffic, but seems to be characteristic of most types (e.g., files, messages), is that a transfer has zero value unless the entire message (file) is delivered.

At the start of the connection for the elastic user, the user guesses the average price and the average Peak Cell Rate (PCR) over the lifetime of the connection and these are denoted by P^* and PCR^* respectively. These could be based on past experience or be given by the network. Assume that $P^* > 0$ and $PCR^* > 0$. Assume also that the connection lifetime is long with respect to the pricing interval. This will avoid the idea of the end effects that might influence the decision of the user, and they will always be willing to wait if the decision is not to send now. Assume that the users think that the price in the next interval will go directly to the average in the next interval, regardless of the actual price now. This is the simplest possible forecasting of the user. The benefit that the user places on the connection is as follows :

$$ben\{x, v\} = v - h(d) \tag{4.10}$$

The value that the user places on the file size x cells is given by v if it were delivered without delay. When a delay of d occurs to the finish then the penalty of $h(d)$ is paid in terms of lost benefit. The elastic user is flexible with respect to delay and so the send decision involves time, in particular their estimate of the effect on overall delay of sending some cells now, as allowed by their current PCR. In general there is a recursive decision making process where either send at current PCR_k or wait till next interval, where: send at current PCR_{k+1} or wait till next interval, where: send at current PCR_{k+2} or wait till next interval, etc.

The model for the delay cost for the user, $h(d)$, would have the following attributes : $h(0) = 0$; $h(\infty) = v$. An example of this could be a function that approaches v asymptotically like $h(d) = v(1 - e^{-d})$ which is shown in Figure 4.14.

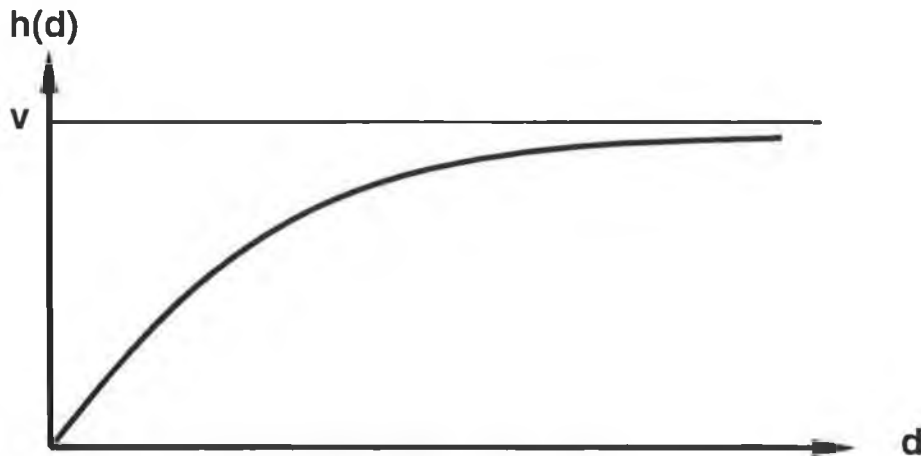


Figure 4.14: Possible User Delay Cost Function

The elastic user will decide whether to send or not in the next interval based on the fact of whether it's better to wait till the next interval to send, or if it's worth sending anything at all. Suppose that the length of the interval is given by t_{int} and that the user has already paid the amount *already-charged*(t) and that the current price is P_t and the allowable peak cell rate from the network is PCR_t . There will be a number of cells remaining to be sent from the elastic user and this is given by $X_R(t)$ which is less than X . There can be two cases as follows :

1. In this case the user can finish the whole message in the interval, so the decision to send is based on if it's better to wait till the next interval to send or not. Of course the user may have run out of time and have to drop the message with no benefit at all. $PCR_t t_{int} \geq X_R(t)$
2. In this case the user can not finish the whole message in the interval, so the decision to send is based on maybe being able to send at the average rates in future intervals as well. $PCR_t t_{int} \leq X_R(t)$

Chapter 4 Pricing In ATM Networks

Case 1

In this case if the user sends now the benefit is :

$$benefit = v - h(t) - already-charged(t) - P_t X_R(t) \quad (4.11)$$

However if the user sends nothing in the interval then there will be $X_R(t)$ left to be sent in the next interval. This will take α more intervals to complete in the forecasted version from the user, where :

$$\alpha \equiv \left\lceil \frac{X_R(t)}{PCR^* t_{int}} \right\rceil \quad (4.12)$$

The benefit that the user expects to get would then be :

$$benefit = v - h(t + \alpha) - already-charged(t) - P^* X_R(t) \quad (4.13)$$

Comparing Equation 4.11 and Equation 4.13, the user will send **now** iff :

$$h(t + \alpha) - h(t) + (P^* - P_t) X_R(t) \geq 0 \quad (4.14)$$

This means that the benefit of waiting and getting the average price and PCR will not be worth the delay in sending. Otherwise the user will wait and send nothing in the interval t .

Case 2

In this case the user can't finish the message in the interval t due to the length of the message left and the PCR available from the network. But the user can send the maximum possible now or wait till the next interval. If the user sends the maximum in this interval then there will be $X_R(t + 1)$ cells remaining at the end of the interval t where :

$$X_R(t + 1) \equiv X_R(t) - PCR_t t_{int} \quad (4.15)$$

The expected benefit to the user of sending the cells now would be :

$$\begin{aligned} benefit = & v - h(t + \beta) - already-charged(t) \\ & - P_t PCR_t t_{int} - P^* X_R(t + 1) \end{aligned} \quad (4.16)$$

where β is the expected number of more intervals that the user will need to send the remaining cells after this interval and is given by :

$$\beta \equiv \left\lceil \frac{X_R(t+1)}{PCR^* t_{int}} \right\rceil \quad (4.17)$$

Now by sending nothing now it will take α more intervals to send the message, where α is as before. The expected benefit to the user of waiting until after this interval to send anything would be :

$$benefit = v - h(t + \alpha) - already-charged(t) - P^* X_R(t) \quad (4.18)$$

Comparing Equation 4.16 and Equation 4.18, the user will send **now** iff :

$$h(t + \alpha) - h(t + \beta) + (P^* - P_t) PCR_t t_{int} \geq 0 \quad (4.19)$$

Otherwise the user will wait and send nothing in the interval t .

