

Number 217



**UNIVERSITY OF
CAMBRIDGE**

Computer Laboratory

Dynamic bandwidth management

Bhaskar Ramanathan Harita

15 JJ Thomson Avenue
Cambridge CB3 0FD
United Kingdom
phone +44 1223 763500
<https://www.cl.cam.ac.uk/>

© Bhaskar Ramanathan Harita

This technical report is based on a dissertation submitted October 1990 by the author for the degree of Doctor of Philosophy to the University of Cambridge, Wolfson College.

Technical reports published by the University of Cambridge Computer Laboratory are freely available via the Internet:

<https://www.cl.cam.ac.uk/techreports/>

ISSN 1476-2986

Summary

Recent advances in semiconductor and optical technologies have contributed greatly to the evolution of broadband integration of multi-service traffic. The asynchronous transfer mode (ATM) has been proposed as the target technique for broadband integrated services digital networks (BISDN's) based on fast packet switching and optical fibre transmission. A primary advantage of ATM is that variable bit rate services can be supported efficiently, which meets the basic needs of flexibility and service independence required of integrated services networks. In order to fully exploit this flexibility and enhance network efficiency by statistical multiplexing, it is important that there be effective methods of bandwidth management and congestion control.

This dissertation describes the use of dynamic bandwidth management to support an ATM overlay superimposed on a public, primary rate ISDN. The overlay architecture provides for the flexible aggregation of switched circuits into larger bandwidth channels. The channels are formatted into a common packet encoding and packets from different sources are statistically multiplexed onto them. In this work, different control schemes that dynamically vary the bandwidth of the channels in a transparent fashion, using out-of-band signalling, are contrasted. The bandwidth is adjusted by adding or deleting circuits in reaction to the traffic rates and the queue sizes at the channels. Performance models of simple bandwidth control schemes as queueing systems are analysed by the use of moment generating functions.

Packet transfer on the overlay is virtual circuit based and connection requests are accepted on the basis of their bandwidth requirements. Dynamic bandwidth management is used to supplement static bandwidth allocations in a congestion control framework presented for the overlay. The cost effectiveness of dynamic bandwidth control is examined for the tariff structure implemented in the underlying public ISDN.

The contributions of this dissertation are the development of schemes for dynamic bandwidth management, their implementation on an ATM testbed and the analysis of performance models for bandwidth control validated by simulation and experiment.

Preface

I would like to thank my supervisor, Dr. Ian Leslie for his advice, help and support during the course of my research. I am grateful to Prof. Roger Needham for his support and encouragement during my stay at the Computer Laboratory.

I have benefited from useful discussions and assistance from many members of the Computer Laboratory: thanks are due to David Tennenhouse, Cosmos Nicolaou, Roger Calnan, Kamal Chaudhary, Derek McAuley, David Greaves, Glenford Mapp, Mathew Doar, Chu Ang, Dr. Arthur Norman and Prof. Jerry Saltzer. John Burren and Chris Cooper at Rutherford Laboratories have always been patient and ready to help and my special thanks to them.

Several people have read draft versions of this dissertation, my thanks and appreciation to Ian Leslie, Derek McAuley, David Tennenhouse, David Greaves, Glenford Mapp, Roger Calnan, Peter Newman, Andy Harter and Liz Cruwys. Kamal Chaudhary helped in the typesetting of this manuscript while Dr. Mike Franklin provided me with access to the excellent computing facilities in Wolfson College.

During my research I have been supported by the Nehru Trust for Cambridge University, the Computer Laboratory and the Committee of Vice Chancellors for which I am grateful. I would also like to thank Wolfson College for the award of a Ribbands Research Studentship and the Senior Tutors, Dr. Rudolf Hanka and Dr. Derek Nicholls for all the help they have extended to me. I am grateful to Dr. Anil Seal of the Cambridge Commonwealth Trust for his help.

Finally thanks are due to all my friends who have been so supportive and made my stay in Cambridge such a memorable one.

Except where otherwise stated in the text, this dissertation is the result of my own work and is not the outcome of work done in collaboration.

I hereby declare that this dissertation is not substantially the same as any I have submitted for a degree or diploma or any other qualification at any other university.

I further state that no part of my dissertation has already been, or is being currently submitted for any such degree, diploma or other qualification.

Contents

List of Figures	viii
List of Tables	ix
Glossary of Terms	x
1 Introduction	1
1.1 Objectives	2
1.2 Outline	2
2 Transfer Modes and Trends in High Speed Networking	4
2.1 Multiplexing Techniques and Transfer Modes	4
2.2 Switching Techniques	6
2.3 Asynchronous Transfer Mode	9
2.4 Resource Allocation and Congestion Control for Broadband Integration . .	9
2.5 ISDN	10
2.6 Developments in High Speed Networking	11
2.7 Summary	12
3 Network Implementations	13
3.1 The Cambridge Fast Ring	13
3.2 The Universe Project	14
3.3 The Unison Project	15

3.4	Summary	25
4	Bandwidth Management and Congestion Control	26
4.1	Introduction	26
4.2	Admission Control and Bandwidth Reservation Systems	27
4.3	Dynamic Bandwidth Management	32
4.4	Algorithms for Dynamic Bandwidth Management	33
4.5	Bandwidth Monitoring and Bandwidth Enforcement	36
4.6	Related Work in Dynamic Channel Management	38
4.7	Summary	39
5	Previous Work	41
5.1	Congestion Control in ATM Networks	41
5.2	Performance Modelling of Integrated Services Networks	45
5.2.1	Performance Modelling of Hybrid Voice-Data Integration Schemes	45
5.2.2	Performance Modelling of Integrated Packet Networks	47
5.2.3	Performance Modelling of ATM Networks	48
5.3	Summary	51
6	Performance Models 1	52
6.1	Introduction	52
6.2	Modelling Assumptions	54
6.3	Solution Techniques	54
6.4	Model A : Hysteresis Control, Poisson Arrivals and Delayed Bandwidth Change (In)	55
6.5	Model B : Hysteresis Control, Poisson Arrivals and Delayed Bandwidth Change (In and Out)	63
6.6	Summary	66
7	Performance Models 2	67

7.1	Introduction	67
7.2	The MMPP Arrival Process	68
7.3	Assumptions	68
7.4	Model C : Hysteresis Control, MMPP Arrivals and Delayed Bandwidth Change (In and Out)	69
7.5	Model D : Hysteresis Control, Poisson Arrivals, Delayed Bandwidth Change (In and Out) and Sampling	73
7.6	Model E : Hysteresis Control, MMPP Arrivals, Delayed Bandwidth Change (In and Out) and Sampling	77
7.7	Analytic Properties of Moment Generating Functions — a Discussion	80
7.8	Numerical Computation and Iterative Methods of Solution	81
7.9	Graphical Analysis of Queue-Based Hysteresis Control	83
7.10	Simulation Models	85
7.11	Related Work on Queues with Fluctuating Parameters	86
7.12	Summary	88
8	Experimental Programme 1	90
8.1	Introduction	90
8.2	Experimental Configuration	90
8.3	Choice of Parameters in the Experimental Programme	92
8.4	Performance Analysis of Queue-Based Bandwidth Algorithms	94
8.5	A Cost Function for Dynamic Bandwidth Management	103
8.6	A Distributed Computing Application — Microemac's Make	104
8.7	Summary	108
9	Experimental Programme 2	109
9.1	Introduction	109
9.2	Bandwidth Switching — Queue and Rate-Based Algorithms	109
9.3	Cost and Traffic Performance of Bandwidth Management Algorithms	116

9.4	Rate and Queue-Based Algorithms — a Summary	122
9.5	Applications of Dynamic Bandwidth Management	123
9.6	Summary	130
10	Conclusions and Further Work	131
10.1	Conclusions	131
10.2	Suggestions for Further Work	133
A	Analysis 1	135
B	Analysis 2	137
B.1	Model C	137
B.2	Model D	139
B.3	Model E	141
C	Additional Experiments	145
C.1	Multiplexing a Bursty Traffic Stream onto a U Channel	145
C.2	Bandwidth Switching Times	149
	References	152

List of Figures

3.1	CFR Packet Format	14
3.2	The Unison Network Architecture	17
3.3	Protocol Stack : Unison and OSI	24
4.1	Dynamic Bandwidth Management and Bandwidth Enforcement in the Congestion Control Framework	28
6.1	Model A : Queue-Based Hysteresis Control — Poisson Arrivals and Delayed Bandwidth Change (In)	56
6.2	Model B : Queue-Based Hysteresis Control — Poisson Arrivals and Delayed Bandwidth Change (In and Out)	64
7.1	Model C : Queue-Based Hysteresis Control — MMPP Arrivals and Delayed Bandwidth Change (In and Out)	70
7.2	Model D : Queue-Based Hysteresis Control — Poisson Arrivals, Delayed Bandwidth Change (In and Out) and Sampling	74
7.3	Model E : Queue-Based Hysteresis Control — MMPP Arrivals, Delayed Bandwidth Change (In and Out) and Sampling	78
7.4	Instantaneous Queue Length vs. Time. Graphical Analysis of Queue-Based Hysteresis Control — Point Thresholding, Poisson Arrivals and Delayed Bandwidth Change (In)	84
8.1	The Experimental Configuration	91
8.2	Average Queue Length vs. Average Arrival Rate for Poisson Arrivals (QI Algorithm) — Comparison of Theoretical Analysis (Model D), Simulation and Experiment. Parameter Set 1	93

8.3	Average Queue Length vs. Average Arrival Rate for Poisson Arrivals (QI Algorithm) — Comparison of Theoretical Analysis (Model D), Simulation and Experiment. Parameter Set 2	95
8.4	Average Queue Length vs. Traffic Intensity for Poisson Arrivals (QI Algorithm) — Point Thresholding vs. Hysteresis Control	97
8.5	Average Queue Length vs. Sampling Interval for Poisson Arrivals (QI Algorithm) — Point Thresholding vs. Hysteresis Control	97
8.6	Average Queue Length vs. Mean Arrival Rate for MMPP Arrivals (QI Algorithm) — Comparison of Theoretical Analysis (Model E), Simulation and Experiment. Parameter Set 3	99
8.7	Average Queue Length vs. Mean Arrival Rate for MMPP Arrivals (QI Algorithm) — Comparison of Theoretical Analysis (Model E), Simulation and Experiment. Parameter Set 4	100
8.8	Average Queueing Delay vs. Mean Arrival Rate for MMPP arrivals (QI Algorithm) — Experiment. Parameter Sets 3 and 4	101
8.9	Average Queue Length vs. Upper Threshold for MMPP arrivals (QI Algorithm) — Experiment. Parameter Set 5	101
8.10	Survivor Functions for Poisson Arrivals from Theoretical Analysis (QI algorithm — Model D). Parameter Set 1	102
8.11	Average Queue Length vs. Queue Capacity for Poisson Arrivals from Theoretical Analysis (QI algorithm — Model D). Parameter Set 1	102
8.12	Cost and Performance Comparison of Rate and Queue-Based Algorithms — Microemac's 'Make' Test.	107
9.1	Tracking a Large Two-Phase Bursty Load with a RAT Algorithm . Traffic Load 1	110
9.2	Tracking a Two-Phase Bursty Load with QAT and RAT Algorithms . Traffic Load 2	112
9.3	Tracking a Three-Phase Bursty Load with QAT and RAT Algorithms . Traffic Load 3	114
9.4	Switching In Bandwidth — RI vs. QI Algorithms	115
9.5	Bandwidth and Queue Length Variations of Rate-Based Algorithms — Traffic Load 4	117

9.6	Bandwidth and Queue Length Variations of Queue-Based Algorithms — Traffic Load 5	121
9.7	Packet Voice Stream — a Run with 4 Voice Sources	124
9.8	Packet Voice Stream — a Run with 10 Voice Sources	124
9.9	Run of a Voice Multiplexer Operating with Dynamic Bandwidth Manage- ment	126
9.10	Dynamic Bandwidth Sharing between Two U Channels	128
A.1	Model A : Queue-Based Hysteresis Control — Poisson Arrivals and De- layed Bandwidth Change (In). Equating the Steady State Flows.	135
C.1	Packet Loss Percentage vs. Logarithm of Continuance of the Active Period. Queue of Limited Capacity. Parameter Set 6	146
C.2	Packet Loss Rate vs. Logarithm of Continuance of the Active Period. Queue of Limited Capacity. Parameter Set 6	146
C.3	Packet Loss Percentage vs. Ratio of the Peak Rate to Assigned Channel Bandwidth. Queue of Limited Capacity. Parameter Set 7	148
C.4	Bandwidth Switching Times (In and Out) vs. Number of B Channels — RAL	150
C.5	Bandwidth Switching In Times vs. Number of B Channels to RAL and Cambridge — a Comparison	150
C.6	Ramping Effect to RAL	151

List of Tables

8.1	Cost and Performance Comparison of Rate and Queue-Based Algorithms — Microemacs 'Make' Test	105
9.1	Cost and Steady State Performance of Rate-Based Algorithms — Traffic Load 4	118
9.2	Cost and Steady State Performance of Queue-Based Algorithms — Traffic Load 4.	118
9.3	Cost and Steady State Performance of Rate-Based Algorithms — Traffic Load 5	120
9.4	Cost and Steady State Performance of Queue-Based Algorithms — Traffic Load 5	122

Glossary of Terms

The number of the page on which the term is introduced appears in parentheses after each term.

ISDN	Integrated Services Digital Network (1).
BISDN	Broadband Integrated Services Digital Network (1).
ATM	Asynchronous Transfer Mode (1).
CFR	Cambridge Fast Ring (3).
FDM	Frequency Division Multiplexing (4).
WDM	Wavelength Division Multiplexing (4).
TDM	Time Division Multiplexing (4).
STDM	Synchronous Time Division Multiplexing (5).
STM	Synchronous Transfer Mode (5).
MAC	Media Access Control (5).
PTM	Packet Transfer Mode (5).
ATDM	Asynchronous Time Division Multiplexing (6).
FPS	Fast Packet Switching (7).
FCS	Fast Circuit Switching (7).
MRCS	Multi-Rate Circuit Switching (8).
SMDS	Switched Multi-Megabit Data Service (11).
MAN	Metropolitan Area Network (11).
DQDB	Distributed Queue Dual Bus (11).
SONET	Synchronous Optical Network (11).

UDL	Unison Data Link (14).
LVC	Lightweight Virtual Circuit (15).
RAL	Rutherford Appleton Laboratories (15).
AHSN	Alvey High Speed Network (16).
TSSI	Time Slot Sequence Integrity (18).
RPC	Remote Procedure Call (21).
IP	ARPA Internet Protocol (24).
CBO	Continuous Bit Stream Oriented (28).
BD	Bursty Data (28).
CTMC	Continuous Time Markov Chain (46).
MMPP	Markov Modulated Poisson Process (47).
MGF	Moment Generating Function (54).
NFS	Network Filing Service (105).
CLR	Cell Loss Rate (145).

Chapter 1

Introduction

Over the last decade there have been increasing efforts worldwide in the development of integrated services networks. Unlike traditional networks, such as public switched telephone networks for voice telephony and packet switched networks for data communication which were engineered for specific applications, integrated services networks are designed to support more than one class of traffic. The first implementation of an integrated services network in the public domain, was the Integrated Services Digital Network (ISDN). ISDN has both circuit and packet switching capabilities that are separated within the network, but provides integrated access to the user.

The growing need to interconnect local area networks and the emergence of new services have stimulated the need to develop a Broadband Integrated Services Digital Network (BISDN) that would offer increased flexibility and service independence. Advances in semiconductor and optical technologies have made broadband architectures based on fast packet switching and optical fibre transmission, an increasingly attractive and technologically viable alternative. The Asynchronous Transfer Mode (ATM) has been proposed by the CCITT¹ and other standard bodies as the target transfer technique for BISDN's which would offer flexibility for existing and future services. The main advantage of the ATM scheme is the flexibility in bandwidth allocation, which allows it to carry a wide range of services on a single network with common switching and transmission facilities.

For the early deployment of ATM services in advance of BISDN, an obvious solution is to run an ATM overlay over a synchronous circuit switched network. An ATM overlay run over an ISDN was the approach adopted in the network architecture that is the basis for the experimental work in this dissertation.

¹Comité Consultatif International Télégraphique et Téléphonique.

1.1 Objectives

The main advantage of ATM access is that variable rate services can be supported efficiently. In order to preserve this advantage when running an ATM overlay over a circuit switched network, such as primary rate ISDN, flexible and dynamic control of the underlying circuit bandwidth is necessary. The first issue addressed in this dissertation is the use of dynamic bandwidth management for the control of variable rate circuits to support ATM access. Dynamic bandwidth management is based on algorithms which control the flexible aggregation of ISDN circuits into larger bandwidth channels. The algorithms react to the traffic state on a channel by adding or deleting circuits from it to maintain constraints on traffic performance and to provide cost-effective site interconnection. Cost reduction is an important objective when using dynamic bandwidth management and is based on a tariff structure, similar to voice telephony, with a usage charge made on a per call basis.

Variable bandwidth channels that will be shared by traffic from many sources were adopted in the architecture used in this work to provide access to larger bandwidths and the benefits of statistical multiplexing. This approach can lead to network congestion and the second issue addressed in this dissertation is its control, based on dynamic bandwidth management. The dynamic variation of bandwidth represents a new approach to controlling congestion although it cannot prevent it from occurring; for that, it is necessary to control the traffic rate of individual sources. The second objective of this dissertation is to propose schemes for bandwidth reservation and bandwidth enforcement to be used with dynamic bandwidth management in a congestion control framework for the ATM overlay.

The performance modelling and analysis of dynamic bandwidth management schemes is important to derive insights into the control of a variable bandwidth channel and this is the third objective addressed in this dissertation. It was intended that models would be developed in an incremental fashion by the successive inclusion of factors that characterise the physical system, such as bandwidth switching latencies and sampling and the use of different arrival processes that emulate a bursty traffic source or stream.

The final objective of this dissertation is the implementation of dynamic bandwidth management algorithms on a channel controller working on the ATM testbed. The comparison of traffic performance and cost efficiency of different bandwidth algorithms, the validation of experimental results against those predicted by theoretical and simulation models, and the use of dynamic bandwidth management in applications involving bandwidth sharing are seen as useful investigations in the experimental programme.

1.2 Outline

Developments in multiplexing and switching and their related transfer modes are reviewed in the next chapter. Information transfer based on ATM is introduced and some proposals

for near-term solutions in the evolution towards broadband integration are presented.

The third chapter describes three network architectures — the Cambridge Fast Ring (CFR), the Universe network and the Unison network which is the network on which the current work is based. While site interconnection strategies in the Universe project have greatly influenced the design of the Unison network, the CFR forms the central component of the Unison architecture in supporting the ATM overlay.

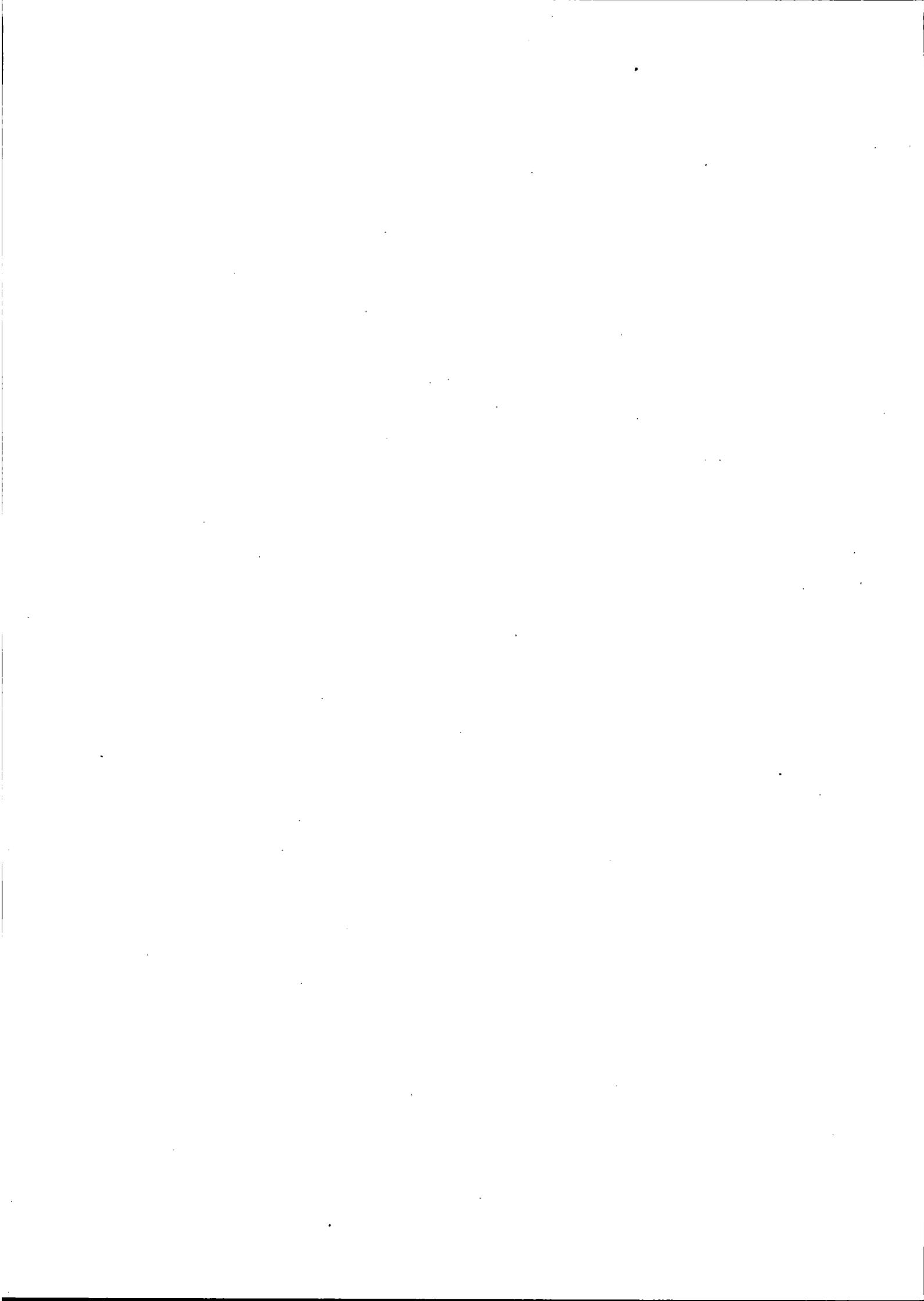
The fourth chapter presents a congestion control framework for the ATM overlay. The first part of this chapter describes bandwidth reservation schemes for the framework. The second part presents algorithms for dynamic bandwidth management. The chapter concludes with a proposal for bandwidth monitoring and bandwidth enforcement on the overlay.

A review of previous work in the performance modelling and analysis of integrated networks is described in the fifth chapter. The chapter also presents recent research on the congestion control of ATM networks.

The sixth and seventh chapters contain a set of performance models in increasing order of complexity for a variable bandwidth channel operating under a simple dynamic bandwidth management algorithm. For the models in the sixth chapter, methods of deriving the stability conditions and steady state performance metrics such as the average queueing delay are presented. The seventh chapter includes a section on the simulation models developed in the current work. Numerical computation problems which were encountered when solving for the steady state performance metrics are also discussed.

The next two chapters — the eighth and the ninth, present results from an experimental programme conducted on the ATM testbed provided by the Unison network. The eighth chapter contains the validation of results predicted by simulation and theoretical models against those obtained from experiment. In the ninth chapter, experiments comparing the performance and cost efficiency of dynamic bandwidth management schemes, and their applications, are discussed.

The final chapter contains conclusions and recommendations for future work. Further details of the theoretical analyses presented in Chapters 6 and 7 are contained in the Appendices A and B. Finally, Appendix C describes some additional experiments on the effect of multiplexing a bursty traffic stream onto a queue of limited capacity and measurements of bandwidth switching times on the testbed.



Chapter 2

Transfer Modes and Trends in High Speed Networking

Developments in multiplexing and switching techniques have contributed greatly to advances in the transfer of multi-service traffic and this chapter reviews a variety of such techniques and their associated transfer modes. The position of ATM in the spectrum of transfer modes is established. The present implementation of ISDN is described and proposals for interim solutions in the evolution towards broadband integration are briefly examined. An introduction to congestion control and resource allocation in ATM networks is followed by a summary.

2.1 Multiplexing Techniques and Transfer Modes

Multiplexing is the partitioning of a transmission channel into logical subchannels which are used to support communication between peer entities. There are many multiplexing methods of which three widely recognised kinds are Frequency Division Multiplexing (FDM), Wavelength Division Multiplexing (WDM) and Time Division Multiplexing (TDM).

In FDM the communication channel is divided into a number of different frequency bands each of which can be used to support a separate channel. For example, analog voice channels were multiplexed using FDM with each channel being transmitted on a different carrier frequency. WDM is an analogous technique currently being used for development in optical communications, in which the channels are carried on different optical wavelengths. In TDM the time domain of the transmission channel is divided into separate intervals called time slots. Two basic approaches to TDM can be distinguished by whether a

common time reference exists between the transmitting sources or not.

Synchronous Time Division Multiplexing

In the first approach, a synchronous frame structure is imposed on the channel to format it into periodically repeated sequences of time slots. Frames are of a fixed duration and individual time slots appear at fixed positions within each frame. For example, in Europe primary rate ISDN has transmission channels of 2.048 Mbps divided into 125 μ sec frames. Each frame in turn is composed of 32 time slots, each of width 1 octet, which results in 64 Kbps sub-channels or circuits. A source which is allocated a time slot fills the assigned time slot in every frame with an octet of data. The frame boundaries can be viewed as common time references between the transmitting sources.

The principle of such a time division multiplexing strategy when abstracted is commonly called Synchronous Time Division Multiplexing (STDM). The mode of operation is referred to as the Synchronous Transfer Mode (STM). STM is suited to synchronous or constant bit rate sources; the 64 Kbps sub-channel transmission rate commonly used matches the bit rate of digitised voice generation at 8 Khz. However, STDM suffers from the disadvantages of reduced flexibility and bandwidth efficiency when handling other synchronous services transmitting at a different bit rate, or asynchronous services which have variable bit rates.

Packet Multiplexing

In packet multiplexing, also called *statistical multiplexing*, there is no common time reference between the transmitting sources. Transmitting sources can use the entire bandwidth of the channel when it appears idle, and such transmission continues until the source relinquishes the channel. Protocols for Media Access Control (MAC) are necessary to resolve possible contention and to limit the delay experienced by other transmitting sources. The mode of operation based on packet multiplexing is called the Packet Transfer Mode (PTM).

In packet multiplexing, there are no invariant associations between time slots and transmitting sources as in STDM. This means that explicit addressing or labelling information must be provided with every packet (a group of bits) transmitted enabling the destination of the packet to be determined. This results in a reduction in transmission bandwidth.

The label information can refer to either a global or a relative address. When relative addresses are used, the mode of transfer is *virtual circuit* oriented and would involve the overhead of circuit set up to establish the mappings between relative addresses at intermediate switches. No prior circuit establishment is necessary when using global addresses since they contain sufficient information necessary for routing through the network. In this *datagram* mode of transfer however, the larger labels result in a reduced data payload. In datagram transfer, neither the loss of sequentiality nor the loss of packets can be wholly avoided. Virtual circuit transfer can ensure that sequentiality will be preserved though

packets could still be lost due to overflowing buffers. An example of a connection oriented packet transfer protocol is X.25 [CCI85], which is widely used in public packet switched data networks.

The main advantages of PTM are its flexibility in handling bursty sources with variable bit rates and its efficiency in bandwidth utilisation — a source only uses bandwidth when it has packets to send. These advantages are gained at the expense of packets facing variable delays and jitter because of the asynchronous access and variable sized packets. This is in contrast to the fixed delay transmissions with very low jitter which characterise STM networks.

Asynchronous Time Division Multiplexing

The two transfer modes presented so far, PTM and STM, differ in the multiplexing principle used (label vs. position), the transport unit employed (variable length packets vs. synchronous slots) and in their circuit set up times (not present for datagram packet transfer but significant for STM). A third transfer mode, which has been proposed by the Standards Bodies [CCI89] as the target transfer mode solution for implementing BISDN is ATM [Minzer89]. The main advantage of the ATM technique is its flexibility in bandwidth assignment, which makes it promising for the integration of multi-service traffic.

This mode of operation is based on Asynchronous Time Division Multiplexing (ATDM), which combines features of both packet multiplexing and STDM. The transmission channel is formatted into recurrent time slots of fixed size which are dynamically assigned to sources on demand on a frame by frame basis. In ATDM, the basic unit of transfer is a short packet of fixed length equal to the length of a slot, with each slot carrying labelling information.

The dynamic assignment of slots to sources necessitates the use of control schemes to resolve contention and constrain access delays. An example of a load balancing protocol is the source release protocol used in a slotted ring, like the CFR [Hopper88]. In this scheme, a source can transmit into a passing empty slot, but cannot reuse it on its return after a ring revolution.

ATDM has evolved as a compromise between STDM and conventional packet multiplexing techniques; by combining slotted transmission formats and label multiplexing it offers both high speed transfer capabilities and bandwidth flexibility.

2.2 Switching Techniques

The switching function supports the exchange of information between source and destination entities across a network. Switching nodes within the network provide the flexible concatenation of transmission channels attached to their input and output ports.

The two main methods of interconnecting the incoming and outgoing links are by space division switching and time division switching. In space division switching, a short physical path is allocated across the switch fabric. In time division switching a mapping is set up across the switch between incoming and outgoing time slots. The mapping is used to copy bits from an incoming time slot into a buffer and then into the corresponding slot on the outgoing link.

Circuit and Packet Switching

The various switching techniques can be studied in a spectrum ranging from circuit switching to conventional packet switching. Circuit switching concatenates channels in a fashion that offers guaranteed bandwidth and a fixed quality of service with small, fixed switching and transmission delays and negligible jitter. However, there is a call set-up delay, since mappings between incoming and outgoing slots have to be established at all the intermediate circuit switches. Once established, a circuit is available for the duration of the call. Circuit switching is suitable for STM operation with constant bit rate sources.

In conventional packet switching a switch copies packets from transmission links connected to its input ports to those connected to its output ports without any relation to frame and slot structures. In connectionless, store and forward packet networks the switching process is asynchronous and variable length packets can arrive at the ports at any time. The switch processor waits for a complete packet to be assembled at an input port and then, based on the labelling information, forwards the packets to an appropriate output port. In virtual cut through switching [Kermani79], forwarding begins as soon as the header information is received without waiting for the assembly of the rest of the packet, provided the output port is free.

The primary advantage of packet switching is its flexibility in supporting a continuum of bit rates up to the switching capacity. The bandwidth allocated to a particular source can vary with the rate at which the source generates packets as long as there is sufficient bandwidth available. Packet switches also perform *rate adaptation* between input and output channels, working at different speeds, through buffering.

Methods of Circuit Switching

Switching disciplines such as Fast Packet Switching (FPS) and Fast Circuit Switching (FCS) represent statistical techniques which attempt to combine aspects of both conventional circuit and packet switching in an effort to achieve enhanced flexibility.

FCS [Gruber81] addresses this problem by extending the concept of circuit switching to handle bursty traffic. Circuits are set up whenever bandwidth is needed and disconnected when there is no information to send, thus freeing resources during idle periods. Circuit set-up and tear-down takes place through a fast signalling mechanism. In Burst Switch-

ing [Amstutz83, Amstutz89, Hasleton83, O'Reilly86] which is a variant of FCS, connection set-up does not take place from scratch each time there is a burst of information. Instead, a virtual circuit is set up through the network when the initial call request is made by a source to a peripheral switch. When there is a burst of information from the source, the virtual circuit is activated and each switch along it assigns bandwidth on the outgoing virtual circuit to establish the real connection. Input buffering would have to be provided to queue the input burst when one of the switches has insufficient bandwidth.

With Multi-Rate Circuit Switching (MRCS) [Kulzer84], which was proposed as a means of improving the flexibility of conventional circuit switching, circuit set-up is not dynamic as in FCS. Instead the technique permits the allocation of multiple time slots to a single transmitting source. The problems with this method are the choice of a basic rate and the need to effect synchronisation between the constituent sub-channels of the logical multi-rate channel.

A more complex proposal for switching [Kulzer84] which includes elements of both MRCS and FCS involves the dynamic set-up and tear-down of multi-rate channels. In this proposal, FCS gives an alternative method of providing very low rate channels and eliminates the need for a small basic rate.

Fast Packet Switching

In FPS [Kulzer84, Turner86a, Oie90] the the transfer of information is characterised by smaller delays and jitter than conventional packet switching. Fast packet switches have reduced internal functionality with simplified protocols and routeing schemes which can be implemented in hardware. Advances in semiconductor technology have made possible the implementation of high capacity fast packet switches based on space division switching.

The basic components of a fast packet switch are the switching fabric, input and output ports which are connected to external transmission links and switch and port controllers. To build a high capacity switch, a multi-path design for the fabric is necessary. In many switches this is implemented by designing the fabric as a self routeing multi-stage interconnection network of simple switching elements.

Fast packet switches may be classified as blocking and non-blocking. In the former there can be internal blocking when more than one packet competes concurrently for the same internal link. In the latter there is no internal blocking, requiring internal buffering within the fabric.

Routeing within a fast packet switch is simple and efficient, being based on virtual circuit identifiers and routes which are set up through the switch at call set-up time. In this way the sequentiality of packets belonging to the same call is maintained. The labels in each packet are small, being sufficient only to distinguish the different virtual circuits. Routeing at each switch is reduced to a simple table look-up, which could have a fast implementation in hardware.

The Cambridge Fast Packet Switch [Newman88] is an example of a high capacity switching fabric based on a multi-stage interconnection network. This blocking switch has no internal buffering, but has input port buffering and has a simple implementation in current hardware technology. Such switches are to be interconnected to form an ATM network in the Fairisle project at the University of Cambridge Computer Laboratory. The CFR, being based on small, fixed slots and asynchronous access, has also been used as a fast packet switching fabric [Tennenhouse89a]. This is described in more detail in Chapter 3.

2.3 Asynchronous Transfer Mode

In its simplest form, ATM is a flexible multiplexing technique for multiplexing a number of virtual circuits onto a common transmission channel. The primary advantage of ATM is its flexibility in bandwidth allocation, through the assignment of small, fixed size packets to virtual circuits whenever required. This enables it to handle a wide variety of heterogeneous sources with variable bit rates. Through the statistical multiplexing of bursty sources in ATM, there are possible advantages to be gained in bandwidth efficiency at the expense of packet delay and packet loss.

The packet size used in an ATM network will typically be in the range of 32 to 120 octets, of which 2 to 8 constitute the labelling information that identifies a virtual circuit. Datagrams and other packet traffic would be fragmented into ATM packets before transfer across the network. Fragmentation and reassembly functions would be carried out at the end points of the network in an adaptation layer, which represents an additional layer of processing compared to conventional packet networks. To minimise the overhead due to processing by network protocols, switching within an ATM network is likely to be based on FPS.

In handling synchronous traffic, ATM networks with statistical multiplexing cannot completely match the constant delay, low-jitter performance of STM networks. However, the connection orientation and the call set-up phase allow for explicit resource allocation to be provided. The simplified switching protocols and the small fixed size packets imply that the absorption of jitter at the receiver buffers is difficult only when there are large queueing delays at contention points.

A study of design issues and principles for ATM transfer over broadband network architectures can be found in [Wernik88].

2.4 Resource Allocation and Congestion Control for Broadband Integration

The main advantages to be derived from broadband integration based on ATM have been defined by [Woodruff88] as *flexibility* in supporting any bandwidth required by a variety

of service types and mixes, *efficiency* gained by statistically multiplexing bursty traffic streams and *simplicity* by the use of a single transfer mode for all services.

In order to achieve these objectives, efficient mechanisms of resource allocation are necessary to strike a suitable performance balance and to ensure that the individual requirements of each traffic type are met. To derive the possible benefits of statistical multiplexing and to cope with periods of heavy traffic load, congestion control is necessary to supplement resource allocation. Resource allocation and congestion control mechanisms should be designed in a simple fashion to avoid over-specifying the control for particular service types, which would contradict the principle of service integration.

There have been several proposals for resource allocation and congestion control for the integration of multi-service traffic using ATM [Woodruff88, Gersht89, Sallberg90, Woodruff90]. A common theme that can be abstracted from these proposals is the emphasis on preventive control schemes consisting of admission control and flow enforcement. Admission control is responsible for accepting or rejecting incoming connection requests on the basis of their bandwidth and other *quality of service* requirements. Acceptance will be based on whether or not, the addition of the new connection compromises the service requirements of the existing connections. Flow enforcement or policing at the network periphery is necessary to ensure that sources correspond to their traffic descriptors specified at connection set-up.

The use of end-to-end congestion control mechanisms and the reduced internal functionality of an ATM network is in contrast to the heavyweight link-by-link error recovery and flow control mechanisms which characterise PTM. Resource allocation and congestion control mechanisms for ATM transfer are seen as working on a *best effort* basis. Consequently end-to-end recovery at the higher levels is necessary because packets may be lost or delayed.

2.5 ISDN

ISDN networks currently implemented in Europe and North America provide 64 Kbps (B channels) circuits between subscribers with a TDM frame format imposed on the carrier. The B channel is a full duplex 64 Kbps channel delivering octets of data to each end synchronously once every 125 μ secs. In Europe, subscribers are provided with *primary rate* access at 2.048 Mbps configured as 30 B (Bearer) channels and a signalling D (Data) channel, with the remaining bandwidth used for synchronisation. The *basic rate* interface consists of 2 B channels and a 16 Kbps D channel at an overall rate of 144 Kbps.¹

Each frame of a primary rate TDM link in Europe has 32 octets. Slot 0 is used to identify the start of a frame while slot 16 is reserved for the signalling D channel, which is packet

¹Other channel types that have been identified by CCITT ISDN standards [CCI88] are H0 – 6 B, H11 – 24 B and H12 – 30 B channels.

switched. A 64 Kbps signal is sent through the network as a series of octets inserted into the same slot of successive frames. Data in a particular slot passes from its source to destination through one or more circuit switches. Octets placed in the same slot in successive frames travel through the network in exactly the same fashion and arrive at the destination as octets in a given slot in successive frames. However, in general the slot position at the destination will not be the same as the slot position at the source.

The switching of B channels is based on out-of-band signalling using ISDN signalling protocols [CCI88] on the D channel. Apart from signalling information, the D channel may also be used for the transport of packetised data. A standard which has been issued for ISDN packet switching is X.31 [CCI84] which allows an X.25 overlay on ISDN, and permits the transfer of packetised data on both the D and B channels.

2.6 Developments in High Speed Networking

Existing and emerging demands for broadband services and the increasing installation of optical fibre links have stimulated the need for the deployment of high speed networks in advance of BISDN based on the ATM technique. In this section some of these plans for the deployment of provisional networks will be considered.

Frame-based proposals to meet high speed data transfer requirements are the Frame-Relay service and the Switched Multi-Megabit Data Service (SMDS). Frame relaying [Chen89] is an ISDN packet mode bearer service for data transfer that is based on the LAPD protocol. LAPD is used over the D signalling channel which handles packetised data; frame relaying extends its use over the bearer channels.

SMDS [Hemrick88] is a public, connectionless packet switched data service that was proposed by Bell Communications Research for high speed data transfer. One of its important applications is extending the functionality of LAN's into the wide area. Information transfer is by datagrams that will be independently switched with no necessity of virtual circuit establishment. The size of the datagrams will be large enough to accommodate packet sizes from most LAN's.

The initial support of SMDS is planned over Metropolitan Area Networks (MANs). The advances in and maturity of MAN technology make it the favoured technology for being the first stage in supporting public broadband services. MANs provide for the extension of high speed connectionless data transfer services beyond the local area. An example of a MAN implementation is the Cambridge Backbone Ring [Greaves90] based on slotted ring technology, a prototype version of which is currently operational at 500 Mhz.

A MAN standard that is soon to be adopted is the IEEE 802.6 standard for Distributed Queue Dual Bus (DQDB) [IEE88]. In DQDB, which is a hybrid network providing integrated circuit and packet switching, the underlying bus bandwidth is formatted into fixed length frames, which in turn consist of a number of slots. At any time, some of these

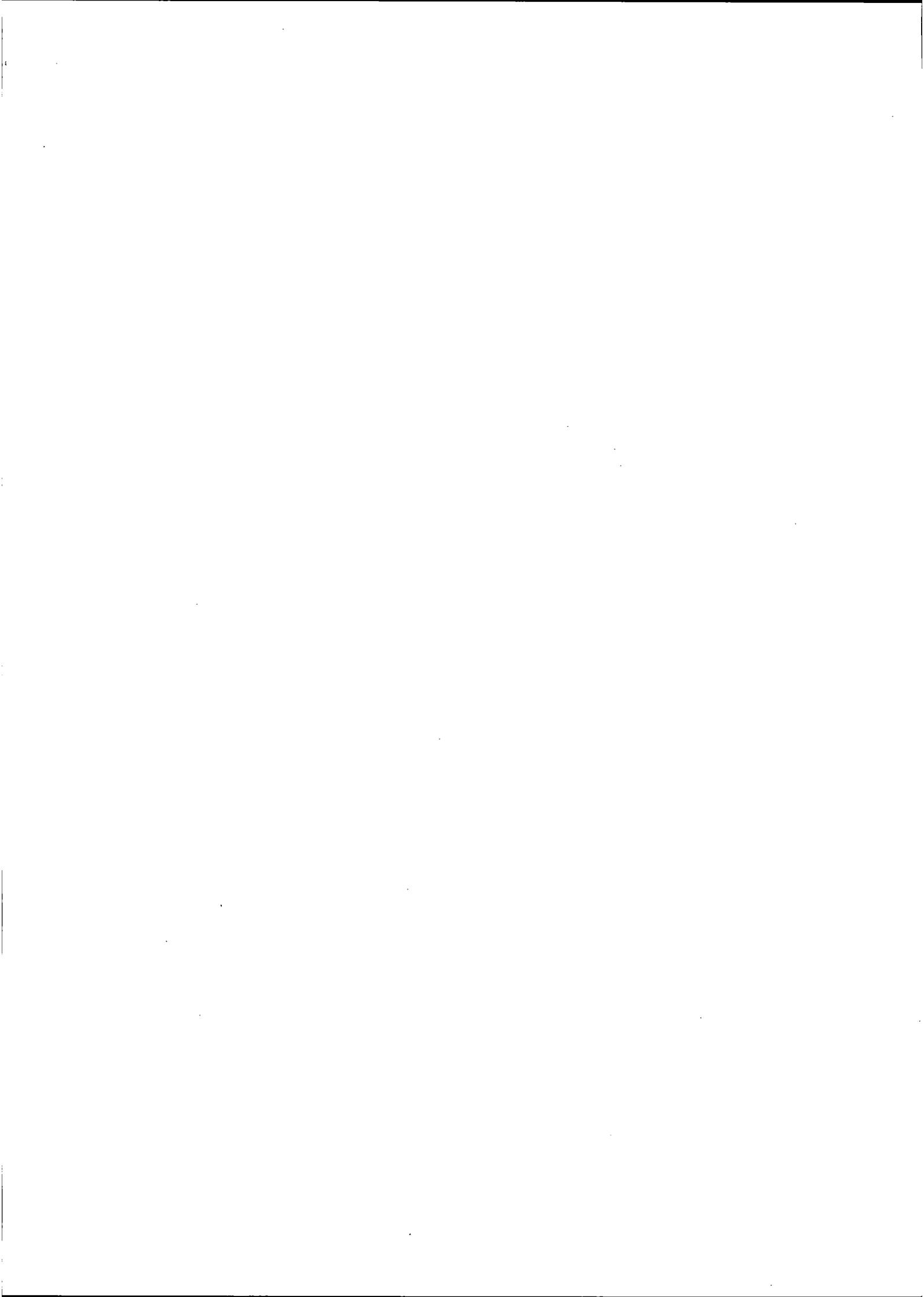
slots (termed *isochronous*) are allocated to synchronous circuits, while the others (termed *asynchronous*) can be used for packetised data. Media access for data is asynchronous and is controlled by a distributed queueing algorithm.

The use of synchronous transmission carriers like Synchronous Optical Network (SONET) [Ballart89], will facilitate the introduction of high speed data transfer before BISDN. SONET is a standard that defines a hierarchy of interfaces for optical signal transmission. In the longer term synchronous carriers are seen as supporting the introduction of broadband services based on ATM as an overlay with ATM cells being transmitted in synchronous payloads. Ultimately ATM switches will be introduced into the network to provide a 'true' BISDN network architecture.

The work in this dissertation is based on an ATM overlay that runs over a synchronous circuit switched carrier, as exemplified by primary rate ISDN. A site interconnection architecture [Tennenhouse89a] which supports the overlay is described in detail in the next chapter.

2.7 Summary

Two established modes for information transfer in digital networks are PTM and STM. ATM which is seen as a single transfer mode suitable for both current and future services was described as a compromise between them. ATM is characterised by its simplicity, flexibility in allocating bandwidth and potential for greater efficiency when statistically multiplexing bursty sources. Congestion control is necessary in ATM networks to maintain the ATM advantages and to ensure that the individual requirements of each user are met. Short-term solutions based on MAN technology and overlay networks to meet current demands for high speed data transfer in advance of BISDN were briefly described.



Chapter 3

Network Implementations

This chapter reviews three network implementations; a local area network used as a packet switch — the CFR; a wide area network based on satellite links — the Universe network; and a terrestrial ISDN interconnecting LAN's — the Unison network. These network architectures are of relevance to the work in this dissertation and exemplify some of the transfer concepts presented in the previous chapter.

3.1 The Cambridge Fast Ring

The CFR [Hopper88] is a slotted ring local area network operating at 40 Mbps. It was developed from the Cambridge Ring [Wilkes79], a slower network running at 10 Mbps. Each slot can contain a CFR packet which is 38 bytes long consisting of 4 bytes of source and destination addressing, 2 bytes of CRC and control information and a 32 byte data payload arranged as shown in Figure 3.1.

The load balancing scheme used on the CFR is *source release* and works as follows. A transmitting station fills an empty slot and sets the full/empty flag. A receiver copies the contents of the slot into its buffer and sets the response bits. When the slot returns the transmitter passes it on after having reset the full/empty flag. If the transmission was unsuccessful, either due to the destination station not being able to receive the packet or if there was a CRC error, retransmission of the packet would continue up to a retry limit.

Source release and the passing on of a slot ensures fairness in bandwidth sharing between transmitters — in a single slotted ring with N transmitting stations all active together, every transmitting station is guaranteed an empty slot at least once every $(N+1)$ slot times. For a ring with S slots the maximum duration between transmissions, assuming

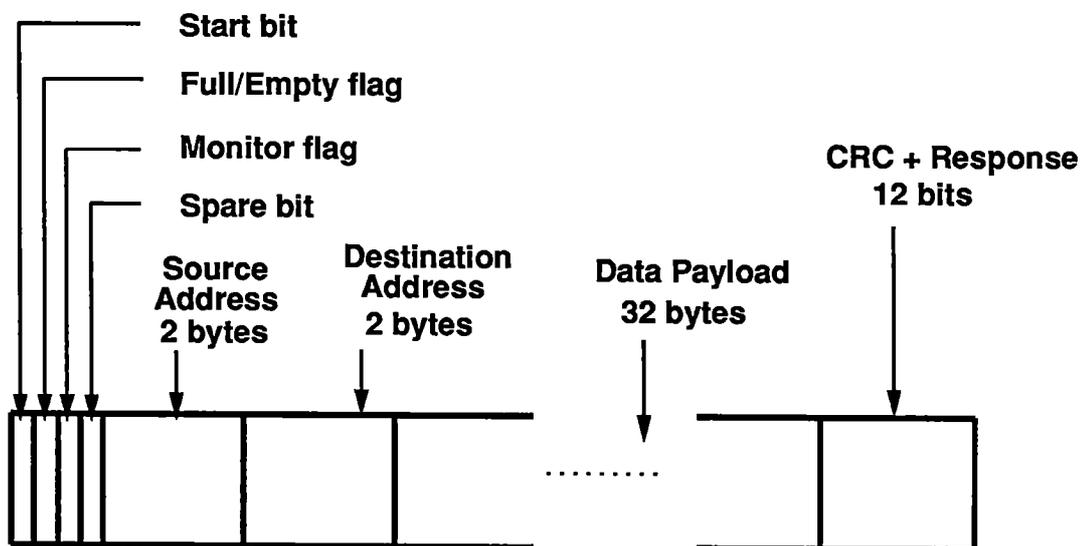


Figure 3.1: CFR Packet Format

all N stations are in transmit mode, is $(S+N)$ slot times. The minimum period between transmissions assuming a single active station is $(S+1)$, so the maximum jitter observable in slot access time when there are S slots is $(N-1)$ slot times. For sources with greater bandwidth requirements the original design of the CFR contained a proposal for *channel slots*, which transmitters could reuse instead of passing on free [Hopper86].

At the data-link level the main protocol used over the CFR is the Unison Data Link (UDL) service [Tennenhouse86] which supports the transfer of UDL blocks over lightweight virtual circuits.

The CFR has been used for the transfer of multi-service traffic; in the Pandora project video streams are transferred across it [Hopper90]. Other studies detailing the use of a slotted ring for integrated services have been described in the ISLAND project [Calnan89] and in the Orwell Ring project [Falconer85]. Performance modelling of slotted rings may be found in [Vukotic88, King82] and a less rigorous analysis in [Tennenhouse89a].

3.2 The Universe Project

The Universe Project [Burren89b, Leslie84] used a 1 Mbps satellite broadcast channel to interconnect homogeneous LAN's composed of Cambridge Rings. The aim of the project was to extend the use of LAN communication techniques into the wide area.

Connecting Cambridge Rings within a site were ring-ring bridges, while inter-site communication was effected across satellite bridges connecting earth stations to one of a site's rings. The Cambridge Ring supported the transfer of mini-packets (with a 2 byte data payload) between stations. The basic block protocol was built on the mini-packet exchange layer at the data-link level. This was the protocol over which almost all communication over the ring was done because of the small data payload in a mini-packet. The disadvantage of this protocol was that a receiver could not multiplex block receptions between transmitters so there was receiver contention and interference between parallel traffic streams.

The primary service offered by the Universe network architecture for communication between hosts was Lightweight Virtual Circuits (LVCs). LVCs preserved the sequentiality of basic blocks that were sent over them, but did not provide end-to-end flow control or error recovery. LVCs were set up between entities on different sites by the establishment of duplex connections across intermediate bridges. Virtual circuit set-up involved initial transactions with a name server to resolve service names into global addresses necessary for routing across bridges through the network. This combined use of bridges and name servers in circuit set-up was designed to provide end users with the same view of communication across the wide area network as they had across a single Cambridge Ring.

The presence of a satellite hop with long round trip delays and the need to achieve simplicity within the network (and not to overburden components like satellite bridges) were reasons why error recovery and flow control schemes on an end-to-end basis were preferred over hop-by-hop schemes.¹ Flow control on an end-to-end basis was also identified as most suited for the transfer of multi-service traffic.

Methods like call blocking at request time and limiting the bandwidth on virtual circuits at the network periphery were proposed as means of limiting the admission of traffic into the network. During information transfer flow control on virtual circuits was thought to be best handled by rate control between the end users, given the satellite hop with long round trip delays. However, these schemes were not implemented on the Universe network. In practice, congestion was approached by discarding packets on heavily used circuits at the bridges. Another method employed was for a bridge to block input from the Ring by the use of acknowledgements at the mini-packet level. This could be performed selectively if necessary.

3.3 The Unison Project

The Unison project [Clark86, Tennenhouse87, Tennenhouse89a, Tennenhouse89b, Griffiths89] is a collaborative venture under the Alvey programme between the University of Cambridge Computer Laboratory, Rutherford Appleton Laboratories (RAL), Loughborough University of Technology, Logica and Acorn Computers. The project, which was

¹There is error detection and correction in the hop across a Ring.

fuelled by experiences in the Universe project, was concerned with the interconnection of LAN's in offices and multi-media office experiments.

While interconnection could have been achieved with dedicated leased 2 Mbps links between sites, a more dynamic interconnection architecture using circuit switched ISDN was preferred [Tennenhouse87]. This choice was seen as a logical step in the migration towards public ISDN networks and access to a universal transmission medium with a large spectrum of end applications and services. A prototype, private ISDN called the Alvey High Speed Network (AHSN) with the ACE circuit switch was used until the middle of 1989. Since then, the British Telecom primary rate ISDN with Multiline-IDA and System X switching exchanges [Newman86] providing primary rate access at 2.048 Mbps has been employed.

The use of a public network influenced the decision to maintain site independence. This was achieved by using deferred binding between service names and addresses and through a flexible interconnection strategy. The latter was provided by the dynamic switching of circuits between sites whenever required.

Network Architecture

The Unison architecture is a hybrid one, with an ATM overlay run over the STM of a circuit switched carrier. At each site in the Unison network (Figure 3.2) is a CFR² called the *exchange*. Exchanges are attached to primary rate ISDN links by transputer based devices called *ramps*. Ramps support the flexible aggregation of ISDN B channels into $N \times 64$ Kbps *U channels*; the bandwidth of U channels can be varied dynamically by adding or deleting B channels. A common packet format is used over the channels and the exchanges at the data-link level. This is based on the CFR packet format which was illustrated in Figure 3.1. Packets are transmitted along virtual circuits called *associations*.

Bridging client distribution networks to the exchange are *portals*. The exchange switches packets between its local portals to support intra-site communication. Inter-site communication is effected by packets being extracted from the local exchange by a ramp and sent down a variable rate U channel to a peer site, where they are injected into the remote exchange. When the communicating entities are not directly connected to the exchange, packets may have to be multiplexed through portals to and from client distribution networks.³

All packets between the same pair of sites are transported on a single U channel to derive the advantages of statistical multiplexing and to provide flexibility in bandwidth allocation useful for the transfer of variable rate services. The bandwidth of the U channel can be varied dynamically depending on the state of its buffers and the traffic rates through it. The management of U channels and the dynamic variation of their bandwidth is performed

²Configured in byte-wide parallel mode.

³Additional processing will be necessary if the client distribution network is not a CFR.

out-of-band and in a transparent fashion.

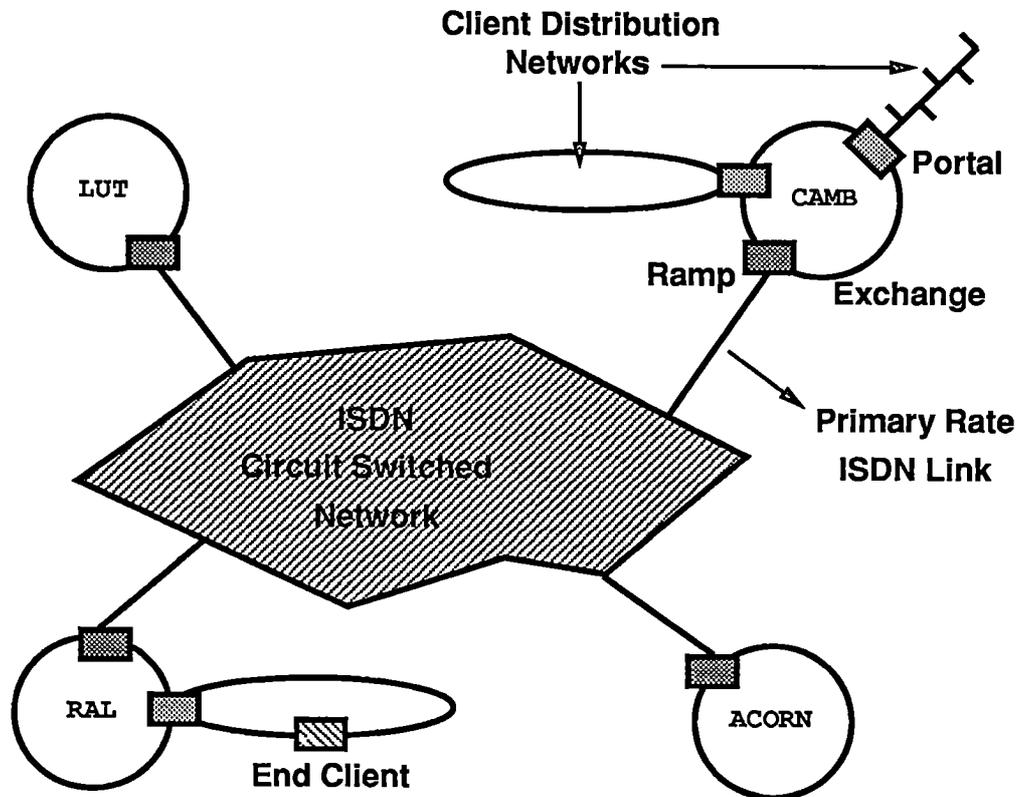


Figure 3.2: The Unison Network Architecture

Ramps and the Flexible Aggregation of Bandwidth

The ISDN ramp can support a number of switched variable bandwidth U channels to dynamically selected peer sites.⁴ The ramp runs algorithms that allow a number of B channels to be flexibly aggregated into a larger bandwidth U channel [Burren89a].

The main problem when aggregating bandwidth is that B channels through the ISDN between the same pair of ramps need not necessarily follow the same route. This causes *frame skew*, where slots that started within the same frame at the source ramp arrive at the destination ramp in different frames. Frame skew is network dependent and is related to the differences in propagation times and the number of switches traversed. Even when

⁴Up to a maximum of 30 different sites can be addressed directly at the same time. Additional sites can be accessed indirectly by *relaying* through an intermediate site.

time slots in a frame follow the same route through the network, there could be *time slot reordering* with a slot position within a frame changed between the source and destination. As a result of frame skew and time slot reordering octets can arrive out of sequence at the receiving ramp. This is called the Time Slot Sequence Integrity (TSSI) problem.

Sequentiality is restored by the use of *recombination* maps at the receiving ramp. To enable recombination, the transmit half of each ramp sends *marker* or synchronisation signals in a distinguished sequence (markers include the slot number on which they are transmitted) on slots belonging to a U channel in the same frame. The receiver reconstructs the correct sequence order by noting the frame and slot on which the markers arrive to form the recombination map. Even after the recombination map has been set up, it needs to be refreshed because uncontrolled frame slips can effect the ordering of time slots. At periodic intervals (once every 5 msecs), the transmit half of the ramp sends markers in-band on all slots belonging to a U channel.

Bandwidth Adjustment

The bandwidth of a U channel can be varied dynamically in a transparent or *non-invasive* fashion. B channels can be added or removed from a channel without disturbing the flow of in-band application traffic. This dynamic configuration of U channels is an out-of-band activity which is directed by an element of the network management called the *channel service*.

The first stage when increasing the bandwidth is *call signalling*, when a B channel is set up across the network between the two ramp halves. The ramps then transmit idle octets with markers at fixed intervals in the new slot to effect synchronisation. When ramps receive markers on the slot, they set a *slot synchronised* flag in their subsequent markers to acknowledge this to each other. On receiving the acknowledgement, each ramp is now ready to transmit on the slot, indicated by the setting of *slot ready* flags in their transmitted markers. This is followed immediately by the U channel being expanded to include the new slot to carry traffic.

The bandwidth of a channel is varied in this slot-by-slot fashion, although due to pipelining of signalling requests adjust requests for multiple slots are more efficient. For example, the time taken to add 2 B channels to a U channel with a single request is less than when making two successive requests for a single B channel each. Some measurements of bandwidth adjustment times are provided in Section C.2 of Appendix C.

A marker also contains a flag to delete slots. The channel service at either end of a connection can direct its local ramp to drop a slot. When asked to delete a B channel, the ramp sets the *slot delete* flag in the transmitted marker for the slot. The slot then ceases to be part of the U channel. The peer ramp receives the slot delete indication and sets the slot delete flag in its own transmitted marker for that slot. The B channel can then be closed through signalling.

In the current version of the ramp software, a logical token passing scheme has been implemented to resolve contention between ramps over the adjustment of bandwidth on their common channel. Contention is caused when both ends make simultaneous attempts to vary the bandwidth or if one end makes an adjust attempt when the other is in the process of altering the bandwidth. In the scheme, a ramp needs to acquire the token in order to initiate adjustment. Having acquired it (possibly by request from the other end), the ramp retains it at least until the bandwidth adjustment has been completed.⁵

Schemes for the aggregation of ISDN B channels into larger bandwidth channels have also been proposed by Boltz and Altarah [Boltz89, Altarah89]. Boltz suggests three possible schemes — the use of switched multi-slot channels where the time slots are chosen such that TSSI is preserved; the end-to-end synchronisation of independent B channels; and the use of higher level protocols to form logical channels between end applications. Both the second scheme which is favoured by Boltz and the scheme proposed by Altarah are based on the use of a synchronisation pattern to maintain TSSI, similar in principle to the method used at the ramps.

Ramp Queuing

The ramp consists of distinct transmit and receive halves, which work in a pipelined fashion to support a duplex channel. The ramp CFR receiver station extracts appropriately addressed packets from the exchange ring and routes them onto a transmit queue.

To cater to traffic services that are delay sensitive there is an *expedited transfer function* available at each set of transmit queues. Packets from these services are routed onto a high priority queue with access to the common carrier on a non pre-emptive priority basis over those routed onto a low priority queue. Expedition ensures constraints on jitter and the *insulation* of high priority traffic from traffic multiplexed onto the low priority queue. The high priority queue is of limited capacity and overload is handled by discarding packets. The low priority queues are elastic — they can store upto 20000 CFR packets across all active channels. Rate adaptation is performed at this point when the input traffic rate temporarily exceeds the channel bandwidth and the queue absorb asynchronous bursts of traffic.

At the receive half of a ramp, CFR packets are assembled from the incoming ISDN stream and multiplexed onto a common queue at the CFR station for transmission onto the exchange. Rate adaptation can take place here, with the transmission rate matching the input rate of a receiving station on the exchange. This could cause interference between parallel traffic streams with a slow receiver holding up other transmissions. However, with receive halves which implement expedited transfer, delay sensitive traffic can be insulated from such effects.

⁵At the start of day, the site which creates a U channel also creates a token for further use with that channel.

Packet transfer through the transmit half of a Unison ramp has been modelled by Tennenhouse [Tennenhouse89a] as three pipelined stages. The first is a CFR *interface* stage, where packets are received from the CFR and put on to the ramp queues. In the *transmit channel* stage, buffered packets are segmented and *quad frames*⁶ are assembled, while in the *frame processing* stage, quad frames are separated into individual ISDN frames which are transmitted onto the ISDN.

The transmit channel stage was in turn modelled by three components. The first is an alignment gate, which delays the asynchronously arriving packets by upto 500 μ secs. The second component is the frame assembly stage which takes 500 μ secs to assemble a quad frame. Once assembled, the quad frames are transmitted onto the ISDN, except at 5 msec intervals, when the synchronisation sequence is transmitted instead.

Packet Encoding at the Ramps

U channels provided between ramp pairs ensure an error protected packet stream over the common carrier. The transmit half of a ramp segments a CFR packet into bytes and places these into successive time slots of a U channel in carrier frames. Decoding and reassembly functions are performed at the receive half — incoming frames with reordered and time skewed slots are stored in a frame buffer. The receive half uses its recombination map to cycle through the buffer and extract bytes to reassemble the packet. There is no error correction or retransmission at this level and corrupted packets are discarded.

Since the frame repetition rate is fixed at 8 Khz, the time taken to transmit a packet onto a U channel is inversely proportional to the number of slots in the U channel. If N is the size of a packet in octets⁷ and TS is the number of time slots in the U channel, then the transmission time is given by $\lceil N/TS \rceil \times 125\mu$ secs.⁸

⁶Quad frames consist of four ISDN frames (4 x 30 octets) and were used because the transputers in the ramp can complete most 32 bit (word) operations in the time required to complete the equivalent 8 bit (octet) operation [Tennenhouse89a].

⁷The packet transmitted between ramps is 40 bytes long with 32 bytes of data, 4 bytes of addressing information and 4 bytes of control information including CRC. Thus the effective transmission rate on a B channel is 200 packets/sec.

⁸A second scheme [Tennenhouse89a] is one in which octets of a packet are confined to be transmitted on a specific time slot of a U channel. In this case, the U channel consists of several time slots each independently carrying a concurrent byte stream corresponding to a different packet. Although sequentiality of the bytes within a packet is maintained in this scheme, the receive half must check for sequentiality between packets of a single stream. The transmission delay for packets is now independent of channel width. It is only dependent on the length of the packet and is given by $N \times 125\mu$ secs (assuming a synchronisation sequence is not transmitted in between). The advantage of this scheme is that the jitter on delay is constrained.

Address Management

Each Unison site represents an addressing domain. Inter-site communication is effected by address translation at the boundary of the domains. Ramps perform address translation using bindings that are set up during association establishment. The 4 bytes of addressing information in the header of a CFR packet (Figure 3.1) are divided equally between source and destination address fields. These 2 byte address fields are in turn divided into window and station sub-fields.

Within a site, stations are referred to by a local window value and a station number. Remote sites are referred to by a window value that is dynamically assigned by the local management, for the duration of communication with that site. The same window value can be used to refer to different sites at different times. Using the combination of window and station fields, a large number of external services can be addressed. Since all traffic between sites is statistically multiplexed onto a single U channel, the set of window values should be sufficient to maintain communication simultaneously with a number of remote services at different sites.

When a U channel is to be set up between sites, peer management services exchange window values; each having independently assigned a window value to refer to the other site. Window values are bound to U channels at the ramps for use in address translation, routing and expedited transfer.

Addressing and window mapping are also important in the *relaying* function of the network management [Tennenhouse89a]. Relaying involves a concatenation of U channels through a relay centre to transfer traffic between a pair of end sites. The end sites may lack direct network connectivity or may wish to maximise the use of a leased link. The relay centre will exchange window values with each end site and set up appropriate mappings on its ramp. Packets sent out from an end site undergo address translations both at the local ramp and the ramp at the relay centre.

Management Hierarchy

The management services on the Unison network support site interconnection by implementing functions such as association establishment, naming and addressing, window and channel management (including dynamic bandwidth management) and monitoring. Management functions are largely performed out-of-band in a transparent fashion to cause minimum interference with in-band transfer. These functions are implemented over a Remote Procedure Call (RPC) mechanism with appropriate semantics.

There are 4 components in the management hierarchy (arranged in a layered but non-embedded fashion). These are the secretary service, the window service, the channel service and the signal service.

The Secretary Service

The main function of the secretary service is in establishing associations. Local associations are set up when an end user contacts the secretary at a public address with the name of a service with which it wishes to connect. The secretary resolves the service name to a local address and notifies the destination of the service request. If the destination wishes to provide the service, it returns a private port number. As soon as the initiator receives the local address and port number from the secretary, it can begin transmission.

Remote associations are threaded across the local exchange, the ISDN and the remote exchange. As in local association establishment, the secretary may use a directory service to resolve the service name supplied by the initiator into an ISDN address for the remote site at which the service is located. The secretary forwards the service request to its peer at that site. The remote secretary contacts the service to obtain a private port number for the association.

The Window Service

The main function of the window service is address space management to support remote transfer. During association set-up, each secretary asks its window service to assign a window value to support the association. Window values that are exchanged between peer sites are bound to an existing U channel between the sites. Once a window has been established on a U channel, the establishment of additional associations may proceed without further recourse to the window manager. Multiple windows can be bound to a single U channel and expedited transfer would be based on the window value.

The Channel Service

The primary function of the channel service is the monitoring and dynamic bandwidth management of U channels by interacting with the ramp. On being informed of a new association to be bound to a U channel the channel service can take independent decisions on whether to increment the existing bandwidth. When a channel does not exist between sites at association set-up, the channel service would set up and synchronise a U channel of the required bandwidth.

The state of the high and low priority queues and the traffic rates through the U channels are used to drive reactive control algorithms which form part of the channel service. These algorithms control the dynamic adjustment of U channel bandwidth to provide effective site interconnection both in terms of cost and traffic performance. The design, modelling and performance analysis of dynamic bandwidth control schemes forms the central theme of this dissertation.

The Signalling Service

The signalling service resident on the ramp implements the DASS2 signalling protocol [Newman86] across the D signalling channel, to communicate with the System X switching exchanges in setting up B channels. The other management services may access the user data field in the D channel through the signalling service for the out-of-band transfer of control information. For example when a U channel is created window values to be bound to it are exchanged in this fashion.

The Unison Protocol Stack

To support the transfer of information over associations, the Unison network has a protocol stack whose relation to the ISO OSI [Halsall88] model is shown in Figure 3.3. The distinguishing feature of the protocol design was the use of receiver multiplexing at the data-link level as opposed to the Universe network. This was useful in providing the fine degree of sharing necessary for the transfer of multi-service traffic.

The physical layer is responsible for the transmission of symbols along the medium. Supported over this layer is the M-access or packet layer, which supports point to point exchange of CFR packets, as for example, between stations on the exchange ring or between peer ramps across the ISDN. The packet layer provides error detection and maintains the sequentiality of packet transfer.

UDL constitutes the protocol at the data-link level layer and operates over associations which support the transfer of UDL blocks. UDL supports the functions of association dependent multiplexing at its lower layers and those of segmentation/reassembly at its upper layers. It is the only layer to support multiplexing between associations in the stack.

UDL blocks to be segmented or reassembled can vary in size from 28 bytes to 7 Kbytes so that multiplexing can be as low as at the level of a single CFR packet. UDL blocks are segmented into packets for transfer in the data payloads of CFR packets. Each packet contains 4 bytes of header, reassembly and sequence fields and upto 28 bytes of data. The header includes a port number which is used to identify the association and as an aid in demultiplexing in the receiver. The reassembly and sequence fields identify the position of a UDL block within an association and the position of a packet within a UDL block, respectively. Packets belonging to different associations may be multiplexed and presented to the packet layer for transmission to the destination entities. At the destination UDL layer, the port values are used to immediately demultiplex packets belonging to different associations for reassembly into UDL blocks.

Continuing upwards through the stack, the absence of explicit network and transport layers is to be noted — this was because many functions that are implemented at these layers are now implemented in the lower layers, like multiplexing at the UDL level and

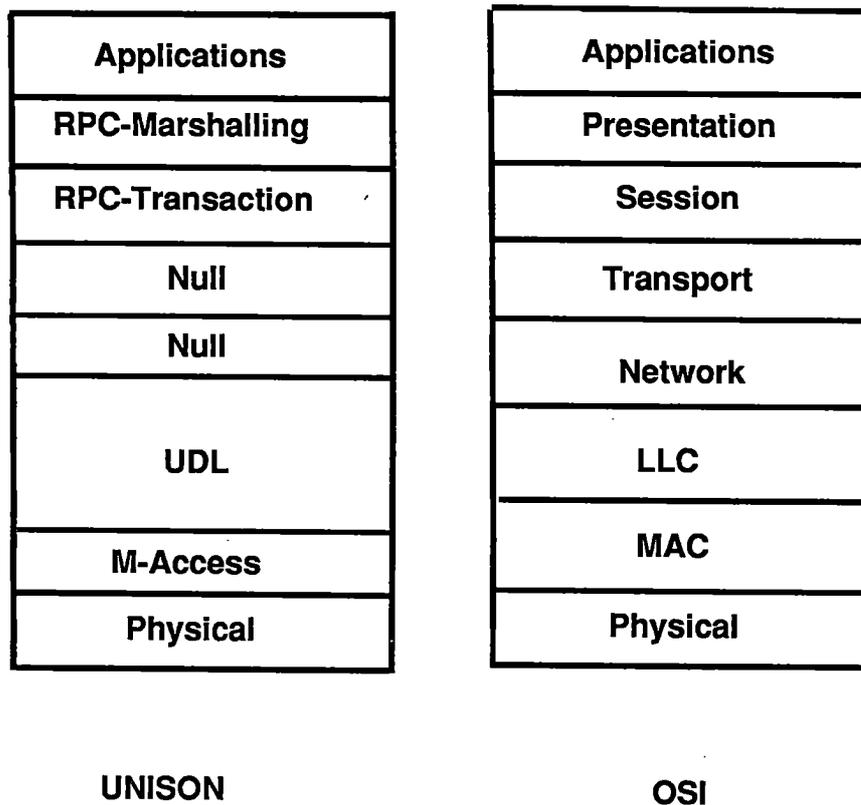


Figure 3.3: Protocol Suite : Unison and OSI

relaying and switching at the packet layer. The flow control and error recovery functions of the network/transport layers are left to be performed on an end-to-end basis by the upper layer applications. However, the equivalent of session and presentation layers is provided by a remote procedure call mechanism called UNITY [McAuley87].

RPCs with proper semantics are used by the management services given their need for reliable transactions. For example, the channel service uses an RPC based interface to control the ramp service in the flexible aggregation of bandwidth. Application level entities can also use RPC for their out-of-band transactions with management components, for instance in association establishment.

The revamped and latest version of the Unison protocol suite is called the Multi-Service Network (MSN) protocol suite and includes a network layer (MSNL) that supports *liaisons* which are concatenations of associations [McAuley90].

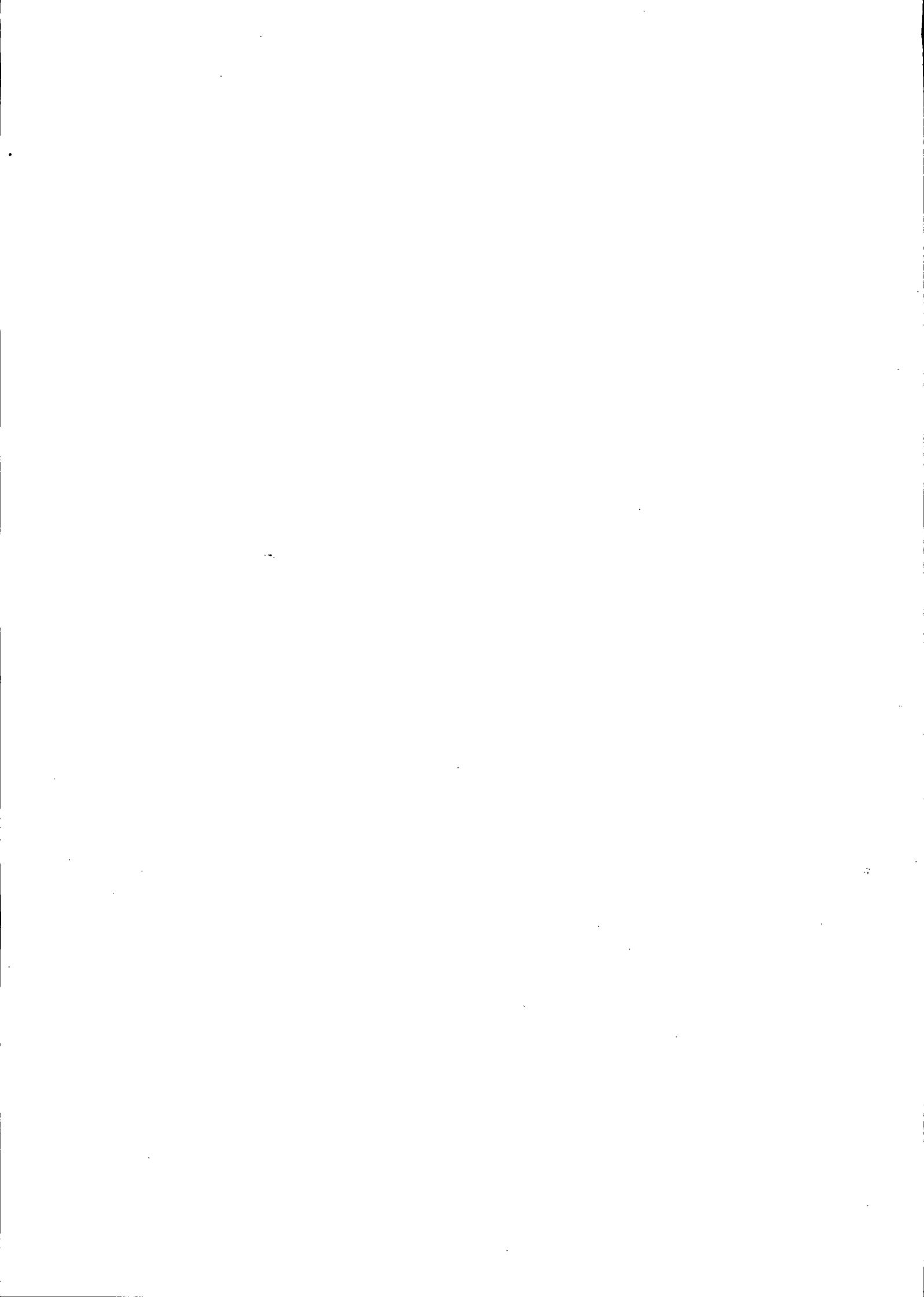
Portals

Portals [Tennenhouse89a] are bridges that connect client LAN's to the exchange. Peer portals can be viewed as supporting communication between applications on their networks using the site interconnection service provided by the exchange architecture. Portals are required to encode information from their local networks into CFR packets which are then transmitted onto the exchange.

The portals used in the Unison project include CFR, Cambridge Ring and ARPA Internet Protocol (IP) portals. The CFR-to-CFR portal functions at the UDL layer so that no encoding is necessary. However, associations still need to be set up across such portals, details of which are given in [Tennenhouse89a]. IP portals [McAuley90] act at the network level, and support communication between host systems on Ethernets.

3.4 Summary

This chapter has considered three network implementations that are of relevance to the work in this dissertation. The CFR which uses asynchronous time division multiplexing has the capability to function as a packet switch. The Universe project was a successful attempt at extending the functionality of LAN's to the wide area, under the constraints of using a satellite link with long round trip delays. Experiences gained from this project contributed to the design of the Unison project which examines the provision of a site interconnection service over a 2 Mbps public ISDN. The Unison exchange-based architecture supports a common ATM packet format at the data-link level and a flexible interconnection strategy for the transfer of multi-service traffic. Flexibility is provided by the dynamic bandwidth management of U channels performed transparently and in an out-of-band fashion.



Chapter 4

Bandwidth Management and Congestion Control

This chapter describes a congestion control framework for an ATM overlay network running over a public primary rate ISDN, as exemplified by the Unison network. The framework includes both static and reactive elements. The static functions are admission control and bandwidth reservation for incoming call requests. The reactive components monitor the flow of traffic and dynamically adjust the bandwidth in response to the traffic load. To ensure the conformance of sources to their specified bit rates, the framework also includes schemes for bandwidth monitoring and bandwidth enforcement implemented at the network periphery. The use of simple admission control and bandwidth allocation schemes, coupled with the dynamic variation of bandwidth to prevent congestion and reduce costs are the novel features of the proposed framework.

4.1 Introduction

The Unison network supports a packet overlay on an underlying circuit switched ISDN. The ramps in the exchange-based architecture preserve the ATM advantage of handling variable bit rates by supporting the flexible amalgamation of the underlying 64 Kbps circuits into larger $N \times 64$ Kbps U channels. This allows individual applications which are statistically multiplexed onto the channels to achieve instantaneous bandwidths of upto 30×64 Kbps.

U-channels between sites are dynamic — both in bandwidth and in the fact that they may or may not exist at a particular time. The basic problem to be addressed by *dynamic bandwidth management* is how wide a U channel should exist between pairs of sites at any

given time. Such dynamic control is important to the performance and cost-effectiveness of the system.

All packets bound to the same remote destination are multiplexed onto either the high or low priority queue of a U channel and experience queueing delays when the channel bandwidth is momentarily insufficient. Dynamic bandwidth control schemes sense the overload from the traffic rate and the size of the queues and make increments in the channel bandwidth. During periods of underload, the bandwidth of a channel would be decreased to reduce costs. Algorithms for implementing dynamic bandwidth management form the central theme of this chapter.

While dynamic bandwidth management is important in alleviating conditions of overload and reacting to burstiness in the input traffic stream, it is essential that it be combined with *admission control* and static bandwidth allocations for incoming call requests. The service requirements specified in a call request are used to decide whether to accept the call and to allocate bandwidth on a U channel, or establish one if necessary.

To supplement the reactive and static elements in controlling congestion, the final component of the framework is one that implements bandwidth monitoring and *bandwidth enforcement* at the network periphery. This component monitors individual traffic streams and ensures that traffic sources conform to the bit rates that were specified at call set-up. The bandwidth enforcement and dynamic bandwidth management components of the framework are illustrated in Figure 4.1.

The emphasis in the design of the congestion control framework has been on simplicity, and the control is intended to function on a best-effort basis. During congestion, packets could be lost and applications demanding reliable transfer would have to implement end-to-end recovery schemes above the framework. In the remainder of this chapter, the following three elements of the congestion control scheme are described in more detail :

- Admission control and static bandwidth allocation.
- Reactive control based on dynamic bandwidth management.
- Bandwidth monitoring and bandwidth enforcement.

4.2 Admission Control and Bandwidth Reservation Systems

Admission control schemes are based on reservation systems and blocking. For example, when a telephone call is made, a reservation for 64 Kbps can be registered with the exchange management — if the reservation cannot be fulfilled, the call fails.

Reservation systems and bandwidth allocation at call set-up will depend upon the traffic types to be transferred over the network. Traffic types can be distinguished on the basis of

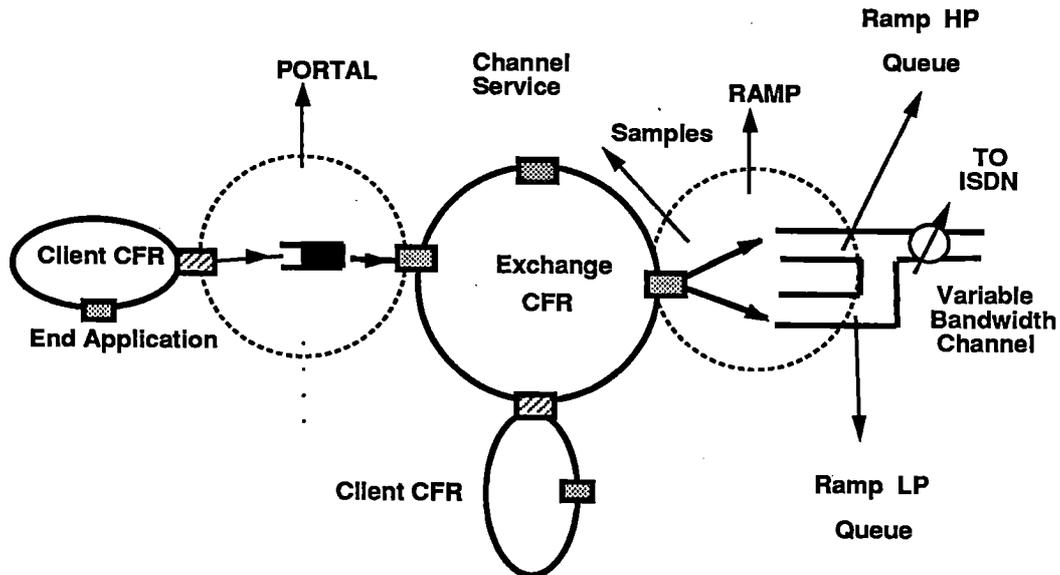


Figure 4.1: Dynamic Bandwidth Management and Bandwidth Enforcement in the Congestion Control Framework

their burstiness and on their delay and reliability requirements. In the present discussions, the two main classes of services considered will be Continuous Bit Stream Oriented (CBO) and Bursty Data (BD) services. CBO sources produce a synchronous stream of packets that are generated at regular intervals and have stringent requirements on delay and jitter. They are exemplified by constant bit rate voice or video calls. BD services consist of an asynchronous or bursty stream of packets and include traffic generated by applications in distributed computing and data retrieval. The two classes can be further distinguished on whether they require guaranteed service or not as will be discussed in more detail in the next section.

In designing an admission control scheme for the packet overlay there are two sharing problems to be considered. The first is the sharing of the ISDN link bandwidth between the traffic classes; the *access control* problem. The second is that of dividing the bandwidth of the ISDN link between multiple U channels emanating from a site.

Access Control

When integrating different traffic types, it is essential that there be mechanisms for controlling their relative access to the common transmission channel.

The simplest method of access control is a *complete sharing* scheme [Schwartz82] with the traffic types competing for the channel bandwidth on a first come first served (FCFS) basis. Incoming traffic requests are accommodated for as long as there is enough remaining bandwidth. A disadvantage of the complete sharing scheme for heterogeneous traffic requests with widely varying bandwidth requirements, as pointed out by Kraimeche [Kraimeche84] and Gopal [Gopal83] was the 'instability' effect as the frequency of call requests increased. The channel throughput could decrease and call requests for larger bandwidths could experience higher blocking than those which demand less bandwidth — traffic types could 'crowd out' one another. Kraimeche studies the use of *restricted access* strategies which approach the problem by grouping together sources with similar bandwidth requirements and partitioning the channel bandwidth between the groups, so that they cannot affect each other.

At the other end of the spectrum of access control strategies is the *complete partitioning* scheme [Schwartz82] where the link bandwidth is divided statically between the traffic types, each with its own dedicated access privilege. The drawbacks of this scheme are the possible inefficiencies in network utilisation and that partitioning the bandwidth between specific traffic types are contrary to the ATM principles of flexibility and service independence. The access performance experienced by the various traffic types in the complete partitioning scheme with call blocking can be quantified by the Erlangian blocking formula [Cooper81].

For example, let the connection requests for a particular traffic type be assumed to arrive at a rate of λ calls per second and let the holding times for each call be exponentially distributed with mean value $1/\mu$ secs. Then if the request intensity be defined as $\rho = \frac{\lambda}{\mu}$ and if N be the number of circuits (bandwidth) assigned for the traffic type, the call blocking probability P_B experienced by incoming traffic requests is given by :

$$P_B = \frac{\rho^N / N!}{\sum_{i=0}^N \rho^i / i!}$$

Given the traffic load and blocking probability constraints for each traffic class, the formula can be used to calculate the link bandwidth that should be allocated for the class. A fair allocation scheme would be one which equalises the blocking probability across the set of traffic classes [Kraimeche84].

Various other hybrid schemes have been proposed for blockable traffic requests, both heuristic schemes [Schwartz82] and those derived analytically from optimal control formulations [Kraimeche84, Gopal83, Gersht88]. The criterion of optimality is generally to minimise a function of the blocking probabilities, such as the weighted sum. Access control strategies have also been considered for other traffic environments; these include a

mixed blocked-queued traffic environment [Kraimeche85, Serres88] and an all-queued traffic environment [Kraimeche86], where the blocking and queueing capabilities refer to the input access requests.

The approach to access control in the current work has been to use a complete sharing scheme with CBO and BD call requests accessing the ISDN link bandwidth in a FCFS fashion and handled on a blocking basis. The instability effect described above was for requests with widely differing bandwidth requirements — Gopal and Kraimeche considered traffic classes with bandwidth requirement ratios greater than 1:15. In the primary rate network considered in the current work the bandwidth requirements of CBO services (generally 64 Kbps voice) and BD services (distributed computing traffic) are unlikely to be so different.

Bandwidth Sharing between U Channels

The approach to sharing the bandwidth of the ISDN link between multiple U channels in the current framework is to rely on the access control scheme defined above. This avoids having to define new allocation strategies at this level. A possible disadvantage of this scheme is that a heavy traffic load on a particular channel may 'hog' the bandwidth of the ISDN link. On the other hand, partitioning the link bandwidth statically between the channels could result in inefficiencies in bandwidth utilisation which is an important consideration when using primary rate ISDN. In an experiment to be described in Section 9.5 in Chapter 9 the dynamic sharing of ISDN link bandwidth between U channels by using dynamic bandwidth management will be considered.

With larger bandwidth networks, where bandwidth utilisation is not critical, explicit strategies for the sharing of link bandwidth between channels would be important. Apart from simple partitioning, hybrid methods suggested by Schwartz [Schwartz82] for the sharing problem between traffic classes can be adapted for use in sharing bandwidth between channels. These include schemes that allocate a fraction of the bandwidth of the link to each channel and leave the remainder in a free pool to be accessed by all channels on a FCFS basis (the *unrestricted shared pool* scheme). A conservative variant of this scheme is to further control channel access to the free pool by restricting each channel to the use of only a pre-allocated portion of the free bandwidth (the *restricted shared pool* scheme).

A Mechanism for Call Admission

Call requests corresponding to CBO and BD services are both admitted on a blocking basis in the current framework. However, the criteria on which admission decisions are made and the static bandwidth allocated can be different for the two services depending on the *service descriptors* specified in the call requests. In the current framework, call requests will be confined to service descriptors that specify the peak and average bit rates

and whether guaranteed bandwidth is required.

Requests for guaranteed bandwidth (from CBO or BD services) are accepted on the basis of the peak bit rate specified. If the sum of the bandwidth of the existing connections and the peak bit rate of the new call request does not exceed the ISDN link bandwidth, the call is accepted; otherwise it is rejected. Packets from such sources will be multiplexed onto the high priority queue of a U channel.

The rationale behind peak rate allocations is to emulate the performance of circuit switched transfer. The use of peak rate allocations and expedited transfer at intermediate contention points ensure that such sources receive the same quality of service (jitter and delay constraints) from packet based transfer on the overlay as they would have with dedicated ISDN circuits. The probability of packet loss for these sources within the network will be extremely low, assuming that the peak rate is well known and that the traffic sources are well behaved. The effect of 'malicious' sources who transmit at a different rate from that specified at call set-up will be discussed in Section 4.5.

Call requests not requiring guaranteed bandwidth service (generally of the BD type) are handled differently. When explicit bandwidth allocations are made for such requests, they are done on the basis of the average bit rate specified in the request and the call is admitted if there is enough bandwidth available on the ISDN link. This allocation strategy relies on dynamic bandwidth management to switch in additional bandwidth, if necessary, during the course of the call. Another objective is to exploit the residual bandwidth that may be available due to the granularity in bandwidth adjustments (64 Kbps) and from the statistical multiplexing of other bursty sources, especially those which have been granted peak rate bandwidth allocations. Bandwidth utilisations can be estimated by examining the U channel statistics — their bandwidths, input traffic rates and queue lengths.

A third class of traffic, control traffic, also needs to be transferred across the network. Some control transactions for dynamic bandwidth management and channel establishment are performed in-band. Control packets are multiplexed onto the high priority queue of a U channel and use pre-assigned window bindings set up at channel establishment time. Such control traffic can be transmitted without an explicit static bandwidth allocation since its volume is very low.

Call Admission Control in Practice

In the Unison network, the establishment of virtual circuits between end applications follows a chain of control that is supported by the components of the exchange management — the secretary, the window and channel services. The initial call request from a client application is made through an RPC interface to the secretary service on its local exchange. The subsequent actions of the secretary in setting up the association, including its interactions with the window service have been described in Section 3.3 in the previous chapter.

The channel service would be supplied with the service descriptors specified by the client and can make independent decisions on call acceptance and bandwidth allocation based on the admission control scheme described above. When the call request does not explicitly specify any requirements, the channel service may either try to estimate them in making a static bandwidth allocation or rely on reactive control to adjust the bandwidth during transmission. Bandwidth estimates could be based on hints supplied by the secretary, from its knowledge of the service to which the application is trying to connect.

During association establishment, the remote channel service can refuse to accept incoming requests, if the bandwidth and quality of service parameters are unacceptable, so that there is no danger of compromising the independence of a site in controlling its ISDN bandwidth. However, in the interests of greater reactivity, the sites can negotiate mutually acceptable upper and lower bounds on bandwidth adjustment at the start of day.

Emulating Connectionless Transfer

An interesting problem in call admission, is how connection requests which demand a very low set up delay can be handled on the packet overlay. Connectionless emulation for datagram transfer over ATM networks has been addressed by Prycker [Prycker88] who suggests the use of permanent virtual circuits.

In our framework, it is proposed that low call set-up delays be provided by using pre-assigned window values and enabling transmission without allocating bandwidth as for control traffic. The number of transactions in call set up is now limited, with the secretary service returning a window value (and port number) to the client soon after the initial call request has been made.¹ However, unlike with control traffic, bandwidth needs to be allocated after the onset of transmission, either by an explicit bandwidth allocation or by relying on reactive control to dynamically switch in bandwidth during transmission. Until the bandwidth is increased, the new connection may cause interference with existing traffic.

4.3 Dynamic Bandwidth Management

There are occasions when the bandwidth required by an application is unknown or else so variable that prior reservation and bandwidth allocation makes no sense. It has already been pointed out that for applications which require low call set up delays, the overhead of static bandwidth allocation and connection establishment may not be suitable. A poor estimate of the bandwidth to be allocated at call set-up, inaccurate service requirements specified in a call request or a malicious user transmitting at a different bit rate from the specified one, are all conditions which cannot be solved by static bandwidth allocation alone and can cause congestion to occur. In the present framework, such situations are

¹This may include a transaction between the secretary and the channel service.

seen as being remedied by the dynamic adjustment of bandwidth, performed transparently during the course of packet transfer. Such reactive bandwidth control is also seen as important in improving the effectiveness of statistical multiplexing.

Dynamic bandwidth management schemes react to the traffic load by adding capacity when needed to reduce delays and congestion in the network, and to reduce capacity whenever possible to reduce costs. In doing so, the control seeks to maintain a reasonable tradeoff between the complementary factors of cost reduction and traffic performance.

The problem addressed by dynamic bandwidth management is how, from the state of the network — in the form of queue sizes at or traffic rates through the ramps, the channel service can determine whether to add or subtract bandwidth from a U-channel. The channel service interacts with the ramp associated with the U-channel through an RPC interface across the CFR, both in order to update its state table (the *monitoring* function) and to configure the ramp to set up and delete calls (the *connectivity* function). As part of the monitoring function, the channel service receives samples detailing the state of the U channels that are generated at the ramp at regular intervals. The connectivity function is implemented by bandwidth algorithms that are driven by the samples and are based on overload control, a term taken from the paper by Li [Li89b].

4.4 Algorithms for Dynamic Bandwidth Management

This section describes algorithms that are used for dynamic bandwidth management in the congestion control framework.

One of the most important considerations in designing a dynamic bandwidth control algorithm is the time taken to switch-in bandwidth. Switching-in delay can be quantified by the *slot latency*, which is the time between overload occurring on a channel and when the new circuit(s) can be used. Slot latency includes sampling, transmission, control and bandwidth switching delays. Sampling delays occur because the queue sizes at the ramps are sampled only at discrete epochs. Samples of the queue state are subject to transmission delays across the CFR which separates the channel service and the ramp. The control delay is the time taken by the channel service to update its state tables and use the bandwidth algorithm in deciding whether to adjust the bandwidth. A further transmission delay is incurred across the CFR before the decision reaches the ramp. Finally, the bandwidth switching delay is the time taken at the ramp to increase the bandwidth of the channel — composed of call signalling and channel synchronisation components.

During the switching-in period, the system will be running at overload and it is desirable that switching be effected as soon as possible, especially if it is sensed by the controller that the overload is prolonged. Having experienced a possible performance degradation and incurred a cost because of switching in circuits, it is advantageous to delay the reclamation of the newly switched in bandwidth. This is *hysteresis* control — a mechanism which can be also thought of as delaying the reclamation of excess resources during temporary lulls

in channel activity.

In two-phase hysteresis, control is implemented across a pair of thresholds. A decision to increase the bandwidth is made when the system state (as perceived by the controller) exceeds the upper threshold and the additional capacity is retained until the state drops below the lower threshold. In multi-phase hysteresis, there can be many threshold pairs that may overlap. *Point thresholding* is a degenerate form of hysteresis in which the thresholds of a pair are identical.

Types of Dynamic Bandwidth Algorithms

Two distinct classes of dynamic bandwidth algorithms can be distinguished on the basis of the state information that is used to drive them; queue-based algorithms are driven by the size of the queues of the U channels, rate-based algorithms are driven by the traffic rates into the channels. In practice, samples of the ramp's state received by the bandwidth controller contain both the input traffic rates and the instantaneous values of the queue length at the sampling instant. However, the discussions below assume that 'pure' queue-based algorithms do not use the rate samples.

Queue-Based Algorithms

The simplest queue-based algorithm is one which uses a hysteresis mechanism and a single threshold pair. The algorithm functions as follows: when the sampled queue size is observed to have crossed the upper threshold, a B channel is switched in. The bandwidth is retained until the state has crossed a lower threshold, when it is switched out. In practice, additional limits are imposed on both the maximum and minimum bandwidths permissible. The lower limit is the minimum bandwidth that is to be retained on a channel — usually 64 Kbps. An explicit upper limit² is necessary because the algorithm can switch in more B channels than is necessary during the slot latency.

An improved queue-based algorithm is one in which the control is effected across multiple threshold pairs, with each pair having the same lower threshold value. Multiple thresholds provide an indication of the traffic rate — for lower traffic loads, a smaller number of thresholds will be crossed in a given time interval than for larger loads. If the amount of bandwidth to be switched in is non-decreasing with the threshold value, then the algorithm makes more accurate adjustments of the bandwidth than the previous algorithm.

Queue-based algorithms can use the samples of queue length to estimate the traffic rate. Since the sampling rate at the ramp is known, consecutive samples of queue size and a knowledge of the existing channel bandwidth can be used to estimate the input traffic rate. This estimate can be used to compute the bandwidth increment or decrement.

²The default upper limit is the capacity of the ISDN link.

Rate-Based Algorithms

Rate-based algorithms are driven by samples of the input traffic bit rates into the U channels at the ramps. The difference between the input traffic rate and the existing bandwidth on a channel provides the required bandwidth increment or decrement.

The simplest algorithm in the rate-based family, functioning with hysteresis control across a single threshold pair, does not suffer some of the disadvantages experienced by the corresponding queue-based algorithm. Since the appropriate bandwidth increment is directly calculable, there is no danger of constantly adding bandwidth once the threshold has been crossed, unless the traffic load keeps increasing. Rate based algorithms with multiple thresholds are examined in the next section.

Explicit knowledge of the input traffic rate implies that rate-based algorithms can be more reactive than queue-based ones as shall be demonstrated in Section 9.2 of Chapter 9. Having switched in circuits, rate-based algorithms will not decrease the channel bandwidth as long as the input traffic rate remains unchanged. This is in contrast to queue-based algorithms which decrease the channel bandwidth when the sampled queue length falls below the lower threshold. In keeping with the general principle of delaying the reclamation of resources, decreasing the bandwidth by both rate and queue-based algorithms is done in an incremental fashion — one B channel at a time.

Other classes of bandwidth algorithms may be defined. These can be hybrid algorithms including features of both rate and queue-based algorithms or can be based on different measures of the U channel state. For example, packet loss statistics can supplement the queue size as an index of the traffic state. Hybrid algorithms will be further discussed in Section 9.4 of Chapter 9.

Switching Thresholds

The choice of thresholds has an important effect on the performance of rate and queue-based algorithms when operating in a bursty traffic environment.

In rate-based control, the choice of thresholds is straightforward — the permissible channel bandwidths ($64 \times N$ Kbps, $N = 1, 2, \dots, 30$) represent natural boundaries at which switching can be implemented for both point thresholding and hysteresis control mechanisms.

In queue-based control, the setting of the threshold values is not so clear. In theory, threshold values and bandwidths to be switched-in can be derived from optimal control formulations, but only for simple loads [Lu84]. This will be discussed in Section 7.11 of Chapter 7. Another method is to compute the thresholds from the steady state analyses of queueing models such as an $M/D/1$ queueing system³ [Kleinrock75]. In practice, it

³For example for a given utilisation and Poisson arrivals, the upper threshold U , can be chosen such

is simpler to compute thresholds from approximations of the average queueing delay that would result if bandwidth adjustment were to be initiated at them using the mean values of the bandwidth switching and sampling latencies and the traffic bit rates. This can be guided by heuristic rules. For example, to avoid switching on transient variations in the input traffic stream and in the interests of cost a threshold cannot be very low. On the other hand, a very high threshold will increase the queueing delay.

State Filtering

When operating in a bursty, multi-service environment, the state information that drives the bandwidth algorithms can be highly variable. The information could also be incomplete or outdated, conditions that are exacerbated by the channel service being remote from the ramps. Under such conditions, it is more suitable to use filtered values of the sampled state than instantaneous values to drive the algorithms. Filtering of state is also important in decreasing the reactivity of an algorithm; experiments describing how filtering improves the performance of a rate-based algorithm will be described in Section 9.3 of Chapter 9. In the present discussions filtering is confined to a simple average over a suitably defined *history* interval.

Time Lagging

Hysteresis in time (also called *time lagging*) is another means of decreasing the reactivity of a bandwidth algorithm. This is applicable because of the tariff structure implemented on the ISDN in the Unison network. Charges for switched circuits, as in voice telephony, depend on the destination and duration of the call. Charging instants are discrete and charges are levied at the beginning of a charging interval; having switched in a B channel it is sensible to retain it for the duration of the charging interval. With time lagging, the decision to decrease the bandwidth is still made on the state (queue/rate, filtered/unfiltered) but only at epochs just before the end of a charging interval.

4.5 Bandwidth Monitoring and Bandwidth Enforcement

Bandwidth monitoring and bandwidth enforcement functions are important components of a congestion control framework for an ATM network. Bandwidth monitoring consists of measuring that individual users conform to the bit rates that were specified at call set-up. Bandwidth enforcement is the action taken to control non-conforming users.

Dynamic bandwidth control at the ramps can remedy congestion caused by non-conforming users but cannot prevent it from occurring. Individual traffic streams are not distinguished at the ramps; during congestion packets from all users are discarded at the ramp

that $P[\text{Queue length} > U] \leq (\text{say}) 0.05$.

queues, which is unfair on the 'well-behaved' ones.⁴ Rate control on individual streams is best implemented at other contention points, preferably at the periphery of the network. In the present framework, it is proposed that rate control on individual traffic streams be implemented at the portals.

Methods of Bandwidth Monitoring and Enforcement

Various methods have been proposed for bandwidth monitoring and enforcement in ATM networks [Akhtar85, Joos89, Jacobsen90, Rathgeb90, Gallassi89, Hirano89]. The schemes are based on the rate control of traffic sources at the network periphery or the user-network interface. The underlying principle is to monitor the number of packets emitted by a traffic source during a time period, both parameters being negotiated at call set-up. Control is generally enforced by discarding packets from a violating source.