So in summary the elastic user acts as follows :

$h(t + \alpha) - h(t) + (P^* - P_t) X_R(t) \geq 0, \text{ for } PCR_t t_{int} \geq X_R(t)$ $h(t + \alpha) - h(t + \beta) + (P^* - P_t) PCR_t t_{int} \geq 0, \text{ for } PCR_t t_{int} \leq X_R(t)$
--

Otherwise wait and send nothing in interval t .

4.7.2.2 Inelastic User

This user has a delay bound on the traffic, but can tolerate only sending a fraction of the cells that are ready to go in the interval in question. Assume that if they are not sent in the interval then there are discarded and useless to the user. This might be like the second layer of a video codec, which only enhances the picture, but are useless if they arrive after the frame has been shown. Assume that the user has a concave benefit function, $benefit(X)$, on the number of cells sent, X . This is fixed for each interval but might vary from one interval to another or be fixed. If the user sends in interval t the number of cells X_t then the benefit to the user

is $benefit_t(X_t)$. How the user decides on the number of cells to send in interval t depends on the price given to the user P_t . The user solves the following problem :

$$\frac{\partial benefit_t(X)}{\partial X} = P_t \quad (4.20)$$

The user then solves for X and this value is the number of cells X_t sent in the interval t .

4.7.3 Another Distributed Pricing Algorithm

A distributed iterative pricing algorithm has been developed [56]. The distributed nature of the pricing algorithm suggests that it may be possible to meet the real-time feedback requirement. In addition, the computation required per iteration at each user and ATM access switch is simple, which suggests that inexpensive processing elements may be sufficient in executing the algorithm.

In a feedback-based fast reservation scheme, it may be necessary to enforce the user requests to avoid unnecessarily blocking other users. One suggestion to provide financial incentives to users to encourage truthful traffic descriptors (for both adaptive and non-adaptive users) was presented in [40]. In the scheme shown in Figure 4.15, the network operator could charge each elastic user according to their request if their actual cell input is less than that.

It is possible to extend the pricing algorithm of Figure 4.15 to include the adjustment of VP capacities using a similar iterative scheme. The VP adjustments are not purely local, since contention between users is not confined to the network access points as before. On the other hand the VP adjustments have a longer time period in which to carry out their computations.

4.7.4 Simulation Details

The simulation details are similar to those discussed in Sub-Section 4.5.2. In the proposed scheme here a price is generated by the network based on the present state of the network buffer, and all the sources adapt their demands based on this.

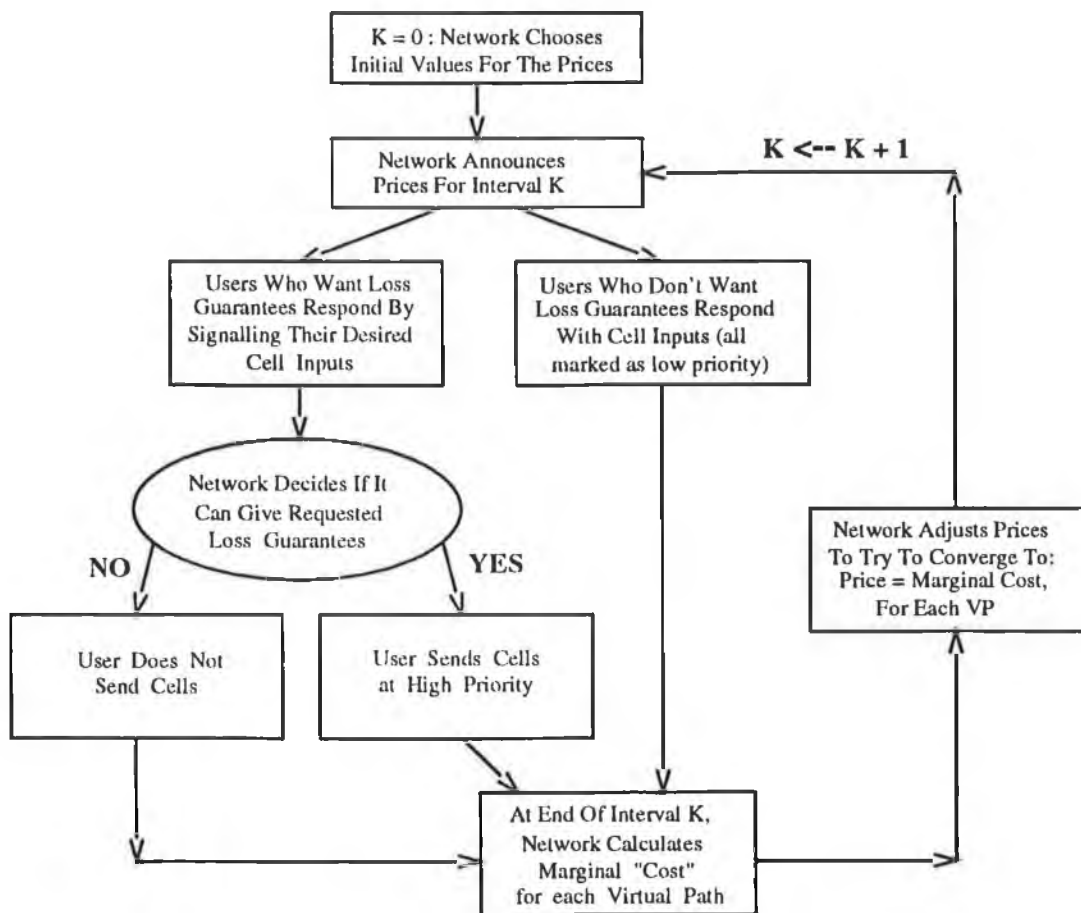


Figure 4.15: Distributed Iterative Pricing Algorithm

What is proposed and simulated here adheres to the UNI 3.0 specification from the ATM Forum [4]. A leaky bucket can also be implemented on top of the scheme here so that if there is cell loss it is possible to discard the marked cells first.

The model takes in cells over a pricing interval and gives a price to all the sources sharing the link. The price reflects the congestion (if any) in the buffer and hence on the virtual path. The pricing interval is short compared to the video frame time : a value of about 0.05 of a frame time was chosen. To achieve feasible run times cell scale effects are neglected. This neglecting of cell scale effects is critical to the speed up of the simulations. The total utilisation of the link is then high, at a value of around 0.85. It is possible to compare the performance of the pricing scheme to other access control schemes by looking at the loss and at the value of the traffic carried.

The initial set up is a high speed ATM 155 Mbps link with two types of users

Chapter 4 Pricing In ATM Networks

connected, inelastic and elastic. The inelastic users are modelled as video sources and the elastic users are data sources. The link model is shown in Figure 4.16.

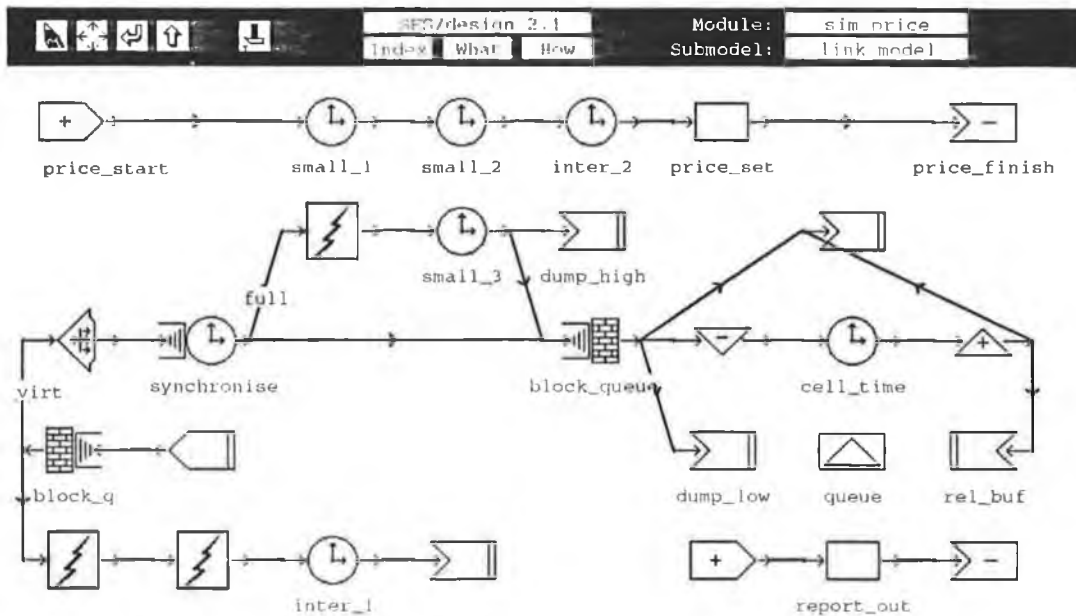


Figure 4.16: Simulation Model For Economic Efficiency

The video source model that is used here is a standard one for video conferencing [87]. The codec is a speeded up one layer model. Each of the 20 video sources have a mean of 2.3 Mbps and a peak of 5 Mbps. These are input at a rate of 30 frames per second, all synchronised together, so that the resulting inelastic users are (more or less) stationary on the milli-second scale, that is with respect to the pricing interval.

The elastic users can be thought of as one user with a lot of files to transfer independently, many users each with one file, or some combination of these types. The negotiations for file transfer or connection set up will only occur when a new video frame is to be sent, i.e. every 1/30 second. This makes the modelling for high speed easier and is not restrictive. Therefore the network re-negotiates the PCR every 1/30 second with the elastic users. A data source is one of the most difficult sources to model as the source type depends on what applications are being run and on what systems. A file transfer application is modelled. This captures the bursty nature of data communications as well as its looser delay requirements relative to voice and video. A model was built based on transferring files from one

computer to another. An empirical distribution for file size ranges was obtained from actual files stored on one of the computers in DCU. In the simulations a range was chosen according to this empirical distribution, and then a file size was chosen from a uniform distribution within this range. The amount of data transfer can be varied depending on how many files are to be transferred. The peak-to-mean ratio of this source can be high with values up around 1000. There are on average 20 data sources in use and these are taken from a uniform distribution between 1 and 39 sources. Each file to be sent is also taken from a uniform distribution between 20 and 660 cells. Therefore the average bit rate of a single data source is about 4.3 Mbps.

The simulation was speeded up by use of time stepping, where the time step was chosen to be the frame time. This was possible as everything was known at the start of the frame time, except the pricing, but if the price was low and the link un-congested then the pricing could be skipped. There was also an approximation to the price that the network generated. If Equation 4.8 is computed on the cell level then the price generated will be between 0 and 1 for all values of $b_q \leq B_q$ but when $b_q = B_q$ then the price given by the network is ∞ . To avoid this problem the equation is changed slightly to give Equation 4.21 :

$$Buffcost_q = \frac{1}{B_q - b_q + 1} - \frac{1}{B_q + 1} \quad (4.21)$$

The next detail is how to calculate the peak cell rate that the network allows the elastic users. This is calculated based on the current level of the buffer and the number that can be served in an interval. If the buffer is empty then it is desirable that the *PCR* to be equal to the number that can be served in an interval, and then find out how many active sources there are and divide that amount up equally into all of them. However when the buffer starts to fill it is desirable to not only reduce the *PCR* but also to clear the buffer. To do this subtract the number in the buffer multiplied by 1.5 from the total number that can be served. This is shown in Equation 4.22 :

$$PCR_{for\ each\ source} = \frac{number\ that\ can\ be\ served - 1.5\ b_q}{number\ of\ active\ sources} \quad (4.22)$$

Chapter 4 Pricing In ATM Networks

There is the possibility of oscillation in the price and this can be tuned by changing Equation 4.22. The average price, P^* , that the user thinks that they will get from the network and the average peak cell rate, PCR^* , are given by :

$$P^* = 0.98 P^* + 0.02 P \quad PCR^* = PCR \quad (4.23)$$

In the simulations here, resource costs were chosen as barrier functions [49] to give a smooth increase as resource usage approached its limit. The simulations were carried out for 20 seconds of real time, which translates into 7.3 million cell times.

4.7.5 Results

The results in Table 4.5 show the difference between using adaptive pricing and no pricing at all. What can be seen is that both the network efficiency and the

Table 4.5: Results Of Both Network And Economic Efficiency

Source Type		% Loss	User Value	% Dec. Loss	% Inc. Value
Unpriced	Inelastic	0	240	91.0	14.8
	Elastic	30.4	146		
	Combined	19.1	386		
Priced	Inelastic	4.4	239		
	Elastic	0.1	204		
	Combined	1.7	443		

economic efficiency can increase by using pricing at the same time. Here there was an increase of 14.8% in the economic benefit to the users as well as a decrease of 91% in the number of cells lost.