A specific proposal to control the input rate is the leaky bucket scheme [Turner86b, Akhtar85] which uses a counter, the state of which reflects the short term bandwidth requirements of the source. Incoming packets cause unit increments of the counter. Decrements of the counter take place periodically at the service rate assigned to the traffic source. When the counter exceeds a pre-defined threshold, incoming packets are discarded. In the virtual leaky bucket scheme [Hirano89, Gallassi89], violating packets are marked and buffered. If there is congestion within the network, then marked packets are preferentially dropped.

Another scheme for bandwidth enforcement has been characterised in recent attempts to apply *fair queueing* algorithms to ATM networks [Demers89, Nagle87], with separate queues maintained at gateways for each individual traffic stream. Fair queueing disciplines can ensure fairness in bandwidth allocation and penalise malicious users selectively in contrast to simple FCFS disciplines.

While these proposals are interesting in theory, there are many issues which have to be addressed before they can be put into practice. The choice of a suitable monitoring interval (threshold value) and the practicalities of implementing complex queueing disciplines and rate control mechanisms for each individual connection are open to question. Woodruff [Woodruff88] states the need for traffic descriptors that can be both accurately monitored and enforced; the use of 'deterministic bounded quantities (like maximum burst lengths or minimum inter-arrival times) which when violated can trigger enforcement' is emphasised.

Verbiest [Verbiest88a] has approached the problem of defining a suitable integration interval to detect the violation of statistical quantities like the average bit rate. The length of the integration interval needs to be sufficient to distinguish between statistical fluctuations of a traffic stream and deliberate violations by traffic sources. Although the use

⁴Having to control individual traffic streams would interfere with the ramp's main function of U channel management and bandwidth aggregation.

of deterministically bounded quantities may be preferred over traditional statistical moments, Verbiest and [Forgie77] have suggested the use of a moving window mechanism that can signal violations even of statistically averaged quantities. One of the disadvantages of this method is that the integration periods are not constant — since at the beginning of a call only small fluctuations are allowed, sources would have to implement ‘slow start’ or fixed bit rate transmissions.

Portals and Bandwidth Enforcement

Portals, as bridges between client distribution LAN’s and the exchange CFR, are more suited to exercising specific control on the traffic streams which are multiplexed through them than the ramps. In forwarding call requests from sources on client LAN’s to the secretary service on the exchange, the portals form part of the admission control framework.⁵ In the following discussion, it is examined how the functionality of the CFR–CFR portals can be extended to include bandwidth monitoring and bandwidth enforcement.

At primary rate access speeds with a sparse number of streams each of moderate bit rate and multiplexed through a portal it is reasonable to assume that the individual streams can be monitored by the portals for conformance. Monitoring would be simple, either by keeping a moving average of traffic rates or by detecting peak rate violations from measuring packet inter-arrival times. Although these schemes involve per-packet processing⁶ they can be implemented in hardware. There will be a need for additional protocols between the channel service and the portals, to exchange permissible traffic rates at call set-up.

At CFR-CFR portals consistently aberrant sources may be penalised by the use of the CFR hate list [Hopper86], so that the portals stop receiving packets from them. Another method of bandwidth enforcement is to discard packets selectively within the portals from sources that contravene average and peak rate specifications. A study of a congestion control scheme at the MAC level in CFR-CFR bridges is to be found in [Porter90].

This proposal is to be treated as providing a flavour of how bandwidth monitoring and enforcement can be implemented at the portals in the Unison network with certain assumptions on traffic flow, rather than as the definitive solution.

4.6 Related Work in Dynamic Channel Management

The problem of the dynamic variation of channel bandwidth in a primary rate ISDN has been identified in papers by [Knight87, Deniz89, Deniz90]. Knight and Deniz

⁵ Association set-up is a rudimentary form of flow enforcement — since no source can start transmitting packets without first contacting the secretary and receiving a window value and port number.

⁶These are CFR packets since the proposed scheme is for a CFR-CFR portal.

[Knight87] present issues in the design of gateways from local networks to a primary rate ISDN. The problem of the design and analysis of circuit management algorithms that are to be implemented at the gateways to cope with varying traffic loads is identified.

In a later paper [Deniz89], Deniz and Knight provided details of an Ethernet-ISDN gateway that supports a connectionless network layer. The authors identified the gateway as having to manage a variable number of B channel connections depending on the system state and traffic variations. Channel management was seen as having to satisfy conflicting criteria, described as 'bandwidth utilisation, user perceived response, data throughput, delay and real communication costs'. It was suggested that the optimisation of these criteria may only be possible by tuning the channel management strategies to suit the end applications or to parameters of the transport and network layer protocols such as time-outs. In the description of an ISDN emulator used in the work, the use of Channel and Circuit Control Status tables are described which maintain the status of channels at the ISDN interface through monitoring. The tables are to be consulted by a Channel Management Routine to determine whether a new circuit should be opened or an existing one closed. However, no experimental results or explicit algorithms for channel management were presented.

Recently, Deniz and Knight [Deniz90] considered the formation of superchannels, similar to U channels, in a primary rate ISDN. This paper suggested modifications to existing ISDN protocols, including the D channel signalling protocol, to allow for superchannel formation and its possible dynamic variation of bandwidth. No schemes for implementing the dynamic bandwidth variation were described.

Gerla and Pazos-Rangel [Pazos-Rangel82, Gerla83] consider the use of 'express' pipes in supporting packet transfer over a circuit switched network. Express pipes are circuit switched channels created between every pair of nodes in the network which can support the transfer of asynchronous traffic between the nodes by statistical multiplexing. The use of express pipes was seen as reducing the end-to-end packet delay and storage and processing overheads. The bandwidth of a pipe is sized on creation according to the traffic volume from source to destination nodes. Gerla suggests that the dynamic variation of bandwidth on the pipe would be more suitable for bursty traffic flows.

4.7 Summary

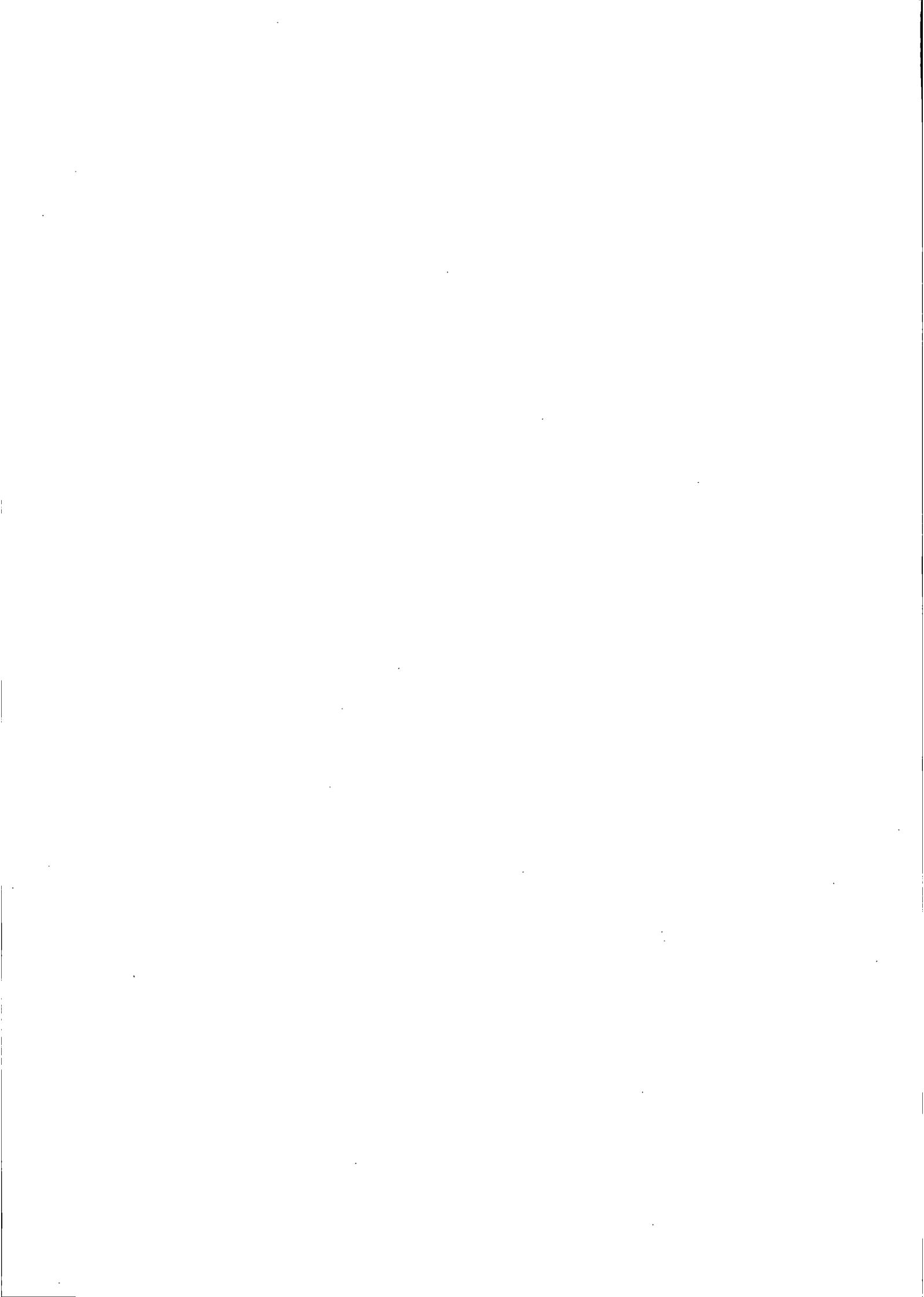
In this chapter a congestion control framework for an ATM overlay running over a public, primary rate ISDN has been described. The framework is based on admission control, dynamic bandwidth management and bandwidth monitoring and enforcement functions. The novel feature of the framework is the use of dynamic resource allocation to control congestion and reduce costs. 'Soft' guarantees on packet loss are provided; in order to ensure reliable transfer, applications would have to run end-to-end recovery schemes.

The basic theme of the framework has been the use of simple admission control schemes and

to rely on reactive bandwidth control to cope with burstiness in admitted traffic streams. Applications which demand guaranteed service are allocated bandwidth on their peak rates and are provided expedited transfer at contention points. For other applications, allocations when made are on the basis of their average bit rates. One of the advantages of a congestion control framework based on dynamic bandwidth management is that traffic requirements do not have to be precisely defined. This is particularly important when they cannot be estimated accurately.

Reactive control schemes use the general strategy of adding bandwidth quickly when the system enters a phase of overload and then delaying its reclamation. The tradeoff between cost reduction and traffic performance is central to the design of dynamic bandwidth management schemes and is studied in detail in the experiments in Chapters 8 and 9.

Tennenhouse and Leslie [Tennenhouse87] originally proposed the use of dynamic bandwidth management in the Unison network. The design of the dynamic bandwidth control schemes and the congestion control framework described in this chapter are the contributions of the author.



Chapter 5

Previous Work

The first part of this chapter reviews developments in the congestion control of ATM networks. The rest of the chapter contains a review of performance modelling and analyses of integrated services networks including those based on ATM.

5.1 Congestion Control in ATM Networks

This section will begin with a description of three proposals for congestion control frameworks. This is followed by a discussion of some general design issues for congestion control schemes.

Proposal A : Packetised Virtual Circuit Network

Forgie and Nemeth [Forgie77] were among the first to suggest a packetised network for integrating voice and data traffic that combined elements from both packet and circuit switching. This farsighted proposal was motivated by a desire to handle both types of traffic in a uniform way and to exploit the statistical multiplexing advantages that are inherent in packet transfer. Another goal was to achieve flexibility and responsiveness in adapting to different traffic mixes. The use of short, fixed length packets with relative labels and virtual circuit oriented transfer was the approach adopted in the network.

New connections were assigned to transmission links on the basis of the average bandwidth requirements specified in the connection request and the maintenance of a statistical safety margin on the link (by accepting connections only when the average bandwidth required was a small fraction of the total link bandwidth). This maintained constraints on packet loss and the probability of overflow. This bandwidth allocation scheme at connection set

up was combined with a rate control mechanism at the network periphery to monitor peak and average traffic rates. Conventional flow control mechanisms, operating at the data-link level, were considered inappropriate for the transfer of packetised voice in the scheme.

For peak rate violations, it was proposed that packets from the offending connection be dropped at the periphery. A moving window mechanism was used to detect infringements on the average bit rate. Infringements on the average rate were to be dealt with by either dropping packets from the connection or by increasing the bandwidth allocated on the associated transmission link. At the onset of congestion it was proposed to discard packets when the buffers were full and rely on end-to-end retransmission and recovery for maintaining loss free data transfer.

The performance analysis which accompanied the proposal was for a voice-data packet multiplexer, with voice having prioritised access, and was a combination of simulation and queueing analysis. Simulations were used to calculate the statistics of the time between empty slots, when a number of talkspurt voice sources were multiplexed together. Since data packets could be transmitted on idle (empty) voice slots, the empty slot distribution was used as the service distribution in an $M/G/1$ queueing model. This was solved to calculate the steady state performance characteristics of the data traffic process. The results of the simulation and analysis demonstrated that acceptable voice and data traffic performance could be achieved within the proposed scheme.

Proposal B : A Congestion Control Framework for ATM Networks

More recently, Gersht and Lee [Gersht89] have proposed a congestion control framework for an ATM based network which assumes two service types — express class service corresponding to voice and video traffic and first class service corresponding to data traffic. Express class traffic is given prioritised access over first class traffic at contention points along the route. The framework has two control levels — a virtual circuit control level and a packet control level, both of which operate at the network periphery.

Virtual circuit control is done on a blocking basis at call set up — express class requests are accepted on the basis of their peak bit rates, while first class connection requests are accepted on the basis of a guaranteed bandwidth requirement, which lies between their peak and average bit rates.

Packet control is exercised during packet transfer at the network interface to choke and relieve only the first class packets during congestion. In an effort to maximise the network utilisation even though first class sources were allocated bandwidth on their guaranteed bandwidth requirements they were permitted to transmit at their peak rates. At the onset of congestion along a route, detected by the queue lengths crossing *onset* thresholds, all first class connections multiplexed on the route would have to reduce their transmission rates. This was to be done by sending out choking packets to each first class source to

reduce their transmission rate to their guaranteed one. When the queue subsequently falls below an *abatement* threshold, relieving packets are sent to each of the sources, permitting them to resume transmission at their peak rates. The choking and relieving scheme is equivalent to a hysteresis mechanism. Methods for flow monitoring and enforcement were proposed to be implemented by rate counters.

The congestion control scheme guarantees the performance of express class traffic through peak rate bandwidth allocations and prioritised access. However, the disadvantage of the scheme is that to function effectively at broadband speeds for which it was intended, the choking/relieving strategy would require large buffers. This is to keep the loss probability of first class packets constrained, but could increase the average delay for this class considerably. Another disadvantage with the scheme was that the guaranteed bandwidth requirement for first class traffic was not defined.

Proposal C : Traffic Control for Guaranteed Performance

A more feasible design of a congestion control scheme for an ATM network is presented in [Woodruff90, Koistipaiboon90]. In their design, the authors emphasize the need for simple and *robust* control in order to cope with traffic uncertainties and to ensure guaranteed traffic performance. Based on simulation and theoretical performance results, a simple congestion control scheme is proposed to handle multiple classes of traffic. The scheme is based on preventive control and bandwidth enforcement.

The preventive component, based on admission control, is concerned with accepting calls and allocating bandwidth on the basis of a traffic flow description provided by an end application at connection set-up. Constant bit rate sources and those with stringent performance requirements are to be allocated bandwidth according to their peak rate requirements (the *non-statistical* mode). Only bursty traffic sources with low peak rate to link bandwidth ratios¹ and those with burst lengths that were short in relation to the buffer sizes at the nodes were accommodated in the *statistical* mode. In this mode bandwidth was allocated according to a virtual bandwidth measure that lay between the peak and average bit rates of a traffic source. Applications with longer bursts were to be handled in the non-statistical mode on the basis of their peak rate bandwidth requirements unless the peak rate to link bandwidth ratios were very low.

No explicit definition of the virtual bandwidth was made in the model — it was proposed that it be calculated from experiment and simulation. Assuming that it can be defined, for the statistical mode of operation, admission decisions for a new connection were based on whether the sum of the virtual bandwidths of existing connections and that for the new request was less than the channel capacity reserved for the statistical mode. A similarly computed sum of peak rates was used for accepting a new connection in the non-statistical mode of operation.

¹A similar principle to the statistical safety margins in Forgie and Nemeth's proposal.

Non-statistical mode traffic would have prioritised access over statistical mode traffic at contention points, to ensure a low packet loss and minimal packet delay. More than two levels of priority were seen as unnecessary at the broadband speeds the framework was meant for and as adding to the complexity. Buffers for non-statistical mode traffic were small to constrain the delay while those for the statistical mode would be larger. It was proposed that flow enforcement schemes be used at the statistical mode buffers in selectively dropping tagged packets from them (that were tagged by rate monitoring schemes at the network periphery) during congestion.

Although the congestion control framework for the Unison network was proposed independently, there are similarities between the two schemes — in the choice of only two traffic classes, prioritised access for one of them and admission control based on relatively simple rules. An advantage of the Unison framework is that there is no necessity to define a virtual bandwidth measure. Low priority traffic requests are admitted on the basis of their average bit rates. Burstiness in such traffic is absorbed by statistical multiplexing when there are a sufficient number of sources, or reacted to by the dynamic adjustment of bandwidth.

Design Issues in Congestion Control

This section describes some design principles for congestion control schemes in ATM networks that are based on the work of [Woodruff88, Woodruff90, Sallberg90]². Woodruff emphasizes that the design of a congestion control framework for high speed packet transfer should be simple, effective, fair and not optimised for any specific service type. The control should be designed to meet the packet loss and packet delay requirements of all traffic, and bandwidth should be allocated fairly among all classes. The use of preventive control is preferred to reactive control based on feedback during congestion. The disadvantages with the latter are pointed out as being the large buffers required at broadband speeds and the unfairness in back pressuring all sources during a period of congestion rather than just the offending ones.

Preventive methods are seen as being based on access control and a connection oriented approach. Three levels of access control are distinguished — route control, admission control and bandwidth enforcement. Woodruff states that admission control 'needs to be based on the current network loading, the new connection's anticipated characteristics and its delay/loss performance requirements.' Sallberg [Sallberg90] emphasizes the need to maintain the quality of service requirements for existing connections when accepting a new one.

The benefits of using statistical multiplexing to improve the bandwidth utilisation when a large proportion of the traffic is bursty, has been recognised by Woodruff. It is stressed that the tradeoff between improved utilisation and the increased complexity in traffic control must be carefully considered. The use of peak rate bandwidth allocations are suggested

²Unless otherwise stated the descriptions are from [Woodruff88, Woodruff90].

whenever rigid performance guarantees are to be met. The use of multiple traffic classes each with its own guaranteed quality of service requirement is discouraged as adding to the complexity and over specifying the control for particular traffic types.

Bandwidth monitoring and bandwidth enforcement principles for ATM networks were presented in Section 4.5 of Chapter 4. Sallberg distinguishes bandwidth enforcement into flow enforcement and flow throttling. Flow enforcement is implemented by the network and consists of simple packet discarding, while flow throttling (or traffic shaping as referred to by Woodruff) is self-imposed rate control by the traffic source, to optimise its own throughput.

5.2 Performance Modelling of Integrated Services Networks

Interest in integration began with developments in hybrid networks, using the two conventional traffic types — packet switched data and circuit switched voice. Later the emphasis shifted to packet switched networks with a single fast channel handling statistically multiplexed traffic from all sources, because of the inherent flexibility of the packet transfer mode. This advantage, combined with advances in optical and semiconductor technologies, have led to the current concept of broadband integration based on fast packet switching and ATM. The performance modelling of integrated services networks has been the focus of much research and some of the relevant work will be discussed in the rest of this chapter.

5.2.1 Performance Modelling of Hybrid Voice-Data Integration Schemes

The performance modelling and analysis of hybrid networks, especially voice-data networks, has been extensively studied in the literature. Reviews of the field are contained in [Saito90b, Gruber81, Chen88]. Most of the work has been concentrated on modelling voice-data integration over a fixed frame TDM link. The commonly used performance objectives were to maximise the bandwidth utilisation while satisfying constraints on the call blocking for voice traffic and the average queueing delay experienced by data traffic.

In hybrid multiplexing schemes with static bandwidth allocation, circuit switched voice calls and packet switched data share the time slots of the underlying TDM frame in a fixed fashion. In contrast to this fixed boundary approach, Kummerle [Kummerle74] was the first to propose the use of a more dynamic multiplexer in which data packets could use idle voice slots arising from on-off variations in voice calls, a scheme which was shown to provide more efficient channel utilisation. A *movable boundary* scheme in which both on-off and talkspurt-silence variations (with speech activity detection) in the voice process were exploited for the transfer of data packets was modelled by Sriram [Sriram83].

In movable boundary schemes, the performance of the circuit switched voice traffic can be studied separately as a loss system and characterised by the Erlang B formula [Cooper81]. From simulations, Weinstein [Weinstein80] observed that the number of voice calls in

progress was highly correlated in time and could lead to excessive data queues in a movable boundary multiplexer. The correlation between voice calls makes the performance of data traffic difficult to analyse and has made it the subject of much research (listed in [Saito90b]).

Performance Analysis Techniques

The performance evaluation techniques used in the analysis of hybrid voice-data integration schemes includes simulations, 'exact' analysis and approximate analysis [Williams84, Li85]. In exact analysis, the voice and data states are generally represented by a two dimensional Continuous Time Markov Chain (CTMC) model, which is solved using moment generating functions [Kleinrock75]. However, the common problem with this method was one of deriving the roots of the denominator of the generating function to solve the model. In 'approximate' analysis, simplifications are made at the outset — for example with the quasi-static approximation based on the relative time scales of voice calls and data traffic; the data process is assumed to settle into the steady state as the voice process remains unchanged for a relatively longer time.

CTMC models for a hybrid multiplexer assume exponential distributions for the service times (packet lengths) and neglect the discrete slot structure of the underlying TDM frame, resulting in over-estimates for the average data queueing delays. Discrete time models for the multiplexer [Sriram83], improve on these limitations by assuming general distributions for the packet lengths. However they still over-estimate the data queueing delay because of the assumption on 'gating' — a packet that arrives during a frame has to wait until the beginning of the next frame for service, even though there may be idle slots in the current one. Recently Chang [Chang88] presented a discrete time queueing analysis of an integrated voice-data system using a movable boundary scheme. In this model the gating assumption is relaxed for data packets which can be served in idle slots in the current frame.

In this dissertation, the models for Unison ramp queueing are represented as two or three dimensional CTMCs (Chapters 6 and 7) with exponential distributions assumed for the packet service times. Consequently, the average queueing delays will be over-estimated.

Flow Control in Hybrid Networks

In hybrid voice-data networks, particularly those of the movable boundary kind, flow control is necessary because data traffic can experience congestion due to the long holding times and the correlation of voice calls.

A method for rate control of a voice-data multiplexer was proposed in [Williams84], where the data arrival rate is adjusted with each voice call connect and disconnect and is implemented by selective throttling among a number of prioritised data classes. Gafni [Gafni84]

presents a distributed iterative algorithm for voice and data networks that can dynamically adjust the input rate of virtual circuits and ensure fairness in sharing the link bandwidth between them. Adjustment is based on feedback information that is periodically received at the network access point, about the level of congestion along a circuit.

Network flow control as part of a unified control strategy that includes bandwidth management and routing was presented in [Gerla84b, Gerla84a] for a hybrid network with circuit and packet switched traffic. The scheme was based on the use of a bandwidth routing algorithm [Gerla81, Gerla83], which computes for each node pair in the network two sets of paths; one ranked in order of increasing delay while in the second set they were ranked in order of increasing residual bandwidth. The second set was then used to formulate three separate procedures for the flow control of datagrams, virtual circuit and circuit switched traffic [Gerla83].

The flow of datagrams into the network is regulated on the bandwidth availability to a given destination being above a certain threshold. Within the network, datagrams can be further controlled by conventional mechanisms like buffer management. Virtual circuit calls are controlled at the network access point, by establishing the circuit if the average bandwidth requested on it can be assigned on intermediate links. If a request is rejected, the user has the option to resubmit the request with a reduced bit rate. Admission control for circuit switched traffic would be made at call set up time on the basis of peak rate requirements. Traffic rates would be metered by the network controller and compliance methods effected.

5.2.2 Performance Modelling of Integrated Packet Networks

The flexibility in bandwidth allocation and the benefits of statistical multiplexing stimulated considerable interest in the use of the packet transfer mode for integrated networks. One of the early proposals for a packet based integrated network was by Forgie [Forgie77] which was discussed in Section 5.1 of this chapter.

The performance modelling of packet multiplexers in integrated voice-data networks has concentrated mainly on the characterisation of the aggregate voice arrival process. For a single talkspurt voice source, packet arrivals are not Poisson, because packets are generated at deterministic intervals during the talkspurt phase. Even when packets from several voice sources are multiplexed together, it has been shown [Sriram86] that large correlation exists in the packet arrivals that manifests itself in periods of burstiness. Sriram analysed packet multiplexers fed by both talkspurt voice and Poisson data streams and found that the voice correlation effects were important in determining the performance of the multiplexer, especially under heavy loads where the Poisson approximation was inadequate. This was also pointed out earlier by Jenq [Jenq84]. Recently, Sriram [Sriram89] considered the use of bit dropping schemes to smooth the superposed voice arrival process and showed that a Poisson approximation was then suitable for use in a packet voice multiplexer with bit dropping. The multiplexer was modelled as an $M/\bar{D}/1/K$ system

in which the service rate was deterministic, but state dependent in order to model bit dropping.

Heffes [Heffes86] presented a model of a voice-data packet multiplexer in which the aggregate voice stream from talkspurt voice sources was approximated by a Markov Modulated Poisson Process (MMPP) (described in detail in Section 7.2 in Chapter 7); by matching the characteristics of the superposed voice stream with those of the MMPP. Matrix geometric methods [Neuts81] were used to solve the resultant MMPP/G/1 queueing system. Since the superposition of a Poisson process and an MMPP is again an MMPP [Heffes86], the system could model voice and data streams incident on the multiplexer. The model with MMPP arrivals was shown to correspond to simulations, unlike a model with Poisson arrivals which displayed deviations at large loads.

Three models for packet voice multiplexers with incident talkspurt voice streams were presented in [Daigle86]. These are a Semi-Markov process model, a CTMC model and a uniform arrival and service model based on a fluid flow approximation.

In the CTMC model proposed by Daigle, each voice source varied between talkspurt and silence phases with exponential distributions assumed for the duration of each phase. However, during an active phase, unlike the other models in which the packets were generated deterministically, the CTMC model assumed Poisson arrivals during the phase. An exponential distribution was assumed for the packet service times. It was expected [Daigle86] that these assumptions would lead to conservative measures for performance metrics predicted by the model like the average waiting times. This was indeed substantiated in comparisons of the theoretical results with simulations, where it was also noted that the over-estimate increased with the number of voice sources.

5.2.3 Performance Modelling of ATM Networks

There has been considerable interest recently in the performance modelling and analysis of various aspects of ATM networks. A short review and bibliography of research in the field can be found in [Kawashima89].

Most of the research has been on modelling the statistical multiplexing of bursty traffic streams onto a broadband channel. Other work has been on analysing admission control schemes and on bandwidth allocation strategies for ATM networks.

Statistical Multiplexing

The statistical multiplexing of bursty traffic streams at an ATM multiplexer has been considered by several authors [Maglaris87, Kawarazaki90, Dittmann88, Hirano89, Decina90a]. The work by [Sriram86, Daigle86, Heffes86], which was described in the last section also falls into this category. In [Maglaris87] the statistical multiplexing of bursty video sources was studied, using fluid flow analysis techniques and the fitting

of experimental data to a continuous time Markov process to model the video streams. In [Kawarazaki90], MMPP arrival models were used for both packetised voice and video streams. Kawarazaki exploited the additive property of MMPPs [Hellstern89] in modelling the statistical multiplexing of the voice and video streams as an MMPP/ E_L /1/K queueing system, where E_L is the L^{th} order Erlangian distribution [Kleinrock75]. The performance analysis of the system emphasized the sensitivity of bandwidth utilisation to the peak bit rate of the multiplexed video sources.

The use of input queues with limited capacities has been emphasized in performance models of ATM multiplexers as being necessary for limiting the maximum queueing delay at the multiplexer but as resulting in packet loss during congestion. In [Dittmann88], a uniform arrival and service model was used to model the multiplexing of homogeneous bursty sources in an ATM multiplexer with a finite input buffer. The performance study quantified the variation of packet loss probability with the burstiness parameters. Hirano [Hirano89] modelled the statistical multiplexing of bursty traffic through a packet multiplexer and used a discrete time analysis to quantify the relation between the packet loss rate and burstiness. Hirano observed that to maintain a given packet loss rate for a fixed average bit rate of the bursty traffic stream, the size of the buffer required is more sensitive to variations in the peak rate than to the *continuance* or duration of the burst. This was also observed by Decina [Decina90a] in simulations of a finite buffer multiplexer with bursty input streams.

Verbiest and Joos [Verbiest88a, Verbiest88b, Joos89] have considered the use of variable bit rate coding for video sources, as this flexibility is offered by ATM. In the statistical multiplexing of variable bit rate video streams onto a channel with a queue of limited capacity, the authors noted the multiplexing gains that could be achieved even when multiplexing a moderate number of video sources. A gain of 2 was possible when 16 video sources were multiplexed together.³

Admission Control

Admission control schemes based on maintaining packet loss probability standards are analysed in [Saito90a]. These schemes work without actual traffic monitoring and are based on traffic descriptors specified in the connection request — the average and peak bit rates or the average and variance of bit rate. A connection request is accepted, if an admission function that estimates the system's packet loss probability with the new connection's traffic descriptors, is comparable with a loss probability standard. In [Kamitake89], an admission control scheme based on a different metric, the instantaneous packet loss rate, is proposed. A discrete time analysis is used for calculating the instantaneous packet loss rate in a finite buffer system.

³The gain was defined the ratio of peak bit rate of a source to the bandwidth actually allocated for it. This allocation was such that a packet loss probability standard of 10^{-8} was satisfied [Verbiest88b].

Bandwidth Allocation

A conservative method of allocating bandwidth to a traffic source that offers performance guarantees in an ATM network is to use the peak bit rate. However, with the gains possible in statistical multiplexing, allocating less bandwidth may be sufficient. The definition of a virtual (equivalent) bandwidth measure that lies between the peak and average traffic rates of a bursty source has been recognised as an open problem [Woodruff88].

Decina, Toniatti and Gallassi [Decina90a, Decina90b, Gallassi89] have considered how the equivalent bandwidth assigned to traffic sources varies with their burstiness. The maximum number of bursty sources that could be multiplexed together while satisfying constraints on the packet loss probability and packet delay was obtained from simulations. The equivalent bandwidth was then obtained by dividing the link bandwidth by this number. The authors observed that for traffic with large peak rate to link bandwidth ratios (highly bursty), the equivalent bandwidth approaches the peak bit rate for each source. Under these conditions, using a bandwidth allocation that assigns less than the peak rate (i.e. trying to use statistical multiplexing) would degrade the system performance. For low values of the ratio, the equivalent bandwidth is relatively insensitive to the average burst length, which in general would be larger than the buffer sizes, but depends more on the peak bit rate to average bit rate ratio and substantial statistical multiplexing gains are to be had if this ratio is high. It is only when the burst and buffer lengths are comparable that the burst length has a large effect on the equivalent bandwidth measure.

Verbiest and Joos [Verbiest88a, Verbiest88b, Joos89] considered a static bandwidth allocation procedure for variable rate encoded video sources. This scheme was based on the principle that when a sufficient number of video streams were multiplexed together, the probability density function of the multiplex could be approximated by the *normal* distribution. Bandwidth allocation and call acceptance were based on the statistics of the standard normal distribution, the link bandwidth and a packet loss rate standard.

Other work

Multi-level congestion and resource allocation strategies for ATM networks have been proposed and analysed by [Hui88, Filipak89]. Hui suggests that congestion measures based on blocking probabilities should be evaluated at three levels — the packet, burst and call levels, which can then be used in a multi-level bandwidth allocation algorithm. A similar consideration of congestion and identification of resource allocation schemes at the packet and call levels was presented in [Saito89, Kawashima89]. In this work, Saito and Kawashima distinguish call level and packet level dimensioning. Call level dimensioning is similar to that in STM networks and gives the maximum number of virtual circuits that can be accommodated while satisfying a specified call blocking probability constraint. Packet level dimensioning considers the assignment of bandwidth to these circuits and the buffer size necessary to satisfy packet level grade of service metrics i.e. the packet loss probability and maximum packet delay.

The performance modelling of fast packet switches is well described in [Hui90]. An exposition of protocol requirements for ATM networks that are important for the design of congestion control schemes is presented in [Rider89].

5.3 Summary

In this chapter some approaches to congestion control for networks with ATM transfer were described. Proposals for congestion control commonly consist of call admission control and bandwidth monitoring and enforcement based on rate control methods implemented at the network periphery. An open problem when seeking to exploit the advantages of statistical multiplexing was seen in defining a virtual bandwidth for bursty sources that is between the average and peak bit rates.

The chapter presented a brief review of relevant developments in the performance modelling of integrated services networks — such research has been largely concentrated on the analysis of the multiplexing discipline and the characterisation of the traffic arrival processes.

Chapter 6

Performance Models 1

This chapter and the next describe performance models for a variable bandwidth U channel controlled by a simple queue-based bandwidth algorithm. The algorithms are driven by samples of the instantaneous state of the queue which is associated with the channel. The development of the models is done in an incremental fashion through the inclusion of factors which characterise the control — the bandwidth switching latencies and the sampling rates. The arrival processes which are used to model the traffic flow into the channels are a Poisson arrival process and the MMPP. The main objectives in this study were the formulation of queueing models for dynamic bandwidth control and their steady state analyses to obtain performance metrics and stability conditions.

6.1 Introduction

The models in this dissertation emulate the control of a packet multiplexer connected to a variable bandwidth channel at a Unison ramp. The channel bandwidth is adjusted by a remote controller (channel service) on the basis of feedback received about the state of the ramp queues. A simple queue-based algorithm with hysteresis control across a threshold pair is assumed to be implemented on the controller. If the instantaneous queue length in the feedback sample exceeds a switching-in threshold then the controller decides to increase the bandwidth. This decision is transmitted to the ramp which initiates the bandwidth adjustment that is completed after the bandwidth switching-in latency. When the queue length falls below a switching-out threshold the channel bandwidth is decreased in a similar fashion.

The various delays that are experienced in the bandwidth adjustment process have been described in Section 4.4 of Chapter 4 as sampling, transmission, control and bandwidth switching delays. In practice, the bandwidth switching and sampling delays are much

larger than the transmission and control delays (at least an order of magnitude more). In the current models the transmission and control delays will not be modelled explicitly but are assumed to be included in the sampling delay. Bandwidth switching delays are modelled separately.

The Models

The models in this chapter emulate a variable bandwidth channel controlled by a simple queue-based algorithm. It is assumed that the controller has *perfect knowledge* of the state of the queue at every instant in time and sampling and transmission delays are not modelled. However, the system takes a finite time to adjust the bandwidth (increase or decrease) which is the bandwidth switching latency.

The variable bandwidth channel is modelled as a single server queue with a variable service rate. The service rate fluctuates between two values giving rise to a two-phase switching process. The state of the system is the number of packets that are queued for service and the switching phase the system is in. In the models in this chapter, input packets arrive according to a Poisson arrival process. The average arrival rate is chosen to be greater than the lower of the service rates so that there is a need to increase the bandwidth. The rationale behind this choice is to model a simple system in which bandwidth is increased in response to an increase in traffic load. In the next chapter the variation of service rate with a more complex arrival process – the MMPP will be considered.

Bandwidth switching is state dependent to model the queue-based hysteresis control. When the channel is functioning with the lower service rate the queue begins to build up as the system is running in a state of overload. When the queue length exceeds an upper threshold, the channel switches to the higher service rate after a finite time which is the bandwidth switching-in latency. The second service rate is assumed to be greater than the average arrival rate, so that the system enters a period of underload and the queue decays. When the queue length falls below a lower threshold, bandwidth decrease is initiated. The channel resumes operation at the lower service rate after the bandwidth switching-out latency. The cycle then repeats — this is the oscillatory behaviour characteristic of a queue-based algorithm.¹ The choice of simple queue-based control in the modelling was dictated by reasons of analytical tractability as were the choice of the arrival processes and service disciplines used in this work.

This chapter presents two models for the control of a variable bandwidth channel. The first model (Model A) assumes that bandwidth increase is delayed but that bandwidth decrease takes place instantaneously. In the second model (Model B), both the switching in and the switching out of bandwidth are delayed.

¹Oscillations between the two service rates takes place even when the traffic rate remains unchanged.

6.2 Modelling Assumptions

The queueing system is modelled as a continuous time Markov chain (CTMC) whose state transitions are governed by the hysteresis control and the bandwidth switching latencies. The models neglect the framing effects observed in practice (Section 3.3 of Chapter 3) at the ramp. The packet service times in switching phase $i, i = \{0, 1\}$ are assumed to be exponentially distributed, independently of each other with mean values μ_i^{-1} secs and without any loss of generality it may be assumed that $\mu_1 > \mu_0$.

The bandwidth switching-in and switching-out latencies are also assumed to be exponentially distributed with average values γ_{in}^{-1} and γ_{out}^{-1} secs. The bandwidth switching latencies observed in practice are random variables with small variances. This is because of the variations in call signalling and channel synchronisation times. Because of the transparent nature of bandwidth adjustment, there are no service interruptions when the service rate is changed. A packet that is being serviced at one rate, would continue to be serviced at the second rate if switching occurs during the service.

Packet arrivals are assumed to follow a Poisson process with rate λ where it is assumed that $\lambda > \mu_0$ but $\lambda < \mu_1$. This is the necessary and sufficient condition for stability; i.e. the average arrival rate should be less than at least one of the service rates in the switching process. The input queues are assumed to be infinitely large.²

6.3 Solution Techniques

The steady state queueing analyses of the performance models in the current work are by the use of Moment Generating Functions (MGFs) [Kleinrock75, Schwartz87] computed from the steady state balance equations. General solution trends can be abstracted for the models and are described below.

The formulation of the MGF, generally involves the formulation of partial MGFs one for each switching (and arrival) phase of the system. The expressions for these partial functions contain unknown boundary (zero-state) probabilities which need to be solved for by the simultaneous solution of a sufficient number of independent equations. Some of the required equations can be extracted from the original balance ones. The others are obtained from a consideration of the properties of the MGF — its limiting value at unity and its *analytic* properties.

The MGF is defined as $\Pi(z) = \sum_{i=0}^{\infty} z^i P_i$ where P_i is the steady state probability of the queue length being i . The MGF when evaluated at $z = 1$ is the sum of the steady state probabilities $\sum_{i=0}^{\infty} P_i$. This sum should be unity provided the steady state exists.³ As shall

²Although in practice, the input queues at the ramps can accommodate only 20000 packets, for small values of the arrival and service parameters this assumption is valid.

³This is referred to as the *Conservation of Probability Sum* in later discussions.

be seen in the analyses, this equation is used to derive the necessary and sufficient conditions for stability, as is done for example in an $M/M/1$ queueing system [Kleinrock75] where the utilisation has to be less than one for the infinite sum of the steady state probabilities to converge.

However, in the evaluation of the MGF at unity, a frequently occurring problem is that the MGF is usually indeterminate because it contains $(z - 1)$ as a factor. It is then necessary to extract this root from both the numerator and denominator or to apply L'Hopital's rule [Kleinrock75]. The other required equations are derived from the analytic properties of the MGF. This method, exemplified in [Schwartz87], uses the property that a generating function is analytic in the unit circle, so that a real root of its denominator that is less than one will also be a root of its numerator.

Having computed the unknowns, various performance measures of the queue can then be calculated from an application of the limiting properties of the MGF. For example, the mean queue length is given by the expression :

$$E[X] = \lim_{z \rightarrow 1} [\Pi'(z)] \quad (6.1)$$

The mean queueing delay can be calculated from the mean queue length by an application of Little's Law [Kleinrock75]; for a Poisson arrival process with rate λ and mean queue length $E[X]$, the mean queueing delay is given by $E[X]/\lambda$.

6.4 Model A : Hysteresis Control, Poisson Arrivals and Delayed Bandwidth Change (In)

- Queue-based algorithm with hysteresis control
- Poisson arrival process
- Exponentially distributed service times
- Exponentially distributed bandwidth switching-in times
- Instantaneous bandwidth decrease

The first model in this chapter is for a variable bandwidth channel, where an increase in bandwidth (μ_0 to μ_1) is delayed but a decrease in the bandwidth is assumed to take place instantaneously. A simple queue-based algorithm with hysteresis control across a single threshold pair is modelled. Packet arrivals occur according to a Poisson process. The bandwidth switching-in time is assumed to be exponentially distributed with mean γ^{-1} secs.

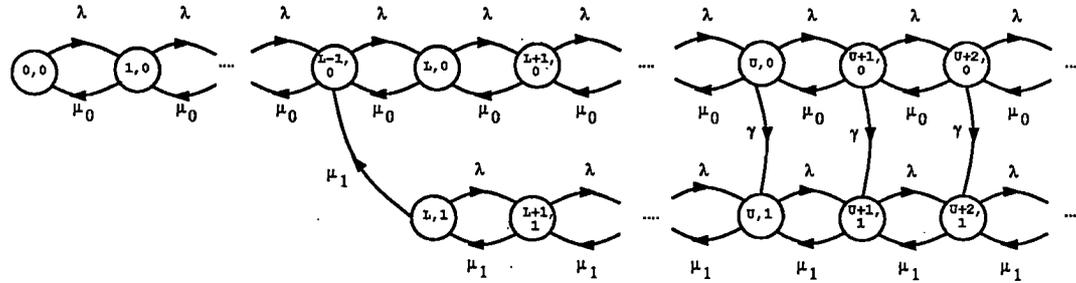


Figure 6.1: Model A : Queue-Based Hysteresis Control — Poisson Arrivals and Delayed Bandwidth Change (In)

Model Description

The state space of the system is defined by the two dimensional vector $[\{X(t), Y(t)\}]$, $X(t) \geq 0$, $Y(t) = \{0, 1\}$, $t \geq 0$, where $X(t)$ is the queue length at time t , and $Y(t)$ is the switching phase. The system is modelled as a CTMC with state transitions and transition rates detailed in Figure 6.1. The upper and lower layers in the diagram correspond to the switching phases 0 and 1 respectively.

When the system state exceeds the upper threshold state $(U, 0)$, the bandwidth switching-in process is initiated and is completed after an exponentially distributed delay. This is represented by the set of transitions from phase 0 to phase 1, with rate γ . An assumption made here is that the bandwidth is only switched in when the state of the queue is greater than or equal to the switching-in threshold. The switching out of the bandwidth is instantaneous and is represented by the transition from state $(L, 1)$ in phase 1 to state $(L - 1, 0)$ in phase 0.

Let $\Pi = \{\Pi_0, \Pi_1\}$ represent the stationary distribution for the process $[X(t), Y(t)]$, where $\Pi_0 = \{P_{0,0}, P_{1,0} \dots\}$ and $\Pi_1 = \{P_{L,1}, P_{L+1,1} \dots\}$. $P_{j,i}$ is the probability of having j packets in the queue when the system is in switching phase i . The steady state analysis for this model is obtained by using moment generating functions and will be presented in full to emphasise the general solution procedure that was described earlier.

In the steady state, the balance equations written down from the state transition diagram are :

Phase 0 :

$$\lambda P_{0,0} = \mu_0 P_{1,0} \quad (6.2)$$

$$(\lambda + \mu_0) P_{i,0} = \lambda P_{i-1,0} + \mu_0 P_{i+1,0} \quad (6.3)$$

$$i = 1, 2, \dots, L-2$$

$$(\lambda + \mu_0) P_{L-1,0} = \lambda P_{L-2,0} + \mu_0 P_{L,0} + \mu_1 P_{L,1} \quad (6.4)$$

$$(\lambda + \mu_0) P_{i,0} = \lambda P_{i-1,0} + \mu_0 P_{i+1,0} \quad (6.5)$$

$$i = L, L+1, \dots, U-1$$

$$(\lambda + \mu_0 + \gamma) P_{i,0} = \lambda P_{i-1,0} + \mu_0 P_{i+1,0} \quad (6.6)$$

$$i = U, U+1, \dots$$

Phase 1 :

$$(\lambda + \mu_1) P_{L,1} = \mu_1 P_{L+1,1} \quad (6.7)$$

$$(\lambda + \mu_1) P_{i,1} = \lambda P_{i-1,1} + \mu_1 P_{i+1,1} \quad (6.8)$$

$$i = L+1, \dots, U-1$$

$$(\lambda + \mu_1) P_{i,1} = \lambda P_{i-1,1} + \mu_1 P_{i+1,1} + \gamma P_{i,0} \quad (6.9)$$

$$i = U, U+1, \dots$$

Solution Procedure

Let the partial MGFs $\Pi_i(z)$, $i = 0, 1$ and the MGF, $\Pi(z)$ be defined as :

$$\Pi_0(z) \stackrel{\text{def}}{=} \sum_{i=0}^{\infty} z^i P_{i,0}$$

$$\Pi_1(z) \stackrel{\text{def}}{=} \sum_{i=L}^{\infty} z^i P_{i,1}$$

$$\Pi(z) \stackrel{\text{def}}{=} \Pi_0(z) + \Pi_1(z).$$

Evaluating the Partial MGFs :

Phase 0 :

The partial MGF $\Pi_0(z)$ at phase 0 is formed from the set of balance equations 6.2 through to 6.6 as follows :

Multiplying the i^{th} equation in the set of equations 6.6, by z^i and summing over relevant i , ($i = U, U+1, \dots$) results in :

$$(\lambda + \mu_0 + \gamma) \sum_{i=U}^{\infty} z^i P_{i,0} = \lambda \sum_{i=U}^{\infty} z^i P_{i-1,0} + \mu_0 \sum_{i=U}^{\infty} z^i P_{i+1,0}$$

which can be rewritten after extracting the term $\Pi_0(z)$ as :

$$\begin{aligned}
 (\lambda + \mu_0 + \gamma)[\Pi_0(z) - \sum_{i=0}^{U-1} z^i P_{i,0}] &= \lambda z[\Pi_0(z) - \sum_{i=0}^{U-2} z^i P_{i,0}] \\
 &+ \frac{\mu_0}{z}[\Pi_0(z) - \sum_{i=0}^U z^i P_{i,0}] \quad (6.10)
 \end{aligned}$$

Now consider the set of equations 6.5. Multiplying the i^{th} equation in this set by z^i and summing over relevant i , ($i = L, L+1, \dots, U-1$) yields :

$$(\lambda + \mu_0) \sum_{i=L}^{U-1} z^i P_{i,0} = \lambda z \sum_{i=L-1}^{U-2} z^i P_{i,0} + \frac{\mu_0}{z} \sum_{i=L+1}^U z^i P_{i,0} \quad (6.11)$$

Next consider the balance equation 6.4. Multiplying both sides by z^{L-1} gives :

$$(\lambda + \mu_0) z^{L-1} P_{L-1} = \lambda z z^{L-2} P_{L-2,0} + \frac{\mu_0}{z} z^L P_{L,0} + \mu_1 z^{L-1} P_{L,1} \quad (6.12)$$

Multiplying the i^{th} equation in the set of equations 6.3 by z^i and summing over i , ($i = 1, 2, \dots, L-2$) yields :

$$(\lambda + \mu_0) \sum_{i=1}^{L-2} z^i P_{i,0} = \lambda z \sum_{i=0}^{L-3} z^i P_{i,0} + \frac{\mu_0}{z} \sum_{i=2}^{L-1} z^i P_{i,0} \quad (6.13)$$

Finally, equation 6.2 can be rewritten as :

$$(\lambda + \mu_0) P_{0,0} = \frac{\mu_0}{z} z P_{1,0} + \mu_0 P_{0,0} \quad (6.14)$$

Adding equations 6.11, 6.12, 6.13 and 6.14 results in a new equation :

$$\begin{aligned}
 (\lambda + \mu_0) \sum_{i=0}^{U-1} z^i P_{i,0} &= \lambda z \sum_{i=0}^{U-2} z^i P_{i,0} + \frac{\mu_0}{z} \sum_{i=1}^U z^i P_{i,0} \\
 &+ \mu_1 z^{L-1} P_{L,1} + \mu_0 P_{0,0} \quad (6.15)
 \end{aligned}$$

Adding equation 6.10 to equation 6.15 results in :

$$\begin{aligned} \Pi_0(z)[(\lambda + \mu_0 + \gamma) - \lambda z - \frac{\mu_0}{z}] &= \gamma \sum_{i=0}^{U-1} z^i P_{i,0} - \frac{\mu_0}{z} P_{0,0} \\ &+ \mu_0 P_{0,0} + \mu_1 z^{L-1} P_{L,1} \end{aligned}$$

which can be rewritten to give the expression for the partial MGF at phase 0 :

$$\Pi_0(z) = \frac{\mu_0(1-z)P_{0,0} - \mu_1 z^L P_{L,1} - \gamma z \sum_{i=0}^{U-1} z^i P_{i,0}}{[\lambda z^2 - (\lambda + \mu_0 + \gamma)z + \mu_0]} \quad (6.16)$$

Phase 1 :

In a similar fashion by considering the balance equations 6.7 to 6.9 and summing them over relevant i , the partial MGF for phase 1 is given by :

$$\Pi_1(z) = \frac{\mu_1 z^L P_{L,1} + \gamma z \sum_{i=0}^{U-1} z^i P_{i,0} - \gamma z \Pi_0(z)}{[\lambda z^2 - (\lambda + \mu_1)z + \mu_1]} \quad (6.17)$$

or equivalently by substituting the expression for $\Pi_0(z)$ from equation 6.16 :

$$\begin{aligned} \Pi_1(z) &= \\ &\frac{\{[\mu_1 z^L P_{L,1} + \gamma z \sum_{i=0}^{U-1} z^i P_{i,0}][\lambda z^2 - (\lambda + \mu_0)z + \mu_0] \\ &- \gamma z \mu_0(1-z)P_{0,0}\}}{[\lambda z^2 - (\lambda + \mu_0 + \gamma)z + \mu_0][\lambda z^2 - (\lambda + \mu_1)z + \mu_1]} \quad (6.18) \end{aligned}$$

Evaluating the Unknowns :

The partial MGFs expressed in equations 6.16 and 6.18, involve $U + 1$ unknown steady state probabilities : $P_{i,0}$, $i = 0, 1, \dots, U - 1$ and $P_{L,1}$. From the balance equations 6.2 through to 6.9 at phase 0 and phase 1, can be extracted $U - 1$ independent linear equations which would enable $U - 1$ of these probabilities to be expressed in terms of the boundary or zero-state probabilities — $P_{0,0}$ and $P_{L,1}$. The remaining two equations required to solve for these boundary values, are obtained from the value of the MGF $\Pi(z)$ at $z = 1$ and from the analytic properties of the partial MGF $\Pi_1(z)$.