4.8 How Pricing Fits Into The ATM Controls

There are two main functions in the area of congestion and control of ATM sources, namely connection admission control, (CAC), and usage parameter control,

(UPC), as is shown in Figure 4.17. ATM is connection based and the understanding is that each connection will be able to describe itself and will require a certain quality of service, (QOS). The description of the connection will be by using some parameters that are useful to the network in deciding if it will accept the connection into the network and give it the required QOS while still maintaining the QOS to all the other connections in the network. Therefore the CAC problem is multilevel and distributed across the whole network. Many methods have been proposed to break this problem into smaller more manageable ones. The most promising solution is that proposed by effective bandwidths (see Sub-Section 1.5.3) which effectively describe a connection independently of the other connections sharing the network and so decouples the problem. There is a further advantage of effective bandwidths in that the constraint on the network resources is linear in the effective bandwidths.

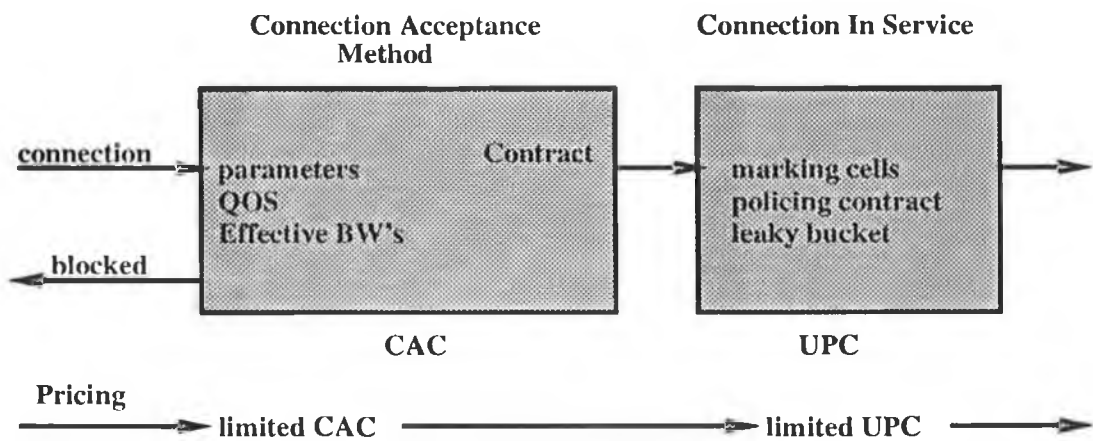
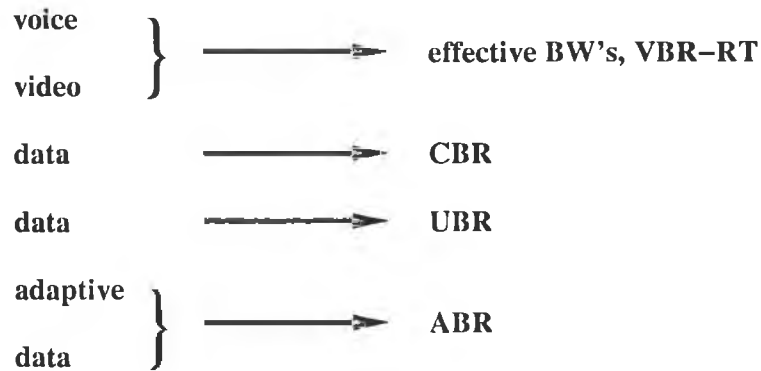


Figure 4.17: CAC, UPC & Pricing

The overall picture of the ATM admission scheme and congestion control might be as follows. The various sources like voice and video that are compressed and variable bit rate would have an associated effective bandwidth and would be admitted based on a CAC and then policed by a leaky bucket scheme to ensure compliance with the contract. These sources should efficiently add together to give network efficiency. The constant bit rate sources would similarly be admitted to the network by a CAC and be policed. However there are a number of sources that will not be able to undergo the above schemes. The source that is in mind is

that of a data connection where there may be no description in advance in terms of effective bandwidth and so it may not be possible to include the above connection in this scheme. The other type of source is that of an adaptive source, where it is possible for the source to adapt it's behaviour to network or other conditions without degrading the QOS of the connection. This is shown in Figure 4.18.



VBR-RT variable bit rate; CBR constant bit rate; UBR unspecified bit rate; ABR available bit rate

Figure 4.18: Which Sources Pricing Applies To

Pricing would be affected by the CAC and UPC functions of the network. It is possible that the CAC would limit the number of adaptive sources into the network, or similarly the CAC would take into account how many priced source were admitted, but this is not necessary for the scheme to work effectively. A limited CAC might be to impose a peak cell rate, (PCR), on the adaptive sources admitted. There might also be a limited amount of UPC on these sources then to ensure that they comply with this PCR. The cells from the priced sources might be policed to ensure that they are low priority and hence would not interfere with the other cells that have been given guarantees by the network.

4.9 Discussion

This chapter found that there is a class of users, called adaptive users, that may require guarantees from the networks and this may be possible by allowing the users and network to co-operate. This chapter suggests pricing and user self-regulation

Chapter 4 Pricing In ATM Networks

as a means of allocating bandwidth in ATM networks. The pricing scheme is independent of the traffic arrival model and does not require the network to police and enforce user commitments in real-time as cells arrive from the sources. The pricing algorithm is distributed and local to the ATM access nodes, suggesting that it may be able to meet the tight time requirements of an admission control scheme. There is no need for the network to define and support traffic classes. The users have more flexible access to network resources than in other CAC schemes, but they take the responsibility of determining how much network resources to acquire in order to meet their particular needs. It should be noted that this view of the user-network relationship should not be judged based on the feasibility or otherwise of the pricing scheme here. Other CAC schemes may also be suitable and it remains to be seen whether the scheme here is compatible with other approaches, perhaps in a multi-level solution approach. Similarly the pricing scheme here offers users the possibility of taking more control over their service than they have under traditional system models, but it does not *require* them to. Users whose applications have inflexible bandwidth demands, or who set a threshold price above which they demand zero bandwidth, can be accommodated in this formulation. In order to pursue an inflexible strategy, users may have to pay premium prices at times of heavy demand; but it is believed here that simply being forced to examine the relative importance of their usage will give rise to a variation in user valuations that a scheme like ours can exploit. The success of this approach here depends on having a sufficient number and range of flexible applications, although how much is enough is as yet unknown.

It has also been shown that it is possible to maximise both the network and economic efficiency of the network at the same time. However it is felt here that it is still possible to use economics in networks even if it is not part of a pricing scheme, for example in a private network. In a VPN pricing can be used to choose the desired application mix and vary this mix dynamically, as demonstrated in a simple case by the results here. Another use of pricing is in selectively dropping cells during congestion. Suppose that an application uses Ethernet frames and one cell is dropped from a frame. Then the benefit of the rest of the frame should be

set to zero if the whole frame has to be retransmitted. Similarly, if there are a number of video calls in progress, it may be better to drop all the cells from one connection during congestion rather than to spread out the loss. The feasibility of using this pricing approach here as a dynamic adaptive priority scheme in this way is one avenue for further research.

The pricing scheme here also applies to general ATM networks in which the user benefit functions are unknown to the network operator. There are many issues which would have to be resolved in a practical implementation. For example, the assumption here is that the benefit of making a connection is associated entirely with the sender. But in practice the benefit is shared, and some pre-connection negotiation may be necessary to agree on connection parameters acceptable to all parties. If intermediate ATM nodes switch on the VCI as well as on the VPI [20], users at different ATM nodes have to compete for shared resources, which complicates the algorithm and may increase the minimum length of a pricing interval. Whether the prices used in the scheme here could be used as part of an ATM network tariff is not clear. However it is believed here that the paradigm suggested is an interesting one and may be helpful in solving these difficult network management problems.

It has been shown here that pricing – and feedback in general – may be useful in resource allocation and congestion control for integrated-services networks. At the very least, such mechanisms deserve to be investigated : as an alternative to more “traditional” allocation and control mechanisms, or in tandem with them.

Chapter 5

Conclusions

There are a number of conclusions that can be drawn from the resource allocation methods in ATM networks that have been discussed here. The issues involved in modelling and simulation were presented in Chapter 2. There are multiple levels of resource allocation in ATM networks but the two that have been discussed here are the cell level, in Chapter 3, and the connection level, in Chapter 4.

After studying various methods of modelling and simulation techniques in Chapter 2, and why these are important in ATM, it is seen that there is a need to present specialised models for ATM network performance. These specialised models are needed both to speed up the simulations and to decrease the number of simultaneous events taking place in the model. Speed up factors of over **6000** have been achieved in the satellite simulation in Chapter 3 which was achieved by careful design of the problem being pursued and taking into account the decomposition methods as well as the cell rate and step-by-step methods. The end results are simplified approximate models for performance analysis that have been implemented and that are verified, validated and credible to the user.

At the cell level it has been shown that the way that the cells are put together and also the order and timing between them are important considerations in some cases, as discussed in Chapter 3. The way the cell is formed for an application is the ATM adaptation layer, (AAL) and for satellite links there are problems when using the AALs that have been standardised so far. The problems are in terms of error correction, error detection, the percentage throughput and the maximum

Chapter 5 Conclusions

throughput. In Chapter 3 after analysing and modelling the errors on the satellite links in the Deep Space Network a new AAL is proposed to allow data transmission that incorporates selective retransmission techniques. This new AAL allows both data to be guaranteed delivery, even with errors, and also improves the efficiency of the AAL by 7.5% over the AAL 3/4. The order and the timing of the cells is also shown to be important when looking at the performance of the connection policer called the leaky bucket in Chapter 3. The issue of the worst case traffic from a number of sources using the leaky bucket is an open problem. It was believed that the greedy on-off source was the worst case, even for two identical sources. It is shown in Chapter 3 that neither of the two traditional worst case sources produces the highest cell loss all the time. There may be of course an as yet unknown source that is worst case over all permutations of the sources.

The connection level resource allocation issues are presented in Chapter 4. It is shown that there is a need to have a new class of user called an adaptive user. This comes about because some users will not be catered for properly when looking at the current connection admission control schemes. There is a gap in user types when it is either not possible to describe the source or if it is possible to adapt the behaviour of the source automatically or very quickly. The possible types of users are categorised in terms of :

1. the ability to provide a traffic description that is useful to the network
2. the need for guarantees in terms of loss and / or delay
3. whether the source is adaptive or not.

The adaptation can be either in terms of delay or loss. After considering this, it is then possible to give sources that do not have a traffic descriptor guarantees on their connection, while still providing all the previously held guarantees and contracts of the other users.

There is a proposal also on how to achieve this adaptation in Chapter 4. It is shown both that there is nothing inherently monetary in applying economic principles to networks and that a possible feedback signal for adaptive users is price.

Chapter 5 Conclusions

A mathematical model is proposed for pricing, in which the congestion problem is turned into a constrained optimisation problem. It is then found that by the users acting independently, but optimising their benefits from the network, that the system optimisation is achieved. There is therefore no need for a system manager or overseer of the network, all the network has to calculate is the price to give the user, and this is found from the constraint of the resources. By using the results from Chapter 2 it is possible to carry out simulations. The results show that with pricing the cell loss is reduced and the buffer occupancy is also reduced. Without pricing there was over 7% cell loss and that with pricing there was no cell loss.

In Chapter 4 it is also shown how a pricing scheme could be implemented and what issues need addressing. A more detailed examination of efficiency is carried out with the conclusion that there are two types of efficiencies :

1. network efficiency
2. economic efficiency

There is a detailed mathematical model of the different types of expected adaptive users, where the adaptation can be either by time or by loss. A framework for the pricing and the effect on the efficiencies is formulated. By simulation it is possible to examine the effect of pricing on the different types of efficiencies. The conclusion is that it is possible to increase both the economic and network efficiency at the same time by using pricing. By using pricing a 91% decrease in cell loss and also a 14.8% increase in economic benefit to the users was achieved simultaneously. Therefore it has been shown that pricing can be used to utilise adaptive user behaviour and can lead to an increase in the efficiency of the network.

Currently there are problems with resource allocation methods in ATM networks – however it has been shown (through mathematical modelling and simulation) that pricing, and feedback in general, can be helpful in solving them.