Balance Equations :

From the balance equations, $\sum_{i=0}^{U-1} P_{i,0}$ can be expressed in terms of $P_{0,0}$ and $P_{L,1}$ as follows where $\rho_0 = \lambda/\mu_0$ the traffic intensity in the first switching phase (details in Appendix A)

$$\sum_{i=0}^{U-1} P_{i,0} = P_{0,0} \left\{ \frac{(1-\rho_0^U)}{(1-\rho_0)} \right\} - P_{L,1} \left\{ \frac{\mu_1}{\mu_0(1-\rho_0)^2} [(U-L)(1-\rho_0) - \rho_0(1-\rho_0^{U-L})] \right\} \quad (6.19)$$

Conservation of Probability Sum :

The equation derived from the conservation of probability sum property is $\Pi(1) = \Pi_0(1) + \Pi_1(1) = 1$. Now the value of the partial MGF for phase 0 at $z = 1$ is :

$$\Pi_0(1) = \frac{\mu_1}{\gamma} P_{L,1} + \sum_{i=0}^{U-1} P_{i,0}$$

and that of the partial MGF at phase 1 is indeterminate.

Resolving the Indeterminacy :

The indeterminate form of $\Pi_1(1)$ can be resolved by factoring the term $(z-1)$ from both the numerator and denominator of $\Pi_1(z)$ in equation 6.18 which is then expressed as :

$$\begin{aligned} \Pi_1(z) &= \frac{\{[\mu_1 z^L P_{L,1} + \gamma z \sum_{i=0}^{U-1} z^i P_{i,0}][\lambda z - \mu_0] + \gamma z \mu_0 P_{0,0}\}}{[\lambda z - \mu_1][\lambda z^2 - (\lambda + \mu_0 + \gamma)z + \mu_0]} \end{aligned} \quad (6.20)$$

The value of $\Pi_1(1)$ is then :

$$\Pi_1(1) = \frac{\mu_1 P_{L,1}(\lambda - \mu_0) + \gamma \sum_{i=0}^{U-1} P_{i,0}(\lambda - \mu_0) + \gamma \mu_0 P_{0,0}}{\gamma(\mu_1 - \lambda)}$$

$\Pi(1)$ is therefore :

$$\begin{aligned} \Pi(1) &= \frac{\mu_1(\mu_1 - \mu_0)}{\gamma(\mu_1 - \lambda)} P_{L,1} + \frac{(\mu_1 - \mu_0)}{(\mu_1 - \lambda)} \sum_{i=1}^{U-1} P_{i,0} + \frac{\mu_1}{(\mu_1 - \lambda)} P_{0,0} \\ &= 1 \end{aligned} \quad (6.21)$$

Substituting the expression for $\sum_{i=0}^{U-1} P_{i,0}$ from equation 6.19 in the above equation gives the first of the equations involving the two unknowns $P_{0,0}$ and $P_{L,1}$.

Stability Condition :

The condition for stability, that is the condition for the steady state probability distribution to exist is derived from equation 6.21. Since the steady state probabilities are positive, and since by initial assumption, $\mu_1 > \mu_0$, it is necessary that $\mu_1 > \lambda$. That is the necessary condition for stability is that the average arrival rate be less than at least one of the service rates. This is also a sufficient condition.

Analytic Properties of the Partial MGF — $\Pi_1(z)$:

The second equation necessary to solve for the unknowns $P_{0,0}$ and $P_{L,1}$ is obtained by the analytic properties of the function $\Pi_1(z)$. which can be expressed as $\Pi_1(z) \stackrel{\text{def}}{=} N_1(z)/D_1(z)$. From equation 6.20, $N_1(z)$ and $D_1(z)$ are given by :

$$N_1(z) = \{\mu_1 z^L P_{L,1} + \gamma z \sum_{i=0}^{U-1} z^i P_{i,0}\}[\lambda z - \mu_0] + \gamma z \mu_0 P_{0,0}$$

$$D_1(z) = [\lambda z^2 - (\lambda + \mu_0 + \gamma)z + \mu_0](\lambda z - \mu_1)$$

Considering the function $D_1(z)$, it is evident that $D_1(0) = -\mu_0\mu_1 < 0$ and $D_1(1) = -\gamma(\lambda - \mu_1) > 0$, since μ_0, μ_1 are positive quantities and $\mu_1 > \lambda$ from the stability condition. Therefore, there are one or three roots of $D_1(z)$ in the interval $[0, 1]$. Now consider $z = \frac{\mu_1}{\lambda} > 1$. As $D_1(\frac{\mu_1}{\lambda}) = 0$ this implies that there is a root of $D_1(z)$ at $\frac{\mu_1}{\lambda} > 1$. Hence the number of roots in the interval $[0, 1]$ equals one, i.e. there is a unique root in $[0, 1]$. Let this root be denoted as z^* .

Since $\Pi_1(z)$ is analytic within $z < 1$ (by the definition of the MGF, as an infinite series which converges for all $|z| < 1$), a root of its denominator is also a root of its numerator and hence the desired second equation between the two boundary probabilities is :

$$N_1(z^*) = 0$$

or

$$(\lambda z^* - \mu_0)[\mu_1 (z^*)^L P_{L,1} + \gamma z^* \sum_{i=0}^{U-1} (z^*)^i P_{i,0}] + \gamma z^* \mu_0 P_{0,0} = 0 \quad (6.22)$$

where z^* is computed from,

$$z^* = \frac{(\lambda + \mu_0 + \gamma) - \sqrt{(\lambda + \mu_0 + \gamma)^2 - 4\lambda\mu_0}}{2\lambda}$$

Equivalently the analytic properties of the functions $\Pi_0(z)$ or $\Pi(z)$, could have been considered with a similar argument to derive the necessary equation.

Steady State Performance Metrics :

From the two equations 6.21 and 6.22, $P_{0,0}$ and $P_{L,1}$ can be calculated and substituted in the partial MGFs for the two phases. The lack of closed form expressions for the boundary probabilities is to be noted. Various steady state performance measures may then be derived from $\Pi(z)$ in a straightforward fashion. For example, the mean queue size $E[X]$ is given by :

$$E[X] = \lim_{z \rightarrow 1} \Pi'(z)$$

which after substitution and some algebra is given by :

$$\begin{aligned} E[X] = & P_{L,1} \{ \mu_1 (\mu_1 - \mu_0) (\lambda - \mu_0) (\mu_1 - \lambda) \\ & + \lambda \gamma \mu_1 (\mu_1 - \mu_0) \\ & + (\mu_1 - \mu_0) (\mu_1 - \lambda) \mu_1 \gamma [U - 1] \} \\ & + \sum_{i=0}^{U-1} P_{i,0} \{ \gamma (\mu_1 - \mu_0) (\lambda - \mu_0) (\mu_1 - \lambda) + \lambda \gamma^2 (\mu_1 - \mu_0) \} \\ & + \sum_{i=0}^{U-1} i P_{i,0} \{ \gamma^2 (\mu_1 - \mu_0) (\mu_1 - \lambda) \} \\ & + P_{00} \{ \gamma \mu_0 (\mu_1 - \mu_0) (\mu_1 - \lambda) + \lambda \gamma^2 \mu_0 \} \\ & \div \gamma^2 (\lambda - \mu_1)^2 \end{aligned} \quad (6.23)$$

In the limit as γ tends to zero, it can be seen that $E[X] \rightarrow \infty$. In words this means that as the switching latency increases and transitions from phase 0 (overload) to phase 1 (underload) become less frequent, the queue size increases without bound. Further when $\lambda \rightarrow \mu_1$, $E[X] \rightarrow \infty$; an intuitive result which states that as the average arrival rate approaches the average service rate in the second switching phase (λ_1 being $> \mu_0$), the queue again increases without any bound. The variance of queue size $var[X]$ may be computed from the equation :

$$var[X] = \lim_{z \rightarrow 1} \Pi''(z) + E[X] - E^2[X] \quad (6.24)$$

Model A with Point Thresholding

The analysis just presented was for a model with hysteresis control of a variable bandwidth channel. When using point thresholding control, the steady state solution procedure is

similar and can be obtained as a degenerate case of the present model. For example, the average queue size in the point thresholding model can be obtained by setting $L = U = K$, where K is the point threshold, in equation 6.23 and is given by :

$$\begin{aligned}
E[X] = & P_{K,1}\{\mu_1(\mu_1 - \mu_0)(\lambda - \mu_0)(\mu_1 - \lambda) \\
& + \lambda\gamma\mu_1(\mu_1 - \mu_0) \\
& + (\mu_1 - \mu_0)(\mu_1 - \lambda)\mu_1\gamma[K - 1] \\
& + \sum_{i=0}^{K-1} P_{i,0}\{\gamma(\mu_1 - \mu_0)(\lambda - \mu_0)(\mu_1 - \lambda) + \lambda\gamma^2(\mu_1 - \mu_0)\} \\
& + \sum_{i=0}^{K-1} iP_{i,0}\{\gamma^2(\mu_1 - \mu_0)(\mu_1 - \lambda)\} \\
& + P_{00}\{\gamma\mu_0(\mu_1 - \mu_0)(\mu_1 - \lambda) + \lambda\gamma^2\mu_0\} \\
& \div \gamma^2(\lambda - \mu_1)^2
\end{aligned} \tag{6.25}$$

which is identical to the expression obtained from a complete solution of the point thresholding model, starting from the balance equations that was presented in [Harita89].

6.5 Model B : Hysteresis Control, Poisson Arrivals and Delayed Bandwidth Change (In and Out)

- Queue-based algorithm with hysteresis control
- Poisson arrival process
- Exponentially distributed service times
- Exponentially distributed bandwidth switching-in and switching-out times

The second and final model in this chapter is for a variable bandwidth channel where both an increase (μ_0 to μ_1) and decrease (μ_1 to μ_0) in bandwidth are delayed. A simple queue-based algorithm with hysteresis control across a single threshold pair is modelled. Packet arrivals are according to a Poisson process. The bandwidth switching-in and switching-out times are assumed to be exponentially distributed with mean values of γ_{in}^{-1} and γ_{out}^{-1} secs respectively.

Model Description

As in Model A, the state of the system is defined by the two dimensional vector $[\{\mathbf{X}(t), \mathbf{Y}(t)\}]$, $\mathbf{X}(t) \geq 0, \mathbf{Y}(t) = \{0, 1\}, t \geq 0$, where $\mathbf{X}(t)$ is the queue length at time t , and $\mathbf{Y}(t)$ is the

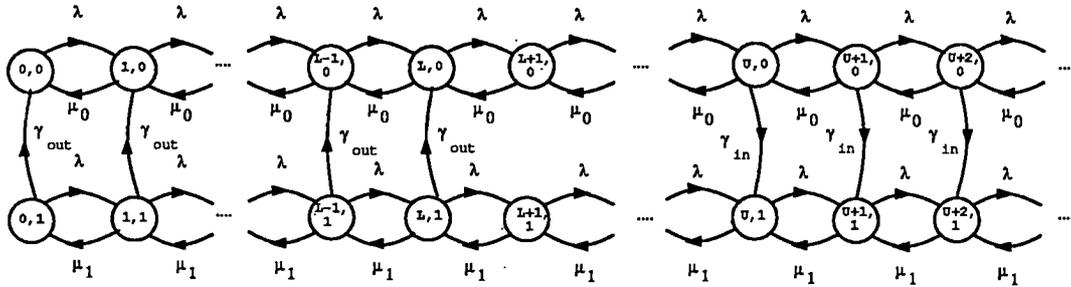


Figure 6.2: Model B : Queue-Based Hysteresis Control — Poisson Arrivals and Delayed Bandwidth Change (In and Out)

switching phase. The system is modelled as a CTMC with state transitions and transition rates detailed in Figure 6.2. The upper and lower layers in the diagram as before correspond to the switching phases 0 and 1. Delay in switching out the bandwidth implies that the system functions with the higher service rate for a longer duration than the previous model. In Model A, as soon as the queue length crossed the lower threshold state, the service rate changed to its lower value.

The switching in latencies are represented by the infinite set of transitions from phase 0 to phase 1, beginning with the state $(U, 0)$ with rates γ_{in} . A finite set of transitions from phase 1 to phase 0 represents the switching-out latencies, starting at the state $(L, 1)$ and ending with the state $(0, 1)$, with transition rates γ_{out} . As in Model A, an assumption made is that the bandwidth is only increased when the queue length is greater than or equal to the switching-in threshold and is conversely decreased only when the queue length is less than or equal to the switching-out threshold.

Let $\Pi = \{\Pi_0, \Pi_1\}$ represent the stationary probability distribution for the process $[X(t), Y(t)]$ where $\Pi_i = \{P_{0,i}, P_{1,i}, \dots\}$, $i = \{0, 1\}$. $P_{j,i}$ being the probability of having j packets in the queue when the switching phase is i . In the steady state, the balance equations are :

Phase 0 :

$$\begin{aligned} \lambda P_{0,0} &= \mu_0 P_{1,0} + \gamma_{out} P_{0,1} \\ (\lambda + \mu_0) P_{i,0} &= \lambda P_{i-1,0} + \mu_0 P_{i+1,0} + \gamma_{out} P_{i,1} \\ & \quad i = 1, 2, \dots, L \end{aligned} \tag{6.26}$$

$$\begin{aligned}
(\lambda + \mu_0)P_{i,0} &= \lambda P_{i-1,0} + \mu_0 P_{i+1,0} \\
&\quad i = L + 1, L + 2, \dots, U - 1 \\
(\lambda + \mu_0 + \gamma_{in})P_{i,0} &= \lambda P_{i-1,0} + \mu_0 P_{i+1,0} \\
&\quad i = U, U + 1, \dots
\end{aligned}$$

Phase 1 :

$$\begin{aligned}
(\lambda + \gamma_{out})P_{0,1} &= \mu_1 P_{1,1} & (6.27) \\
(\lambda + \mu_1 + \gamma_{out})P_{i,1} &= \lambda P_{i-1,1} + \mu_1 P_{i+1,1} \\
&\quad i = 1, 2, \dots, L \\
(\lambda + \mu_1)P_{i,1} &= \lambda P_{i-1,1} + \mu_1 P_{i+1,1} \\
&\quad i = L + 1, L + 2, \dots, U - 1 \\
(\lambda + \mu_1)P_{i,1} &= \lambda P_{i-1,1} + \mu_1 P_{i+1,1} + \gamma_{in} P_{i,0} \\
&\quad i = U, U + 1, \dots
\end{aligned}$$

Solution Procedure :

The solution procedure for evaluating the steady state probability distribution is similar to Model A and only brief details will be given here. The partial MGFs are defined as $\Pi_0(z) = \sum_{i=0}^{\infty} z^i P_{i,0}$ and $\Pi_1(z) = \sum_{i=0}^{\infty} z^i P_{i,1}$.

$\Pi_0(z)$ and $\Pi_1(z)$ are computed as in Model A by summing the relevant balance equation sets 6.26 and 6.27 over appropriate i and adding the resultant equations together. The expressions for the partial MGFs are :

$$\Pi_0(z) = \frac{\mu_0(1-z)P_{0,0} - \gamma_{in}z \sum_{i=0}^{U-1} z^i P_{i,0} - \gamma_{out}z \sum_{i=0}^L z^i P_{i,1}}{[\lambda z^2 - (\lambda + \mu_0 + \gamma_{in})z + \mu_0]} \quad (6.28)$$

$$\Pi_1(z) = \frac{\mu_1(1-z)P_{0,1} + \gamma_{in}z \sum_{i=0}^{U-1} z^i P_{i,0} + \gamma_{out}z \sum_{i=0}^L z^i P_{i,1} - \gamma_{in}z \Pi_0(z)}{[\lambda z^2 - (\lambda + \mu_1)z + \mu_1]} \quad (6.29)$$

The similarity of the above equations to the expressions for the partial MGFs in Model A (equations 6.16 and 6.17 in this chapter) can be observed. The equations 6.28 and 6.29 involve $U + L + 1$ unknowns: $P_{i,0}$, $i = 0, 1, \dots, U - 1$ and $P_{i,1}$, $i = 0, 1, \dots, L$. From the balance equations can be extracted $U + L - 1$ independent linear equations. The two remaining equations are obtained as before from the value of $\Pi(z)$ at unity and by considering the analytic properties of the function $\Pi_1(z)$.

At $z = 1$, $\Pi_1(z)$ is indeterminate as it contains the factor $(z - 1)$. After factoring $(z - 1)$ from its numerator and denominator the resultant value of the MGF $\Pi(z) = \Pi_0(z) + \Pi_1(z)$ at $z = 1$ is :

$$\begin{aligned}
\Pi(1) = & \frac{\mu_1}{\mu_1 - \lambda} P_{0,0} + \frac{\gamma_{out}(\mu_1 - \mu_0) + \gamma_{in}\mu_1}{\gamma_{in}(\mu_1 - \lambda)} P_{0,1} \\
& + \frac{(\mu_1 - \mu_0)}{\gamma_{in}(\mu_1 - \lambda)} \left(\gamma_{in} \sum_{i=1}^{U-1} P_{i,0} + \gamma_{out} \sum_{i=1}^L P_{i,1} \right) \quad (6.30)
\end{aligned}$$

From the above equation 6.30 given the initial assumption that $\mu_1 > \mu_0$ and that the steady state probabilities are positive, the necessary and sufficient condition for stability is $\mu_1 > \lambda$ that is the arrival rate is to be less than the higher of the two service rates.

By using an argument, similar to that employed in the analysis of Model A, it can be shown that $D_1(z)$ has a unique root z^* in $[0,1]$. The second desired equation between the boundary probabilities is then given by $N_1(z^*) = 0$ or

$$\begin{aligned}
N_1(z^*) = & -\mu_1 P_{0,1}(\lambda z^{*2} - (\lambda + \mu_0 + \gamma_{in})z^* + \mu_0) + \gamma_{in} z^* \mu_0 P_{0,0} + \\
& (\gamma_{in} z^* \sum_{i=0}^{U-1} z^{*i} P_{i,0} + \gamma_{out} z^* \sum_{i=0}^L z^{*i} P_{i,1})(\lambda z^* - \mu_0) \\
= & 0
\end{aligned}$$

Once the boundary probabilities are determined, the steady state performance measures can be derived in a straightforward fashion — the average and variance of queue length can be calculated from the formulas in equations 6.1 and 6.24.

6.6 Summary

In this chapter were presented two performance models for a variable bandwidth channel operating with simple queue-based hysteresis control. The control algorithms were driven by the instantaneous state of the queues and ‘perfect’ knowledge of the queue state was assumed. The first model assumes that bandwidth increase is delayed while in the second model both bandwidth increase and bandwidth decrease are assumed to be delayed. A complete analysis based on MGFs was presented for the Model A. In the next chapter, the basic model will be further extended to include the effect of an imperfect knowledge of the queue state — the sampling problem and the use of a more complex arrival process — the MMPP.

Chapter 7

Performance Models 2

A second set of performance models for a variable bandwidth channel working under simple queue-based hysteresis control is presented in this chapter. When the controller is separated from the channel, perfect knowledge of the queue state can no longer be assumed. The bandwidth algorithms at the controller are driven by samples of the instantaneous queue state that are generated by the ramp at discrete instants; two of the models in this chapter model the sampling process. This chapter also considers the use of the MMPP arrival process to model a bursty traffic source or stream. An approximate analysis for queue-based bandwidth control based on a graphical method is presented. Numerical computation problems encountered when solving the models are also described. A note on the simulation models used in the current work and related research on queues with fluctuating parameters concludes the chapter.

7.1 Introduction

The models C and E described in this chapter are driven by an MMPP arrival process while the third model D is driven by a Poisson arrival process. Model D extends on the models presented in the previous chapter by modelling both bandwidth switching latencies and the sampling of the queue state. The use of an MMPP process with two arrival phases and the modelling of sampling increases the number of phases of the system. Models C and D each have four phases while Model E, the most complex of those presented, has eight.

7.2 The MMPP Arrival Process

The Poisson arrival process used in the previous chapter is useful in studying bandwidth switching and in analysing the oscillatory behaviour of queue-based algorithms¹. However while it is useful to model the aggregate arrival process when a large number of traffic sources (possibly bursty) are multiplexed together, it is not suitable to model a sparse traffic stream.

The MMPP arrival process [Heffes86, Daigle86, Hellstern89] is finding increasing application in the performance modelling of multi-service traffic because of its analytical tractability and because it provides a useful characterisation of both bursty traffic streams and sources. The MMPP is a Poisson process whose instantaneous rate varies randomly according to an irreducible N-state Markov chain. When the MMPP is in phase i , arrivals occur according to a Poisson process with rate λ_i . A superposition of independent MMPP's is also an MMPP [Hellstern89] so that a single MMPP can model a multiplex of bursty sources, each varying between N levels.

The MMPP used in the current work is a two-phase one which varies between arrival phases 0 and 1 with mean arrival rates λ_0 and λ_1 which are assumed to be distinct. A two phase MMPP in which one of the arrival rates is zero is called an Interrupted Poisson Process. This process generates packets during the active phase and remains dormant during the idle phase and can be used to model a talkspurt voice source.

The sojourn time in each phase of an MMPP is determined by the characteristics of the underlying Markov chain. If these times are assumed exponentially distributed with average values $1/\rho_0$ and $1/\rho_1$ seconds, the mean arrival rate λ^* of the MMPP is given by $(\rho_0\lambda_1 + \rho_1\lambda_0)/(\rho_0 + \rho_1)$. A two state MMPP characterised by the four parameters $\lambda_0, \lambda_1, \mu_0, \mu_1$ has been used to approximate a voice arrival stream by matching four of its statistical characteristics with the measurable characteristics of the arrival stream [Heffes86, Zukerman88]. The two state MMPP has also been used to approximate a multi-state MMPP by Hellestern [Hellstern89].

7.3 Assumptions

The basic model of the variable bandwidth channel continues to be a queueing system with a single server that can operate at one of two service rates. The system is modelled as a CTMC whose state transitions are governed by the hysteresis control, the bandwidth switching latencies and the sampling latencies.

The packet service times are assumed to be exponentially distributed, with mean μ_i^{-1}

¹Rate-based algorithms do not display oscillations with Poisson arrivals; once bandwidth is switched in, it is retained except for occasional switching transients caused by statistical fluctuations in the Poisson arrival process.

secs, $i = \{0, 1\}$ and without any loss of generality it may be assumed that $\mu_1 > \mu_0$. This assumption on service time is necessary for reasons of analytical tractability. The bandwidth switching-in and switching-out times are also assumed to be exponentially distributed with mean values γ_{in}^{-1} and γ_{out}^{-1} secs. As before, there are no interruptions of service during switching because of the transparent nature of bandwidth change. When a packet is in service and the rate is changed, service continues at the new rate.

The sampling interval when modelled is assumed to be exponentially distributed with mean value δ^{-1} seconds. Although samples are generated periodically at the ramp the sampling intervals observed at the channel controller tend to be non-deterministic with a small variance. Since samples are sent out over the CFR to the controller there is a transmission jitter.² In addition, process scheduling effects on the ramp transputers and on the controller system can contribute more jitter.

In the two-phase MMPP used in Models C and E, the arrival rates do not vary with the change in service rate. Without any loss of generality it is assumed that $\lambda_1 > \lambda_0$. The necessary and sufficient condition for stability in the models fed by MMPP packet arrivals will be proved to be $\lambda^* < \mu_1$. For Model D which is fed by a simple Poisson process, packet arrivals occur according to a Poisson process with rate λ where $\lambda > \mu_0$. The necessary and sufficient condition for stability for this model is $\mu_1 > \lambda$.

The formulation of the models, their state descriptions and the balance equations will be presented in this chapter. Some details of the steady state analyses and the derivation of stability conditions are given in Appendix B.

7.4 Model C : Hysteresis Control, MMPP Arrivals and Delayed Bandwidth Change (In and Out)

- Queue-based algorithm with hysteresis control
- MMPP arrival process
- Exponentially distributed service times
- Exponentially distributed bandwidth switching-in and switching-out times

The first model in this chapter is for a variable bandwidth channel with MMPP arrivals and finite bandwidth switching latencies. A simple queue-based algorithm with hysteresis control implemented across a single threshold pair is modelled. A perfect knowledge of the queue state is assumed at every instant so that sampling is not modelled. The model is similar to Model B in the previous chapter except that the present model assumes MMPP packet arrivals.

²The sampling delay is assumed to include the transmission delay as described in Section 6.1 of the previous chapter.

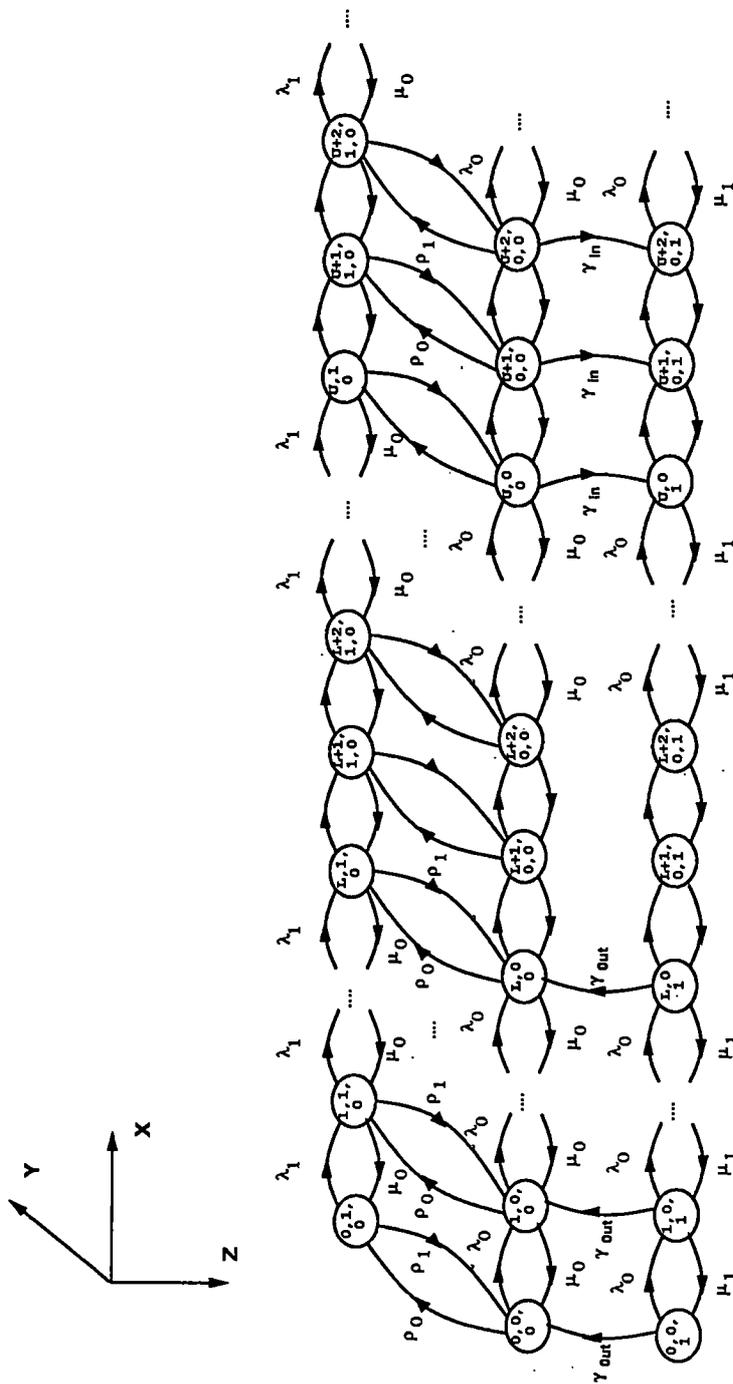


Figure 7.1: Model C : Queue-Based Hysteresis Control — MMPP Arrivals and Delayed Bandwidth Change (In and Out)

Model Description

The state transition diagram for the model is shown in Figure 7.1 and has a three dimensional representation. The state along the X axis represents the number of packets queued for service in the system while that along the Z axis represents the switching phase — as in Model B. In addition along the Y axis are represented the two arrival phases of the MMPP. The transitions between the switching phases (with rates γ_{in} and γ_{out}) are similar to Model B and model the bandwidth switching latencies. The transitions between the arrival phases are with rates ρ_0 and ρ_1 as indicated on the figure.

Depending on the values of the packet arrival and service rates and the sojourn times in each arrival phase, for which the stability condition is satisfied, various modes of system behaviour can be observed, two of which will now be described.

The stability condition for the system is that the mean arrival rate λ^* should be less than μ_1 . For this condition to be true the smaller of the arrival rates λ_0 will have to be less than μ_1 .³ However, the larger arrival rate λ_1 can be greater than or less than μ_1 , giving rise to two modes of system behaviour which are described below. It is assumed in this description that the channel is initially operating at the lower service rate μ_0 (also called the *base bandwidth*) and that $\lambda_0 < \mu_0$.

Case 1 : $\mu_0 < \lambda_1 < \mu_1$

When the channel is operating with the base bandwidth and the arrival process switches to the rate λ_1 the queue begins to increase and switching to the higher service rate (μ_1) will soon take place. If the duration of this arrival phase is prolonged, then oscillations will be observed. When the arrival rate switches back to λ_0 oscillations will cease and the channel will resume operation at the base bandwidth.

Case 2 : $\mu_0 < \mu_1 < \lambda_1$

In this case when the arrival process switches to the rate λ_1 the queue will continuously build up even when the channel has switched to the higher service rate μ_1 . When the arrival process subsequently switches back to λ_0 , the queue would begin to decay and the system would go back to operating at the base bandwidth. The system is stable in the steady state sense, it is just that there are alternating phases of overload and underload.

³In practice it is quite likely that λ_0 will also be less than μ_0 . If this were not the case the lower service rate would be less than both the arrival rates and the mean arrival rate λ^* value which is not a sensible allocation of bandwidth. The better choice would be for μ_0 to be greater than λ^* . Then with a proper choice of queue threshold, bandwidth will be switched in only when the arrival process switches to its overload phase (with arrival rate $\lambda_1 > \mu_0$), for an extended duration.

State Description

The three dimensional vector $[\{X(t), Y(t), Z(t)\}, X(t) \geq 0, Y(t) = \{0, 1\}, Z(t) = \{0, 1\}, t \geq 0]$, describes the state space of the system where $X(t)$ is the queue length, $Y(t)$ is the arrival phase and $Z(t)$ is the switching phase at time t .

Arrival and switching planes can be distinguished on the state transition diagram in Figure 7.1; the former are parallel to the XZ plane while the latter are parallel to the XY plane. Switching plane 0 lies along the XY plane while switching plane 1 is parallel to it but with a positive displacement along the Z axis. Arrival plane 0 lies along the XZ plane while arrival plane 1 is parallel to it with a positive displacement along the Y axis.

The bandwidth switching-in latencies are represented by the set (infinite) of transitions from switching plane 0 to switching plane 1, beginning with the state pair $\{(U, 0, 0), (U, 1, 0)\}$ with transition rates γ_{in} . The bandwidth switching-out latencies are represented by a finite set of transitions from switching plane 1 to switching plane 0, starting with the state pair $\{(L, 0, 1), (L, 1, 1)\}$ and ending with the state pair $\{(0, 0, 1), (0, 1, 1)\}$ with transition rates γ_{out} . As in Models A and B an assumption made with the current model is that the bandwidth is increased only when the queue length is greater then or equal to the switching-in threshold or conversely decreased only when the queue length is less then or equal to the switching-out threshold.

Let $\Pi = \{\Pi_{0,0}, \Pi_{0,1}, \Pi_{1,0}, \Pi_{1,1}\}$ represent the stationary distribution for the process $\{X(t), Y(t), Z(t)\}$ where $\Pi_{i,j} = \{P_{0,i,j}, P_{1,i,j}, \dots\}, i = \{0, 1\}$ and $j = \{0, 1\}$. $P_{x,y,z}$ is the probability of having x packets queued in the system when the arrival phase is y and the switching phase is z . Four levels of the system can be defined according to the vector (a, b) , where $a = \{0, 1\}$ is the arrival phase and $b = \{0, 1\}$ is the switching phase. Level $(1, 1)$ is not visible in the figure.

In the steady state the balance equations are :

Level (0,0) :

$$\begin{aligned}
 (\lambda_0 + \rho_0)P_{0,0,0} &= \mu_0 P_{1,0,0} + \gamma_{out} P_{0,0,1} + \rho_1 P_{0,1,0} & (7.1) \\
 (\lambda_0 + \mu_0 + \rho_0)P_{i,0,0} &= \lambda_0 P_{i-1,0,0} + \mu_0 P_{i+1,0,0} + \gamma_{out} P_{i,0,1} + \rho_1 P_{i,1,0} \\
 & \quad i = 1, \dots, L \\
 (\lambda_0 + \mu_0 + \rho_0)P_{i,0,0} &= \lambda_0 P_{i-1,0,0} + \mu_0 P_{i+1,0,0} + \rho_1 P_{i,1,0} \\
 & \quad i = L + 1, \dots, U - 1 \\
 (\lambda_0 + \mu_0 + \gamma_{in} + \rho_0)P_{i,0,0} &= \lambda_0 P_{i-1,0,0} + \mu_0 P_{i+1,0,0} + \rho_1 P_{i,1,0} \\
 & \quad i = U, U + 1, \dots
 \end{aligned}$$

Level (1,0) :

$$\begin{aligned}
 (\lambda_1 + \rho_1)P_{0,1,0} &= \mu_0 P_{1,1,0} + \gamma_{out} P_{0,1,1} + \rho_0 P_{0,0,0} & (7.2) \\
 (\lambda_1 + \mu_0 + \rho_1)P_{i,1,0} &= \lambda_1 P_{i-1,1,0} + \mu_0 P_{i+1,1,0} + \gamma_{out} P_{i,1,1} + \rho_0 P_{i,0,0}
 \end{aligned}$$

$$\begin{aligned}
& i = 1, \dots, L \\
(\lambda_1 + \mu_0 + \rho_1)P_{i,1,0} &= \lambda_1 P_{i-1,1,0} + \mu_0 P_{i+1,1,0} + \rho_0 P_{i,0,0} \\
& i = L + 1, \dots, U - 1 \\
(\lambda_1 + \mu_0 + \gamma_{in} + \rho_1)P_{i,1,0} &= \lambda_1 P_{i-1,1,0} + \mu_0 P_{i+1,1,0} + \rho_0 P_{i,0,0} \\
& i = U, U + 1, \dots
\end{aligned}$$

Level (0,1) :

$$\begin{aligned}
(\lambda_0 + \rho_0 + \gamma_{out})P_{0,0,1} &= \mu_1 P_{1,0,1} + \rho_1 P_{0,1,1} & (7.3) \\
(\lambda_0 + \mu_1 + \rho_0 + \gamma_{out})P_{i,0,1} &= \lambda_0 P_{i-1,0,1} + \mu_1 P_{i+1,0,1} + \rho_1 P_{i,1,1} \\
& i = 1, \dots, L \\
(\lambda_0 + \mu_1 + \rho_0)P_{i,0,1} &= \lambda_0 P_{i-1,0,1} + \mu_1 P_{i+1,0,1} + \rho_1 P_{i,1,1} \\
& i = L + 1, \dots, U - 1 \\
(\lambda_0 + \mu_1 + \rho_0)P_{i,0,1} &= \lambda_0 P_{i-1,0,1} + \mu_1 P_{i+1,0,1} + \rho_1 P_{i,1,1} + \gamma_{in} P_{i,0,0} \\
& i = U, U + 1, \dots
\end{aligned}$$

Level (1,1) :

$$\begin{aligned}
(\lambda_1 + \rho_1 + \gamma_{out})P_{0,1,1} &= \mu_1 P_{1,1,1} + \rho_0 P_{0,0,1} & (7.4) \\
(\lambda_1 + \mu_1 + \rho_1 + \gamma_{out})P_{i,1,1} &= \lambda_1 P_{i-1,1,1} + \mu_1 P_{i+1,1,1} + \rho_0 P_{i,0,1} \\
& i = 1, \dots, L \\
(\lambda_1 + \mu_1 + \rho_1)P_{i,1,1} &= \lambda_1 P_{i-1,1,1} + \mu_1 P_{i+1,1,1} + \rho_0 P_{i,0,1} \\
& i = L + 1, \dots, U - 1 \\
(\lambda_1 + \mu_1 + \rho_1)P_{i,1,1} &= \lambda_1 P_{i-1,1,1} + \mu_1 P_{i+1,1,1} + \rho_0 P_{i,0,1} + \gamma_{in} P_{i,1,0} \\
& i = U, U + 1, \dots
\end{aligned}$$

The derivation of the stability condition and some details of the steady state analysis of the model are to be found in Section B.1 of Appendix B.

7.5 Model D : Hysteresis Control, Poisson Arrivals, Delayed Bandwidth Change (In and Out) and Sampling

- Queue-based algorithm with hysteresis control
- Poisson arrival process
- Exponentially distributed service times
- Exponentially distributed bandwidth switching-in and switching-out times
- Exponentially distributed sampling intervals

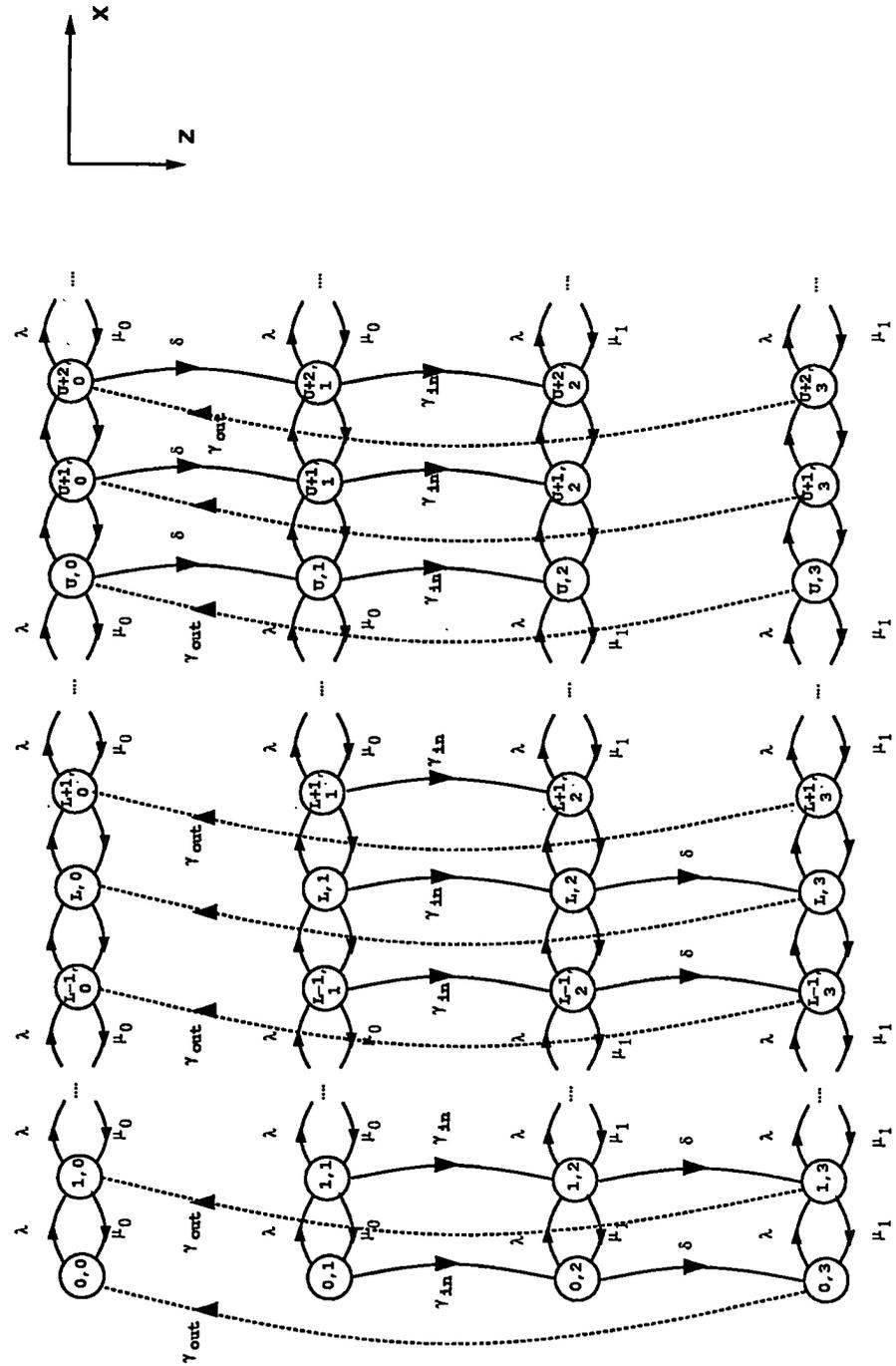


Figure 7.2: Model D : Queue-Based Hysteresis Control — Poisson Arrivals, Delayed Bandwidth Change (In and Out) and Sampling

Model D is for a variable bandwidth channel with Poisson arrivals and delayed bandwidth change. A simple queue-based algorithm with hysteresis control implemented across a single threshold pair is modelled. An imperfect knowledge of the queue state is assumed and the sampling process is emulated. Model D is the most sophisticated of the models with Poisson arrivals and is used in the performance validation against experimental and simulation results presented in Section 8.4 of Chapter 8.

Model Description

The state transition diagram for the model is shown in Figure 7.2. The state along the X axis represents the number of packets queued for service in the system while that along the Z axis represents the switching phase. There are four distinct switching phases in the model numbered from zero to three starting at the top. The state descriptor (i, j) indicates that i packets are queued in the system when the switching phase is j , $j = \{0, 1, 2, 3\}$. Unlike the earlier models where each switching phase had a distinct service rate, in the present model switching phases 0 and 1 have the same service rate μ_0 while phases 2 and 3 have a service rate μ_1 .

When the system is in phase zero it is in a state of overload because of the assumption on the average arrival rate λ being $> \mu_0$. The controller lacking perfect knowledge about the state of the queue will not know about the overload until it receives a sample of the queue length which is greater than or equal to the switching-in threshold state $(U, 0)$. Samples of the queue that are less than the threshold do not initiate switching and hence there are no transitions shown on Figure 7.2 from these states to the corresponding ones on the second phase. Sampling is hence represented by the set (infinite) of transitions from switching phase 0 to switching phase 1 from state $(U, 0)$ onwards with transition rates δ .

A sampling transition represents an initiation of the bandwidth adjustment process. During bandwidth adjustment the system can transit to any state (although it is more likely to move to the higher states because $\lambda > \mu_0$). This is represented by the set (infinite) of transitions with rate γ_{in} from every state in switching phase 1 to the corresponding one in switching phase 2.

When transition takes place to phase 2 the channel operates with the higher service rate $\mu_1 > \lambda$. The build up in the channel queue observed in the previous two phases is arrested and the queue starts to decay. As in the switching-in process only sample values that are less than or equal to the switching-out threshold state $(L, 2)$ can initiate bandwidth adjustment. This is represented by the finite set of transitions with sampling rates δ from switching phase 2 to switching phase 3, starting with the state $(L, 2)$ and ending with the state $(0, 2)$. The final set of transitions from switching phase 3 back to switching phase 0 with rate γ_{out} models the bandwidth switching-out latency.

State Description

The state of the system is described by the two dimensional vector $[\{X(t), Y(t)\}, X(t) \geq 0, Y(t) = \{0, 1, 2, 3\}, t \geq 0]$, where $X(t)$ is the queue length and $Y(t)$ is the switching phase at time t . Let $\Pi = \{\Pi_0, \Pi_1, \Pi_2, \Pi_3\}$ represent the stationary distribution for the process $[X(t), Y(t)]$ where $\Pi_i = \{P_{0,i}, P_{1,i}, \dots\}, i = \{0, 1, 2, 3\}$. $P_{x,y}$ is the probability of having x packets queued in the system when the switching phase is y .

In the steady state the balance equations are :

Switching Phase 0 :

$$\begin{aligned} \lambda P_{0,0} &= \mu_0 P_{1,0} + \gamma_{out} P_{0,3} & (7.5) \\ (\lambda + \mu_0) P_{i,0} &= \lambda P_{i-1,0} + \mu_0 P_{i+1,0} + \gamma_{out} P_{i,3} \\ & i = 1, 2, \dots, U-1 \\ (\lambda + \mu_0 + \delta) P_{i,0} &= \lambda P_{i-1,0} + \mu_0 P_{i+1,0} + \gamma_{out} P_{i,3} \\ & i = U, U+1, \dots \end{aligned}$$

Switching Phase 1 :

$$\begin{aligned} (\lambda + \gamma_{in}) P_{0,1} &= \mu_0 P_{1,1} & (7.6) \\ (\lambda + \mu_0 + \gamma_{in}) P_{i,1} &= \lambda P_{i-1,1} + \mu_0 P_{i+1,1} \\ & i = 1, 2, \dots, U-1 \\ (\lambda + \mu_0 + \gamma_{in}) P_{i,1} &= \lambda P_{i-1,1} + \mu_0 P_{i+1,1} + \delta P_{i,0} \\ & i = U, U+1, \dots \end{aligned}$$

Switching Phase 2 :

$$\begin{aligned} (\lambda + \delta) P_{0,2} &= \mu_1 P_{1,2} + \gamma_{in} P_{0,1} & (7.7) \\ (\lambda + \mu_1 + \delta) P_{i,2} &= \lambda P_{i-1,2} + \mu_1 P_{i+1,2} + \gamma_{in} P_{i,1} \\ & i = 1, 2, \dots, L \\ (\lambda + \mu_1) P_{i,2} &= \lambda P_{i-1,2} + \mu_1 P_{i+1,2} + \gamma_{in} P_{i,1} \\ & i = L+1, L+2, \dots \end{aligned}$$

Switching Phase 3 :

$$(\lambda + \gamma_{out}) P_{0,3} = \mu_1 P_{1,3} + \delta P_{0,2} \quad (7.8)$$

$$\begin{aligned}
(\lambda + \mu_1 + \gamma_{out})P_{i,3} &= \lambda P_{i-1,3} + \mu_1 P_{i+1,3} + \delta P_{i,2} \\
& \quad i = 1, 2, \dots, L \\
(\lambda + \mu_1 + \gamma_{out})P_{i,3} &= \lambda P_{i-1,3} + \mu_1 P_{i+1,3} \\
& \quad i = L + 1, L + 2, \dots
\end{aligned}$$

The derivation of the stability condition and some details of the steady state analysis for Model D will be described in Section B.2 of Appendix B.

7.6 Model E : Hysteresis Control, MMPP Arrivals, Delayed Bandwidth Change (In and Out) and Sampling

- Queue-based algorithm with hysteresis control
- MMPP arrivals
- Exponentially distributed service times
- Exponentially distributed bandwidth switching-in and switching-out times
- Exponentially distributed sampling intervals

Model E is the most sophisticated of the models presented in the two chapters and combines elements from both Models C and D. It models a variable bandwidth channel with MMPP arrivals and bandwidth switching latencies. A simple queue-based algorithm with hysteresis control implemented across a single threshold pair is modelled. An imperfect knowledge of the queue state is assumed and the sampling process is emulated. Model E is used in the performance validation against experimental and simulation results presented in Section 8.4 of Chapter 8.

Model E has an expanded state space (Figure 7.3) with a three dimensional representation with two arrival phases as in Model C and four switching phases as in Model D to model the bandwidth switching and sampling latencies. The model operates in a similar manner to Model D except that there are now arrival transitions which take place independently of the switching ones.

State Description

The three dimensional vector [{ $\mathbf{X}(t)$, $\mathbf{Y}(t)$, $\mathbf{Z}(t)$ }, $\mathbf{X}(t) \geq 0$, $\mathbf{Y}(t) = \{0, 1\}$, $\mathbf{Z}(t) = \{0, 1, 2, 3\}$, $t \geq 0$], describes the state of the system where $\mathbf{X}(t)$ is the queue length, $\mathbf{Y}(t)$ is the arrival phase and $\mathbf{Z}(t)$ is the switching phase at time t . The system is modelled as a continuous time Markov chain with four switching phases along the Z axis and two arrival phases along the Y axis as in Figure 7.3.

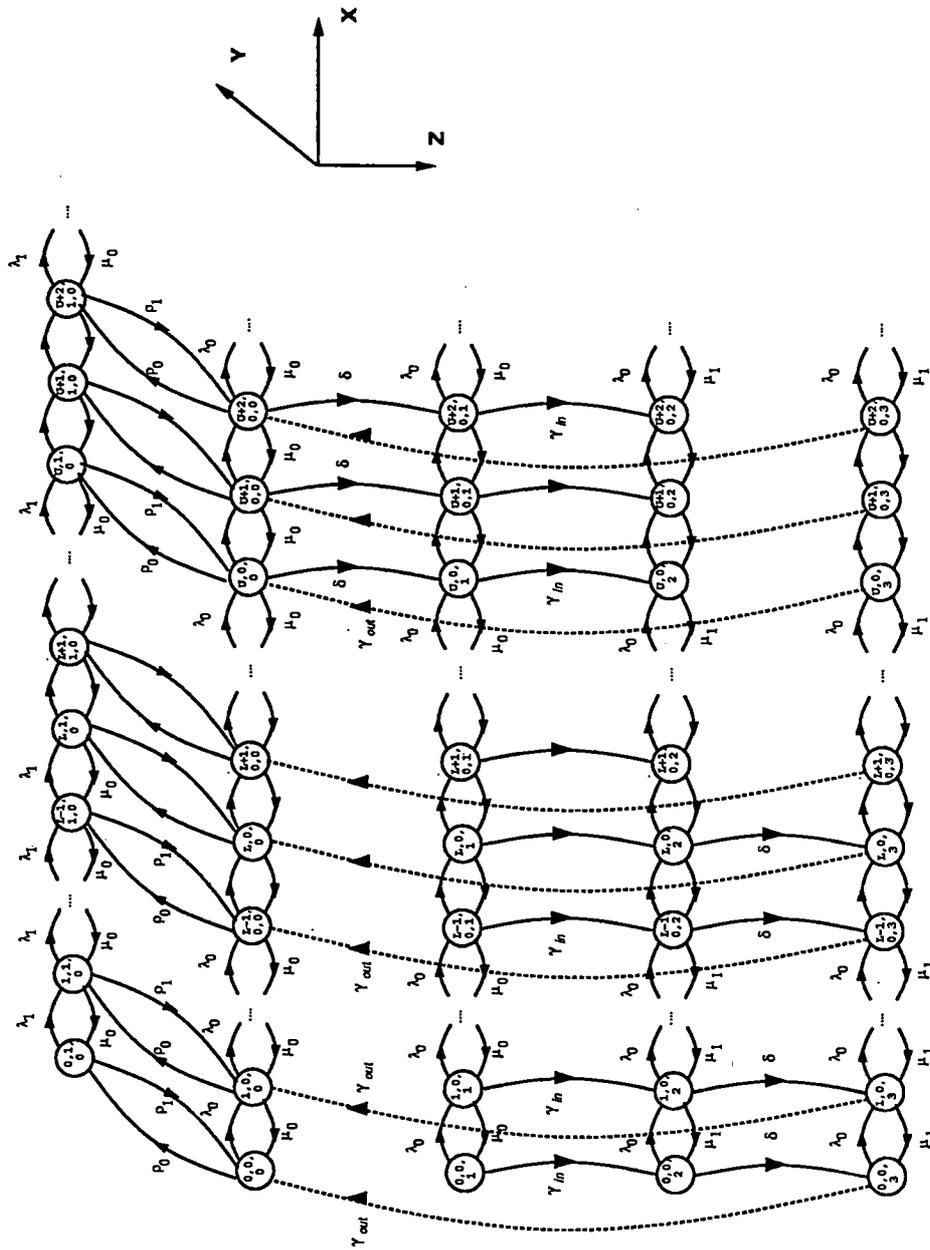


Figure 7.3: Model E : Queue-Based Hysteresis Control — MMPP Arrivals, Delayed Bandwidth Change (In and Out) and Sampling

Arrival and switching planes can be distinguished on the state transition diagram as in Model C. The two arrival planes are parallel to the XZ plane while the four switching planes are parallel to the XY plane. Switching plane 0 lies along the XY plane while switching planes 1,2 and 3 are parallel to it but with positive displacements along the Z axis. Arrival plane 0 lies along the XZ plane while arrival plane 1 is parallel to it with a positive displacement along the Y axis.