References

- [1] S. Akhtar, 'Congestion Control In A Fast Packet Switching Network', Master's thesis, Washington University, 1987.
- [2] R. Ahmad, F. Halsall, K. Lam and J. Zhang, 'Performance Evaluation of Private ATM Networks', *10th UK IEE Teletraffic Symp.*, BT Laboratories, Martlesham Heath, April 1993.
- [3] G. J. Armitage and K. M. Adams, 'Using the Common LAN to Introduce ATM Connectivity', *IEEE 18th Conf. on Local Computer Networks*, Sept. 1993.
- [4] *ATM Forum*, 'ATM User-Network Interface Specification', Prentice Hall, 1993.
- [5] V. E. Benes, 'Programming and control problems arising from optimal routing in telephone networks', *Bell Syst. Tech. J.*, Vol. 45, pp 1373-1438, 1966.
- [6] D. Bertsekas, and R. Gallager, *Data Networks*, 2nd ed., Prentice Hall, 1992.
- [7] E. W. Biersack, 'Performance Evaluation of Forward Error Correction in an ATM Environment', *IEEE J. Select. Areas Commun.*, Vol. 11, No. 4, 1993, pp. 631-640.
- [8] M. Bocci, J. M. Pitts and L. G. Cuthbert, 'Exploiting concurrency through knowledge of event propagation in cell rate ATM simulation', *Twelfth UK IEE Teletraffic Symp.*, Windsor, UK, 15-17 March 1995.
- [9] M. Bocci, J. M. Pitts and E. M. Scharf, 'Performance of Time Stepping Mechanism for Parallel Cell Rate Simulation of ATM Networks', *Eleventh UK IEE Teletraffic Symp.*, Cambridge, UK, 23-25 March 1994.

References

- [10] F. Bonomi and K. Fendick, 'The Rate-Based Flow Control Framework for the Available Bit Rate ATM Service', *IEEE Network*, Vol. 9, No. 2, March 1995.
- [11] D. D. Botvich, T. Curran, A MacFhearraigh and S. Murphy, 'Hierarchical Approach to Video Source Modelling', *Eleventh UK IEE Teletraffic Symp.*, Cambridge, UK, 23-25 March 1994.
- [12] P. T. Brady, 'A Statistical Analysis of On-Off Patterns in 16 Conversations', *Bell Syst. Tech. J.*, pp 73-91, 1968.
- [13] J. Burgin and D. Dorman, 'Broadband ISDN Resource Management : The Role Of Virtual Paths', *IEEE Commun. Mag.*, Vol. 29, No. 9, pp 44-48, Sept. 1991.
- [14] E. T. Chow, and R. W. Markley, 'Asynchronous Transfer Mode Link Performance Over Ground Networks', *TDA Progress Report 42-114, JPL*, pp. 185-189, August 1993.
- [15] D. Corry, 'Simulation and Performance of a Statistical Multiplexor in an ATM Network', *Master of Engineering*, DCU, 1993.
- [16] R. L. Cruz, 'A Calculus for Network Delay, Part I: Network Elements in Isolation', *IEEE Trans. on Information Theory*, Vol. 37, No. 1, pp 114-131, Jan. 1991.
- [17] D. de Burca, 'An Investigation into Computer Simulation of High-Speed Networks', *B.E. Electronic Engineering Project Report*, DCU 1995.
- [18] P. Fonseca, J. Pitts and L. Cuthbert, 'Implementing a Space Priority Mechanism in a Cell-Rate ATM Network Simulator', *Eleventh UK IEE Teletraffic Symp.*, Cambridge, UK, 23-25 March 1994.
- [19] A. de Fontenay, M.H. Shugard and D. S. Sibley, *Telecommunications Demand Modeling : An Integrated View*, North-Holland, 1990.
- [20] M. de Prycker, *Asynchronous Transfer Mode : Solution for Broadband ISDN*, 2nd Ed., Ellis Horwood, 1993.

References

- [21] G. de Veciana, C Courcoubetis and J. Walrand, 'Decoupling Bandwidths For Networks: A Decomposition Approach to Resource Management', *UCB / ERL Technical Report M93 / 50, Submitted to IEEE / ACM Trans. Networking*, June 1993.
- [22] M. de Prycker, R. Peschi and T. van Landegem, 'B-ISDN and the OSI Protocol Reference Model', *IEEE Network*, Vol. 7, No. 2, pp 10-18, March 1993.
- [23] B. T. Doshi, 'Deterministic Rule Based Traffic Descriptors For Broadband ISDN: Worst Case Behaviour And Connection Acceptance Control', *International Teletraffic Congress, ITC-14, Vol 1a*, pp 591-600, Antibes, France, June 1994.
- [24] B. Erimli, J. Murphy and J. Murphy, 'On Worst Case Traffic in ATM Networks', *Twelfth UK IEE Teletraffic Symp.*, Windsor, UK, 15-17 March 1995.
- [25] D. Ferrari, J. Ramaekers and G. Ventre, 'Client-Network Interactions in Quality of Service Communication Environments', *IFIP Transactions C (Communication Systems)*, Vol. C-14, pp 221-234, 1993.
- [26] R. F. Garzia and M. R. Garzia, *Network Modeling, Simulation, and Analysis*, Marcel Dekker, NY, 1990.
- [27] H. Gilbert, O. Aboul-Magd and V. Phung, 'Developing a Cohesive Traffic Management Strategy for ATM Networks', *IEEE Commun. Mag.*, Vol. 29, No. 10, pp 36-45, Oct. 1991.
- [28] Y. Gong and I. F. Akyildiz, 'Dynamic Traffic Control Using Feedback and Traffic Prediction in ATM Networks', *Proc. of IEEE INFOCOM '94*.
- [29] J. G. Gruber, 'A Comparison of Measured and Calculated Speech Temporal Parameters Relevant to Speech Activity Detection', *IEEE Trans. Commun.*, Vol. COM-30, No. 4, pp 728-738, 1982.
- [30] T. Hanson and J. Murphy, 'Internetworking Asynchronous Transfer Mode with the Consultative Committee for Space Data Systems Advanced Orbiting Sys-