Sampling is represented by two sets of transitions, the first from switching plane 0 to switching plane 1, beginning with the state pair $\{(U, 0, 0), (U, 1, 0)\}$ and with rates δ . The second set of transitions is from switching plane 2 to switching plane 3, starting with the state pair $\{(L, 0, 2), (L, 1, 2)\}$ and ending with the state pair $\{(0, 0, 2), (0, 1, 2)\}$ also with the transition rates δ .

The bandwidth switching latencies are modelled by transitions emanating from all states on their originating plane. The switching in transitions are from switching plane 1 to switching plane 2 with rates γ_{in} , while those for switching out are from switching plane 3 to switching plane 0 with rates γ_{out} .

Let $\Pi = \{\Pi_{0,0}, \Pi_{1,0}, \Pi_{0,1}, \Pi_{1,1}, \Pi_{0,2}, \Pi_{1,2}, \Pi_{0,3}, \Pi_{1,3}\}$ represent the stationary distribution for the process $[X(t), Y(t), Z(t)]$ where $\Pi_{i,j} = \{P_{0,i,j}, P_{1,i,j}, \dots\}$, $i = \{0,1\}$ and $j = \{0,1,2,3\}$. $P_{x,y,z}$ is the probability of having x packets queued in the system when the arrival phase is y and the switching phase is z . Eight levels of the system can be defined according to the vector (a, b) where $a = \{0,1\}$ is the arrival phase and $b = \{0,1,2,3\}$ is the switching phase. Only five of these levels are visible on the state transition diagram.

The steady state balance equations are :

Level (j,0), j = 0,1 :

$$\begin{aligned}
 (\lambda_j + \rho_j)P_{0,j,0} &= \mu_0 P_{1,j,0} + \gamma_{out} P_{0,j,3} + \rho_{1-j} P_{0,1-j,0} & (7.9) \\
 (\lambda_j + \rho_j + \mu_0)P_{i,j,0} &= \lambda_j P_{i-1,j,0} + \mu_0 P_{i+1,j,0} + \gamma_{out} P_{i,j,3} + \rho_{1-j} P_{i,1-j,0} \\
 & \quad i = 1, 2, \dots, U-1 \\
 (\lambda_j + \rho_j + \mu_0 + \delta)P_{i,j,0} &= \lambda_j P_{i-1,j,0} + \mu_0 P_{i+1,j,0} + \gamma_{out} P_{i,j,3} + \rho_{1-j} P_{i,1-j,0} \\
 & \quad i = U, U+1, \dots,
 \end{aligned}$$

Level (j,1), j = 0,1 :

$$\begin{aligned}
 (\lambda_j + \gamma_{in} + \rho_j)P_{0,j,1} &= \mu_0 P_{1,j,1} + \rho_{1-j} P_{0,1-j,1} & (7.10) \\
 (\lambda_j + \mu_0 + \gamma_{in} + \rho_j)P_{i,j,1} &= \lambda_j P_{i-1,j,1} + \mu_0 P_{i+1,j,1} + \rho_{1-j} P_{i,1-j,1} \\
 & \quad i = 1, 2, \dots, U-1 \\
 (\lambda_j + \mu_0 + \gamma_{in} + \rho_j)P_{i,j,1} &= \lambda_j P_{i-1,j,1} + \mu_0 P_{i+1,j,1} + \delta P_{i,j,0} + \rho_{1-j} P_{i,1-j,1} \\
 & \quad i = U, U+1, \dots,
 \end{aligned}$$

Level (j,2), j = 0,1 :

$$(\lambda_j + \delta + \rho_j)P_{0,j,2} = \mu_1 P_{1,j,2} + \gamma_{in} P_{0,j,1} + \rho_{1-j} P_{0,1-j,2} \quad (7.11)$$

$$\begin{aligned}
(\lambda_j + \mu_1 + \delta + \rho_j)P_{i,j,2} &= \lambda_j P_{i-1,j,2} + \mu_1 P_{i+1,j,2} + \gamma_{in} P_{i,j,1} + \rho_{1-j} P_{i,1-j,2} \\
& \quad i = 1, 2, \dots, L \\
(\lambda_j + \mu_1 + \rho_j)P_{i,j,2} &= \lambda_j P_{i-1,j,2} + \mu_1 P_{i+1,j,2} + \gamma_{in} P_{i,j,1} + \rho_{1-j} P_{i,1-j,2} \\
& \quad i = L + 1, L + 2, \dots,
\end{aligned}$$

Level (j,3), j = 0,1 :

$$\begin{aligned}
(\lambda_j + \gamma_{out} + \rho_j)P_{0,j,3} &= \mu_1 P_{1,j,3} + \delta P_{0,j,2} + \rho_{1-j} P_{0,1-j,3} & (7.12) \\
(\lambda_j + \mu_1 + \gamma_{out} + \rho_j)P_{i,j,3} &= \lambda_j P_{i-1,j,3} + \mu_1 P_{i+1,j,3} + \delta P_{i,j,2} + \rho_{1-j} P_{i,1-j,3} \\
& \quad i = 1, 2, \dots, L \\
(\lambda_j + \mu_1 + \gamma_{out} + \rho_j)P_{i,j,3} &= \lambda_j P_{i-1,j,3} + \mu_1 P_{i+1,j,3} + \rho_{1-j} P_{i,1-j,3} \\
& \quad i = L + 1, L + 2, \dots,
\end{aligned}$$

The derivation of the stability condition and some details of the steady state analysis of Model E are presented in Section B.3 of Appendix B.

7.7 Analytic Properties of Moment Generating Functions — a Discussion

In the analysis of Model B in the previous chapter, equation 6.22 was derived from a consideration of the analytic properties of a partial MGF and was used in solving for the zero-state probabilities. In the analysis of Models C, D and E evaluating the roots of the denominator polynomials of the MGFs to derive equations to solve for the unknown zero-state probabilities is more complex. The method for doing so would be notionally similar to the argument used in Model B, except that $N - 1$ roots would have to be found in the interval $(0, 1)$, where N is the number of zero-state probabilities (for example $N = 8$ in Model E).

In [Mitrany68] and in [Kraimeche85] a theorem to show that a sufficient number of roots exists in the denominator polynomial of an MGF is presented. The proof of the theorem uses inductive reasoning to show that the denominator polynomial has exactly $N - 1$ distinct real roots in the interval $(0, 1)$. Each of these roots when substituted in the numerator polynomial gives an equation in the zero-state probabilities. However, as was pointed out closed form expressions for the unknown probabilities are difficult to obtain if at all possible. Numerical methods are suggested for solving actual systems. This was also observed by [Yechiali71] and [Serres88], the latter also pointing out problems that may arise when the roots are not distinct. Serres preferred the use of matrix geometric schemes [Neuts81], which avoid the explicit computation of roots and circumvent the non-distinct roots problem.

In the current work, by adapting Mitrany's result it may be theoretically proven that a sufficient number of distinct roots exist in the denominator polynomials of the appropriate MGFs to solve for the unknown zero-state probabilities.

7.8 Numerical Computation and Iterative Methods of Solution

In practice, except for the simplest model — Model A and for small values of the system parameters, steady state performance measures like the average queue length were not computed directly from the MGFs. This was because of floating point computation problems. Since the dual service rate systems modelled could operate temporarily in overload, expressions for the steady state probabilities such as in equation 6.19 of the previous chapter :

$$\sum_{i=0}^{U-1} P_{i,0} = P_{0,0} \left\{ \frac{(1 - \rho_0^U)}{1 - \rho_0} \right\} - P_{L,1} \left\{ \frac{\mu_1}{\mu_0(1 - \rho_0)^2} [(U - L)(1 - \rho_0) - \rho_0(1 - \rho_0^{U-L})] \right\}$$

which involve ρ_0 ($\rho_0 > 1$) raised to a positive power, could result in floating exceptions for typical values of the powers; for example when $U =$ (say) 1000 packets.

Similar numerical stability problems were reported in [Serres88]. Serres modelled the access control of queueable and blockable traffic requests onto a transmission channel operating with a movable boundary scheme. The resultant queueing model was solved by both generating function and matrix geometric methods. Serres noted that as the system size increased, there was a loss of precision in both the solution methods, which manifested itself in the appearance of negative boundary probabilities.

The preferred approach for the steady state solution of the models in this dissertation was to use Gauss-Seidel type iterations on the balance equations [Cooper81] with improved precision arithmetic. The Gauss-Seidel numerical scheme has been used extensively to solve a finite system of equations in large state space systems. Consider the general set of linear equations, $\mathbf{Ax} = \mathbf{B}$ which can also be written as :

$$a_{i0}x_0 + a_{i1}x_1 + a_{i2}x_2 + \dots + a_{iN}x_N = b_i \\ i = 0, 1, 2, \dots, N$$

The Gauss-Seidel iteration technique used to solve for the unknown x_i is :

$$x_i^{(m+1)} = \frac{1}{a_{ii}} [b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(m+1)} - \sum_{j=i+1}^N a_{ij}x_j^{(m)}] \\ i = 0, 1, 2, \dots, N$$

where $x_i^{(m)}$ is the m^{th} iterate of x_i , $i = 0, 1, 2, \dots, N$. The iterations begin with $x^{(0)}$, an arbitrary initial vector. During an iteration, each element $x_i^{(m+1)}$ is calculated from the most recent values of the other components.

Iteration Scheme

This section describes the iteration scheme that was applied to the balance equations of Models D and E to solve for the steady state probabilities. Consider the set of balance equations 7.5 through to 7.8 for Model D presented in Section 7.5 of this chapter. The balance equations for the four level switching process can be rewritten in matrix form as :

$$\begin{aligned}
 P_0 A_0 + P_1 \Delta(\mu) &= \mathbf{0} \\
 P_{i-1} \Delta(\lambda) + P_i A_1 + P_{i+1} \Delta(\mu) &= \mathbf{0} \\
 & \quad i = 1, 2, \dots, L \\
 P_{i-1} \Delta(\lambda) + P_i A_2 + P_{i+1} \Delta(\mu) &= \mathbf{0} \\
 & \quad i = L + 1, L + 2, \dots, U - 1 \\
 P_{i-1} \Delta(\lambda) + P_i A_3 + P_{i+1} \Delta(\mu) &= \mathbf{0} \\
 & \quad i = U, U + 1, \dots, N
 \end{aligned} \tag{7.13}$$

where P_i is a row vector composed of the state probabilities in the four switching phases defined as $P_i = \{P_{i,0}, P_{i,1}, P_{i,2}, P_{i,3}\}$. L and U are the lower and upper queue thresholds while N is the total number of steady state equations iterated over.⁴ $\Delta(\lambda)$ and $\Delta(\mu)$ are diagonal matrices and are defined below along with the matrices A_i , $i = \{0, 1, 2, 3\}$.

$$\Delta(\lambda) = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

$$\Delta(\mu) = \begin{pmatrix} \mu_0 & 0 & 0 & 0 \\ 0 & \mu_0 & 0 & 0 \\ 0 & 0 & \mu_1 & 0 \\ 0 & 0 & 0 & \mu_1 \end{pmatrix}$$

$$A_0 = \begin{pmatrix} -\lambda & 0 & 0 & 0 \\ 0 & -\lambda - \gamma_{in} & \gamma_{in} & 0 \\ 0 & 0 & -\lambda - \delta & \delta \\ \gamma_{out} & 0 & 0 & -\lambda - \gamma_{out} \end{pmatrix}$$

$$A_1 = \begin{pmatrix} -\lambda - \mu_0 & 0 & 0 & 0 \\ 0 & -\lambda - \gamma_{in} - \mu_0 & \gamma_{in} & 0 \\ 0 & 0 & -\lambda - \delta - \mu_1 & \delta \\ \gamma_{out} & 0 & 0 & -\lambda - \gamma_{out} - \mu_1 \end{pmatrix}$$

⁴A typical value used for N was 12000 chosen on the basis of the arrival, service and switching parameters in the models.

$$A_2 = \begin{pmatrix} -\lambda - \mu_0 & 0 & 0 & 0 \\ 0 & -\lambda - \gamma_{in} - \mu_0 & \gamma_{in} & 0 \\ 0 & 0 & -\lambda - \mu_1 & 0 \\ \gamma_{out} & 0 & 0 & -\lambda - \gamma_{out} - \mu_1 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} -\lambda - \mu_0 - \delta & \delta & 0 & 0 \\ 0 & -\lambda - \gamma_{in} - \mu_0 & \gamma_{in} & 0 \\ 0 & 0 & -\lambda - \mu_1 & 0 \\ \gamma_{out} & 0 & 0 & -\lambda - \gamma_{out} - \mu_1 \end{pmatrix}$$

The system of equations 7.13 may be further rewritten as :

$$\begin{aligned} P_0 &= [P_0(A_0 + \Delta_0) + P_1\Delta(\mu)]\Delta_0^{-1} \\ P_i &= [P_{i-1}\Delta(\lambda) + P_i(A_1 + \Delta_1) + P_{i+1}\Delta(\mu)]\Delta_1^{-1} \\ &\quad i = 1, 2, \dots, L \\ P_i &= [P_{i-1}\Delta(\lambda) + P_i(A_2 + \Delta_2) + P_{i+1}\Delta(\mu)]\Delta_2^{-1} \\ &\quad i = L + 1, L + 2, \dots, U - 1 \\ P_i &= [P_{i-1}\Delta(\lambda) + P_i(A_3 + \Delta_3) + P_{i+1}\Delta(\mu)]\Delta_3^{-1} \\ &\quad i = U, U + 1, \dots, N \end{aligned} \tag{7.14}$$

where Δ_i , $i = \{0, 1, 2, 3\}$ are defined as $\Delta_i \stackrel{\text{def}}{=} -\text{diag}(A_i)$. That is Δ_i is a diagonal matrix whose elements are the negative of the corresponding diagonal elements of A_i . Along with the conservation of probability sum equation $\sum_{i=0}^N P_i \mathbf{e} = 1$ where \mathbf{e} is the unit column vector of dimension 4×1 , the set of equations 7.14 is suitable for a Gauss-Seidel type iteration. A similar iteration scheme was recommended by Neuts [Neuts81]⁵.

During an iteration, the conservation of probability sum condition is used to normalise the values of the iterates (steady state probabilities). This scheme was used in evaluating the steady state probabilities and performance metrics for Model D that will be presented in Section 8.4 of Chapter 8. A similar iteration scheme was developed for Model E. The iterations were performed on DEC 3100 (MIPS) workstations but the convergence was slow. The stopping condition for iterations was when successive iterates differed by less than a convergence limit. Faster convergence may be obtained with the use of other iteration schemes such as the Overrelaxation method [Cooper81] which is a problem for further work.

7.9 Graphical Analysis of Queue-Based Hysteresis Control

With a knowledge of the arrival and service rates and the bandwidth switching latencies, a simple graphical analysis can yield approximate values for the steady state measures such

⁵page 279.

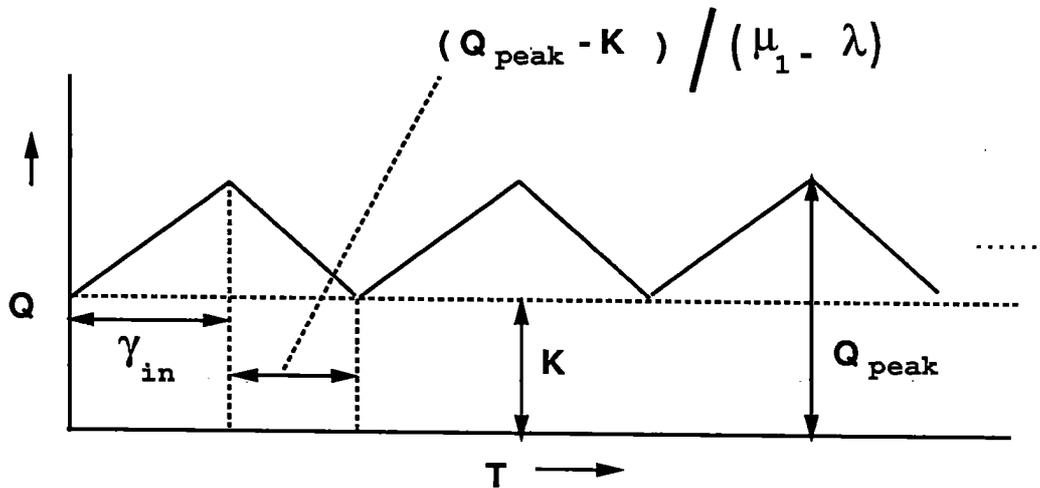


Figure 7.4: Instantaneous Queue Length vs. Time. Graphical Analysis of Queue-Based Hysteresis Control — Point Thresholding, Poisson Arrivals and Delayed Bandwidth Change (In)

as the mean queue size when using a simple queue-based algorithm to control a variable bandwidth channel.

As an example, consider the system described in Model A – with service rates μ_0 and μ_1 , a Poisson arrival process (rate λ , $\mu_0 < \lambda < \mu_1$), delayed bandwidth increase (mean value γ_{in}^{-1}) and instantaneous bandwidth decrease. In the present analysis the use of a queue-based bandwidth algorithm operating with point thresholding control is assumed (threshold value K). The system exhibits oscillations and the mean queue size may be calculated by integrating the variation of the instantaneous queue length. Perfect knowledge of the queue state is assumed so that sampling is not emulated.

Figure 7.4 shows the typical queue size variation of the system with time. The queue will always have a residue of K packets since the bandwidth is switched out instantaneously. When the service rate is μ_0 , the queue builds up and bandwidth adjustment is initiated. In the time (γ_{in}) taken to switch to the higher rate μ_1 the queue reaches a value $Q_{peak} = (K + \gamma_{in} * (\lambda - \mu_0))$. The queue then begins to decay and will reach the threshold K in a time $(Q_{peak} - K) / (\mu_1 - \lambda)$. At this point, the system changes its service rate instantly to μ_0 . The queue starts to build up again and the cycle repeats.

The mean queue size evaluated from Figure 7.4 is given by :

$$E[Q] = K + 0.5[\gamma_{in} + (Q_{peak} - K)/(\mu_1 - \lambda)] * (Q_{peak} - K)$$

This simple analysis assumes linear variations of the parameters with time. In the experimental results similar trends of increase and decrease were noted for queue-based control (Section 9.2 in Chapter 9). A graphical analysis may also be attempted for rate-based algorithms to derive the mean queue size for a varying traffic load.

An equivalent graphical analysis was presented recently for a rate control scheme in a store and forward network where the traffic arrival rate is controlled by feedback detailing the queue state by Bolot [Bolot90].

7.10 Simulation Models

The simulation results that will be presented in Section 8.4 in Chapter 8 along with results from experiment and theoretical analyses are derived from an event driven discrete time simulator developed by the author. The simulator was used to model a variable bandwidth channel controlled by a simple queue-based bandwidth algorithm. The simulator modelled bi-level hysteresis control, the bandwidth switching latencies and the sampling latencies. Poisson and MMPP arrivals were used in the two main simulation models that correspond to the theoretical Models D and E presented in this chapter.

The main difference between the simulation and theoretical models was that the former made less restrictive assumptions on the service discipline used at the ramp. In the simulation models the service times were assumed to be fixed (in a switching phase) in contrast to the exponential distributions that were assumed for them in the theoretical models. This implies that the simulations model fixed size packets. However, the simulation models do not emulate the framing effect at the ramp. The effect of these assumptions will be further described in Section 8.4 of Chapter 8.

Another assumption made in the simulations was that the service time of a packet remains unchanged even if the channel switches to a different service rate during the service. However, succeeding packets on the queue are serviced at the new rate until the switching phase changes again. This assumption is unlikely to cause any pronounced effect on the performance predicted by the simulation models.

Each simulation was run for an initial period before the results were logged; the period commonly used was the time taken to generate 5000 packets. Each run was repeated from 30 to 50 times with different seed values for the random number generator. The 95 percent confidence intervals were obtained on these samples assuming a t-Student distribution [Law82]. The total number of packets that were generated in each simulation run were approximately a million.

7.11 Related Work on Queues with Fluctuating Parameters

One of the earliest studies of a queue with fluctuating parameters was by Eisen who considered a single server queueing system with pairs of arrival and service rates [Eisen63]. The arrival and service rates changed concurrently and the time spent in each phase was assumed to be exponentially distributed. The steady state solution was obtained from the generating function method. The analytic properties of the MGF were examined to derive an equation for the zero-state probabilities.

Yechiali modelled a similar two-phase system [Yechiali71] and pointed out the difficulty in expressing the unknown state probabilities in the MGFs in closed form in terms of the boundary probabilities, except indirectly through recursive formulations. Yechiali demonstrated that slow switching between the arrival and service phases when the traffic utilisation in each phase was less than one could lead to two quasi-equilibrium system states, each of $M/M/1$ type.

This model was then extended by Yechiali [Yechiali73] to an n -phase variation of the arrival and service parameters. Yechiali showed that his earlier model, and the model of [Mitrany68] were special cases of the new one. Mitrany modelled an $M/M/N$ queueing system with service interruptions; the servers could break down and take an exponentially distributed time to be repaired.

[Zukerman85, Zukerman86a] generalised the two-phase model proposed by Yechiali [Yechiali71] to a multi-server system — for both a finite and an infinite number of servers. In a K server system the fluctuations of service phase were such that each of the $j \leq K$ operational servers made identical and concurrent transitions to the new service rate. Zukerman used the MGF approach in a continuous time analysis of the system to solve for the steady state performance metrics like the average queue size. In this performance analysis it was shown how slow switching between the phases could adversely affect the average queue size when one of the phases had a utilisation greater than one. In systems with the utilisations in the two phases both less than unity, Zukerman showed that as the switching rate was increased, the two-phase arrival process (MMPP) became *homogenised* that is converged to a simple Poisson process with rate given by the mean arrival rate of the MMPP. A discrete time queueing analysis for the same (two-phase, multi-server) system was presented in [Zukerman86b].

Other work on queues with dynamic parameters that extend on Zukerman's models comes from Japan [Fukuda87, Sotelo87, Sotelo88]. This work models the synchronous fluctuations of the arrival and service parameters, unlike the previous models which are based on their asynchronous fluctuations. In a synchronous model, phase changes occur only at the beginning of packet service or at a packet arrival, while in the asynchronous case, phase changes can occur even when a packet is being serviced. The synchronous equivalent of the two-phase single server model presented by Yechiali [Yechiali71] was solved by

Fukuda [Fukuda87]. This was extended to the multi-server case by Sotelo [Sotelo87]. Finally Sotelo [Sotelo88] extended the multi-server synchronous system to include n arrival and service phases.

Matrix geometric solution methods for queues operating in random environments have been widely used and are described in detail by Neuts [Neuts81]. Neuts emphasises care when solving for the steady state in such systems because local instabilities caused by temporary phases of overload can distort the steady state performance metrics.

State Dependent Switching

The fluctuations of arrival and service rate described in the work in the previous section are governed by an external random process. In contrast this section considers models with state dependent variations of the arrival and service parameters. State dependent service rate variations were emulated in the models presented in this dissertation.

The analysis of a queueing system with state dependent service rates and a Poisson arrival process was presented by Gross [Gross85]⁶. Both Gross and Li [Li89b] who studied the overload control of a finite buffer queueing system with state dependent parameter variations assumed instantaneous change between the parameter values. Li implemented overload control as a hysteresis mechanism by adapting the arrival and service rates when the queue crossed thresholds. The overload control model was used in simulations to control a voice traffic stream input to a packet multiplexer. Recently Lee [Lee89] has developed a transient analysis of the overload control scheme where the arrival rate of a switched Poisson process is regulated in a state dependent fashion. Such transient analyses are important because of the dynamism of queueing systems with fluctuating parameters.

[Neuts85] has considered the state dependent control of the arrival rate, in a hysteresis fashion, for an unbounded $M/G/1$ queueing system. In this model changes in arrival rate were assumed to take place only after an *actuation* time, which was assumed to be exponentially distributed.

A recent paper by Ohta [Ohta88a] models the effects of control and transmission delays on the performance of a scheme that controls the transmission rate of a traffic source in a communication network. Feedback based on the buffer occupancy at a congested link is used to control the transmission rate of the source. The control delay is the time required for the transmission of a control signal from the network node at which congestion occurs to the traffic source. The transmission delay is the time taken for the transmission of data from the source to the link. Control signals are transmitted to the source when the buffer occupancy exceeds an upper threshold to suppress transmission. Retransmission is enabled either after a fixed interval or after another control signal is received when the buffer occupancy falls below a lower threshold. Ohta solves the model by a fluid approximation method and shows the sensitivity of the packet blocking probabilities at

⁶pages 114-122.

the link to the two delays.

Optimal Control of Queues with State Dependent Switching

The models presented in this dissertation are based on heuristic bandwidth control schemes. An obvious extension would be to theoretically derive the control strategy to be used, from an optimal control formulation of the problem.

In [Yadin67] a class of control policies for changing the service rate in an $M/M/1$ system was considered (not derived). The policies were based on hysteresis – service rates were adjusted when the queue crossed thresholds. Steady state performance measures were evaluated, including the average queue size and the frequency of service rate changes. It was pointed out that the cost of policies could be formulated in terms of costs relating to the packet delay, the service rate and the frequency of service rate changes.

Gebhard [Gebhard67] evaluated the costs of two policies to change the service rate. The policies were proposed for an $M/M/1$ queueing system and were based on bi-level hysteresis control and point thresholding respectively.

[Heyman68] derived that the control policy to change the service rate should be a hysteresis one for an $M/G/1$ system with state dependent service rate. The server was assumed to be either in an on or an off state and start-up and shutdown costs for this variation were modelled. The policy was to turn the server off when the queue length was zero and to turn it on again when the queue had crossed a threshold K . Recently, Lu [Lu84] has derived optimal policies for $M/M/1$ queues where the service rate can vary between many levels and not just between on and off states. A list of other work on state dependent queueing systems, is to be found in a survey by Crabill [Crabill77].

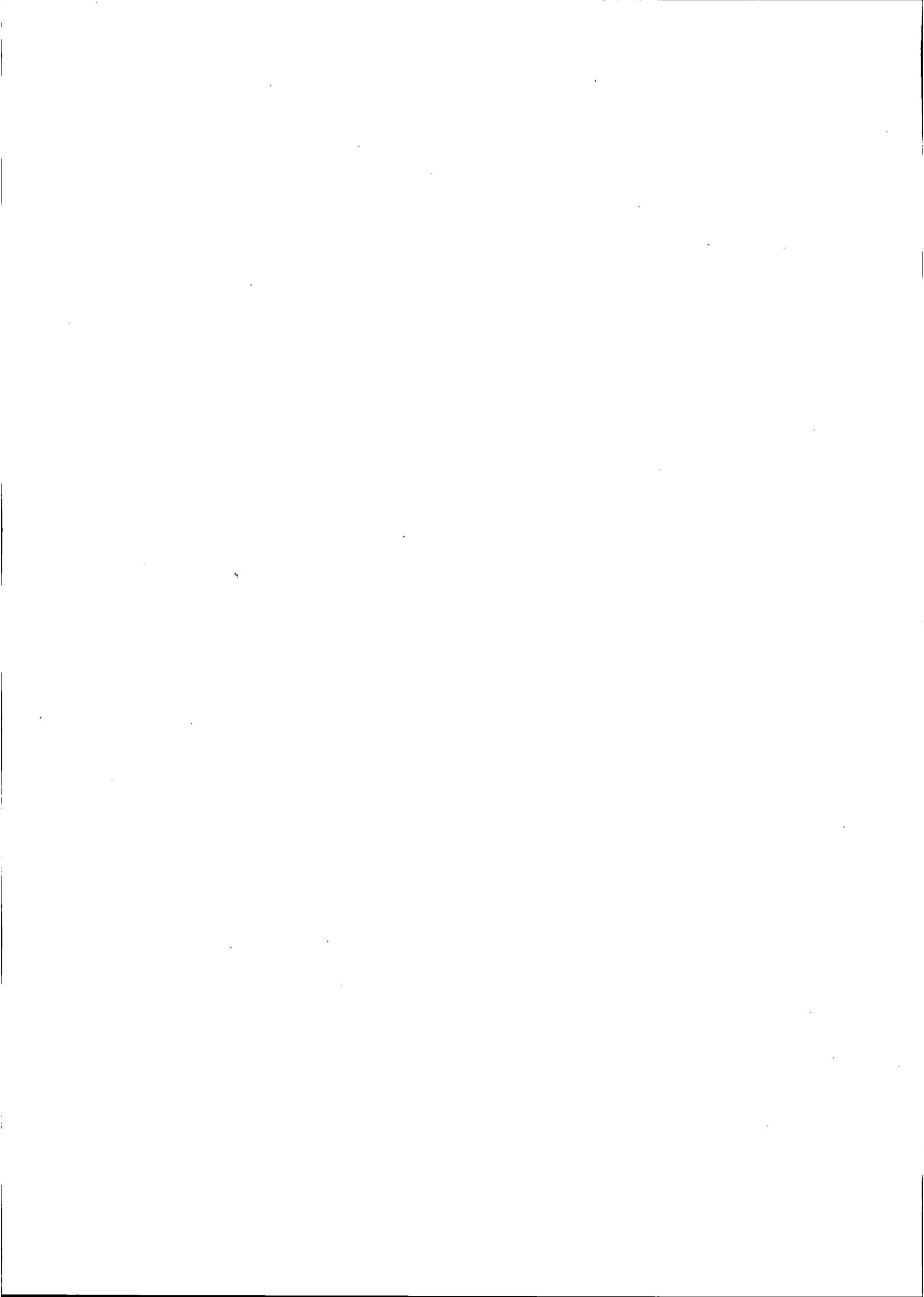
The theoretical derivation of bandwidth adjustment policies for the variable bandwidth U channel considered in this dissertation would be an interesting problem for future work. This would attempt to quantify the tradeoff between bandwidth utilisation (cost reduction) and traffic performance (quantified by metrics like the average queueing delays). The cost function to be optimised could include costs related to bandwidth usage, bandwidth switching frequency, packet queueing delay and the switching latencies.

7.12 Summary

The five models presented in this dissertation are for a variable bandwidth U channel operating with queue-based hysteresis control. The system was modelled as a single server queueing system with a variable service rate, which assumes one of two values. The control of the service rate was modelled by state dependent transitions in the associated CTMC model which was solved by using the MGF approach.

A general solution procedure was applicable to all the models. The MGFs were first formulated from the balance equations. The unknown state probabilities in the MGFs were solved for by extracting a sufficient number of linear independent equations from the balance equations, the conservation of probability sum condition and from the analytic properties of the MGF. The conservation of probability sum condition was used to derive the stability conditions for the models. In practice because of the numerical stability problems and the absence of closed form expressions for the performance metrics, the preferred approach for computing the steady state solutions was to use Gauss-Seidel iterations on the balance equations with improved precision arithmetic.

The necessary and sufficient condition for stability in the models was that the mean arrival rate (Poisson or MMPP) should be less than at least one of the service rates. The stability conditions do not explicitly include the bandwidth switching or sampling latencies. However, these parameters can affect the system performance; for example although the system may be stable in the steady state sense, the effect of a large switching latency can cause local instabilities with prolonged phases of overload and large queue build up.



Chapter 8

Experimental Programme 1

8.1 Introduction

This chapter and the following one present performance studies of dynamic bandwidth management from experiments conducted on an ATM testbed. The aim of the experimental programme is to evaluate the performance of dynamic bandwidth management schemes in reducing costs and controlling congestion in an ATM overlay running over a circuit switched ISDN.

This chapter begins with a description of the configuration used in the experimental programme. The validation of results predicted by theoretical and simulation models against experimental results forms the core of this chapter. In this work the performance of a simple queue-based bandwidth algorithm is compared for both MMPP and Poisson arrival processes. This is followed by the definition of a cost function based on the tariff structure implemented on the ISDN. The cost function is used with steady state performance metrics in differentiating between various rate and queue-based algorithms. This approach is used in the concluding section of the chapter to study the performance of a number of bandwidth algorithms when running a distributed computing application on the overlay.

8.2 Experimental Configuration

Figure 8.1 shows the configuration used in the experimental programme on the Unison testbed. The bandwidth control algorithms are implemented on the channel service (described in Section 3.3 of Chapter 3) running on a 68020 VME based system with a CFR interface. This system supports the TRIPOS [Richards79] operating system which provides a multi-threaded environment. The ramp and the channel service communicate

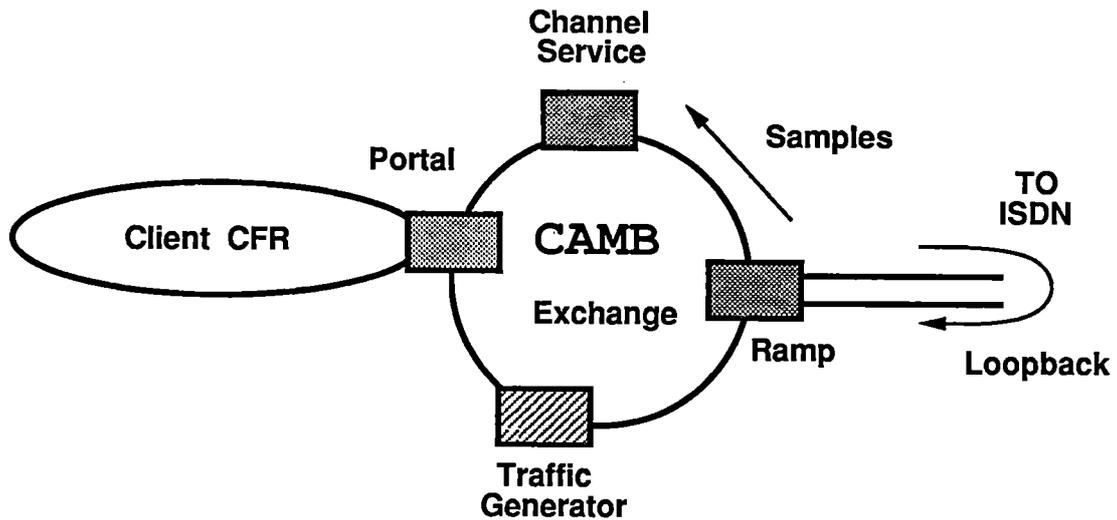


Figure 8.1: The Experimental Configuration

across the CFR through an RPC mechanism. The ramp sends out samples of its state at regular intervals to the channel service. The sampling period can be set from a minimum value of 0.25 seconds (4 Hz). Unless otherwise stated, the sampling rate used in the experiments is 2 Hz and the experiments were performed on British Telecom's ISDN, offering multiline-IDA.

The synthetic traffic load consists of CFR packets transmitted between pairs of machines (either 68020's or Acorn ARM's) each with a CFR UDL driver. The delay between transmitted UDL blocks (which in the experiments were composed of single packets) can be adjusted at the traffic generator to conform to various random distributions and synthetic load patterns.

The traffic generators can lie on either the exchange or client CFR's; when on a client CFR, packets are multiplexed through a CFR-CFR portal. When transmission is over the wide area, packets sent out by the generator are extracted by the ramp CFR receiver station and transmitted onto the ISDN, which in most experiments was configured in loopback mode. Packets are looped back at a System X switching exchange in Cambridge and returning packets at the ramp are sent out onto the exchange ring to be received by a station, either on the exchange ring or on a client CFR. In non-loopback configuration,

packets transmitted to a remote site (RAL) were received by a station on the remote exchange. The use of loopback mode was preferred in the experiments for reasons of cost and for ease of implementation. Except where otherwise stated the experiments to be described were performed in loopback mode.

Each U channel on the ramp has low and high priority queues associated with it. Packets multiplexed on the high priority queue have priority in transmission over those multiplexed on the low priority queue. The low priority queue can accommodate upto 20000 CFR packets while the high priority one has a smaller limit. The capacity of the low priority queue can be altered (made less than 20000 packets) by setting a queue threshold at the start of an experiment.

To control a U channel, the channel service executes a loop¹ with delay and bandwidth control elements. The bandwidth control element updates the state tables maintained at the channel service and makes decisions to adjust the bandwidth. These decisions are based on bandwidth algorithms that are rate or queue-based and can use state filtering and time lagging.²

8.3 Choice of Parameters in the Experimental Programme

The experimental programme was conducted on a public ISDN and for reasons of economy the bandwidths used on the U channels had to be small. For example in many of the experiments the commonly used service rates were 400 (2 B channels) and 800 (4 B channels) packets/sec when in loopback mode. The bandwidth switching latencies were calculated from measurements on the testbed and are further examined in Section C.2 of Appendix C.

The switching thresholds used in the rate-based algorithms were the permissible channel bandwidths ($N \times 200$ packets/sec, $N = 0, 1, \dots, 30$). Unless otherwise stated, point thresholding control was used for the rate-based algorithms. The thresholds for queue-based algorithms were chosen to avoid switching on transient variations in the queue and to limit the queueing delay that would result if switching were to be initiated at them.³

Commonly used values for the state filtering interval in the algorithms were 10 and 15

¹To control multiple U channels, multiple threads are used each running such a loop.

²For notational convenience, the following symbols are used to differentiate the bandwidth algorithms : Q/R — Queue/Rate-Based, I/A — Instantaneous/Averaged State, T — Time Lagging. As an example, a queue-based algorithm with state averaging and time lagging is denoted as QAT.

³As an example consider the parameters used in the experiment with a QI algorithm that are shown in Figure 8.2. The parameters are an average arrival rate of 700 packets/sec, a switching-in latency of 1.5 secs ($1/0.6873$), a bandwidth of 400 packets/sec (μ_0), a sampling interval of 1.0 secs (1 Hz) and an upper queue threshold of 700 packets. For these values the queue would build upto to a length of about 1450 packets before switching takes place. This peak value was computed from the approximation : $700 + (1.5 + 1.0) \times (700 - 400)$. Equivalently this corresponds to a peak queueing delay of ≈ 2.0 secs (using Little's Law).

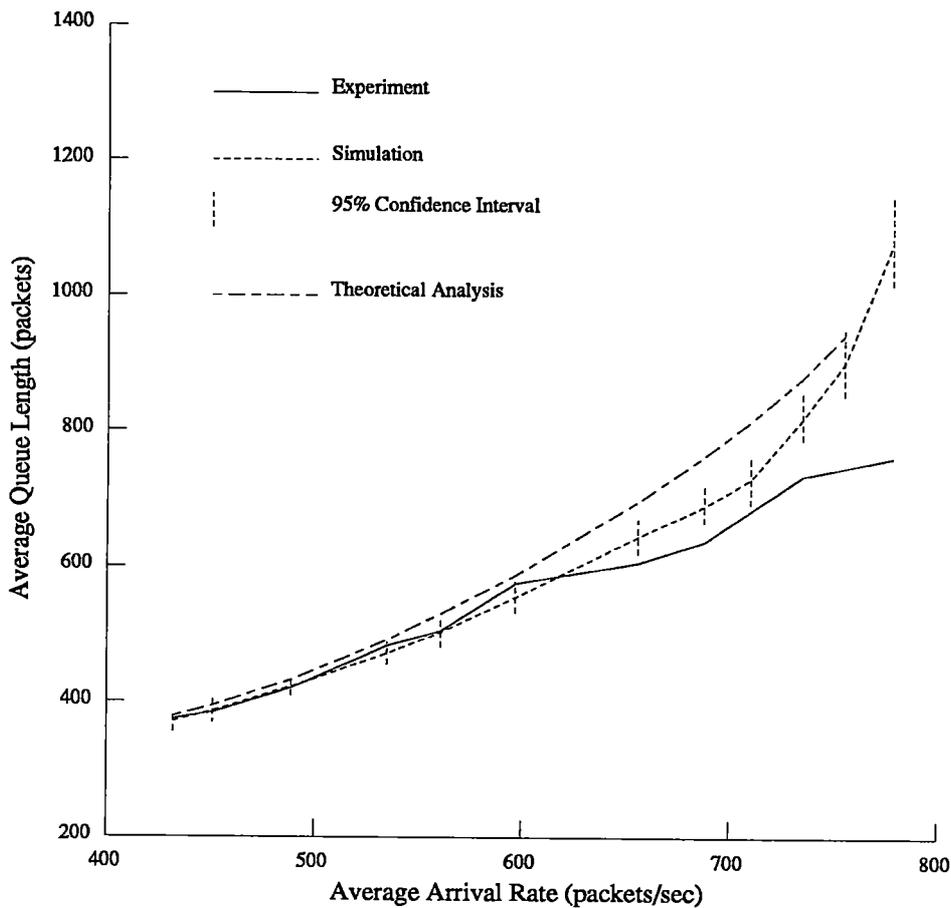


Figure 8.2: Average Queue Length vs. Average Arrival Rate for Poisson arrivals (QI Algorithm) — Comparison of Theoretical Analysis (Model D), Simulation and Experiment. Parameter Set 1 : $\mu_0=400$, $\mu_1=800$, $\gamma_{in}=0.6873$, $\gamma_{out}=2.5907$, $\delta=1.0$, Lower Threshold=50 packets, Upper Threshold=700 packets.

seconds which corresponds to averaging over 20 and 30 state samples when the sampling rate is 2 Hz. The length of the filtering interval could be chosen to vary the reactivity of a bandwidth algorithm. A common value chosen for the time lagging interval was 15 seconds which lies at the lower end of the range of charging intervals for switched B channels on the public ISDN (Loopback mode and to RAL).

Although not considered in the experimental programme a useful extension would be to study the dynamic variation of the control parameters of the bandwidth algorithms in response to the input load characteristics and the charging policies.

8.4 Performance Analysis of Queue-Based Bandwidth Algorithms

The results presented in this section describe the performance of a simple queue-based algorithm driven by the instantaneous length of the channel queue (QI algorithm). Selected experimental results are compared against those obtained from simulations and theoretical analyses of the models D and E described in Sections 7.5 and 7.6 of Chapter 7 for Poisson and MMPP arrival processes respectively. Other results presented in this section include a comparison of point thresholding and hysteresis control mechanisms and graphs of survivor functions [Daigle86]⁴ The average queue length, the average queueing delay and the steady state queue length distribution from the survivor functions are the main metrics used in the performance analyses.

Evaluation of Queue-Based Control with Poisson Arrivals

Comparison of Results from Experiment, Simulation and Theoretical Analysis

- Queue-based algorithm with hysteresis control (QI)
- Poisson arrival process
- Results from experiment, simulation and theoretical analysis (Model D)
- Parameter Sets 1 and 2

This section compares steady state performance results from experiment, simulation and theoretical analysis (Model D) of a variable bandwidth channel controlled by a simple queue-based bandwidth algorithm (QI) and with Poisson arrivals.

Figures 8.2 and 8.3 display the variation of the average queue length with the average arrival rate λ of a Poisson arrival process for two different parameter sets.⁵ A hysteresis control scheme is used, with the same lower threshold of 50 packets in each of the parameter sets but with different upper thresholds of 700 and 1200 packets. It is assumed that the system switches between the service rates, $\mu_0 = 400$ packets/sec and $\mu_1 = 800$ packets/sec, where $\mu_0 < \lambda$ while $\mu_1 > \lambda$, the latter inequality being the condition for stability. The values of the bandwidth switching-in and switching-out latencies γ_{in}^{-1} and

⁴A survivor function details the spread of the steady state queue length distribution. It is a plot of the complementary cumulative distribution function of the queue length.

⁵The symbols used in the parameter sets in this chapter are : λ — Average arrival rate of a Poisson process in packets/sec, λ_0 / λ_1 — Average arrival rates in the two phases (0/1) of an MMPP in packets/sec, μ_0 / μ_1 — Average service rates in s^{-1} , $\gamma_{in} / \gamma_{out}$ — Average bandwidth switching-in/out rates in s^{-1} , δ — Average sampling rate in s^{-1} or Hz, ρ_0 / ρ_1 — Average switching rate from arrival phase 0/1 to arrival phase 1/0 in s^{-1} .

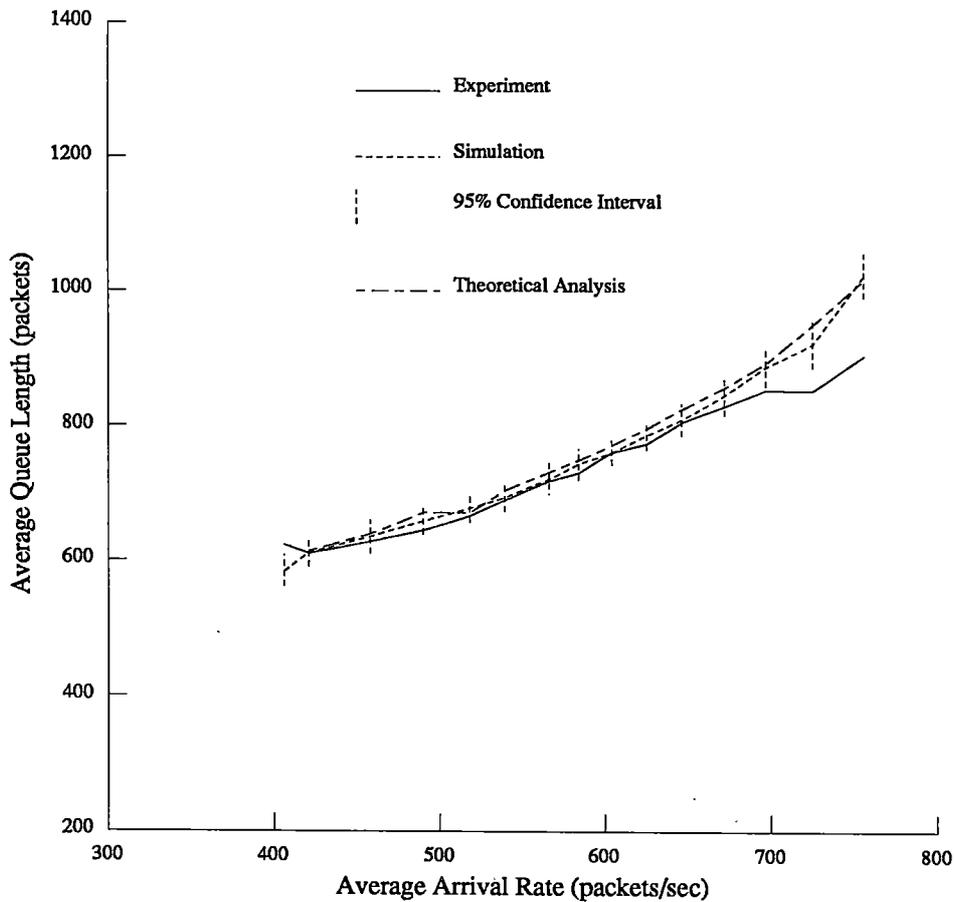


Figure 8.3: Average Queue Length vs. Average Arrival Rate for Poisson Arrivals (QI Algorithm) — Comparison of Theoretical Analysis (Model D), Simulation and Experiment. Parameter Set 2 : $\mu_0=400$, $\mu_1=800$, $\gamma_{in}=0.8045$, $\gamma_{out}=2.809$, $\delta=2.0$, Lower Threshold=50 packets, Upper Threshold=1200 packets.

γ_{out}^{-1} were calculated from the runs of the experiment.⁶ The values of $\delta=1.0$ and 2.0 in the two sets represent sampling frequencies of 1 and 2 Hz, respectively.

Analysis of Results

Figures 8.2 and 8.3 display the increase of average queue length with the average arrival rate λ obtained from experiment, simulations and from the theoretical analysis of Model D.⁷

⁶A faster version of the ramp software was used for the second experiment which explains the higher values of γ_{in} and γ_{out} .

⁷Equivalently the average queue length could have been plotted against the traffic intensity λ/μ_0 .

For small values of the average arrival rate there is a closer correspondence between the traces derived from theoretical analysis, experiment and simulation than at larger values, in both the figures. As the arrival rate increases, there is a fanning out among the traces especially in Figure 8.2 with the experimental one becoming a lower bound. The main reason for the divergence is the assumptions on the service disciplines made in the theoretical and simulation models.

The theoretical model (Model D) assumed a continuous, non-gated service discipline with exponentially distributed packet service times (for reasons of analytical tractability). Exponential distributions were also assumed for the sampling interval and the bandwidth switching latencies but these assumptions are more practical; sampling and switching latencies are indeed random variables, albeit with small variances. The exponential assumptions on service times and the other parameters can be used to explain the conservative performance predicted by the theoretical models, especially at larger loads. The over-estimation of delay in CTMC models with exponentially distributed parameters as in Model D, was pointed out by Sriram [Sriram83] and Daigle [Daigle86] (Sections 5.2.1 and 5.2.2 of Chapter 5).

The trace for the simulation model displays the 95 percent confidence intervals and the average queue length for each arrival rate. The smaller divergence between the simulation and the experimental results at low and medium values of the arrival rate ($\lambda < 650$ on the two figures) compared to the divergence between the theoretical and experimental results is to be noted. Further the simulation trace lies below the theoretical trace for most values of the arrival rate in both the figures. The simulations better estimate the performance of the actual system because they assume deterministic values for the service times and the sampling intervals. It is only the bandwidth switching latencies that are assumed exponentially distributed in the simulation models. The increase in length of the 95 percent confidence interval with the average arrival rate is to be observed in both the figures.