References

- tems Protocols', *Third International Conference on Telecommunication Systems, Modelling and Analysis*, Nashville, USA, 16-19 March 1995.
- [31] T. Hanson, and R. Markley, 'A Proposed Communications Architecture for the First Lunar Outpost', *Proc. of Globecom '93*, Houston, Texas, 29 November - 2 December, 1993.
- [32] T. Hanson, and R. Markley, 'Communications Requirements for the First Lunar Outpost', *Proc. of IEEE Southeastcon '93*, Charlotte, North Carolina, April 1993.
- [33] P. G. Harrison and N. M. Patel, *Performance Modelling of Communication Networks and Computer Architectures*, Addison-Wesley, 1992.
- [34] D. P. Heyman and T. V. Lakshman, 'Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks', *IEEE Trans. Circuits and Systems For Video Tech.*, Vol. 2, No. 1, pp. 49-59, March 1992.
- [35] D. Hong and T. Suda, 'Congestion Control and Prevention in ATM Networks', *IEEE Network*, pp 10-16, July 1991.
- [36] J. Y. Hui, *Switching and Traffic Theory for Integrated Broadband Networks*, 1st. Ed., Kluwer, Boston, 1990.
- [37] *ITU-T I.311*, 'B-ISDN General Network Aspects', ITU 1993.
- [38] H. Kanakia, P. P. Mishra and A. Reibman, 'An Adaptive Congestion Control Scheme for Real-Time Packet Video Transport', *Computer Communication Review*, Vol. 23, No. 4, pp 20-31, Oct. 1993.
- [39] F. P. Kelly, 'Effective Bandwidths at Multi-Class Queues', *Queueing Systems Theory and Applications*, Vol. 9, No. 1-2, pp 5-15, Oct. 1991.
- [40] F. P. Kelly, 'Tariffs, Policing and Admission Control for Multiservice Networks', *10th UK IEE Teletraffic Symp.*, BT Laboratories, Martlesham Heath, April 1993.

References

- [41] D. E. Knuth, *The Art of Computer Programming*, Vol. 2, Semi-numerical Algorithms, 2nd ed., Addison-Wesley, 1981.
- [42] P. Komisarczuk, M. H. Hadjitheodosiou, F. Coakley and C. Smythe, 'B-ISDN Implementation Via Satellite : The Catalyst Project', *10th UK IEE Teletraffic Symp.*, BT Laboratories, Martlesham Heath, April 1993.
- [43] J. Lane, *Asynchronous Transfer Mode : Bandwidth for the Future*, 1st ed., Telco, 1992.
- [44] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*, 2nd Ed., McGraw-Hill International Editions, 1991.
- [45] W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, 'On the Self-Similar Nature of Ethernet Traffic', *Computer Communication Review*, Vol. 23, No. 4, pp 183-193, Oct. 1993.
- [46] I. M. Leslie, D. R. McAuley and D. L. Tennenhouse, 'ATM Everywhere ?', *IEEE Network*, Vol. 7, No. 2, pp 40-46, March 1993.
- [47] S. H. Low and P. P. Varaiya, 'A New Approach to Service Provisioning in ATM Networks', to appear in *IEEE Trans. on Networking*.
- [48] S. Low and P. Varaiya, 'A Simple Theory of Traffic and Resource Allocation in ATM', *Proc. of Globecom '91*, pp 1633-1637, Dec. 1991.
- [49] D. G. Luenberger, *Linear and Non-linear Programming*, 2nd Ed., Addison-Wesley, 1984.
- [50] P. McClure, 'GCF Digital Circuit Reliability', *Final Draft of Proposed TDA Progress Report, JPL*, January 15, 1993.
- [51] J. K. MacKie-Mason and H. R. Varian, 'Pricing the Internet', presented at *Public Access to the Internet*, JFK School of Government, Harvard University, May 26-27, 1993.
- [52] J. K. MacKie-Mason, L. Murphy and J. Murphy, 'The Role of Feedback in Network Allocation', *MIT Workshop*, Cambridge, USA, 7 March 1995.

References

- [53] J. MacKie-Mason, J. Murphy and L. Murphy, 'Using Price Feedback for ATM Efficiency', *Third International Conference on Telecommunication Systems, Modelling and Analysis*, Nashville, USA, 16-19 March 1995.
- [54] J. R. Meyer, R. W. Wilson, M. A. Baughcum, E. Burton and L. Caouette, *The Economics of Competition in the Telecommunications Industry*, Oelgeschlager, Gunn and Hain, 1980.
- [55] J. Murphy, E. Chow and R. Markley, 'ATM Service-Based Selective Retransmission Over DSN Satellite Links', 15th AIAA International Communication Satellite Systems Conference, San Diego, USA, March 1994.
- [56] J. Murphy and L. Murphy, 'Bandwidth Allocation By Pricing In ATM Networks', *IFIP Transactions C-24: Broadband Communications II*, pp 333-351, North-Holland, 1994.
- [57] J. Murphy, L. Murphy and E. C. Posner, 'Distributed Pricing For Embedded ATM Networks', *International Teletraffic Congress ITC-14*, Antibes, France, June 1994.
- [58] J. Murphy, E. Chow and R. Markley, 'Systems Design of a Satellite Link Protocol', *to appear in the American Institute of Aeronautics and Astronautics Journal of Spacecrafts and Rockets, AIAA JSR*, 1995.
- [59] J. Murphy and J. Teahan, 'Video Source Models for ATM Networks', *Eleventh UK IEE Teletraffic Symp.*, Cambridge, UK, 23-25 March 1994.
- [60] J. Murphy and T. Curran, 'Optimising Embedded ATM Networks', *Tenth UK IEE Teletraffic Symp.*, Ipswich, UK, April, 1993.
- [61] J. Murphy, 'Modelling Of High Speed Data Network Protocols', *First SES/european Users Group Conference*, Leatherhead, UK, 16-17th February, 1994.
- [62] J. Murphy, 'Worst Case Traffic in Asynchronous Transfer Mode Networks', *B.E. Electronic Engineering Project Report*, DCU 1995.

References

- [63] L. Murphy and J. Murphy, 'Pricing for ATM Network Efficiency', *Third International Conference on Telecommunication Systems, Modelling and Analysis*, Nashville, USA, 16-19 March 1995.
- [64] I. Norris, J. Roberts, A. Simonian and J. Virtamo, 'The Superposition of Variable Bit Rate Sources in an ATM Multiplexer', *IEEE J. Selected Areas Commun.*, Vol.9, No. 3, April 1991.
- [65] I. Nikolaidis and R. O. Onvural, 'A Bibliography on Performance Issues in ATM Networks' *Computer Communication Review*, Vol. 22, No. 5, pp 8-23, Oct. 1992.
- [66] C. Parris, S. Keshav and D. Ferrari, 'A Framework for the Study of Pricing in Integrated Networks', *Technical Report TR-92-016, International Computer Science Institute*, Berkeley, CA, March 1992.
- [67] C. Partridge, *Gigabit Networking*, Addison-Wesley, 1994.
- [68] K. K. Ramakrishnan and P. Newman, 'Integration of Rate and Credit Schemes for ATM Flow Control', *IEEE Network*, Vol. 9, No. 2, March 1995.
- [69] G. Ramamurthy, R. S. Dighe and D Raychaudhuri, 'An UPC-Based Traffic Control Framework for ATM Networks', *Proc. of Globecom '94*, 1994.
- [70] J. W. Roberts, 'Variable-Bit-Rate Traffic Control in B-ISDN', *IEEE Commun. Mag.*, Vol. 29, No. 9, pp 50-56, Sept. 1991.
- [71] R. M. Rodriguez-Dagnino, M. R. K. Khansari and A. Leon-Garcia, 'Prediction of Bit Rate Sequences of Encoded Video Signals', *IEEE J. Selected Areas Commun.*, Vol. 9, No. 3, pp 305-313, April 1991.
- [72] O. Rose and M. R. Frater, 'A Comparison of Models for VBR Video Traffic Sources in B-ISDN', *IFIP Transactions C-24: Broadband Communications II*, pp 275-287, North-Holland, 1994.
- [73] O. Rose, 'Modeling of MPEG Video Traffic for an ATM Traffic Generator', *COST 242 TD (95)*, Stockholm, May 10-11, 1995.