Comparison of Point Thresholding and Hysteresis Control for Queue-Based Control with Poisson Arrivals

- Queue-based algorithm with hysteresis control (QI)
- Poisson arrival process
- Results from experiment
- Point Thresholding vs. Hysteresis Control
- Alvey High Speed Network

Figure 8.4 displays traces obtained from an experiment on the Alvey High Speed Network (AHSN) to compare the performance of the hysteresis and point thresholding control

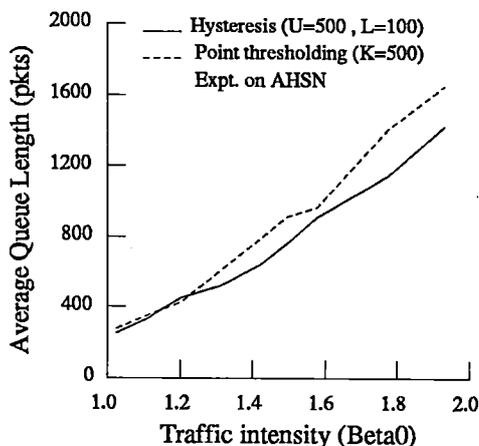


Figure 8.4: Average Queue Length vs. Traffic Intensity β_0 for Poisson Arrivals (QI Algorithm) — Point Thresholding vs. Hysteresis Control : $\mu_1=800$, $\mu_0=400$, $\delta=0.5$.

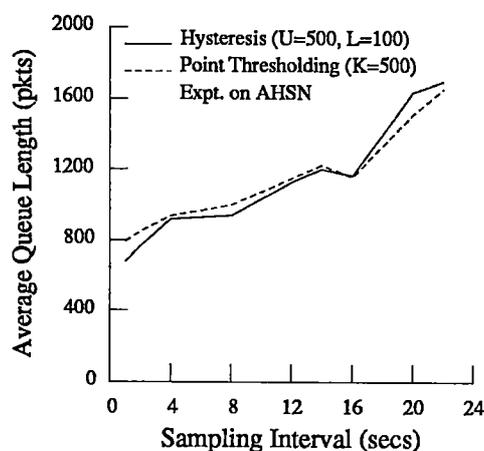


Figure 8.5: Average Queue Length vs. Sampling Interval for Poisson Arrivals (QI Algorithm) — Point Thresholding vs. Hysteresis Control : $\lambda=568$, $\mu_0=400$, $\mu_1 = 800$.

schemes for a variable bandwidth channel. The average queue length for the two control mechanisms is plotted against the traffic intensity β_0 ($\beta_0 = \lambda/\mu_0$) where λ is the average arrival rate of the input Poisson process and μ_0 is the lower of the two service rates used. The upper threshold of the hysteresis control and the point threshold value were both chosen to be 500 packets; the lower threshold for the hysteresis scheme was 50 packets.

Figure 8.4 displays a fanning out in the two traces as β_0 increases. As the load increases, the conservatism of the hysteresis scheme in switching out bandwidth becomes more prominent when compared to the point thresholding one.

Figure 8.5 displays the results of another experiment on the AHSN that compares the performance of the two schemes. The average queue length is plotted against the sampling interval which is the interval between successive queue samples generated at the ramp. A Poisson arrival process with traffic intensities $\beta_0 = 1.42$ and $\beta_1 = 0.5$ in the two switching phases is used to model the input traffic flow. The threshold values used were the same as in the previous experiment.

As the sampling interval increases and samples are received less frequently at the channel service, the state information is refreshed less often. This has the effect of decreasing the reactivity of the controller. The actual bandwidth adjustment latency consisting of call signalling and U channel synchronisation takes the same time as when the sampling interval was small, on average. However, the adjustment process is now generally initiated after a relatively larger delay because the state is updated less often. This results in larger

instantaneous and average queue lengths than when the sampling was more frequent.⁸

From Figure 8.5 it can be seen that there is little difference between the average queue lengths observed for the hysteresis and point thresholding schemes over the variation of the sampling interval. As the state information maintained at the controller becomes less accurate (in time), the effect of hysteresis tends to be negated.⁹

Evaluation of Queue-Based Control with MMPP Arrivals

Comparison of Experiment, Simulation and Theoretical Results

- Queue-based algorithm with hysteresis control (QI)
- MMPP arrivals
- Results from experiment, simulation and theoretical analysis (Model E)
- Parameter Sets 3 and 4

This section compares steady state performance results from experiment, simulation and theoretical analysis (Model E) of a variable bandwidth channel controlled by a simple queue-based bandwidth algorithm (QI) and with MMPP arrivals.

Figures 8.6 and 8.7 display the variation of average queue length with the mean arrival rate of a two-phase MMPP process for two different parameter sets. Bi-level hysteresis control is assumed with the same lower threshold of 50 packets for each parameter set but differing upper thresholds of 1000 and 1600 packets. The service rates can take one of two values — $\mu_0 = 400$ packets/sec and $\mu_1 = 800$ packets/sec. The bandwidth switching rates, γ_{in} and γ_{out} , were obtained from experiment.

For stability, the necessary and sufficient condition derived in in the analysis of Model E was that the mean arrival rate of the MMPP, which is given by $\lambda^* = (\lambda_0 * \rho_1 + \lambda_1 * \rho_0) / (\rho_0 + \rho_1)$, should be less than μ_1 . In the two experiments, the duration of the arrival phases were assumed to be exponentially distributed, each with an average value of 10 seconds so that ρ_0 and ρ_1 are each 0.1 Hz. The average packet generation rate in the first arrival phase, λ_0 , was set at 350 packets/sec. This meant that the average packet generation rate in the second arrival phase λ_1 could be varied up to a maximum value of

⁸Even though the switched bandwidth is retained for a longer time because of the infrequent sampling.

⁹The experiments corresponding to Figures 8.4 and 8.5 were performed on the Alvey High Speed Network. The bandwidth switching-in and switching-out times on the AHSN, were considerably greater than on the current ISDN offering multiline-IDA. Increasing the bandwidth by 2 B channels could take as long as 5.2 seconds, with a signalling component of 3.8 seconds and a synchronisation component of 1.4 seconds. In the experiments over the AHSN the channel service had to maintain separate interfaces to the signalling and ramp services, in contrast to the integrated interface in the current version, so that bandwidth adjustment involved multiple RPC transactions.

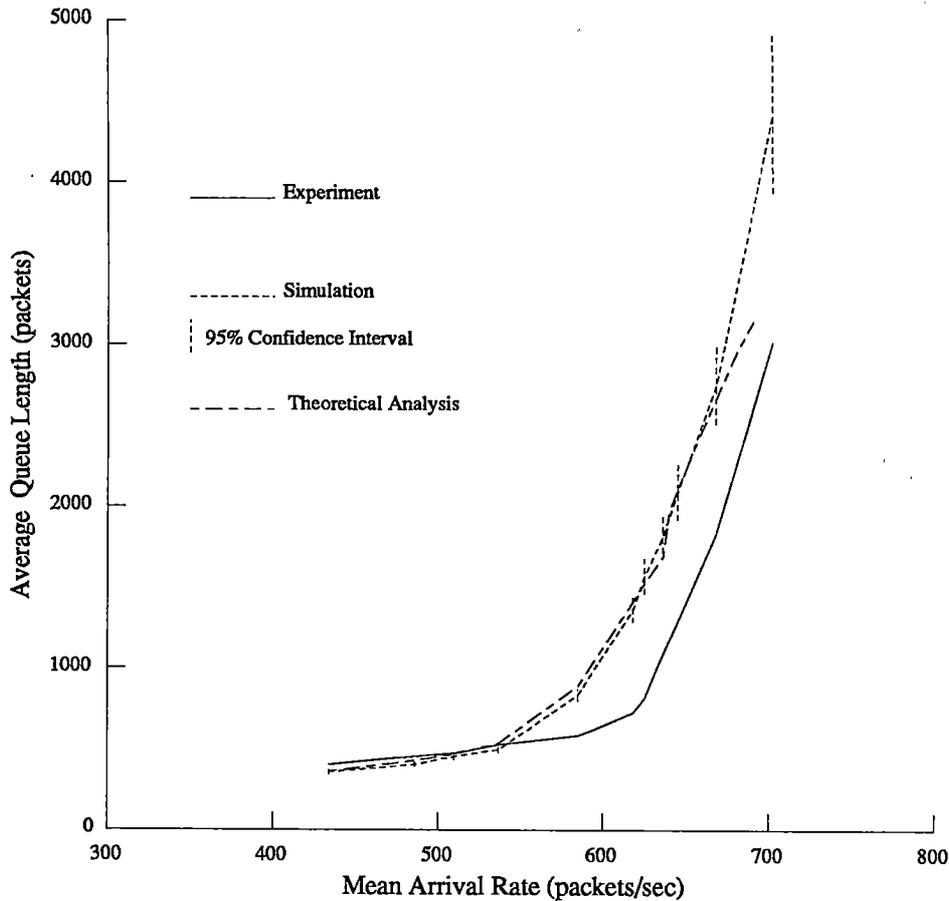


Figure 8.6: Average Queue Length vs. Mean Arrival Rate for MMPP Arrivals (QI Algorithm) — Comparison of Theoretical Analysis (Model E), Simulation and Experiment. Parameter Set 3 : $\mu_0=400$, $\mu_1=800$, $\gamma_{in}=0.7955$, $\gamma_{out}=2.7855$, $\delta=2.0$, Lower Threshold=50 pkts, Upper Threshold=1000 pkts, $\rho_0 = \rho_1=0.1$, $\lambda_0=350$.

1250 packets/sec.¹⁰

Analysis of Results

From the figures it is evident that the growth in average queue length has the same increasing trend for both parameter sets. However, the average queue length for the same value of arrival rate is greater in the second figure — this is because of the higher threshold value, other parameters in the two sets being comparable. The three traces in each figure display a sharp increase in queue size, as the mean arrival rate approaches the stability limit of 800 packets/second. In the experiments, values of the mean arrival rate which were closer to this limit were used. These are not plotted on the figures as in several of the

¹⁰Computed from the stability condition, which for this set of parameter values is $\lambda_1 + 350 < 1600$, where $\lambda_0 = 350$ packets/sec, $\mu_1 = 800$ packets/sec and $\rho_0 = \rho_1=0.1$ Hz.

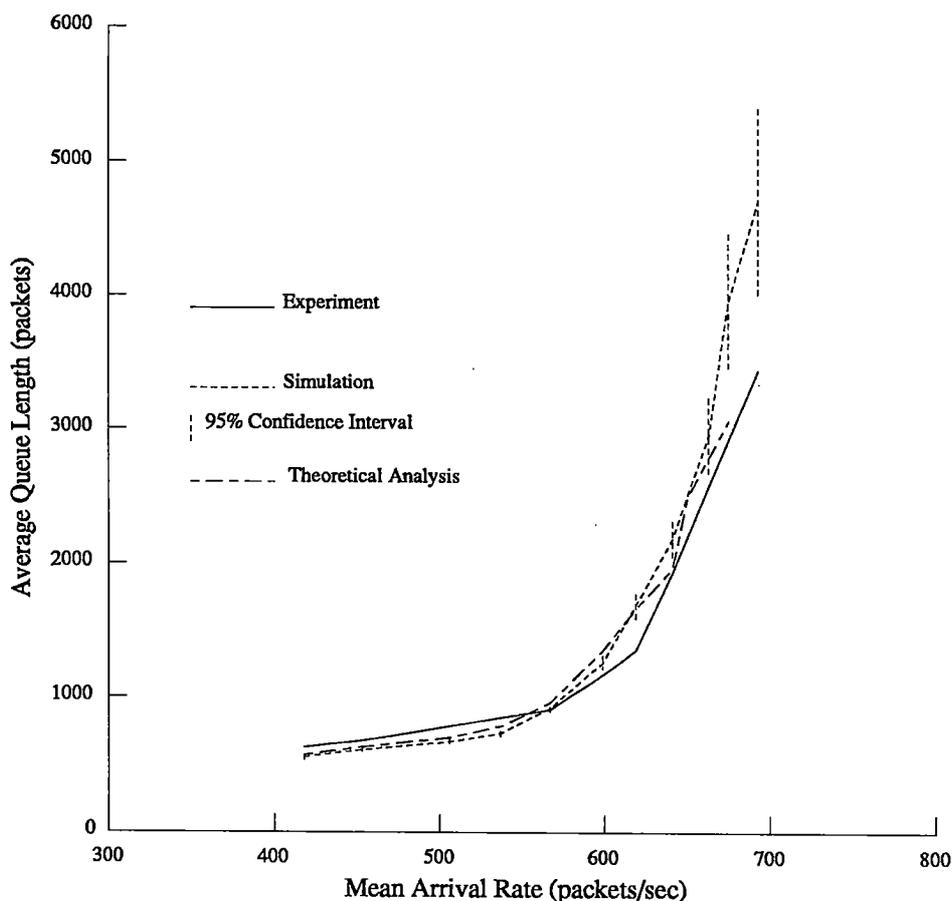


Figure 8.7: Average Queue Length vs. Mean Arrival Rate for MMPP Arrivals (QI Algorithm) — Comparison of Theoretical Analysis (Model E), Simulation and Experiment. Parameter Set 4 : $\mu_0=400$, $\mu_1=800$, $\gamma_{in}=0.7943$, $\gamma_{out}=2.8169$, $\delta=2.0$, Lower Threshold=50 pkts, Upper Threshold=1600 pkts, $\rho_0 = \rho_1=0.1$, $\lambda_0=350$.

runs it was observed that the instantaneous queue length exceeded the maximum buffer limit of 20000 packets and packets were lost.

The deviation between the traces for average queue length for theoretical analysis, simulation and experiment in Figures 8.6 and 8.7 can be explained in a similar fashion to the previous experiments with Poisson arrivals. The limiting assumptions causing the over-estimation of queue length at larger values of the mean arrival rate by the simulation and theoretical models were those made on the service disciplines. Both the simulations and theoretical models assume exponential distributions for the bandwidth switching latencies. In addition, the theoretical models assume exponential distributions for the packet service times and the sampling intervals. However, unlike the previous experiment with Poisson arrivals, the simulation and theoretically derived queue length traces with MMPP arrivals

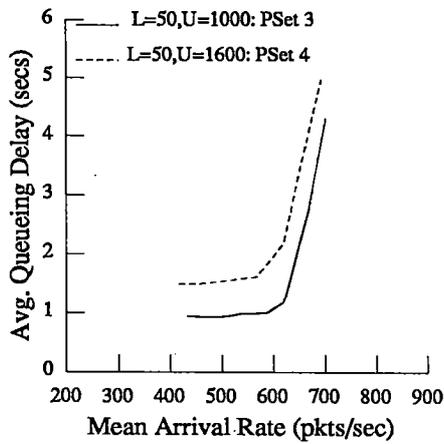


Figure 8.8: Average Queueing Delay vs. Mean Arrival Rate for MMPP arrivals (QI Algorithm) — Experiment. Parameter Sets 3 and 4.

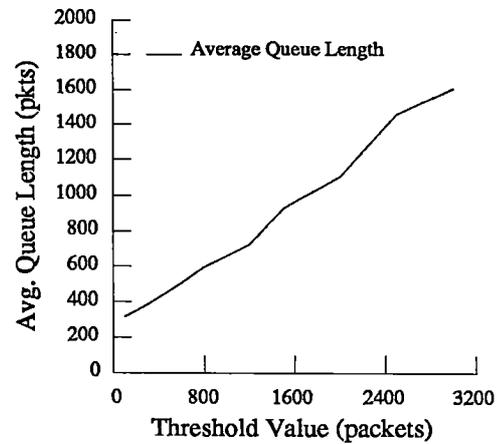


Figure 8.9: Average/Variance of Queue Length vs. Upper Threshold for MMPP arrivals (QI Algorithm) — Experiment. Parameter Set 5 : $\mu_0=400$, $\mu_1=800$, $\gamma_{in}=0.7974$, $\gamma_{out}=2.7855$, $\delta=2.0$, Lower Threshold=50 pkts, $\lambda_{avg}=595$, $\rho_0 = \rho_1=0.1$.

lie much closer together.¹¹ As before for the simulation results the 95 percent confidence intervals increase with the arrival rate.

Average Queueing Delay for MMPP Arrivals

- Queue-based algorithm with hysteresis control (QI)
- MMPP arrivals
- Results from experiment
- Parameter Sets 3 and 4

Figure 8.8 displays the average queueing delay plotted against the mean arrival rate, with values generated from experiment for the two parameter sets (3 and 4) detailed in the

¹¹In the figures it can be noted that the theoretical models tend to underestimate the average queue length for large values of the mean arrival rate when compared to the simulation models. A possible reason for this is that as the mean arrival rate increases so does the queue length. The iteration upper bound (the number of queue states iterated over which was set at 12000) in the Gauss-Seidel type scheme used to solve the Models D and E may be insufficient and would have to be increased at the expense of slower convergence.

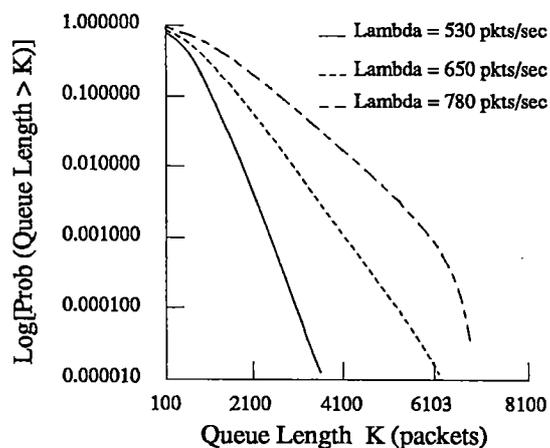


Figure 8.10: Survivor Functions for Poisson Arrivals from Theoretical Analysis (QI algorithm — Model D). Log[Prob(Queue length > K) vs. Queue Length K. Parameter Set 1.

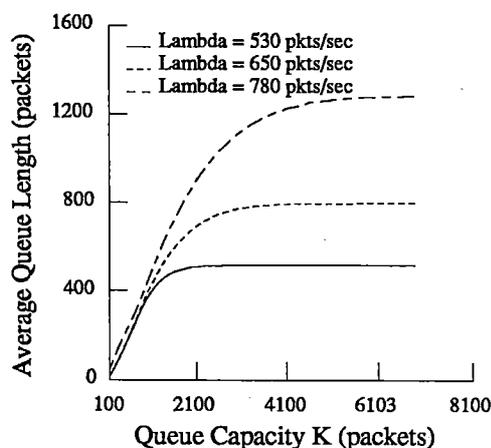


Figure 8.11: Average Queue Length vs. Queue Capacity K for Poisson Arrivals from Theoretical Analysis (QI algorithm — Model D). Parameter Set 1.

previous section. The average queueing delay was calculated from Little's Law — using the average values of queue length and mean arrival rate from the previous section. The same pattern of variation observed in Figures 8.6 and 8.7 is also observed in this figure for each of the traces. The trace corresponding to the higher threshold value (Parameter Set 4) is displaced upwards. Such a linear shift is not surprising given the near linear variation of average queue length with threshold value for a queue-based algorithm and MMPP arrivals which is presented in Figure 8.9. This trace was plotted from experimental results and displays the variation with upper threshold of the average queue length for a hysteresis control mechanism.

Dynamic Bandwidth Control with Queues of Limited Capacity

- Queue-based algorithm with hysteresis control (QI)
- Poisson Arrivals
- Results from theoretical analysis (Model D)
- Parameter Set 1

In the previous experiments described in this chapter, the use of the low priority queue with a capacity of 20000 packets and the choice of parameter values generally ensured that packets were not lost due to overflow when the system was running at overload. This

section examines the effect that a lower limit on the queue capacity can have on queue-based bandwidth control. Limiting the queue capacity is important to satisfy performance constraints on the maximum queueing delay and jitter experienced by packets, but comes at the expense of packet blocking and packet loss. The results in this section were derived theoretically from a consideration of survivor functions.

Figure 8.10 displays survivor functions computed from theoretical analysis (Model D) for queue-based bandwidth control with Poisson arrivals and for the Parameter Set 1. The survivor function displays the complementary cumulative distribution of queue length — the probability that the queue length exceeds K , for each state K of the system and has a decaying characteristic as long as the system is stable. Survivor functions are useful in evaluating the probability of experiencing a certain queueing delay and in calculating the packet loss caused by a limit on the queue capacity. The three traces in the figure, plotted with a logarithmic scale along the Y axis, show how the spread of the queue length distribution increases with an increase in the input arrival rate, λ .

Figure 8.11 shows the variation with the queue capacity of the average queue length for the same parameters used in the previous figure. It is interesting to note how the slopes of the traces decrease with increasing queue capacity, which provides an indication of the queue capacity necessary to maintain constraints on the average queue length.

To study the effect of finite buffers on the bandwidth control requires a consideration of both the figures. If packet loss can be tolerated but has to be constrained, as has the queueing delay, then the variations in Figures 8.10 and 8.11 can be used to estimate the queue threshold to be used. For example, at an average arrival rate of 530 packets/sec and with a cell loss constraint of 0.0001, a threshold value of 2950 packets would give an average queue length of 515 packets. This value would ensure that the maximum queue length would be 2950 packets and the maximum queueing delay at this arrival rate would be 5.64 seconds (from Little's Law).

8.5 A Cost Function for Dynamic Bandwidth Management

The efficiency of a bandwidth algorithm can be measured in terms of the performance it offers to end applications and the cost it incurs in doing so. The former can be quantified by statistical measures like the average queue length, average queueing delay and jitter on the delay. The cost is related to the bandwidth utilisation, the number and frequency of bandwidth changes and the charging interval. The objective of the controller would be to try to minimise the cost while satisfying constraints on the queueing delay and jitter in order to maintain a quality of service to the end user.

In the public ISDN used in the Unison network, charges consist of a fixed sum towards the hire of equipment and a bandwidth usage charge made on a per-call (B channel) basis just as for an ordinary telephone subscriber. The cost depends on the destination of a call, its duration and the time of day at which it was made. The 'switching cost' is the

minimum charge levied per call; depending on the time of the day and the call destination, this provides a certain retention interval. After this interval, further retention costs are imposed at the start of each subsequent charging interval until the circuit is disconnected.

This discreteness in charging influences whether frequent switching is cost effective. When the charging interval is small (eg. on an international call), the switching cost is relatively low but the retention cost is high. However, when the charging interval is large (eg. a local call made in the middle of the night) the retention cost is relatively low but the switching cost is proportionately higher. From a cost perspective, delaying the reclamation of switched bandwidth is a useful policy when the retention interval is large, but may not be cost efficient over shorter charging intervals. When the retention interval is large,¹² switching out bandwidth at the end of the interval costs the same as when switching it out in between, so that it makes sense to keep the connection 'alive' for at least that long, which is the rationale behind time lagging.

From a performance viewpoint, the longer the switched bandwidth is retained, the better the traffic performance and so slow reclamation is a desirable feature of dynamic bandwidth management. Fast reclamation and excessive switching may not be desirable for use in public switching exchanges. Bandwidth control will lead to a tradeoff between traffic performance and cost reduction. In the next section and in the experiments to be described in the next chapter this tradeoff will be examined in more detail.

In the experimental programme in this dissertation, a representative cost function was computed from a history of bandwidth changes during experimental runs. The connection times for each switched B channel is known. The ratio of this time and the applicable charging interval was rounded off to the next higher integer value, giving the number of cost units (CUs) for the circuit. The minimum cost for a switched circuit — the switching cost will be one CU. The costs referred to in this chapter are in terms of CUs and include only the switched bandwidth, not the statically allocated bandwidth (the base bandwidth) which is always maintained on the U channel.

In the experiments to be described in the next section and in Section 9.2 of the next chapter, the cost function is used along with steady state performance metrics to study the performance of the queue and rate-based algorithm families, for different traffic loads and charging intervals.

8.6 A Distributed Computing Application — Microemac Make

This section presents a performance analysis of the rate and queue-based bandwidth algorithm families conducted on the AHSN, for an end user application running over the

¹²Or as in the United States where the retention costs for the first minute are more expensive than for subsequent minutes.

Algorithm	M. Time(s)	BW Util.	Call Conn.	CL	CM	CS
RIT	425	54.8	13	16	41	335
RAT	446	19.1	3	6	20	179
RA	464	14.7	3	6	16	139
RI	476	21.8	28	30	34	147
QIT	482	25.5	6	7	20	146
QAT	511	4.3	2	2	6	47
QI	511	3.1	5	6	6	24
QA	514	5.8	2	2	4	34
2 Slots	538	-	-	-	-	-
4 Slots	419	-	-	-	-	-
30 Slots	353	-	-	-	-	-
CFR/Ethernet	333	-	-	-	-	-
Ethernet	322	-	-	-	-	-

Table 8.1: Cost and Performance Comparison of Rate and Queue-Based Algorithms — Microemacs ‘Make’ Test. AHSN. Legend : **Algorithm** — Bandwidth Algorithm, **M. Time** — ‘Make’ Time (seconds), **BW Util.** — Switched Bandwidth Utilisation or Percentage of the test time for which 4 slots were in, **Call Conn.** — Number of Call Connections made (low to high), **CL** — Cost over Long Charging Interval (40 secs), **CM** — Cost over Medium Charging Interval (10 secs), **CS** — Cost over Short Charging Interval (1 sec).

network.¹³ The application is a ‘make’¹⁴ of microemacs where the host machine (Acorn R140 running the UNIX operating system) has access to a Sun file server running the Network Filing Service (NFS) through a pair of IP portals, Unison exchanges and the ISDN. (The IP gateways lie on the exchange and treat the Unison network as an IP subnet). The ‘make’ process is sensitive to the round trip delay since the NFS interactions are RPC based and it was of interest to study the tradeoff between the increase in ‘make’ time and the reduction in cost for the various algorithms.

Table 8.1 explores this tradeoff for a family of algorithms that are driven by rate or queue size information from the ramps. The algorithms are distinguished on the basis of the state used (rate/queue), whether the state was filtered before use and whether time hysteresis was utilised or not. All the algorithms employed hysteresis in state — the rate-based algorithms used upper and lower threshold values of 400 and 200 packets/s, while the queue-based algorithms used values of 300 and 50 packets. The *time lagging* interval was 15 seconds while the averaging was over 10 sampling intervals, with the sampling rate being 0.5 Hz. A two-phase switching process was used with service rates of 400 (the base

¹³This experiment was done with D. McAuley of the Computer Laboratory. The present author’s contribution was the dynamic bandwidth management.

¹⁴In its simplest form a make is the process of compiling and linking the components of a source program.

bandwidth) and 800 packets/sec.

The table displays the make times taken for the test when dynamic bandwidth management was used. The table also displays the switched bandwidth utilisations which are the percentages of the total test time for which the system runs at the higher service rate — a measure of the bandwidth conservation. The number of call connections made by each algorithm are listed as are the cost functions computed over three charging intervals — 1, 10 and 40 seconds. Other entries in the table include the test times when the bandwidth is statically fixed at 2, 4 or 30 slots. The times in the last two rows are for the experiment carried out directly over the Ethernet and over a CFR and Ethernet (without the ISDN in both cases).

The make times when the test is run over the ISDN span a range of values from 353 seconds when 30 slots (B channels) were used, to 538 seconds when only two were used. The times for the various rate and queue-based bandwidth algorithms lie between those for the static bandwidth allocations of 2 slots (538 seconds) and 4 slots (419 seconds), with the rate-based ones occupying the lower end of the scale.

When interpreting the results in the table and the two graphs in Figure 8.12, it should be noted that they are for a specific application (RPC interactions with a file server) and for a particular choice of thresholds. However, some general characteristics of the performance of the algorithms can be pointed out. Rate-based algorithms have lower make times than the queue-based algorithms (except for the QIT algorithm) because of their relatively larger bandwidth utilisations. Further, the rate algorithms which are driven by the instantaneous rate (RI and RIT) are more reactive than the corresponding members of the queue-based family, as exemplified by the larger number of bandwidth changes.

The use of the cost function is best illustrated by comparing the performance of the RI and RIT algorithms. Although the number of bandwidth changes for the RI algorithm is almost twice that of the RIT algorithm, the latter has a greater switched bandwidth utilisation — 55 against 22 percent for the RI algorithm. When the charging interval is large, the switching cost predominates and the RIT algorithm has a lower cost, but as the charging interval decreases, the switched bandwidth utilisation becomes more important and the cost of the RIT algorithm increases. The relative cost of the RI algorithm improves as the charging interval decreases.

Over the short charging interval, time lagging tends to increase the cost especially if the state is not filtered. This is exemplified by the QIT algorithm which over the large charging interval is efficient both in make times and cost when compared to the other queue-based algorithms. However it is relatively more expensive over the short charging interval. The increase in cost of the QAT algorithm over the short charging interval is not so pronounced because of its decreased reactivity due to filtering which ensures that bandwidth is not switched in often.

From a study of the input traffic patterns it was found that while the traffic was bursty, the bursts were generally short in duration. While the bursts were sufficient to trigger

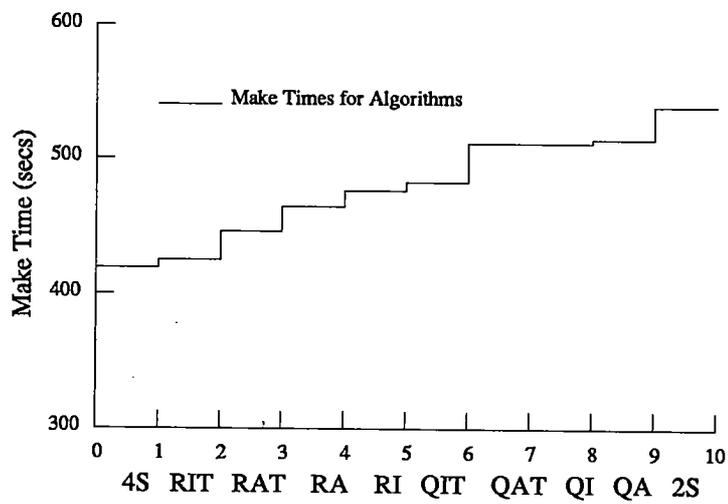
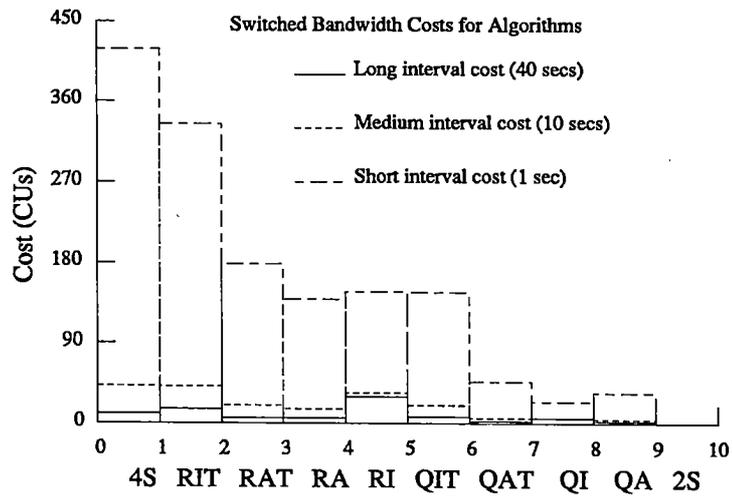


Figure 8.12: Cost and Performance Comparison of Rate and Queue-Based Algorithms — Microemac's 'Make' Test. Graphical Representation.

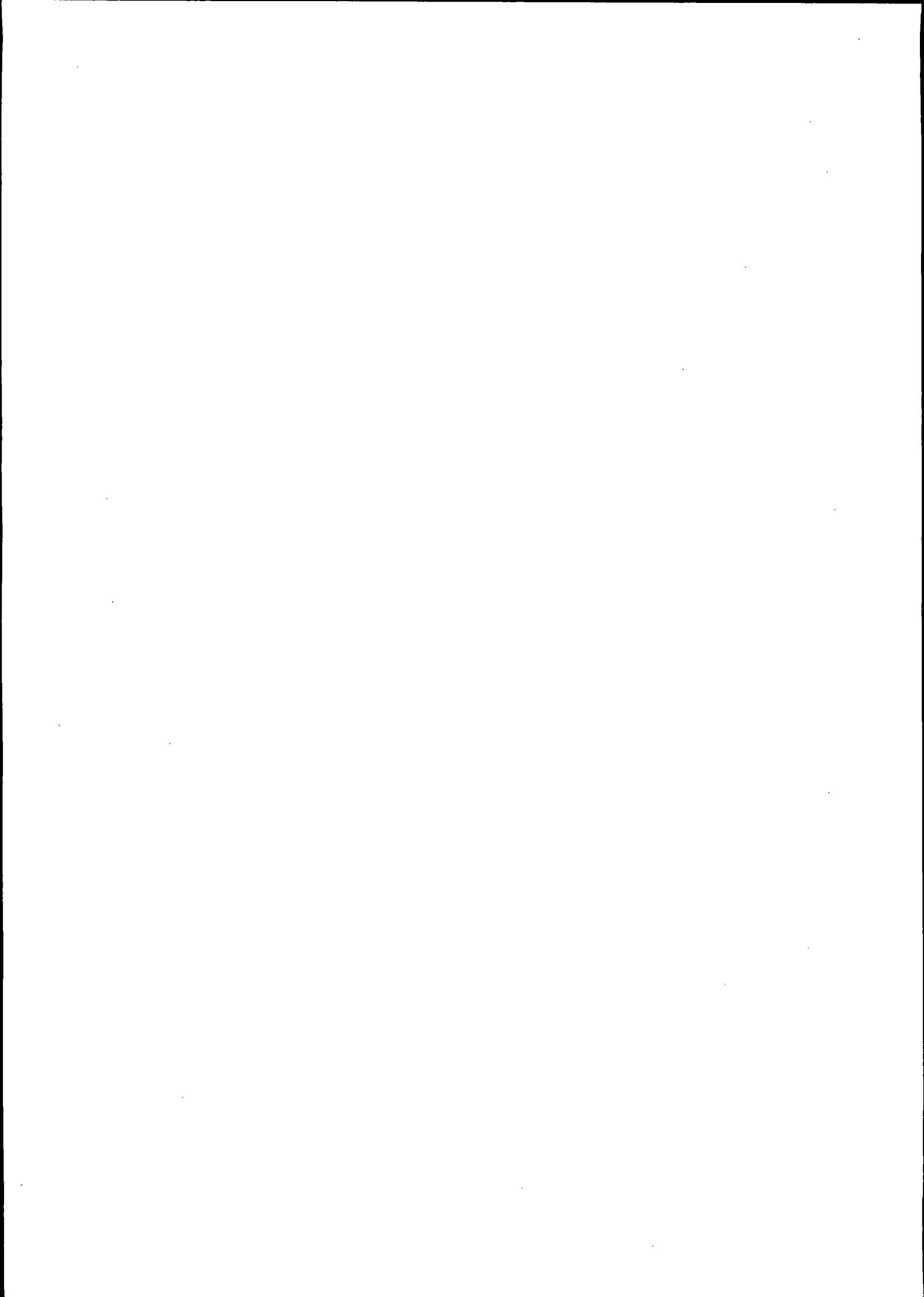
rate-based algorithms, the build up in queue size was often insufficient for the queue-based algorithms to switch in bandwidth, especially when they used state filtering.

8.7 Summary

The validation of experimental results for simple queue-based bandwidth control against those predicted by theoretical analyses and simulations formed the central theme of this chapter. The performance results served to demonstrate the adequacy of the performance models described in chapters 6 and 7. Considering the assumptions that were made on the distributions of the model parameters and the structure of the multiplexer, a sufficiently close correspondence between the theoretical and experimental results was observed. The simulation models with less restrictive assumptions on the service discipline and the sampling intervals provided a better approximation to the experimental results.

A cost function to evaluate the performance of dynamic bandwidth management along with steady state performance metrics was described. This approach was used in studying the performance of a distributed computing application for the rate and queue-based bandwidth algorithm families. A similar approach will be used in the next chapter to study the performance of the algorithms for a variety of traffic loads and charging intervals.

A more general cost formulation than was described in this chapter could be based on a client-server model as follows. The service provided would be wide area site inter-connection to a set of clients. The service provider (channel service and ramp) would purchase bandwidth from a common carrier (with a usage charge and a fixed charge towards maintenance and the hire of equipment). In turn, it would sell the bandwidth to its clients, imposing fixed and usage charges, with the charging policy including factors like the quality of service desired. This can be formulated as an optimisation problem, with the objective being to maximise the revenue earned by the service provider while maintaining constraints on the traffic performance perceived by the clients. If clients are to be charged on a per packet basis, quality of service constraints on delay and jitter can be modelled by the use of penalty functions. The service provider can maximise its revenues by 'hunting' for suitable carriers and optimising the purchase of the long haul bandwidth.



Chapter 9

Experimental Programme 2

9.1 Introduction

This chapter presents further performance studies of dynamic bandwidth management from the experimental programme conducted on the Unison testbed. The first section in this chapter compares the bandwidth switching process in queue and rate-based algorithms. This is studied for different traffic loads and for bandwidth switching between multiple service rates. The second section in this chapter compares the performance of rate and queue-based algorithms for a two-level switching process. The emphasis is on examining the tradeoff between steady state traffic performance and cost reduction using the cost function defined in Section 8.5 of the previous chapter. The final section in this chapter describes two applications of dynamic bandwidth management. In the first application dynamic bandwidth management is used to control congestion in a voice-data multiplexer when data traffic is multiplexed onto the bandwidth assigned to voice traffic. The second application considers the use of bandwidth control in the dynamic sharing of the ISDN link bandwidth between multiple U channels. A summary of the results of the experimental programme will be presented at the end of the chapter.

9.2 Bandwidth Switching – Queue and Rate-Based Algorithms

Bandwidth Switching between Multiple Service Rates

In this section the bandwidth switching process will be examined for a selection of bandwidth algorithms and for three different traffic loads. The loads are relatively larger than

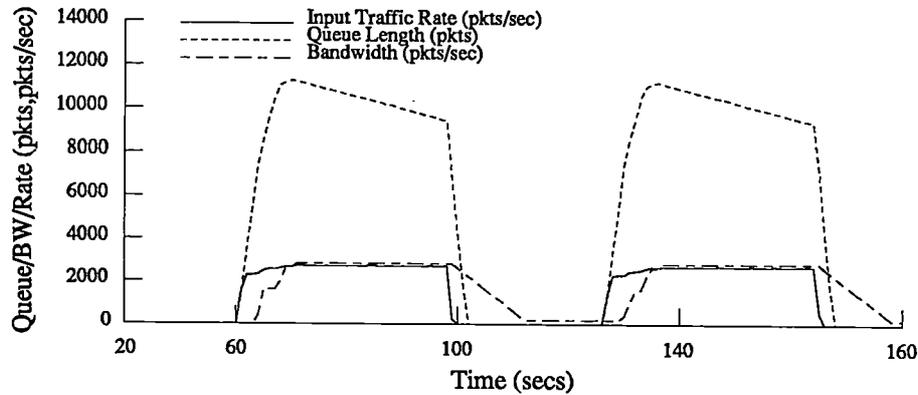


Figure 9.1: Tracking a Large Two-Phase Bursty Load with a RAT Algorithm. Experiment to RAL. Traffic Load 1 : $\lambda_0=0$, $\lambda_1=2800$, Threshold $[i] = 200 \times i$ packets/sec, $i = 1, 2, \dots, 30$.

those used in the previous experiments and necessitate bandwidth switching between multiple service rates. The experimental configuration used in the current experiments is as described in Section 8.2 of the previous chapter.

Traffic Load 1

- RAT algorithm
- Multi-level bandwidth switching
- Two-phase traffic load
- Experiment to RAL

Figure 9.1, displays the performance of an RAT algorithm in tracking a highly bursty two-phase traffic load, described by the solid line in the figure. During its active phase, the source generates packets at a rate of 2800 packets/sec onto a U channel bound to RAL. This rate corresponds to a bandwidth equivalent of 14 B channels and these are switched in by the rate-based control. The RAT algorithm has multiple rate thresholds beginning at 200 packets/sec, and uses state filtering and time lagging intervals of 10 and 40 seconds, respectively.

The switched bandwidth variation follows the input traffic load reasonably closely. The bandwidth trace displays 'soft' edges; this is because of state filtering (the leading edge) and time lagging and the bandwidth reclamation policy (the trailing edge).¹ Large queues² can be observed in the figure, this is because of the large input traffic rate and the switching latency. The slow decay of the queue even when bandwidth has been switched in is because the particular load used in the experiment (2800 packets/sec) was coincident with a rate threshold. It also implies that rate-based algorithms tend to switch in only the necessary bandwidth³ as computed from the input traffic rate.

Traffic Load 2

- RAT and QAT algorithms
- Multi-level Bandwidth Switching
- Two-phase traffic load
- Experiment in loopback mode

Figure 9.2 contrasts the performance of RAT and QAT bandwidth algorithms for a two-phase traffic load. The state filtering and time lagging intervals used in the two algorithms were 10 and 15 seconds respectively. In the first and third graphs in this figure, the solid lines represent the input traffic rate variations and the dashed lines represent the bandwidth variations. The more efficient tracking of the traffic load by the rate-based algorithm when compared to the queue-based algorithm can be observed. While the RAT algorithm switches in an optimal amount of bandwidth, the QAT algorithm is much less accurate. This is because queue-based algorithms can keep increasing the bandwidth as long as the switching threshold is exceeded. For this reason, upper and lower limits on bandwidth adjustment have to be imposed on queue-based algorithms. In this experiment the permissible upper limit on bandwidth adjustment was set at 12 B channels for the QAT algorithm.⁴

The second and fourth graphs in the figure show the variation of the instantaneous queue length with time for the two algorithms. The trace for the QAT algorithm exhibits steep rising and falling edges, but with a larger average than the trace for the RAT algorithm.

¹Although a rate-based algorithm can switch out multiple B channels with a single adjust request, the preferred approach was to switch out bandwidth one B channel at a time, in keeping with the policy of delaying the reclamation of switched bandwidth.

²The queue length variations with time plotted in the figures in this chapter represent instantaneous values. In all figures in this chapter containing more than one graph, it is assumed that they are numbered starting from the top.

³In general a rate-based algorithm would switch in bandwidth to the nearest (highest) 200 packets/sec threshold (400 packets/sec in loopback mode).

⁴A lower limit corresponding to the base bandwidth on a channel was imposed for all the algorithms, rate and queue-based.

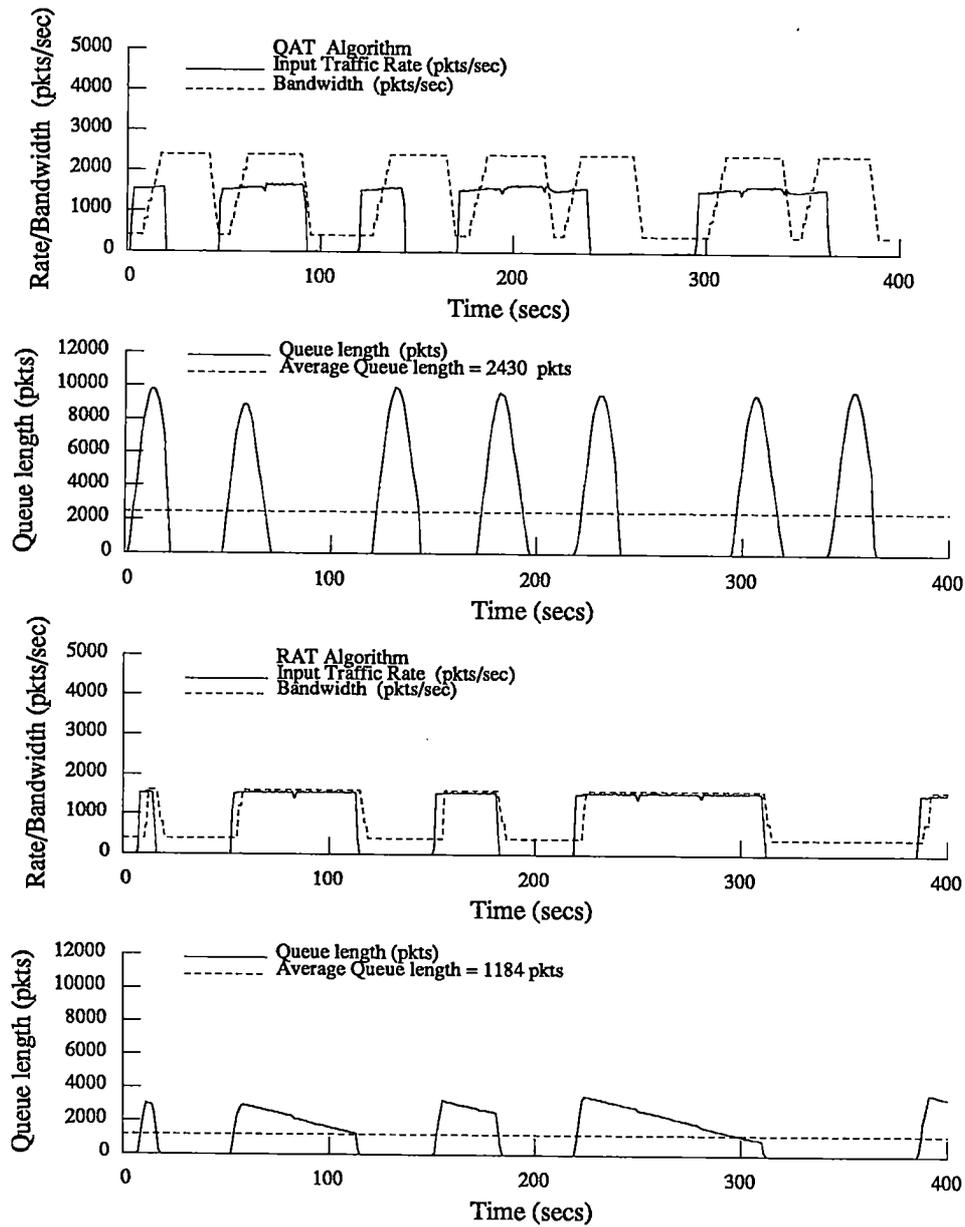


Figure 9.2: Tracking a Two-Phase Bursty Load with QAT and RAT Algorithms. Traffic Load 2 : $\lambda_0=0$, $\lambda_1=1600$.

The trace corresponding to the latter has a greater spread and slower decay because the enhanced service rate (8 B channels) is comparable with the traffic rate (1600 packets/sec). The greater peakedness of the queue length trace for the QAT algorithm is because it switches in bandwidth incrementally — two B channels at a time (since the experiment was in loopback mode). The RAT algorithm makes more efficient adjustments of the bandwidth by making larger increments at a time and deriving the benefits of pipelining in B channel set-up (Section C.2 of Appendix C).

Traffic Load 3

- RAT and QAT algorithms
- Multi-level Bandwidth Switching
- Three-phase traffic load
- Experiment in loopback mode

Figure 9.3 displays the performance of RAT and QAT bandwidth algorithms in tracking a three-phase traffic load with peak arrival rates of 400, 800 and 1600 packets/sec in the three phases. Transitions between the phases take place randomly, with equal transition probabilities from a phase to either of its peers. Within a phase, which has an exponentially distributed duration packets are generated deterministically as for Traffic Loads 1 and 2. The state filtering and time lagging intervals used for the QAT and RAT algorithms were 10 and 15 seconds.

The performance of the algorithms for Traffic Load 3 displays similar characteristics to those observed for Traffic Load 2. The traces describing bandwidth variation further exemplify the close tracking of load that is possible with a rate-based algorithm. In Section 4.4 of Chapter 4 it was pointed out that queue-based control could switch out bandwidth even when the input traffic load remained unchanged. Such oscillations can be observed on the first graph of both Figures 9.2 and 9.3. This is in contrast to the stability displayed by rate-based algorithms when the traffic rate is unchanged.

Bandwidth Switching between Two Service Rates

A Single Switching Operation

- RI and QI algorithms
- Two-level Bandwidth Switching
- Two-phase traffic load

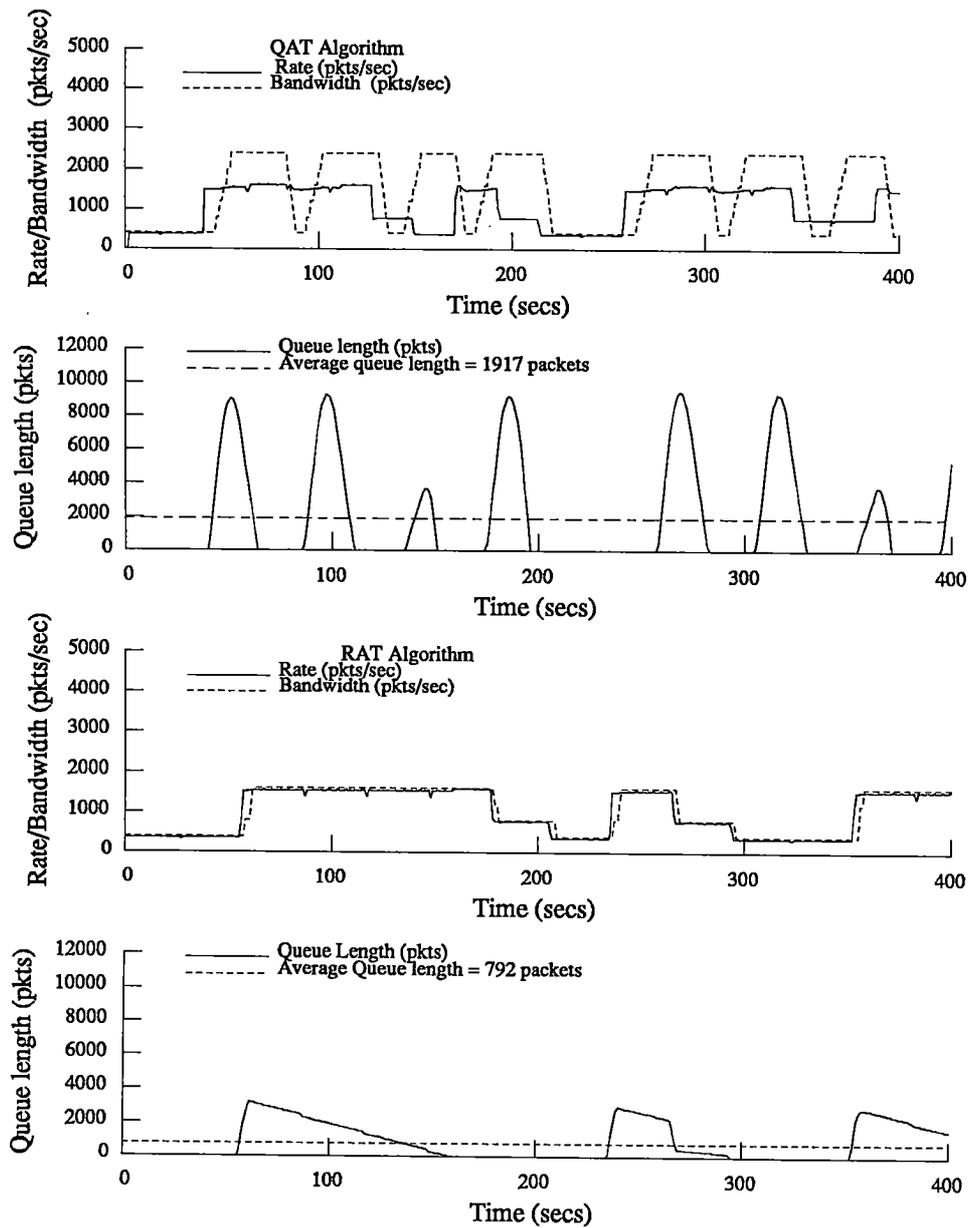


Figure 9.3: Tracking a Three-Phase Bursty Load with QAT and RAT Algorithms. Traffic Load 3 : $\lambda_0=400$, $\lambda_1=800$, $\lambda_2=1600$.