References

- [74] J. M. Schneider, T. Preuss and P. S. Nielsen, 'Management of Virtual Private Networks for Integrated Broadband Communication', *Computer Communication Review*, Vol. 23, No. 4, pp 224–237, Oct. 1993.
- [75] M. Schwartz, *Telecommunication Networks: Protocols, Modelling and Analysis*, 1st ed., Addison–Wesley, 1987, pp.119–160.
- [76] S. Shenker, 'Service Models and Pricing Policies for an Integrated Services Internet', presented at *Public Access to the Internet*, JFK School of Government, Harvard University, May 26–27, 1993.
- [77] R. J. Simcoe and L. Roberts, 'The Great Debate Over ATM Congestion Control', *Data Communications*, September 21st, 1994.
- [78] J. Teahan, 'ATM Video Source Simulation', *B.E. Electronic Engineering Project Report*, DCU 1994.
- [79] L. Trajkovic and S. J. Golestani, 'Congestion Control for Multimedia Services', *IEEE Network*, Vol. 6, No. 5, pp 20–26, Sept. 1992.
- [80] J. S. Turner, 'Managing Bandwidth in ATM Networks with Bursty Traffic', *IEEE Network*, Vol. 6, No. 5, pp 50–58, Sept. 1992.
- [81] S. M. Walters, 'A New Direction for Broadband ISDN', *IEEE Commun. Mag.*, Vol. 29, No. 9, pp 39–42, Sept. 1991.
- [82] W. Whitt, 'Tail Probabilities with Statistical Multiplexing and Effective Bandwidths in Multi-Class Queues', *AT&T Bell Laboratories*, 1992.
- [83] T. Worster, 'Modelling Deterministic Queues: The Leaky Bucket as an Arrival Process', *International Teletraffic Congress, ITC-14*, Vol 1a, pp 581–590, Antibes, France, June 1994.
- [84] N. Yamanaka, Y. Sato and K. Sato, 'Performance Limitation of Leaky Bucket Algorithm for Usage Parameter Control and Bandwidth Allocation Methods', *IEICE Trans. Commun.*, Vol E75–B, No. 2, February 1992.

References

- [85] S. Yazid and H. T. Mouftah, 'Congestion Control Methods for BISDN', *IEEE Commun. Mag.*, Vol. 30, No. 7, pp 42-47, July 1992.
- [86] N. Yin and M. G. Hluchyj, 'On Closed-Loop Rate Control for ATM Cell Relay Networks', *Proc. of IEEE INFOCOM '94*.
- [87] 'Redundancy Reduction Techniques for Coding of Video Signals in Multi-Media Services', COST 211ter, Compendium ATM, Coding Procedures, Simulation Subgroup, pp 1-83, May 1993.
- [88] SES, *SES/workbench Reference Manual*, Release 2.1, Feb. 1992.
- [89] US Department of Commerce, *Telecommunications in the Age of Information*, Oct. 1991.

Publications Arising From This Work

- [24] B. Erimli, J. Murphy and J. Murphy, 'On Worst Case Traffic in ATM Networks', *Twelfth UK IEE Teletraffic Symp.*, Windsor, UK, 15-17 March 1995.
- [30] T. Hanson and J. Murphy, 'Internetworking Asynchronous Transfer Mode with the Consultative Committee for Space Data Systems Advanced Orbiting Systems Protocols', *Third International Conference on Telecommunication Systems, Modelling and Analysis*, Nashville, USA, 16-19 March 1995.
- [52] J. K. MacKie-Mason, L. Murphy and J. Murphy, 'The Role of Feedback in Network Allocation', *MIT Workshop*, Cambridge, USA, 7 March 1995.
- [53] J. MacKie-Mason, J. Murphy and L. Murphy, 'Using Price Feedback for ATM Efficiency', *Third International Conference on Telecommunication Systems, Modelling and Analysis*, Nashville, USA, 16-19 March 1995.
- [55] J. Murphy, E. Chow and R. Markley, 'ATM Service-Based Selective Retransmission Over DSN Satellite Links', 15th AIAA International Communication Satellite Systems Conference, San Diego, USA, March 1994.
- [56] J. Murphy and L. Murphy, 'Bandwidth Allocation By Pricing In ATM Networks', *IFIP Transactions C-24 : Broadband Communications II*, pp 333-351, North-Holland, 1994.
- [57] J. Murphy, L. Murphy and E. C. Posner, 'Distributed Pricing For Embedded ATM Networks', *International Teletraffic Congress ITC-14*, Antibes, France, June 1994.
- [58] J. Murphy, E. Chow and R. Markley, 'Systems Design of a Satellite Link

Publications Arising From This Work

Protocol', to appear in the *American Institute of Aeronautics and Astronautics Journal of Spacecrafts and Rockets, AIAA JSR*, 1995.

[59] J. Murphy and J. Teahan, 'Video Source Models for ATM Networks', *Eleventh UK IEE Teletraffic Symp.*, Cambridge, UK, 23-25 March 1994.

[60] J. Murphy and T. Curran, 'Optimising Embedded ATM Networks', *Tenth UK IEE Teletraffic Symp.*, Ipswich, UK, April, 1993.

[61] J. Murphy, 'Modelling Of High Speed Data Network Protocols', *First SES/-european Users Group Conference*, Leatherhead, UK, 16-17th February, 1994.

[63] L. Murphy and J. Murphy, 'Pricing for ATM Network Efficiency', *Third International Conference on Telecommunication Systems, Modelling and Analysis*, Nashville, USA, 16-19 March 1995.