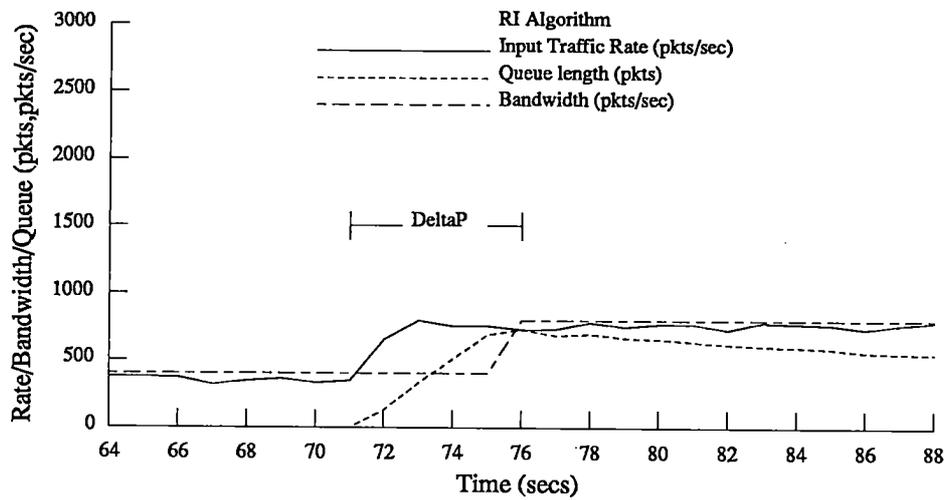
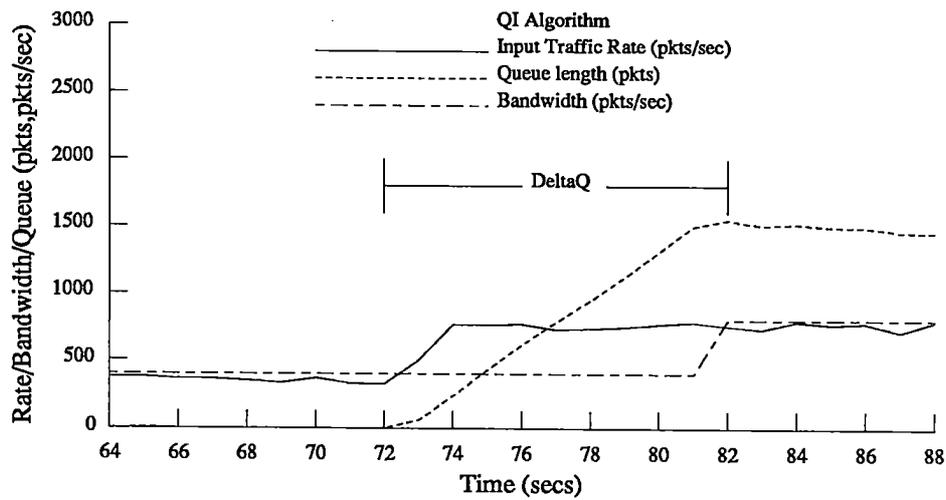


Figure 9.4: Switching In Bandwidth — RI vs. QI Algorithms.

- Experiment in loopback mode

Figure 9.4 focuses on a single switching in operation by RI and QI bandwidth algorithms for a two level switching process with service rates of 400 and 800 packets/sec. The switching thresholds used in this experiment were a point threshold of 400 packets/sec for the rate-based algorithm and an upper hysteresis threshold of 800 packets for the queue-based algorithm. From the figure it can be seen how rate-based control can be more reactive than queue-based control, the latter having to wait until the queue crosses the switching threshold before increasing the bandwidth. The actual bandwidth adjustment time (call signalling and U channel synchronisation delays) is the same in both the cases (on average). The queue-based algorithm just takes longer to initiate the adjustment. Obviously, its initial reactivity does depend on the threshold value; a lower value would reduce the latency, but would increase the possibility of switching in bandwidth on transients.

While queue-based algorithms react to an increase in queue size, rate based algorithms sense an increase in rate and can initiate bandwidth adjustment more quickly, provided the sampling rate is high and for non-trivial queue threshold values. For rate-based algorithms the latency between sensing an overload and initiating bandwidth adjustment can be as low as a single sampling interval. By switching in the necessary amount of bandwidth at one shot, rate-based algorithms also derive the advantages of pipelining available in call set-up (Section C.2 of Appendix C).

9.3 Cost and Traffic Performance of Bandwidth Management Algorithms

In this section the performance of rate and queue-based algorithms in maintaining steady state traffic performance and in reducing bandwidth costs will be evaluated for two traffic loads (Loads 4 and 5) and for three charging intervals.

- Rate and Queue-based algorithms
- Two-level Bandwidth Switching
- Two-phase Traffic Load
- Experiment in loopback mode

Traffic Load 4

Figure 9.5 describes the performance of various rate based algorithms for a bursty traffic load. The traffic load is a two-phase one, with packet generation rates of 300 and 650 packets/sec in the two phases. The duration of each phase is exponentially distributed with an average value of 3 seconds. The first graph in the figure shows the traffic load

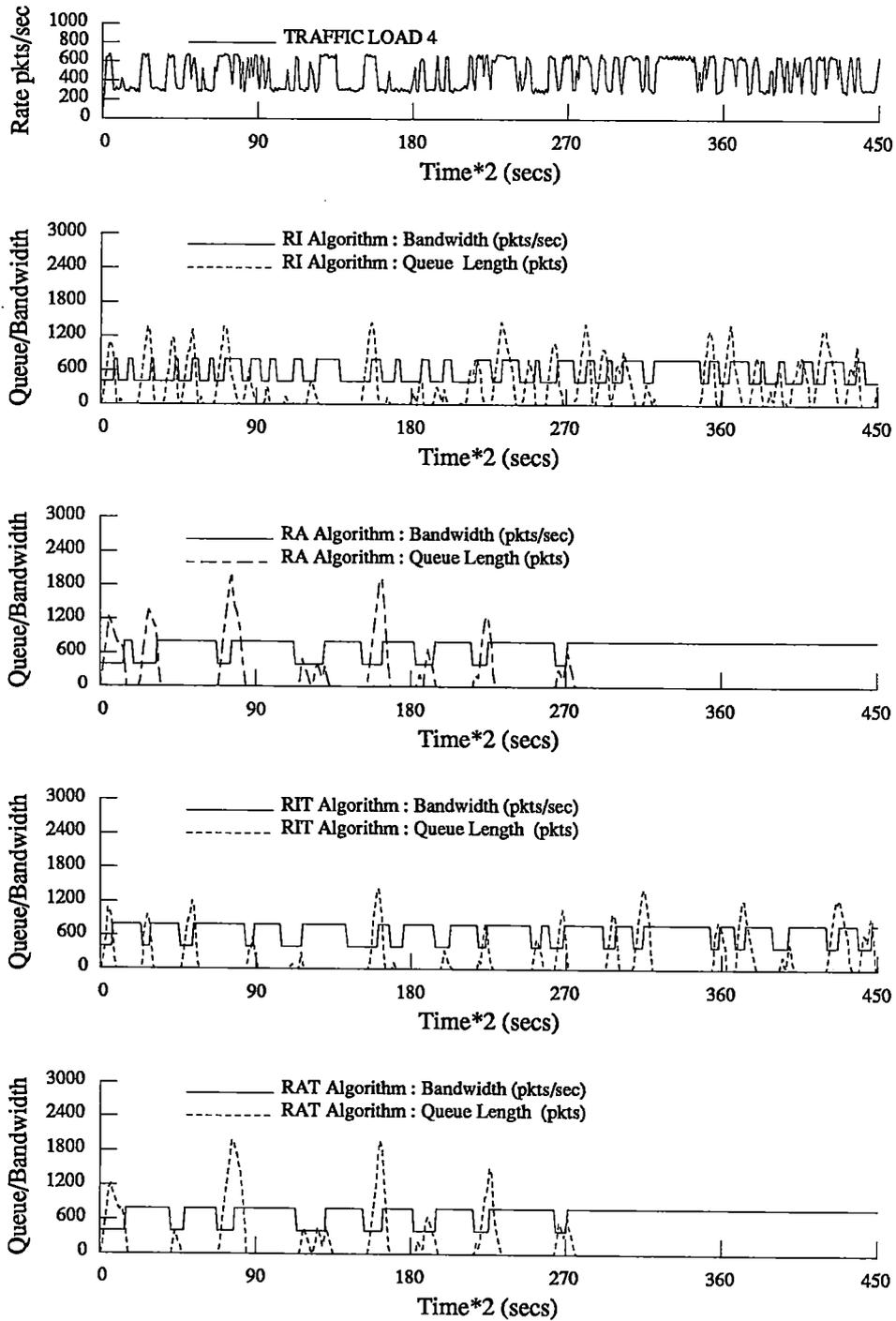


Figure 9.5: Bandwidth and Queue Length Variations of Rate-Based Algorithms — Traffic Load 4.

Alg.	Sw. BW Util.	RS Frac.	Avg. Queue	CL	CM	CS
RI	43.5	1.0	323	32	52	400
RA	78.9	0.28	155	22	75	726
RIT	69.9	0.58	176	24	73	644
RAT	80.3	0.26	148	22	78	742

Table 9.1: Cost and Steady State Performance of Rate-Based Algorithms — Traffic Load 4 : $\lambda_0=300$, $\lambda_1=650$, $\lambda_{avg}=475$, $\mu_0=400$, $\mu_1=800$, $\rho_0 = \rho_1=0.333$, Lower Threshold=320 pkts/sec, Upper Threshold=400 pkts/sec, Time Lagging=15 secs, History Interval=10 secs, $\delta=2.0$ Hz. Cost computed over test interval=920 secs.

Alg.	Sw. BW Util.	RS Frac.	Avg. Queue	CL	CM	CS
QI	29.2	0.94	476	16	31	268
QA	47.4	0.76	630	15	50	436
QIT	57.1	1.0	325	17	65	528
QAT	47	0.71	640	14	48	434

Table 9.2: Cost and Steady State Performance of Queue-Based Algorithms — Traffic Load 4 : $\lambda_0=300$, $\lambda_1=650$, $\lambda_{avg}=475$, $\mu_0=400$, $\mu_1=800$, $\rho_0 = \rho_1=0.333$, Lower Threshold=100 pkts, Upper Threshold=500 pkts, Time Lagging=15 secs, History Interval=10 secs, $\delta=2.0$ Hz. Cost computed over test interval=920 secs.

variation as it fluctuates rapidly between the two phases. The second and subsequent graphs display the variation of the instantaneous queue length (dashed line) and switched bandwidth (solid line) for the different rate-based algorithms.

Tables 9.1 and 9.2 display various performance parameters and costs incurred by rate and queue-based algorithms respectively for the traffic load.⁵ The relative switching fraction is defined as the number of bandwidth changes (low to high) made over the test interval by an algorithm, normalised by the largest such value in the algorithm class it belongs too. The switched bandwidth utilisation is the percentage of the test interval for which the channel was operating with the higher service rate. The costs for each algorithm are computed over a test interval of 920 seconds.

⁵Legend for Tables: Alg. — Bandwidth Algorithm, Sw. BW Util. — Switched Bandwidth Utilisation or Percentage of the test interval in which the channel was operating at the higher service rate, RS Frac. — Relative Switching Fraction, Avg. Queue — Average Queue Length, CL — Cost over Long Charging Interval (40 secs), CM — Cost over Medium Charging Interval (10 secs), CS — Cost over Short Charging Interval (1 sec).

Analysis of Results

Rate-Based Algorithms

The second graph in Figure 9.5 shows how the bandwidth switched in by the RI algorithm fluctuates very rapidly in an effort by the control to track the input load. The use of an RA algorithm with state averaging reduces these fluctuations by dampening the reactivity as can be seen on the third graph. The RIT algorithm, which is driven by the instantaneous samples of the traffic rate but with time lagging, offers less improvement in reducing the reactivity. The switching performance of the RAT algorithm is comparable to that of the RA algorithm. This is evident both from the figure and from Table 9.1. This is because the state filtering and time lagging intervals used in this experiment are comparable — 10 and 15 seconds respectively.

In both the RA and RAT algorithms, filtering of the rate samples decreases the reactivity and increases the peakedness of the queues when compared to the unfiltered rate algorithms. However, the average value of queue size is lower for the two algorithms because of their larger switched bandwidth utilisations (Table 9.1).

From Table 9.1, it can also be observed that the RI algorithm has the lowest bandwidth utilisation and the highest average queue length among the rate algorithms. The larger switching frequency of the RI algorithm is the reason for its relatively higher cost over the long charging interval. Over the short charging interval, its cost effectiveness improves substantially, strengthening the observation that over short charging intervals, bandwidth utilisation becomes a more important factor in determining cost than does the frequency of switching. Another example is provided by the RAT algorithm, which has the best control of queue size and has a relatively low cost over the long charging interval, but the largest cost among all the algorithms over the short charging interval, because of its relatively larger bandwidth utilisation.

Queue-Based Algorithms

Table 9.2 presents equivalent results for the Traffic Load 4, this time for a family of queue-based algorithms. When comparing the Tables 9.1 and 9.2, an immediate difference noticeable is in the average queue size — the values for the queue-based algorithms are greater because of their lower switched bandwidth utilisations. As with the rate-based algorithms (RA and RAT) there is not much difference between the QA and QAT algorithms — either in average queue size or costs. This can be again attributed to the comparable values of the time lagging (15 secs) and filtering intervals (10 secs).

The costs for both the algorithm families can be directly compared — as they have been computed over the same test interval and are driven by the same traffic load. Over all charging intervals, the costs for the queue-based algorithms are lower, because of their lower bandwidth utilisations (effective over the short charging interval) and the frequency

Alg.	Sw. BW Util	RS Frac.	Avg. Queue	CL	CM	CS
RI	62.3	1.0	158	38	92	775
RA	77.7	0.32	180	30	103	990
RIT	78.4	0.61	99	32	99	899
RAT	80.2	0.26	136	29	105	1022

Table 9.3: Cost and Steady State Performance of Rate-Based Algorithms — Traffic Load 5 : $\lambda_0=350$, $\lambda_1=750$, $\lambda_{avg}=550$, $\mu_0=400$, $\mu_1=800$, $\rho_0 = \rho_1=0.1$, Lower Threshold=400 pkts/sec, Upper Threshold=400 pkts/sec, Time Lagging=15 secs, History Interval (up)=10 secs, History Interval (down)=15 secs, $\delta=2.0$ Hz. Cost computed over test interval=1270 secs.

of bandwidth changes (low to high) over the test interval⁶ (effective over the long charging interval). For example the number of bandwidth changes for the RI algorithm was about twice as much as for the QI algorithm over the test interval — a ratio which is directly reflected in the costs for the two algorithms over the long charging interval.

Traffic Load 5

Figure 9.6 and Tables 9.3 and 9.4 display the performance and costs of the rate and queue-based algorithm families for another two-phase load (Traffic Load 5). This load has higher packet generation rates in each of its arrival phases than did Traffic Load 4. The load also has longer burst lengths with the phase durations being exponentially distributed with an average value of 10 seconds. The parameters used in the rate and queue-based algorithms in this experiment are indicated in the tables.

Analysis of Results

Many of the general observations made for Traffic Load 4 are also true for the present one. The increase in packet generation rates and the average durations of the phases increases the switched bandwidth utilisations and the frequency of bandwidth changes (not displayed in the tables) for both the rate and queue-based algorithms. The costs of the two algorithm families are more comparable over the three charging intervals although the costs of the queue-based algorithms are still lower because of their lower switched bandwidth utilisations. This is also the reason why the average queue lengths for the queue-based algorithms are larger.

Graphs describing the performance of the queue-based algorithms for Traffic Load 5 are provided in Figure 9.6. The traces of queue length have steep leading edges, but exhibit a

⁶The ‘absolute’ frequency of bandwidth changes made by each algorithm are not displayed in the tables which display only the relative switching frequencies within a class.

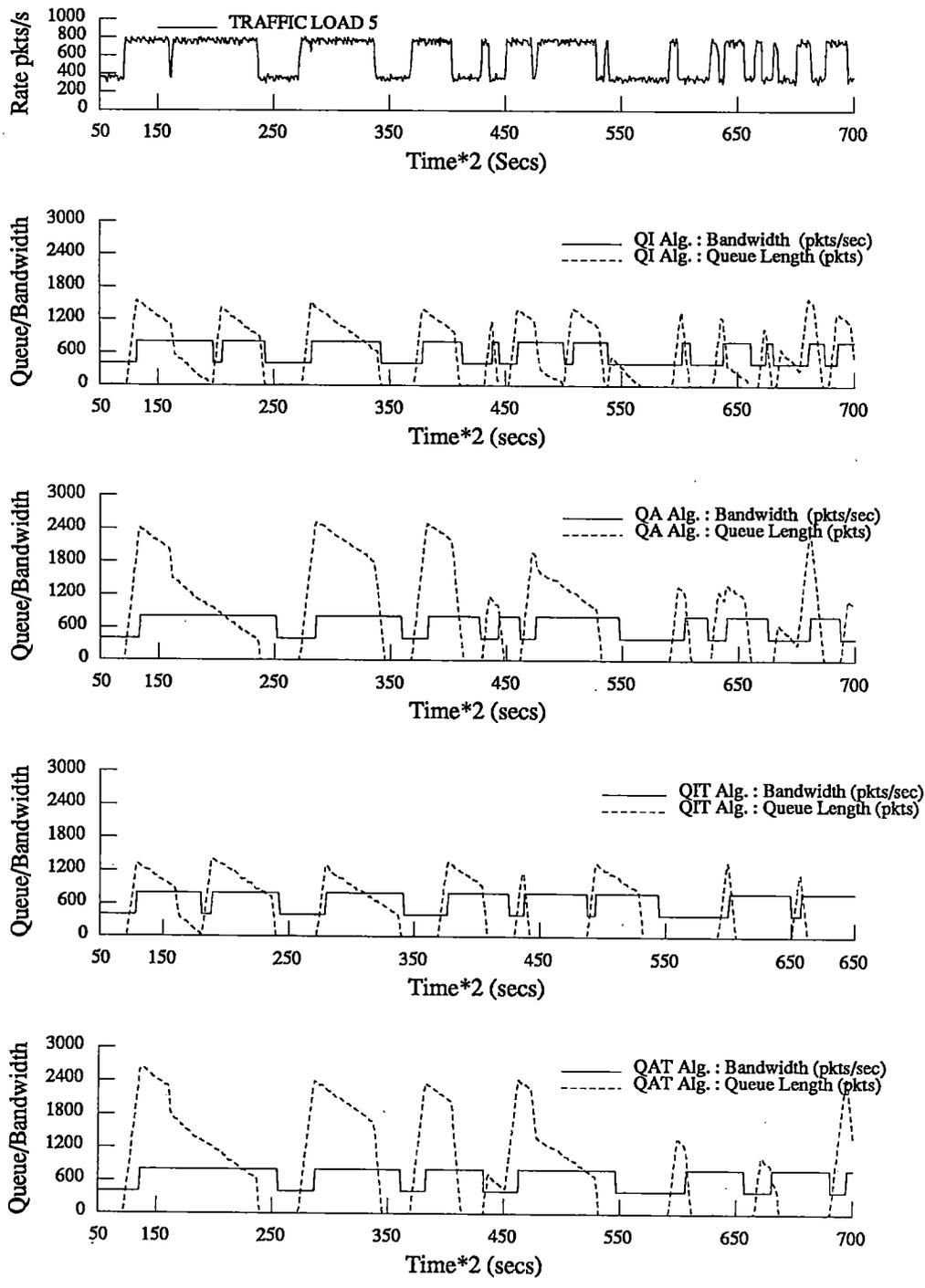


Figure 9.6: Bandwidth and Queue Length Variations of Queue-Based Algorithms — Traffic Load 5.

Alg.	Sw. BW Util	RS Frac.	Avg. Queue	CL	CM	CS
QI	57.5	1.0	629	29	81	727
QA	66.9	0.71	860	28	90	851
QIT	71.4	0.81	483	34	95	907
QAT	70.5	0.62	895	30	96	896

Table 9.4: Cost and Steady State Performance of Queue-Based Algorithms — Traffic Load 5 : $\lambda_0=350$, $\lambda_1=750$, $\lambda_{avg}=550$, $\mu_0=400$, $\mu_1=800$, $\rho_0 = \rho_1=0.1$, Lower Threshold=50 pkts, Upper Threshold=800 pkts, Time Lagging=15 secs, History Interval (up)=10 secs, History Interval (down)=15 secs, $\delta=2.0$ Hz. Cost computed over test interval=1270 secs.

slow decay even when the system is operating at the higher service rate. This is because the traffic rate (750 pkts/sec) in the second phase is close to the higher service rate (800 packets/sec). The QA and QAT algorithms with state averaging have larger average and peak queue lengths than the unfiltered algorithms (QI and QIT). The performance of the QA and QAT algorithms, as for the previous load, are similar because the state filtering interval (for switching out bandwidth) and the time lagging interval are comparable — both being 15 seconds.

9.4 Rate and Queue-Based Algorithms — a Summary

The experimental results in the previous sections have served to reinforce some of the general observations made about queue and rate-based control in Chapter 4.

Rate-based algorithms are more reactive than queue-based algorithms. Algorithms such as RI may switch bandwidth excessively with a bursty load (for example Traffic Load 4), which is undesirable particularly when the charging interval is long. Filtering the rate samples dampens the reactivity and ensures that switching does not take place for transient rate changes. Combined with appropriate time lagging, filtering would ensure that once circuits are switched in, they would be retained for at least a charging interval.

Queue-based algorithms perform better with bursty loads (such as Traffic Load 4) for suitable choices of the queue thresholds. However, they can display oscillations even when the input traffic load remains unchanged. For this reason and because queue-based algorithms do not explicitly use the traffic rate, the use of pure queue-based control is restricted. The oscillatory behaviour of queue-based control can be reduced by filtering and time lagging. However these methods merely increase the time period of the oscillations, they do not prevent them from occurring. Practical queue-based algorithms are likely to be hybrid ones that use the traffic rate in deciding whether to switch out bandwidth.

A hybrid version of the RI algorithm that could offer improved performance is one which uses queue state when switching in bandwidth. For an appropriate choice of queue thresh-

old this would dampen its reactivity to transients in a bursty traffic stream. Queue state is also useful when disconnecting circuits in a rate-based algorithm. Switching out the bandwidth when the input rate decreases may not be suitable since a queue residue could exist from the previous overload phase. Improved traffic performance would be obtained if the bandwidth were retained until the residue was cleared.

9.5 Applications of Dynamic Bandwidth Management

This section describe two applications of dynamic bandwidth management. The first involves a voice-data packet multiplexer while the second application is in the dynamic sharing of ISDN link bandwidth between multiple U channels.

Exploiting the Residual Bandwidth in a Voice-Data Multiplexer

The experiments in this section describe how the residual bandwidth in a talkspurt voice process, with speech activity detection, may be used for the transmission of other traffic. The voice arrival stream, composed of a number of superposed talkspurt streams, is multiplexed onto the high priority queue of a U channel operating under dynamic bandwidth management. Since voice traffic has stringent requirements on delay, voice sources are allocated bandwidth on the basis of their peak rates. Data traffic on the other hand is multiplexed onto the low priority queue of the U channel without any explicit allocation of bandwidth. Instead data packets are transmitted on the residual bandwidth arising from the burstiness of the voice stream. Dynamic bandwidth management is used to increase the bandwidth when the residual capacity is insufficient. In more general terms the problem can be considered as the sharing of bandwidth between two traffic classes; the first a high priority bursty stream requiring guaranteed bandwidth and the second a low priority bursty stream without any guaranteed requirements. The objective of the experiments in this section was to study how effective dynamic bandwidth management was in enabling the sharing of bandwidth (reducing costs) while maintaining traffic performance.

The exploitation of on-off variations in voice calls, both at the call admission and talkspurt levels, has been studied extensively in the context of voice-data multiplexers and movable boundary schemes (Section 5.2.1 in Chapter 5). The present set of experiments are confined to voice talkspurt variations – it is recognised that additional bandwidth exploitation is possible at the call admission level but this is not considered here.

Figures 9.7 and 9.8 display 20 minutes of the bit rate variations of two aggregate voice streams, composed of four and ten talkspurt voice sources, with speech activity detection multiplexed together. Each voice source which is emulated in software and has a bit rate of 56 Kbps, alternates between talkspurt and silence phases, with phase durations

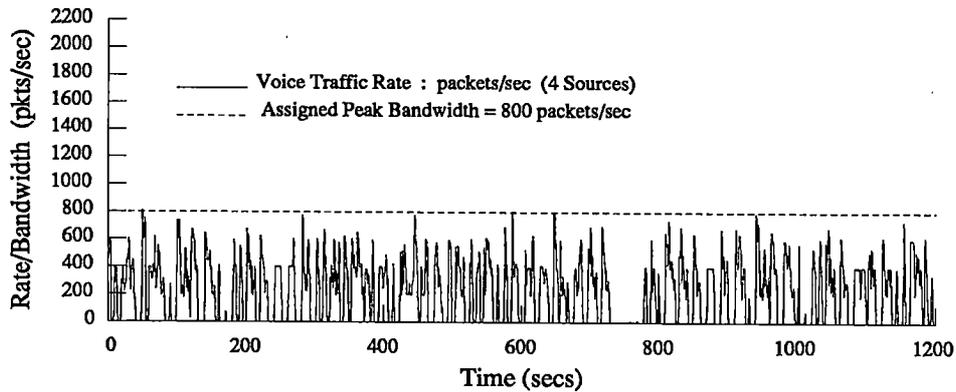


Figure 9.7: Packet Voice Stream — a Run with 4 Voice Sources. $\delta=1$ Hz.

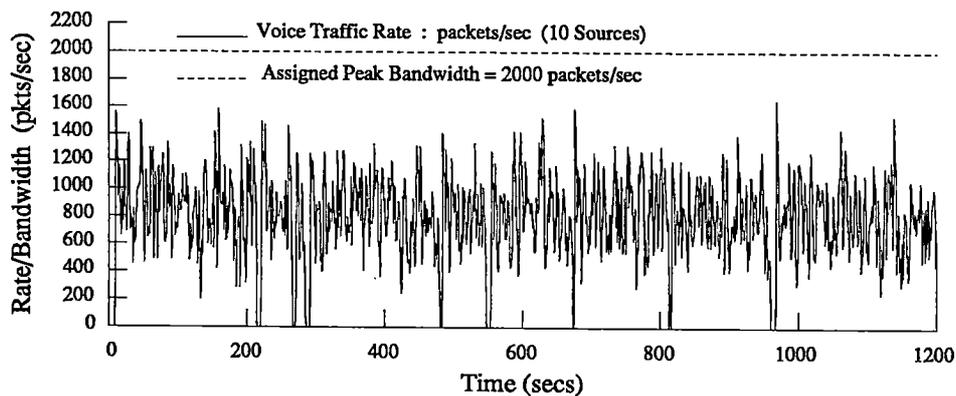


Figure 9.8: Packet Voice Stream — a Run with 10 Voice Sources. $\delta=1$ Hz.

that are exponentially distributed with values 1.2 and 1.8 seconds. These commonly used values were originally calculated from experiment [Brady68]. The dotted line on each figure represents the total bandwidth allocated to the voice stream which is given by the product of the peak rate of each source (200 packets/sec) and the number of sources. The deviation from this line at every time instant K , represents the instantaneous residual bandwidth available — which can be used for the transmission of other traffic. Although a more meaningful exploitation of bandwidth might be had with more voice sources (a figure of more than 10 is suggested [Sriram86]), for reasons of economy, the current experiments were confined to the use of two to eight sources.

Data traffic packets are routed onto the low priority (LP) queue of a U channel and are generated by a source which fluctuates between two arrival phases. The average bit rate

of the data traffic source (the LP source) corresponds to the average residual bandwidth available in the voice stream. When the LP source operates at its larger bit rate or when there is a decrease in the instantaneous residual bandwidth, LP packets could experience queueing delays. The function of the bandwidth controller would be to sense the overload and bring in additional bandwidth to alleviate the condition. It would be desirable if it could be determined whether the condition is likely to be transient or prolonged. This was achieved in some measure by using a modified RAT algorithm with state filtering and time lagging. The algorithm was driven by the samples of the traffic rate into the low and high priority queues at the ramp. The traffic rate into the high priority queue when subtracted from the total bandwidth assigned to the voice stream gives the instantaneous residual bandwidth. Bandwidth adjustment decisions are made when the filtered traffic rate on the LP queue exceeds the value of the filtered residual bandwidth.

Analysis of Results

Figure 9.9, shows a sample run (observed over 10 minutes) for a multiplexer with six voice sources and a bursty LP source. The first graph in the figure displays the variation in filtered residual bandwidth that is used by the channel service; the second displays the data traffic rate; while in the third graph the variation in switched bandwidth is displayed. The durations of the active and silent periods for the LP source were exponentially distributed, each with an average value of 16 seconds. The data packet generation rates in the two phases were 400 and 1060 packets/sec. The average traffic rate of the LP traffic source (730 packets/sec) was comparable with the average residual bandwidth of the voice arrival process (795 packets/sec).⁷ However, overload could occur when the LP source transmitted at its peak rate of 1060 packets/sec and the residual bandwidth available was small.

In the second and third graphs of figure 9.9, it can be observed that during periods of overload, the bandwidth controller switches in bandwidth — an additional 2 B channels. On the average, a saving in bandwidth of 69 percent was made when compared to a static scheme which would have allocated bandwidth to the LP source (800 packets/sec based on its average bit rate) for the duration of the experiment. The tradeoff is against performance and for this particular experiment, the average value of LP queue size observed was 450 packets against a negligible queue size in the static case. The experiment demonstrated the saving in bandwidth that is possible even with a sparse number of voice streams.

As a comparison, when the multiplexer was run with the same parameters as the first experiment, but without dynamic bandwidth management or any additional static bandwidth allocation, large instantaneous queue sizes were observed (queues as large as the low priority queue capacity of 20000 packets). The average LP queue size was in the order of thousands of packets. This degradation in data traffic performance emphasised the need for additional control in such a multiplexer; either by static bandwidth allocations or by dynamic bandwidth control. During this experiment, voice packets were largely insulated

⁷This value was calculated from direct measurements.

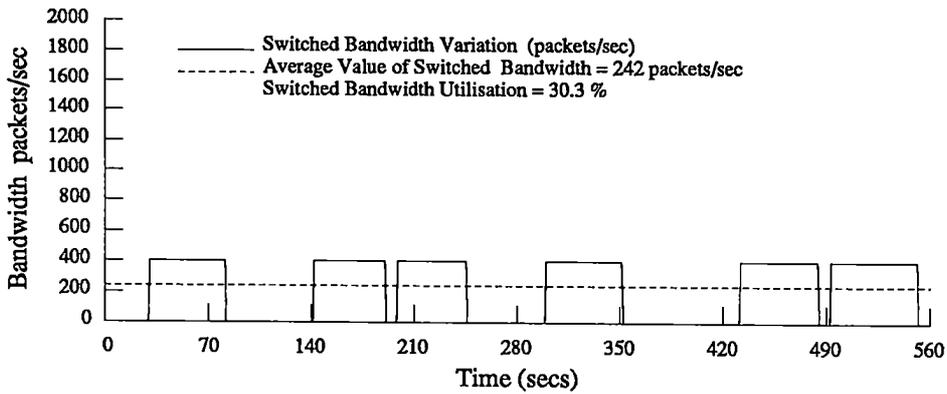
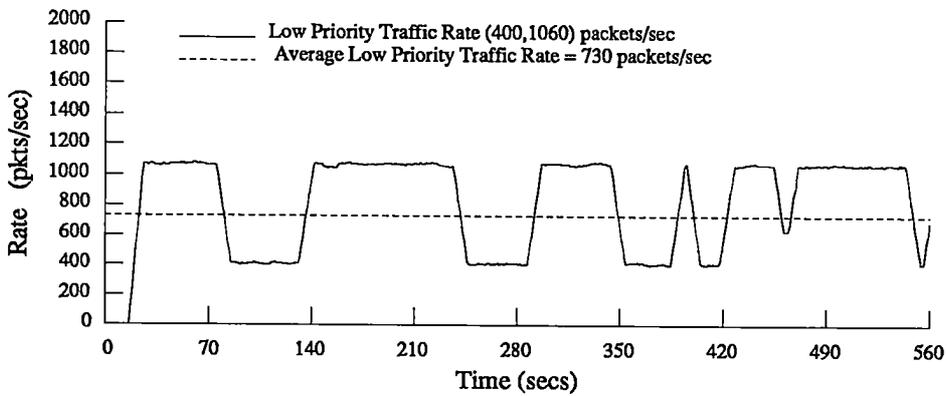
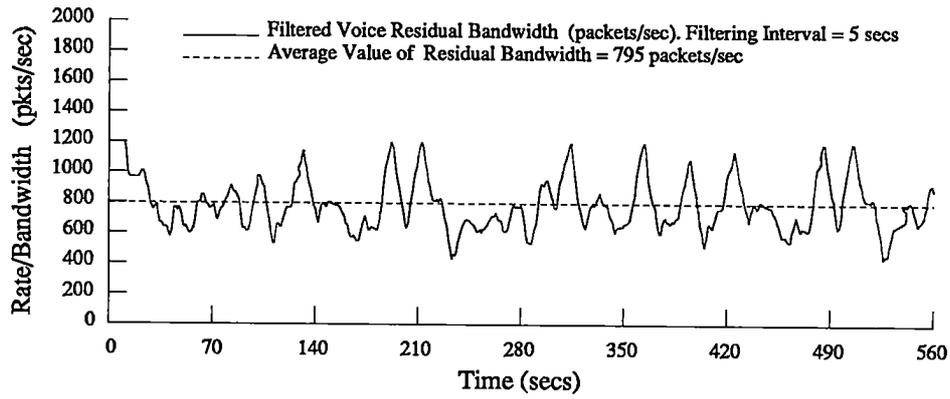


Figure 9.9: Run of a Voice Multiplexer Operating with Dynamic Bandwidth Management. Variation of Residual Bandwidth, Input Traffic Rate and Switched Bandwidth against Time. Modified RAT Algorithm. Parameters : Traffic Rate Filtering Interval=10 secs, Residual Bandwidth Filtering Interval=5 secs, Time Lagging=15 secs, $\delta=2$ Hz, Number of Voice Sources=6, Average Residual Bandwidth=795 pkts/sec, LP Source : $\lambda_{avg}=730$ pkts/sec.

from the overload conditions on the LP queue. The average high priority queue length was negligible but instantaneous queue lengths of upto 16 packets were observed, giving rise to a small jitter.

In the final experiment, a sparse voice stream with two multiplexed sources and an average residual voice bandwidth of 295 packets/sec was used. The two-phase LP source had an average bit rate of 285 packets/sec. The modified RAT algorithm employed had state filtering and time lagging intervals of 10 seconds. The average LP queue size measured was 20 packets, while the average switched bandwidth used was 300 packets/sec which was equivalent to a bandwidth saving of 25 percent. This experiment served to demonstrate that as the voice stream becomes sparser the saving in bandwidth decreases.

The experiments described above demonstrate the advantages of dynamic bandwidth management in conserving bandwidth while maintaining reasonable system performance for data traffic, when multiplexing voice and data traffic together.

Dynamic Bandwidth Sharing between U Channels

The second application of dynamic bandwidth management that is described in this section is in the dynamic sharing of the ISDN link bandwidth at a Unison ramp, between multiple U channels to different sites. In a static sharing scheme fixed portions of the ISDN link bandwidth will be allocated to each U channel. The scheme described in this section is based on dynamic bandwidth management and implements the dynamic sharing of the link bandwidth between two U channels fed by homogeneous traffic sources.⁸ One of the advantages of dynamic sharing is in the conservation of bandwidth which is useful in a primary rate ISDN environment.

This experiment studies how two U channels can dynamically share a single B channel when using dynamic bandwidth management. Assigning the B channel to either U channel for the duration of the experiment would be unfair to the other one. The purpose of the experiment was to explore how the dynamic sharing of the single B channel was fair to both the U channels and could equalise the performance (queueing delay) perceived by each of them.

At the start of the experiment each of the channels is made a static bandwidth allocation of 400 packets/sec (two B channels). This is based on the average bit rate of the bursty two-phase traffic source that feed into them. The average bit rates in the two phases of this source are 200 and 600 packets/sec. The channels share a single B channel (200 packets/sec) which is dynamically switched in by an RAT algorithm, with a filtering interval of 5 seconds and a time lagging interval of 15 seconds — small values chosen to increase the reactivity of the control and enable better sharing.

⁸The problem of allocating ISDN link bandwidth between multiple U channels was discussed in Section 4.2 of Chapter 4.

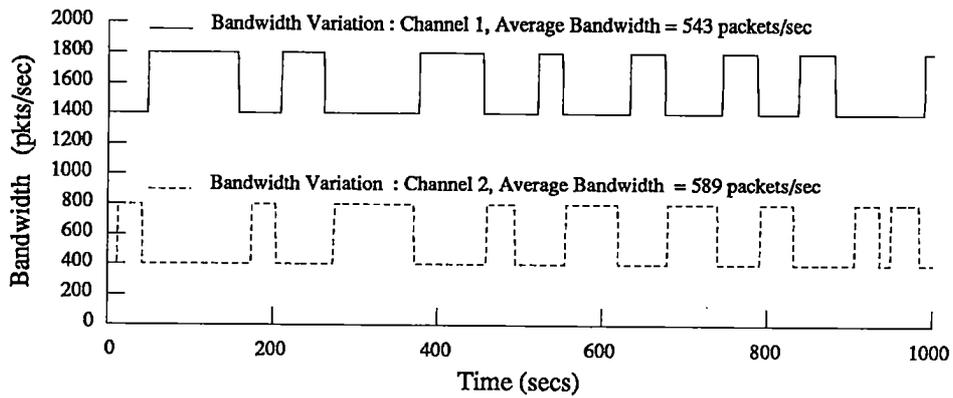
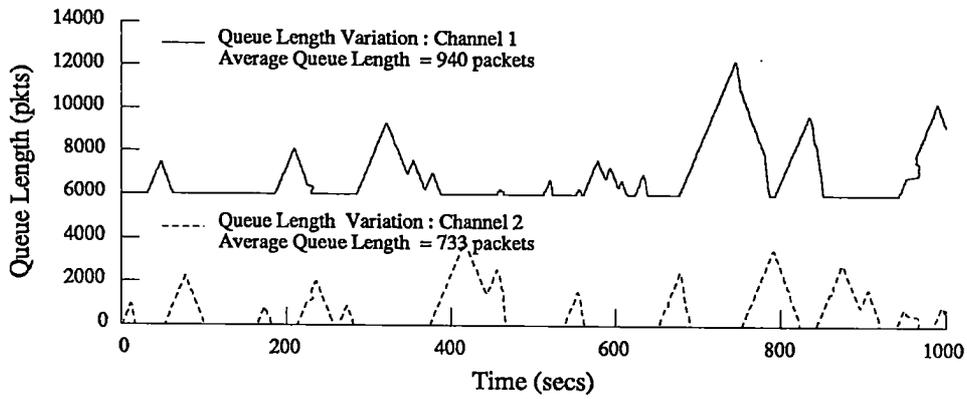
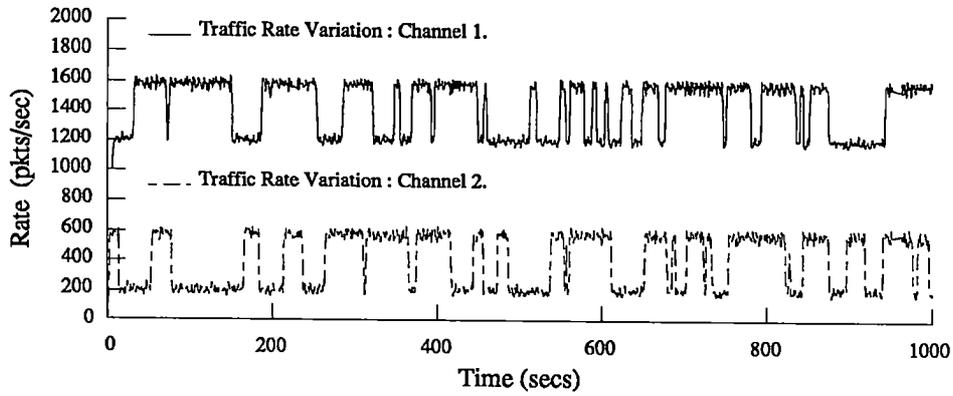


Figure 9.10: Dynamic Bandwidth Sharing between Two U Channels. Variation of Input Traffic Rate, Queue Length and Switched Bandwidth against Time. RAT Algorithm. Rate Filtering Interval=5 secs, Time Lagging=15 secs, $\delta=2$ Hz.

Figure 9.10 displays graphs obtained from the experiment with dynamic bandwidth sharing between the two U channels. Each graph has two traces for the U channels. In the upper traces it is the difference in ordinate values that is important when evaluating the absolute values of the traffic rate, queue length and bandwidth. For example in the first graph the traffic rate of the bursty source multiplexed onto U channel 1, is shown to vary between 1200 and 1600 packets/sec. After subtracting the displacement of 1000 packets/sec this translates to an actual variation of 200 to 600 packets/sec.

When a traffic source transmits at its higher bit rate (600 packets/sec), the associated channel queue begins to build up and the channel controller switches in the free B channel. This takes place in a first come first served fashion, so that if the second channel had also entered a phase of overload but with a lag, it would continue to experience a queue build up. The switched B channel would be released when the traffic source switched back to its lower bit rate (200 packets/sec). The time taken to release the B channel would depend on the time lagging and filtering parameters of the RAT algorithm. The freed B channel is then available for use and can be switched in by the other U channel if needed. Although the two traffic sources are homogeneous i.e. have the same parameters, their variation is not identical because different seed values were used in their random number generators at the start of an experimental run.

Analysis of Results

The average queue lengths for the two channels obtained from experiment were of the same order — 733 and 940 packets and so were the average bandwidths utilised in the two cases — 589 and 543 packets/sec (including the static bandwidth allocation — 400 packets/sec). As a comparison, another experiment was run with the same bandwidth algorithm, but without the sharing of bandwidth. Instead, each U channel had independent access to a single B channel to be switched in dynamically when needed. The average queue lengths obtained in this case were 71 and 117 packets for the two channels. The average bandwidths utilised by the channels had now increased to 681 and 645 packets/sec. Finally, for an experiment in which a B channel was statically allocated to each U channel for the duration of the run (equivalent to assigning bandwidth on the basis of the peak rate of the input traffic source) without dynamic bandwidth management, neither of the queues built up, but the bandwidth usage was now 800 packets/sec for each channel.

These results, while not extensive, provide a flavour of how dynamic bandwidth management could be used to implement sharing and cost reduction in a primary rate ISDN environment while maintaining reasonable traffic performance. It would be of interest to extend this study to larger loads and bandwidths and to more varied traffic patterns.

9.6 Summary

The results of an experimental programme conducted on the Unison testbed have been presented in this chapter and the previous one. The primary objective of the experimental programme was to study the performance of dynamic bandwidth management both in terms of the traffic performance it can offer and the cost reduction it makes in doing so.

The performance of the queue and rate-based algorithm families were studied for various traffic loads that emulate bursty traffic sources and streams. The algorithms were compared on the basis of the steady state performance they offered to traffic, their reactivity and cost efficiency. A cost function was defined that is related to the tariff structure used on the ISDN. The cost function includes both a switching cost and a bandwidth utilisation cost and was used in studying the applicability of different bandwidth algorithms across a spectrum of charging intervals. Algorithms which were highly reactive and changed the bandwidth frequently were cost efficient across short charging intervals, while those with a lower switching frequency and greater retention of bandwidth were better over the longer charging intervals.

Two applications which emphasised the usefulness of dynamic bandwidth management have been described. The first demonstrated how dynamic bandwidth control could reduce congestion in a voice-data multiplexer with a movable boundary. The use of dynamic bandwidth management in the dynamic sharing of ISDN link bandwidth between multiple U channels was also presented — this scheme is relevant to a variety of traffic environments and maintains fairness by equalising the traffic performance.

Although call admission control was not explicitly studied in the experimental programme (i.e. incoming call requests were not modelled), admission control did exist. For example, in the experiment on bandwidth sharing between two U channels — each channel was statically allocated bandwidth at the start of the experiment that was based on the average rate of its input traffic source.

The results of this chapter and the last one emphasise the usefulness of dynamic bandwidth management in controlling congestion and in supporting cost effective packet transfer of multi-service traffic across a circuit switched ISDN.

Chapter 10

Conclusions and Further Work

This dissertation has addressed the use of dynamic bandwidth management in supporting variable rate circuits when running a packet overlay over a circuit switched network. The design of dynamic bandwidth control schemes to provide cost-effective site interconnection while maintaining constraints on the traffic performance perceived by end applications was presented.

Dynamic bandwidth management was described as an important element of a congestion control framework, along with admission control and bandwidth enforcement schemes, for the packet overlay. The formulation of theoretical performance models for simple dynamic bandwidth control schemes and their steady state performance analyses formed another important investigation in this dissertation. Finally an experimental programme to evaluate the performance of dynamic bandwidth management was conducted on an ATM testbed with a prototype implementation of a channel controller.

10.1 Conclusions

From a comparative performance study of rate and queue-based bandwidth algorithms in controlling a variable bandwidth channel in a bursty traffic environment, the use of rate-based schemes for dynamic bandwidth management is preferred. Rate-based algorithms are in general more reactive than queue-based ones and are better in tracking variations in the traffic load. These advantages arise from a better knowledge of the amount of bandwidth to be switched in and their relative stability when compared to the oscillations that characterise simple queue-based control. However, rate-based schemes can be over reactive in bursty traffic environments, behaviour which can be controlled by the use of filtering and time lagging. The latter is the preferred approach because it is directly related to the tariff structure implemented on the underlying circuit switched network.

The novel feature of the congestion control framework proposed for the ATM overlay was the dynamic variation of bandwidth to control congestion occurring within the network. The importance of this scheme is that it simplifies the static allocation of bandwidth at connection establishment time. It was neither essential to know the exact bit rate of a source nor is it necessary to define a virtual bandwidth for a bursty one as in other proposals for admission control in ATM networks. Sources demanding guaranteed bandwidth were admitted with static bandwidth allocations made on their peak bit rates. Experiments demonstrated that traffic handled on the basis of peak rate bandwidth allocations and expedited transfer was well insulated from congestion so that the framework could guarantee a quality of service for such sources. For bursty sources with less stringent requirements, allocations when made, were on the basis of their average bit rates. The emphasis in the framework was on dynamic bandwidth control increasing the bandwidth at the onset of congestion and reclaiming bandwidth to reduce costs when the utilisation decreased.

While dynamic bandwidth management is important in reacting to overload in the aggregate traffic stream within the network, it was recognised that more selective methods were needed to ensure monitoring and enforcement of individual traffic streams. The use of rate control methods, acting at the network periphery to throttle individual sources to their specified connection characteristics, was seen to be an essential component of the framework.

The performance modelling of simple queue-based bandwidth algorithms were as queueing systems with fluctuating parameters and state dependent transitions. The assumptions that were made in these models on the service time and sampling interval distributions and the use of a continuous time analysis meant that performance metrics predicted by the theoretical analyses were more conservative than those observed in practice. Considering the modelling constraints and that the systems modelled were highly dynamic, a sufficiently close correspondence between the theoretical results and the experimental ones was observed, vindicating the use of the models in emulating the real system. The simulation results provided a better approximation to the experimental observations. This was because the simulation models assumed deterministic sampling intervals and service times, the latter to model the fixed length CFR packets actually used on the overlay.

The implementation of a channel controller and the extensive experimental programme on the ATM testbed were important in studying the practical aspects of bandwidth control — particularly those related to cost. A cost function based on the tariff scheme implemented on the underlying public ISDN was defined which included the bandwidth utilisation and the number of bandwidth changes made. The cost function was found useful in studying the applicability of different bandwidth algorithms across a spectrum of charging intervals. Algorithms which were highly reactive were found to be cost efficient across short charging intervals, while those with a greater retention of bandwidth were better over the longer charging intervals.

The experimental programme also investigated applications of dynamic bandwidth man-

agement. The dynamic sharing of transmission link bandwidth between multiple U channels was shown to equalise traffic performance and maintain fairness between the end applications. The use of dynamic bandwidth management in exploiting the residual bandwidth variations in bursty traffic streams and in increasing the bandwidth utilisation and reducing costs, is especially important in a primary rate ISDN environment.

10.2 Suggestions for Further Work

Performance Modelling of Dynamic Bandwidth Management

An immediate area for further work is the extension of the performance models for bandwidth control from a continuous to a discrete time analysis in which both fixed length packets and the frame structure of the underlying ISDN link can be modelled. A general solution technique based on the generating function method for the discrete time queueing analysis of an ATM network has been recently proposed in [Li90, Li89a] and would be useful in this respect. Methods used by Sriram [Sriram83] for the discrete time analysis of a hybrid voice-data multiplexer and by Zukerman [Zukerman86b] for a queueing system operating in a random environment are also relevant.

In this dissertation simple queue-based bandwidth control has been modelled with the service rate varying between two levels. A problem for future work would be the modelling of queue-based control in which the service rate fluctuates between many levels. There is also a need to analyse the performance of rate-based bandwidth control schemes. A starting point could be an approximate analysis of rate based control along the lines of the graphical analysis presented for queue-based control in Section 7.9 of Chapter 7.

Another useful extension to the current models would be to attempt their solution with matrix geometric techniques based on the work of Neuts [Neuts81, Neuts84].

The performance analyses in this dissertation has concentrated on the derivation of steady state performance metrics. For variable bandwidth channels operating in a bursty traffic environment, with fluctuations between periods of overload (with utilisations greater than one), and underload, a dynamic performance analysis of the control would be an appropriate problem for future research.

The theoretical derivation of bandwidth control policies for a variable bandwidth U channel suggested in Section 7.11 of Chapter 7 forms another interesting problem for further work.

Dynamic Bandwidth Management and the Dynamic Control of Virtual Paths

This dissertation has examined the use of dynamic bandwidth management in controlling the bandwidth of U channels to support wide area site interconnection. All associations

(i.e. virtual circuits) between the same end sites are bound to a single U channel. Packets belonging to the different associations are statistically multiplexed onto the channel on the basis of a common window field in the packet address. The use of a common channel provides the benefits of statistical multiplexing and enables new associations to be established with very low connection delays by multiplexing them onto existing bandwidth.

Recently there has been much interest in the use of *virtual paths* in ATM networks [Ohta88b, Sato90, Tirtaatmajda90]. Virtual paths are conceptually similar to U channels, being bundles or collections of virtual circuits with the same origin and destination that are aggregated for purposes of routing, control and management. Packets belonging to a virtual path are identified by a common virtual path identifier (VPI) to be carried in the header of each cell along with a virtual circuit identifier (VCI). The main benefits of the virtual path scheme are the reduced node processing and connection delays and the better network management.

Sato proposes the dynamic allocation of bandwidth on virtual paths as the number of circuits constituting the path are increased or decreased. Sato [Ohta88b] presents a bandwidth control algorithm that operates at the call admission level and accepts call requests either on the basis of the existing bandwidth available or by allocating more bandwidth if necessary. The bandwidth on the path is decreased when circuits are disconnected or according to the utilisation. An interesting problem for future research would be to examine whether rate-based algorithms as described in this dissertation can be applied to the dynamic control of the bandwidth of a virtual path in response to the traffic load through it. Such schemes can work at the origin node of the path and may be useful in supplementing admission level control in improving the transmission efficiency.

In conclusion, dynamic bandwidth management is useful in providing cost effective transfer with suitable traffic performance while supporting an ATM overlay over a circuit switched network, such as primary rate ISDN. It represents a new approach to congestion control and has potential applications in bandwidth sharing schemes.

Appendix A

Analysis 1

In this appendix the equation 6.19 presented in Section 6.4 of Chapter 6 will be derived.

Calculation of $\sum_{i=0}^{U-1} P_{i,0}$

Figure A.1 is the state transition diagram for Model A (identical to figure 6.1 in Chapter 6) showing boundaries (cuts) across which traffic flows will be equated to derive the expression for $\sum_{i=0}^{U-1} P_{i,0}$. Let $\rho_0 = \lambda/\mu_0$ and $\rho_1 = \lambda/\mu_1$.

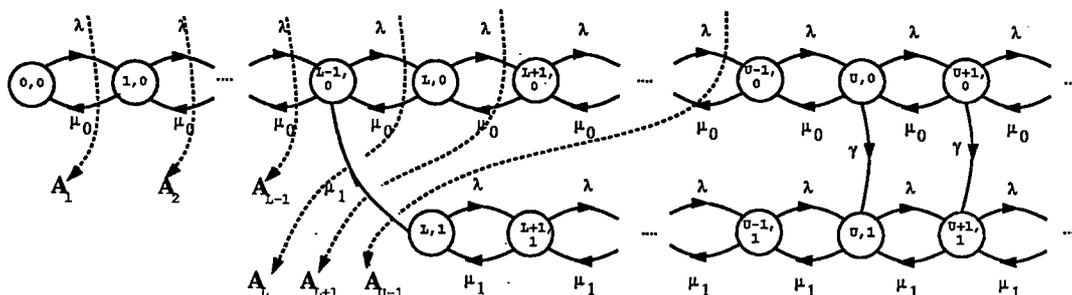


Figure A.1: Model A : Queue-Based Hysteresis Control — Poisson Arrivals and Delayed Bandwidth Change (In). Equating the Steady State Flows.

Equating the steady state traffic flows across the set of cuts A_0, A_1, \dots, A_{L-1} yields the set of equations :

$$P_{i,0} = \rho_0^i P_{0,0} \\ i = 1, 2, \dots, L-1$$

Consider the cut A_L . Equating the steady state flows across it yields :

$$\lambda P_{L-1,0} = \mu_1 P_{L,1} + \mu_0 P_{L,0}$$

or equivalently,

$$P_{L,0} = \rho_0 P_{L-1,0} - \frac{\rho_0}{\rho_1} P_{L,1} \\ = \rho_0^L P_{0,0} - \frac{\rho_0}{\rho_1} P_{L,1}$$

In a similar fashion by considering the traffic flows across the cuts $A_{L+1}, A_{L+2}, \dots, A_{U-1}$ the expressions for $P_{i,0}$ are :

$$P_{i,0} = \rho_0^i P_{0,0} - \frac{\rho_0}{\rho_1} P_{L,1} \frac{(1 - \rho_0^{i-L+1})}{(1 - \rho_0)} \\ i = L+1, L+2, \dots, U-1 \quad (\text{A.1})$$

Therefore,

$$\sum_{i=0}^{U-1} P_{i,0} = P_{0,0} [1 + \rho_0 + \rho_0^2 + \dots + \rho_0^{U-1}] \\ - P_{L,1} \frac{\rho_0}{\rho_1} \left[1 + \frac{1 - \rho_0^2}{1 - \rho_0} + \dots + \frac{1 - \rho_0^{U-L}}{1 - \rho_0} \right] \quad (\text{A.2})$$

which can be rewritten to give the final expression for $\sum_{i=0}^{U-1} P_{i,0}$ as :

$$\sum_{i=0}^{U-1} P_{i,0} = P_{0,0} \left\{ \frac{(1 - \rho_0^U)}{(1 - \rho_0)} \right\} \\ - P_{L,1} \left\{ \frac{\mu_1}{\mu_0 (1 - \rho_0)^2} [(U - L)(1 - \rho_0) - \rho_0(1 - \rho_0^{U-L})] \right\} \quad (\text{A.3})$$

Appendix B

Analysis 2

This appendix contains further details of the analyses of Models C, D and E that were presented in Chapter 7. The necessary and sufficient conditions for the existence of the steady state for each of these models will be derived. A brief outline of the derivation of independent equations necessary to solve for the unknown boundary probabilities is also provided for the models. The analyses are based on the general solution technique that was described in Section 6.3 of Chapter 6.

B.1 Model C : Hysteresis Control, MMPP Arrivals and Delayed Bandwidth Change (In and Out)

Solution Procedure :

The state transition diagram for this model is shown in Figure 7.1 in Chapter 7. Partial MGFs are defined for the 4 arrival-switching phase combinations. The unknown state probabilities in the MGF are evaluated by the simultaneous solution of independent equations derived from the original balance equations, the conservation of probability sum condition from the value at unity of the MGF and its analytic properties.

Let the partial moment generating functions $\Pi_{jk}(z)$, $j = \{0, 1\}$, $k = \{0, 1\}$ where j is the arrival phase and k is the switching phase, be defined as :

$$\Pi_{jk}(z) \stackrel{\text{def}}{=} \sum_{i=0}^{\infty} z^i P_{i,j,k}$$

and the MGF, $\Pi(z)$ as :

$$\Pi(z) \stackrel{\text{def}}{=} \Pi_{00}(z) + \Pi_{10}(z) + \Pi_{01}(z) + \Pi_{11}(z)$$

For notational convenience let the following be defined for $i = \{0, 1\}$:

$$a_{i0} = \mu_0(1 - z)P_{0,i,0} \quad a_{i1} = \mu_1(1 - z)P_{0,i,1}$$

$$f_i = \rho_i z, f_2 = \gamma_{in} z$$

$$\text{Sum}_{i0} = \gamma_{in} z \sum_{j=0}^{U-1} z^j P_{j,i,0}, \quad \text{Sum}_{i1} = \gamma_{out} z \sum_{j=0}^L z^j P_{j,i,1}$$

$$D_{i0} = \lambda_i z^2 - (\lambda_i + \mu_0 + \rho_i + \gamma_{in})z + \mu_0$$

$$D_{i1} = \lambda_i z^2 - (\lambda_i + \mu_1 + \rho_i)z + \mu_1$$

In a similar fashion to the method described for Model A in Section 6.4 of Chapter 6, by summing the balance equations over the relevant values of the indices, the partial MGFs can be written in terms of the above definitions as :

$$\Pi_{00}(z) = \frac{D_{10}(a_{00} - \text{Sum}_{01} - \text{Sum}_{00}) - f_1(a_{10} - \text{Sum}_{10} - \text{Sum}_{11})}{(D_{00}D_{10} - f_0f_1)} \quad (\text{B.1})$$

$$\Pi_{10}(z) = \frac{D_{00}(a_{10} - \text{Sum}_{11} - \text{Sum}_{10}) - f_0(a_{00} - \text{Sum}_{00} - \text{Sum}_{01})}{(D_{00}D_{10} - f_0f_1)} \quad (\text{B.2})$$

$$\begin{aligned} \Pi_{01}(z) &= \frac{D_0[D_{11}(a_{01} + \text{Sum}_{00} + \text{Sum}_{01}) - f_1(a_{11} + \text{Sum}_{10} + \text{Sum}_{11})]}{D_0(D_{01}D_{11} - f_0f_1)} \\ &\quad - \frac{D_{11}f_2N_{00} - f_1f_2N_{10}}{D_0(D_{01}D_{11} - f_0f_1)} \end{aligned} \quad (\text{B.3})$$

$$\begin{aligned} \Pi_{11}(z) &= \frac{D_0[D_{01}(a_{11} + \text{Sum}_{10} + \text{Sum}_{11}) - f_0(a_{01} + \text{Sum}_{00} + \text{Sum}_{01})]}{D_0(D_{01}D_{11} - f_0f_1)} \\ &\quad - \frac{D_{01}f_2N_{10} - f_0f_2N_{00}}{D_0(D_{01}D_{11} - f_0f_1)} \end{aligned} \quad (\text{B.4})$$

where in equations B.3 and B.4, D_0 is the common denominator of $\Pi_{00}(z)$ and $\Pi_{10}(z)$ and N_{00} , N_{10} are the numerators of the two expressions respectively. The equations B.1 through to B.4 involve $2U + 2L + 2$ unknown state probabilities : $P_{i,0,0}$, $P_{i,1,0}$, $i = 0, 1, \dots, U - 1$ and $P_{i,0,1}$, $P_{i,1,1}$, $i = 0, 1, \dots, L$. From the original balance equations 7.1 through to 7.4 in Chapter 7, $2U + 2L - 2$ independent equations can be derived. These are used to express all the other unknown state probabilities in terms of the four boundary probabilities — $P_{0,0,0}$, $P_{0,0,1}$, $P_{0,1,0}$ and $P_{0,1,1}$. Of the four remaining equations, one can be obtained from the conservation of probability sum property of the MGF, while the other three may be obtained from a consideration of the analytic properties of the MGFs as discussed in Section 7.7 of Chapter 7.

Conservation of Probability Sum :

The value of $\Pi(z)$ at $z = 1$ is the sum of the partial MGFs at this value. While $\Pi_{00}(1)$ and $\Pi_{10}(1)$ are directly computable, $\Pi_{01}(1)$ and $\Pi_{11}(1)$ are of indeterminate form. These

indeterminacies are resolved by extracting the factor $(z - 1)$ from the numerators and denominators of the two expressions. This was done using a symbolic algebraic package REDUCE [Hearn85].¹

Defining the partial MGF $\Pi_1(z) \stackrel{\text{def}}{=} \Pi_{01}(z) + \Pi_{11}(z)$ and using the REDUCE program to factorise and evaluate it at $z = 1$ gives :

$$\begin{aligned} \Pi_1(1) = & \frac{\mu_0(\rho_0 + \rho_1)[P_{0,0,0} + P_{0,0,1} + P_{0,1,0} + P_{0,1,1}]}{[(\rho_0 + \rho_1)\mu_1 - \lambda_0\rho_1 - \lambda_1\rho_0]} \\ & + \frac{[\lambda_0\rho_1 + \lambda_1\rho_0 - \mu_0(\rho_0 + \rho_1)](\text{Sum}_{00} + \text{Sum}_{01} + \text{Sum}_{10} + \text{Sum}_{11})}{\gamma_{in}[(\rho_0 + \rho_1)\mu_1 - \lambda_0\rho_1 - \lambda_1\rho_0]}. \end{aligned}$$

so that the value of the MGF, $\Pi(z)$ at $z = 1$ is :

$$\begin{aligned} \Pi(1) &= \frac{\mu_0(\rho_0 + \rho_1)[P_{0,0,0} + P_{0,0,1} + P_{0,1,0} + P_{0,1,1}]}{[(\rho_0 + \rho_1)\mu_1 - \lambda_0\rho_1 - \lambda_1\rho_0]} \\ &+ \frac{[(\mu_1 - \mu_0)(\rho_0 + \rho_1)](\text{Sum}_{00} + \text{Sum}_{01} + \text{Sum}_{10} + \text{Sum}_{11})}{\gamma_{in}[(\rho_0 + \rho_1)\mu_1 - \lambda_0\rho_1 - \lambda_1\rho_0]} \\ &= 1 \end{aligned} \tag{B.5}$$

In equation B.5 since by initial assumption $\mu_1 > \mu_0$ and since the steady state probabilities are to be positive quantities, the necessary and sufficient condition for stability is $(\rho_0 + \rho_1)\mu_1 - \lambda_0\rho_1 - \lambda_1\rho_0 > 0$. The mean arrival rate for the MMPP arrival process was defined earlier as $\lambda^* = \lambda_0 \frac{\rho_1}{\rho_0 + \rho_1} + \lambda_1 \frac{\rho_0}{\rho_0 + \rho_1}$, so that the inequality can be rewritten as $\lambda^* < \mu_1$. In words the condition for stability is that the mean arrival rate of the MMPP should be less than at least one of the service rates.

B.2 Model D : Hysteresis Control, Poisson Arrivals, Delayed Bandwidth Change (In and Out) and Sampling

Solution Procedure

The state transition diagram for this model is shown in Figure 7.2 in Chapter 7. Partial MGFs are defined for the four switching phases. The unknown state probabilities in the MGFs are evaluated by the simultaneous solution of independent equations. These equations are obtained from the original balance equations, the conservation of probability sum property of the MGF and its analytic properties.

¹Reduce is a powerful system for performing symbolic algebra. It is useful in manipulating polynomials - expanding and factoring them. It can also be used for performing analytic differentiation and integration. In this work, the REDUCE package has been used for the factorisation of polynomial expressions and their evaluation at unity.

The partial MGFs $\Pi_j(z)$, $j = 0, 1, 2, 3$ are defined as $\Pi_j(z) \stackrel{\text{def}}{=} \sum_{i=0}^{\infty} z^i P_{i,j}$. The MGF is defined as $\Pi(z) = \Pi_0(z) + \Pi_1(z) + \Pi_2(z) + \Pi_3(z)$. The partial MGFs formed in the usual fashion by summing the balance equations 7.5 through to 7.8 in Chapter 7 over relevant indices are :

$$\Pi_0(z) = \frac{\mu_0(1-z)P_{0,0} - \delta z \sum_{i=0}^{U-1} z^i P_{i,0} - \gamma_{out} z \Pi_3(z)}{[\lambda z^2 - (\lambda + \mu_0 + \delta)z + \mu_0]} \quad (\text{B.6})$$

$$\Pi_1(z) = \frac{\mu_0(1-z)P_{0,1} + \delta z \sum_{i=0}^{U-1} z^i P_{i,0} - \delta z \Pi_0(z)}{[\lambda z^2 - (\lambda + \mu_0 + \gamma_{in})z + \mu_0]} \quad (\text{B.7})$$

$$\Pi_2(z) = \frac{\mu_1(1-z)P_{0,2} + \delta z \sum_{i=0}^L z^i P_{i,2} - \gamma_{in} z \Pi_1(z)}{[\lambda z^2 - (\lambda + \mu_1)z + \mu_1]} \quad (\text{B.8})$$

$$\Pi_3(z) = \frac{\mu_1(1-z)P_{0,3} - \delta z \sum_{i=0}^L z^i P_{i,2}}{[\lambda z^2 - (\lambda + \mu_1 + \gamma_{out})z + \mu_1]} \quad (\text{B.9})$$

The following definitions are made for the sake of notational convenience :

$$a_i = \mu_0(1-z)P_{0,i}, \quad a_j = \mu_1(1-z)P_{0,j}, \quad i = \{0, 1\}, j = \{2, 3\}$$

$$f_0 = \gamma_{out} z, \quad f_1 = \delta z, \quad f_2 = \gamma_{in} z$$

$$\text{Sum}_0 = \delta z \sum_{k=0}^{U-1} z^k P_{k,0}, \quad \text{Sum}_1 = \delta z \sum_{k=0}^L z^k P_{k,2}$$

$$D_0(z) = \lambda z^2 - (\lambda + \mu_0 + \delta)z + \mu_0$$

$$D_1(z) = \lambda z^2 - (\lambda + \mu_0 + \gamma_{in})z + \mu_0$$

$$D_2(z) = \lambda z^2 - (\lambda + \mu_1)z + \mu_1$$

$$D_3(z) = \lambda z^2 - (\lambda + \mu_1 + \gamma_{out})z + \mu_1$$

The partial MGFs in equations B.6 to B.9 can be rewritten in terms of these definitions as :

$$\Pi_3(z) = \frac{a_3 - \text{Sum}_1}{D_3(z)} \quad (\text{B.10})$$

$$\Pi_0(z) = \frac{D_3(z)[a_0 - \text{Sum}_0] - f_0[a_3 - \text{Sum}_1]}{D_0(z)D_3(z)}$$

$$\Pi_1(z) = \frac{D_0(z)D_3(z)[a_1 + \text{Sum}_0] - f_1[D_3(z)(a_0 - \text{Sum}_0) - f_0(a_3 - \text{Sum}_1)]}{D_0(z)D_1(z)D_3(z)}$$

$$\Pi_2(z) = \frac{D_0(z)D_1(z)D_3(z)[a_2 + \text{Sum}_1] - f_2[D_0(z)D_3(z)(a_1 + \text{Sum}_0)]}{D_0(z)D_1(z)D_2(z)D_3(z)}$$

$$- \frac{f_1[D_3(z)(a_0 - \text{Sum}_0) - f_0(a_3 - \text{Sum}_1)]}{D_0(z)D_1(z)D_2(z)D_3(z)}$$

The equations in B.10 involve $U+L+3$ unknown state probabilities, $P_{i,0}$, $i = 0, 1, \dots, U-1$ and $P_{i,2}$, $i = 0, 1, \dots, L$, $P_{0,1}$ and $P_{0,3}$. From the original balance equations 7.5 through to 7.8 in Chapter 7, $U + L - 1$ independent linear equations can be derived. These are used to express all the other unknown state probabilities in terms of the four unknown boundary probabilities — $P_{0,0}$, $P_{0,1}$, $P_{0,2}$ and $P_{0,3}$. One of the four remaining equations is obtained from the conservation of probability sum condition which is also used to derive the condition for stability. The other three may be obtained from a consideration of the analytic properties of the MGF as described in Section 7.7 of Chapter 7.

Conservation of Probability Sum :

Let the following be defined at $z = 1$: $\text{SSum}_0 \stackrel{\text{def}}{=} \sum_{i=0}^{U-1} P_{i,0}$ and $\text{SSum}_1 \stackrel{\text{def}}{=} \sum_{i=0}^L P_{i,2}$. Evaluating the partial MGFs in equations B.10 at $z = 1$ yields :

$$\begin{aligned}\Pi_3(1) &= \frac{\delta}{\gamma_{out}} \text{SSum}_1 \\ \Pi_0(1) &= \text{SSum}_0 + \text{SSum}_1 \\ \Pi_1(1) &= \frac{\delta}{\gamma_{in}} \text{SSum}_1\end{aligned}\tag{B.11}$$

while the value of $\Pi_2(1)$ is indeterminate. This indeterminacy is resolved by factoring $(z - 1)$ from the expression for $\Pi_2(z)$ by the REDUCE program. The expression for $\Pi(1)$ finally is :

$$\begin{aligned}\Pi(1) &= \frac{\mu_0(P_{0,0} + P_{0,1}) + \mu_1(P_{0,2} + P_{0,3})}{(\mu_1 - \lambda)} + \frac{(\mu_1 - \mu_0)}{(\mu_1 - \lambda)} \text{SSum}_0 \\ &+ \frac{(\mu_1 - \mu_0)}{(\mu_1 - \lambda)} \left[1 + \frac{\delta}{\gamma_{in}} \right] \text{SSum}_1\end{aligned}\tag{B.12}$$

In equation B.12 given the initial assumption that $\mu_1 > \mu_0$ for positive state probabilities it is necessary that $\mu_1 > \lambda$ which is also the sufficient condition for stability.

B.3 Model E : Hysteresis Control, MMPP Arrivals, Delayed Bandwidth Change (In and Out) and Sampling

Solution Procedure

The state transition diagram for this model is shown in Figure 7.3 in Chapter 7. Partial MGFs are defined for the eight arrival-switching phases. The unknown state probabilities in the MGF are evaluated by the simultaneous solution of independent equations. These equations are obtained from the original balance equations, the conservation of probability sum property of the MGF and its analytic properties.

The eight partial Z transforms $\Pi_{ij}(z)$, $j = \{0, 1, 2, 3\}$ and $i = \{0, 1\}$ where i is the arrival phase and j is the switching phase are formed in the usual fashion from the balance equations. The following definitions are made for notational convenience :

$$a_{ij} = \mu_0(1 - z)P_{0,i,j}, \quad i = \{0, 1\}, j = \{0, 1\}$$

$$a_{ij} = \mu_1(1 - z)P_{0,i,j}, \quad i = \{0, 1\}, j = \{2, 3\}$$

$$f_0 = \gamma_{out}z, f_1 = \rho_0z, f_2 = \delta z$$

$$f_3 = \gamma_{in}z, f_4 = \rho_1z$$

$$\text{Sum}_{00} = \delta z \sum_{i=0}^{U-1} z^i P_{i,0,0}$$

$$\text{Sum}_{10} = \delta z \sum_{i=0}^{U-1} z^i P_{i,1,0}$$

$$\text{Sum}_{02} = \delta z \sum_{i=0}^L z^i P_{i,0,2}$$

$$\text{Sum}_{12} = \delta z \sum_{i=0}^L z^i P_{i,1,2}$$

$$D_{i0}(z) = \lambda_i z^2 - (\lambda_i + \mu_0 + \delta + \rho_i)z + \mu_0, \quad i = \{0, 1\}$$

$$D_{i1}(z) = \lambda_i z^2 - (\lambda_i + \mu_0 + \gamma_{in} + \rho_i)z + \mu_0, \quad i = \{0, 1\}$$

$$D_{i2}(z) = \lambda_i z^2 - (\lambda_i + \mu_1 + \rho_i)z + \mu_1, \quad i = \{0, 1\}$$

$$D_{i3}(z) = \lambda_i z^2 - (\lambda_i + \mu_1 + \gamma_{out} + \rho_i)z + \mu_1, \quad i = \{0, 1\}$$

The partial MGFs expressed in terms of the definitions are :

$$\begin{aligned} \Pi_{00}(z) &= \frac{a_{00} - f_4 \Pi_{10}(z) - \text{Sum}_{00} - f_0 \Pi_{03}(z)}{D_{00}(z)} \\ \Pi_{10}(z) &= \frac{a_{10} - f_1 \Pi_{00}(z) - \text{Sum}_{10} - f_0 \Pi_{13}(z)}{D_{10}(z)} \\ \Pi_{01}(z) &= \frac{a_{01} - f_4 \Pi_{11}(z) + \text{Sum}_{00} - f_2 \Pi_{00}(z)}{D_{01}(z)} \\ \Pi_{11}(z) &= \frac{a_{11} - f_1 \Pi_{01}(z) + \text{Sum}_{10} - f_2 \Pi_{10}(z)}{D_{11}(z)} \\ \Pi_{02}(z) &= \frac{a_{02} - f_3 \Pi_{01}(z) + \text{Sum}_{02} - f_4 \Pi_{12}(z)}{D_{02}(z)} \\ \Pi_{12}(z) &= \frac{a_{12} - f_3 \Pi_{11}(z) + \text{Sum}_{12} - f_1 \Pi_{02}(z)}{D_{12}(z)} \end{aligned} \tag{B.13}$$

$$\begin{aligned}\Pi_{03}(z) &= \frac{a_{03} - f_4 \Pi_{13}(z) - \text{Sum}_{02}}{D_{03}(z)} \\ \Pi_{13}(z) &= \frac{a_{13} - f_1 \Pi_{03}(z) - \text{Sum}_{12}}{D_{13}(z)}\end{aligned}$$

The equations in B.13 involve $2U + 2L + 6$ unknowns : $P_{i,0,0}, P_{i,1,0}, P_{j,0,2}, P_{j,1,2}$, $i = \{0, 1, \dots, U - 1\}$, $j = \{0, 1, \dots, L\}$ and $P_{0,0,1}, P_{0,1,1}, P_{0,0,3}$ and $P_{0,1,3}$. From the balance equations 7.9 through to 7.12 in Chapter 7 can be derived $2U + 2L - 2$ independent equations. This would leave the 8 boundary probabilities to be solved for. 7 of these equations can be derived from the analytic properties of the MGF as described in Section 7.7 of Chapter 7 while the remaining equation is derived from the conservation of probability sum condition of the MGF as will be presented below.

Conservation of Probability Sum :

Let the following be defined at $z = 1$: $\text{SSum}_{j0} \stackrel{\text{def}}{=} \sum_{i=0}^{U-1} P_{i,j,0}$, $j = \{0, 1\}$ and $\text{SSum}_{j2} \stackrel{\text{def}}{=} \sum_{i=0}^L P_{i,j,2}$, $j = \{0, 1\}$. In terms of these definitions the values of the partial MGFs defined as $\Pi_j(z) \stackrel{\text{def}}{=} \Pi_{0j}(z) + \Pi_{1j}(z)$, $j = 0, 1, 2, 3$ are :

$$\begin{aligned}\Pi_3(1) &= \frac{\delta}{\gamma_{out}} (\text{SSum}_{02} + \text{SSum}_{12}) \\ \Pi_0(1) &= (\text{SSum}_{00} + \text{SSum}_{10} + \text{SSum}_{02} + \text{SSum}_{12}) \\ \Pi_1(1) &= \frac{\delta}{\gamma_{in}} (\text{SSum}_{02} + \text{SSum}_{12})\end{aligned}$$

while $\Pi_2(1)$ is indeterminate. The similarity of the above expressions to the equations B.11 for Model D in the previous section are to be noted. This indeterminacy can be resolved by factoring out $(z - 1)$ from $\Pi_2(z)$ and evaluating the resultant expression at unity by using the REDUCE program. The final expression for the value of the MGF $\Pi(z)$ at unity is displayed below :

$$\begin{aligned}\Pi(1) &= \\ &P_{0,0,0} \left\{ \frac{\mu_0 [(\delta + \gamma_{in} + \rho_0 + \rho_1)\rho_0 + (\delta + \rho_1)(\gamma_{in} + \rho_1) + \rho_0\rho_1]}{(\mu_1 - \lambda^*)(\delta + \rho_0 + \rho_1)(\gamma_{in} + \rho_0 + \rho_1)} \right\} \\ &P_{0,1,0} \left\{ \frac{\mu_0 [(\delta + \gamma_{in} + \rho_0 + \rho_1)\rho_1 + (\delta + \rho_0)(\gamma_{in} + \rho_0) + \rho_0\rho_1]}{(\mu_1 - \lambda^*)(\delta + \rho_0 + \rho_1)(\gamma_{in} + \rho_0 + \rho_1)} \right\} \\ &+ (P_{0,0,1} + P_{0,1,1}) \left\{ \frac{\mu_0}{(\mu_1 - \lambda^*)} \right\} \\ &+ (P_{0,0,2} + P_{0,1,2}) \left\{ \frac{\mu_1}{(\mu_1 - \lambda^*)} \right\} \\ &+ P_{0,0,3} \left\{ \mu_1 \frac{[(\delta + \gamma_{out} + \rho_0 + \rho_1)\rho_0 + (\delta + \rho_1)(\gamma_{out} + \rho_1) + \rho_0\rho_1]}{(\mu_1 - \lambda^*)(\delta + \rho_0 + \rho_1)(\gamma_{out} + \rho_0 + \rho_1)} \right\}\end{aligned}$$

$$\begin{aligned}
& + P_{0,1,3} \left\{ \mu_1 \frac{[(\delta + \gamma_{out} + \rho_0 + \rho_1)\rho_1 + (\delta + \rho_0)(\gamma_{out} + \rho_0) + \rho_0\rho_1]}{(\mu_1 - \lambda^*)(\delta + \rho_0 + \rho_1)(\gamma_{out} + \rho_0 + \rho_1)} \right\} \\
& + (SSum_{00} + SSum_{10}) \left\{ \frac{\mu_1 - \mu_0}{(\mu_1 - \lambda^*)\delta} \right\} \\
& + (SSum_{02} + SSum_{12}) \left\{ \frac{1}{\gamma_{out}} + \frac{(\mu_1 - \mu_0)}{\gamma_{in}(\mu_1 - \lambda^*)} + \frac{(\mu_1 - \mu_0)}{\delta(\mu_1 - \lambda^*)} \right\}
\end{aligned}$$

From this equation given the initial assumption that $\mu_1 > \mu_0$, for positive probabilities the necessary and sufficient condition for the steady state to exist is : $\lambda^* = \lambda_0 \frac{\rho_1}{\rho_0 + \rho_1} + \lambda_1 \frac{\rho_0}{\rho_0 + \rho_1} < \mu_1$ or the mean arrival rate of the MMPP should be less than at least one of the service rates.

Appendix C

Additional Experiments

This appendix describes two sets of experiments, not involving dynamic bandwidth management but that are relevant to the scope of the current work. The first set of experiments is a characterisation of packet loss in terms of the burstiness parameters of a traffic stream. The bursty stream is multiplexed onto the high priority queue of a U channel which is a queue of limited capacity. The second set of experiments describe measurements of the bandwidth switching times at a Unison ramp.

C.1 Multiplexing a Bursty Traffic Stream onto a U Channel

Cell Loss Rate (CLR) ¹ is one of the proposed measures of ATM network performance [Woodruff88], the others including blocking probability, delay and jitter. Recent studies by [Decina90a, Hirano89], reviewed in Section 5.2.3 of Chapter 5, have pointed out how the packet loss rate depends on the burstiness of the input traffic arrival process. In the following set of experiments, it will be studied how for a bursty traffic source multiplexed onto a U channel, the burstiness parameters could adversely effect system performance as quantified by packet loss. The traffic source is allocated bandwidth statically on the channel according to its average bit rate and dynamic bandwidth management is not used. The experiments were performed by multiplexing packets from the bursty source, onto the high priority queue of a U channel (with a capacity of approximately 50 packets) and measuring the instantaneous packet losses.

The traffic source is emulated by a two-phase traffic generator that varies between an active phase, in which packets are transmitted at a constant rate (with deterministic

¹The Packet Loss Rate (PLR) in the following discussions.

inter-generation times) which is the peak rate and an inactive phase, when there is no transmission. The burstiness of the traffic stream can be varied either by changing the durations of the two phases or by varying the packet generation rate in the active phase.

In the first experiment, the variation of packet loss percentage and packet loss rate with the continuance (duration) of the active period was considered while keeping the average and peak rates constant.² In the second experiment, the variation of packet loss percentage is considered against the peak rate again with the average bit rate kept constant. Different peak rates were achieved by varying the packet inter-generation times in the active phase.

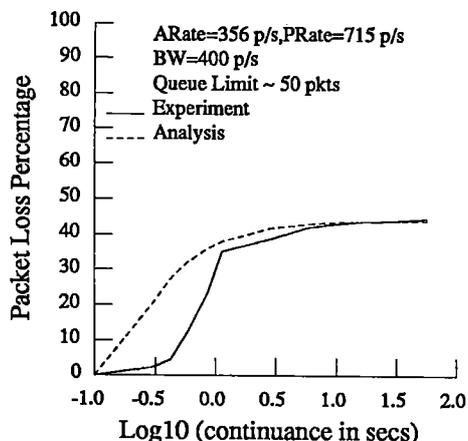


Figure C.1: Packet Loss Percentage vs. Logarithm of Continuance of the Active Period. Queue of Limited Capacity. **Parameter Set 6** : $\lambda_{avg}=356$ packets/sec, $\lambda_{peak}=715$ packets/sec, Bandwidth=400 packets/sec, Buffer limit ≈ 50 packets.

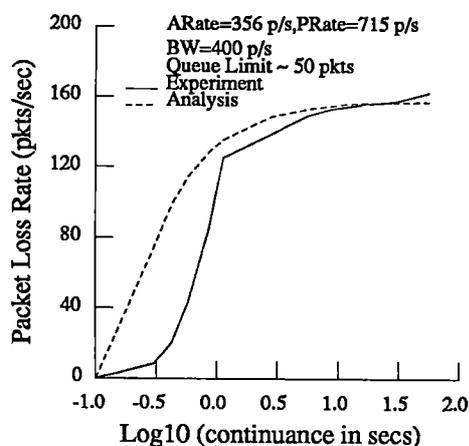


Figure C.2: Packet Loss Rate vs. Logarithm of Continuance of the Active Period. Queue of Limited Capacity. **Parameter Set 6**.

Results

Figure C.1 shows the variation of the *packet loss percentage* (defined as the ratio of the number of packets lost to the total number of packets generated) with the continuance of the active period, expressed on a logarithmic scale. For small values of the continuance, the packet loss percentage is low. The number of packets generated in an active period is

²The average rate is kept constant by keeping the duration of the inactive period equal to that of the active one.

less than or comparable with the queue capacity and the service rate, and packet losses are not pronounced. However, as the continuance increases and burst lengths exceed the queue capacity and service rate, the packet loss percentage increases markedly. It can be noticed from the figure, that the experimental trace follows that obtained from a simple theoretical analysis (that assumes a fixed size queue of 50 packets), better at larger values of the continuance. In Figure C.2 for the same experiment, is plotted the variation of packet loss rate against the logarithm of the continuance period. For increasing values of the continuance, the PLR tends to an asymptotic value, as indeed did the packet loss percentage in Figure C.1.

In the second experiment, the variation in packet loss percentage was studied against the ratio of the peak rate to the bandwidth of the U channel which was set at 400 packets/sec. This allocation of bandwidth was made on the average bit rate of the source — 356 packets/sec (the same as in the previous experiment). The average bit rate was kept constant by simultaneously varying the continuance of the inactive period with the peak rate. The duration of the active period was fixed at 0.2 second — a value that was chosen from the previous experiment as being one that did not cause a significant packet loss³ at the peak rate (715 packets/sec) employed in that experiment.

Figure C.3 shows the packet loss percentage variation with the peak rate to channel bandwidth ratio from experiments and from a simple theoretical analysis based on a fixed size queue of capacity 50 packets.⁴ The packet loss percentage increases with the peak rate, but there is a divergence between the experimental and theoretical values for larger values of the peak rate.

Analysis of Results

In this section, the discrepancies between experimental and theoretical results observed in figures C.1, C.2 and C.3 will be analysed. The main reason for the differences is that the theoretical model assumes that the high priority queue at the ramp can accommodate at the most 50 packets and that once this limit is reached, incoming packets will be lost. In practice however, the queue limit is a 'soft' one and more than 50 packets can be accommodated.

The soft limit on the high priority queue at the ramp is a function of the bandwidth of its outgoing U channel⁵ which in both experiments was kept constant at 2 B channels. If the queue state is continuously monitored, than when the limit is reached, incoming

³In the present context a value of 10^{-1} is used. For future broadband networks, a value as low as 10^{-8} is being considered [Woodruff90].

⁴In the theoretical analysis the number of packets generated during an active phase was calculated from the arrival rate and the duration of the active phase. The number of packets lost was then calculated from this number, the service rate and a queue limit of 50 packets.

⁵'Soft' queue limit = $25 \times$ Number of slots in U channel. 25 was a number chosen by the designers of the ramp as one that would constrain the delay to a reasonable value.

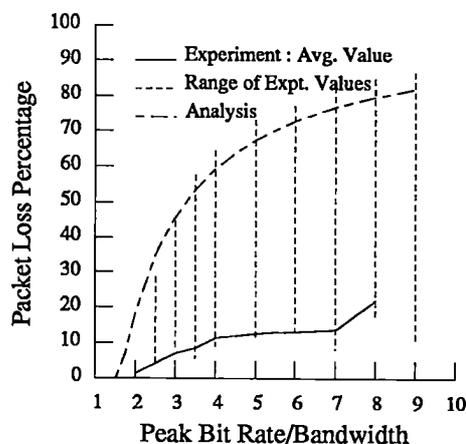


Figure C.3: Packet Loss Percentage vs. Ratio of the Peak Rate to Assigned Channel Bandwidth. Queue of Limited Capacity. Parameter Set 7 : $\lambda_{avg}=356$ packets/sec, Bandwidth=400 packets/sec, Buffer limit ≈ 50 packets, Active period duration=0.2 secs.

packets would be dropped. However, since continuous monitoring of the queues in the ramp is an expensive operation, the queue state is observed only at discrete intervals of 250 ms (4 Hz). At such an observation point, if the queue size is found to exceed the limit, then the ramp inhibits the further flow of incoming packets so that all packets generated upto the next observation point are lost. At that instant, input flow is enabled only if the queue has fallen below the limit. Packets generated during an enabling interval, can be accommodated beyond the soft queue limit. For instance in the second set of experiments, for large values of the peak rate, instantaneous queue sizes of upto 800 packets were observed. Discrete observation points and queues with soft limits will be used to explain the differences between experimental and theoretical results.

In the second experiment, the deviation is best illustrated by a numerical example. Consider the case when the peak rate is 4000 packets/sec (a peak rate/channel bandwidth ratio of 10) and the active period duration is 0.2 seconds. When the active period lies within an observation interval, upto 800 packets can be generated and accommodated in the queue. During this period, upto 100 packets can be transmitted, leaving a backlog of 700 packets which is cleared during the subsequent inactive phase of length 1.8 seconds, so very few packets are lost. In general, the observation point could fall anywhere within an active period and inhibition will occur if this point lies in the range 0.014 seconds to 0.2 seconds (0.014 seconds is the time taken at this input rate to fill the queue upto its soft limit) from the start of the active period. Packet losses thus tend to be less than that predicted by the theoretical analysis which assumes a queue of fixed capacity — 50 packets. It also means that the packet loss is variable and in the Figure C.3 are plotted

both the average value of the packet loss percentage and the range of its values from a set of experimental runs. Because of the large variations in the packet loss percentage and the relatively small number of sample runs made, the results on the figure should be treated as only being indicative of how the queues in a ramp could behave.

In the first experiment, the deviation of the experimental measurements of packet loss rates and packet loss percentages from the theoretical values is noticeable at small continuances. As explained above when the continuance and observation interval lengths are comparable, the packet loss tends to depend on when an observation is made. At the transmission rate of 715 packets/sec, the soft limit of 50 packets would be reached in 0.16 seconds. When the continuance is small — 0.2 seconds, packets would thus be lost only if the observation point falls within the interval 0.16 — 0.20 seconds.

The convergence of the packet loss rates and packet loss percentages to asymptotic values as the duration of the active period increases in the first experiment, can also be explained numerically. When the duration of the active period is much longer than the length of an observation interval, the system alternates between enabling and disabling periods. This is because at the end of an enabling interval of 0.25 seconds, at the transmission rate of 715 packets/second, a backlog of about 78 packets can exist (the difference between the number of packets generated : 178 and the number transmitted during that period : 100). As this backlog exceeds the soft limit, inhibition of input flow takes place for the next 0.25 seconds so that approximately 178 incoming packets are dropped. By the next observation point, the backlog would have been cleared and input flow is enabled again. This implies that the packet loss percentage tends to 50 as the continuance increases. An asymptotic convergence to that value is observed in Figure C.1. The packet loss rate defined as the number of packets lost divided by the sum of the durations of the active and inactive periods has a limiting value of 178 packets/sec. For the largest value of active period continuance used in the experiments, the value obtained was 163 packets/sec.

These experiments demonstrate that packet loss is closely related to the burstiness of the input traffic stream and the bandwidth allocation for such a stream should be greater than the average rate of the stream (a virtual bandwidth measure between the average and peak bit rates). A conservative solution is to allocate bandwidth on the basis of peak rates which would avoid the problem of packet loss. The experiments also demonstrate that the combination of discrete sampling and queues with 'soft' limits at the ramp, does tend to reduce packet losses at the expense of added jitter in delay.

C.2 Bandwidth Switching Times

The experiments in this section describe bandwidth (B channel) switching times on U channels to RAL or locally to the System X exchange and back within Cambridge when operating in loopback mode.

Figure C.4 shows bandwidth switching in times to establish B channels to RAL for a variety

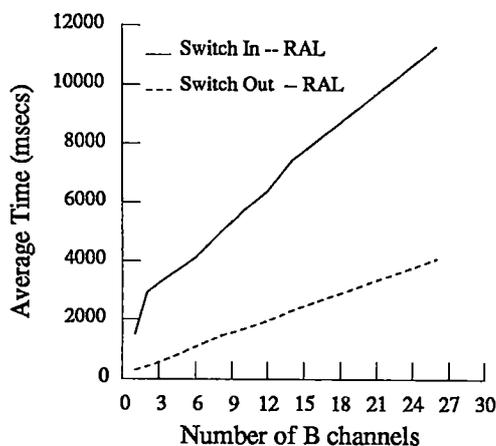


Figure C.4: Bandwidth Switching Times (In and Out) to RAL. Switching Times vs. Number of B Channels.

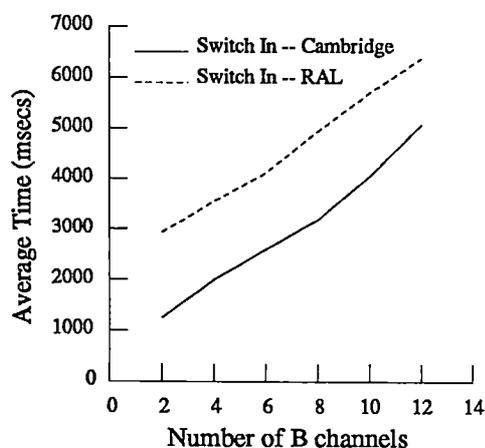


Figure C.5: Bandwidth Switching In Times vs. Number of B Channels to RAL and Cambridge — a Comparison.

of bandwidths ranging from 1 B channel (64 Kbps) to 30 B channels (1920 Kbps). For the switch in curve, the ordinate values represent the time taken to bring in the corresponding bandwidth when starting from a U channel of zero bandwidth. These values include both the call signalling and U channel synchronisation components discussed in Section 3.3 of Chapter 3. The switching out times were computed starting with a U channel of a particular bandwidth and ending with a null one.

Figure C.5 compares switching in times for establishing circuits to RAL and loopback circuits to the Cambridge System X exchange, for a more limited set of bandwidths ranging from 128 Kbps(2 B channels) to 768 Kbps (12 B channels). Each ordinate value as before represents the time to bring in the associated bandwidth on a null U channel. ⁶ As is evident from the figure, the times taken to establish a U channel of comparable bandwidth locally are less than the remote establishment to RAL - the more or less constant difference being mainly due to the differences in signalling times in the two cases.

Figure C.6 demonstrates the *ramping* effect. In this experiment, the time taken to establish a U channel of maximum bandwidth ie. 30 B channels or 1920 kbps starting from a null channel is studied. During the course of the adjust, the ramp returns status messages to the channel service, detailing the active width of the U channel and the time taken as synchronised B channels are added to it ready for transmission upon. The times and active widths are the values plotted on the figure.

⁶Loopback bandwidth increments have to be made in multiples of two. Further the number of loopback circuits is limited because of constraints in the DASS2 signalling protocols.

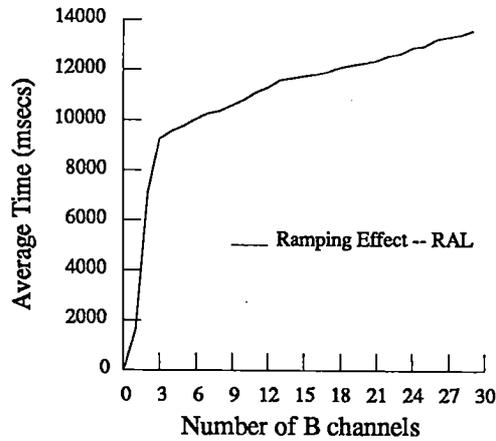


Figure C.6: Ramping Effect to RAL

The experiment demonstrates the pipelining effect in bringing in B channels. Since signalling requests are pipelined, the time taken to set up N calls is not equal to the sum of bringing them in, one at a time. For example making a single adjust for 10 slots takes about 10.8 seconds, doing this with 10 separate adjustments (successive) would have taken as much as $10 \times 1.5 = 15$ seconds. When signalling requests return successfully they do so with an assigned slot which is then amalgamated into an existing U channel by a synchronisation process described in Section 3.3 of Chapter 3. Although multiple slots can be synchronised simultaneously (ie. the synchronisation process can be pipelined) this does not happen in practice, because the time between signalling request returns is usually much greater than the synchronisation time.

Summary

In this appendix the effect of burstiness in an input traffic stream incident on a transmission channel with a queue of finite capacity was analysed. It was seen how queues with soft limits could decrease the packet loss at the expense of additional jitter. The appendix also contained measurements of the bandwidth switching times in the Unison network.



References

- [Akhtar85] S. Akhtar. Congestion Control in a Fast Packet Switching Network. Master's Thesis, Washington Univ., St. Louis, Missouri, 1985.
- [Altarah89] L. Altarah and S. Motard. Dynamic Use of ISDN B Channels. *Proceedings of ISDN in Europe'89*, pages 155-165, April 1989.
- [Amstutz83] S. R. Amstutz. Burst Switching - an Introduction. *IEEE Commun. Mag.*, 21(8):36-42, November 1983.
- [Amstutz89] S. R. Amstutz. Burst Switching - an Update. *IEEE Commun. Mag.*, 27(9):50-57, September 1989.
- [Ballart89] R. Ballart and Y.C. Ching. SONET : Now It's the Standard Optical Network. *IEEE Commun. Mag.*, 27(3):8-15, March 1989.
- [Bolot90] J. Bolot and A.U. Shankar. Dynamical Behaviour of Rate-Based Flow Control Mechanisms. *Computer Commun. Review*, 20(2):35-49, April 1990.
- [Boltz89] B. Boltz, Y. Liang, and J. Roberts. Provision of N x 64 Kbits/s Services in the ISDN. *Proceedings of ISDN in Europe'89*, pages 167-175, April 1989.
- [Brady68] P.T. Brady. A Statistical Analysis of On-Off Patterns in 16 Conversations. *Bell System Technical Journal*, 47:73-91, January 1968.
- [Burren89a] J.W. Burren. Flexible Aggregation of Channel Bandwidth in Primary Rate ISDN. *SIGCOMM'89 Symp., Comp. Commun. Review*, 19(4):191-196, September 1989.
- [Burren89b] J.W. Burren and C.S. Cooper. *Project Universe : An Experiment in High-Speed Computer Networking*. Oxford University Press, Oxford, 1989.
- [Calnan89] R. Calnan. *The Integration of Voice Within a Digital Network*. PhD Thesis, University of Cambridge Computer Laboratory, Cambridge, UK, 1989.
- [CCI84] CCITT. *Recommendation I.462 (X.31) : Support of Packet Mode Terminal Equipment by an ISDN*, red book Edition, 1984.
- [CCI85] CCITT, Geneva. *Recommendation X.25*, red book Edition, 1985. Vol. V111.3.
- [CCI88] CCITT, Melbourne, Australia. *Recommendation of the Series I - ISDN*, blue book Edition, 1988.
- [CCI89] CCITT, Geneva. *Recommendation I.121 : Broadband Aspects of ISDN*, blue book Edition, 1989. Vol. 111.7.
- [Chang88] R. F. Chang and V.K. Li. Analysis of Packet Delay for an Integrated Voice-Data System. In *Proc. IEEE GLOBECOM'88*, pages 1095-1110, November 1988.

- [Chen88] T.M. Chen and D.G. Messerschmitt. Integrated Voice/Data Switching. *IEEE Commun. Mag.*, 26(6):16-26, June 1988.
- [Chen89] K.J. Chen, K.K.Y. Ho, and V.R. Saksena. Analysis and Design of a Highly Reliable Transport Architecture for ISDN Frame-Relay Networks. *IEEE J. Select. Areas in Commun.*, 7(8):1231-1242, October 1989.
- [Clark86] P.F. Clark et al. Unison - Communications Research for Office Applications. *IEE Electronics and Power*, 32:665-668, September 1986.
- [Cooper81] R.B. Cooper. *Introduction to Queueing Theory*. North Holland, New York, 1981.
- [Crabill77] T. Crabill, D. Gross, and M. Magazine. Dynamical Behaviour of Rate-Based Flow Control Mechanisms. *Oper. Res.*, 25(2):219-232, 1977.
- [Daigle86] J.N. Daigle and J.D. Langford. Models for Analysis of Packet Voice Communications Systems. *IEEE J. Select. Areas in Commun.*, SAC-4(6):847-855, September 1986.
- [Decina90a] M. Decina and T. Toniatti. On Bandwidth Allocations to Bursty Virtual Connections in ATM Networks. In *Proc. IEEE ICC'90*, pages 318.6.1-318.6.8, April 1990.
- [Decina90b] M. Decina, T. Toniatti, P. Vaccari, and L. Verri. Bandwidth Assignment and Virtual Call Blocking in ATM Networks. In *Proc. IEEE INFOCOM'90*, pages 881-888, May 1990.
- [Demers89] A. Demers, S. Keshav, and S. Shenker. Analysis and Simulation of a Fair Queueing Algorithm. *SIGCOMM'89 Symp., Comp. Commun. Review*, 19(4):1-12, September 1989.
- [Deniz89] D.Z. Deniz and G.J. Knight. Channel Management Issues in LAN-ISDN Interconnection. In *Proceedings of ISDN in Europe'89*, pages 177-188, April 1989.
- [Deniz90] D.Z. Deniz and G.J. Knight. On the Formation of Superchannel and Dynamic Management of Superchannels in Primary Rate ISDN. In *Proceedings of International Symposium on Local Communications Systems Management*, Canterbury, UK, September 1990. IFIP TC6 WG 6.42.
- [Dittmann88] L. Dittmann and S.B. Jacobsen. Statistical Multiplexing of Identical Bursty Sources in an ATM Network. In *Proc. IEEE GLOBECOM'88*, pages 39.6.1-39.6.6, November 1988.
- [Eisen63] M. Eisen and M. Tainiter. Stochastic Variations in Queueing Processes. *Oper. Res.*, 11:922-927, 1963.
- [Falconer85] R.M. Falconer and J.L. Adams. Orwell : A Protocol for an Integrated Services Network. *British Telecom Technology Journal*, 3(4):27-35, October 1985.
- [Filipak89] J. Filipak. Structured Systems Analysis Methodology for Design of an ATM Network Architecture. *IEEE J. Select. Areas in Commun.*, 7(8):1262-1273, October 1989.
- [Forgie77] J.W. Forgie and A.G. Nemeth. An Efficient Packetized Voice/Data Network using Statistical Flow Control. In *Proc. IEEE ICC'77*, pages 38.2.44-38.2.48, June 1977.

- [Fukuda87] A. Fukuda. Behavior of Queues with Randomly Varying Population Mode. In *International Conference on Communication Technology*, pages 618–621, Nanjing, China, 1987.
- [Gafni84] E.M. Gafni and D.P. Bertsekas. Dynamic Control of Session Input Rates in Communication Networks. *IEEE Transactions on Automatic Control*, AC-29(11):1009–1016, November 1984.
- [Gallassi89] G. Gallassi, G. Rigolio, and L. Fratta. ATM: Bandwidth Assignment and Bandwidth Enforcement Policies. In *Proc. IEEE GLOBECOM'89*, pages 49.6.1–49.6.6, November 1989.
- [Gebhard67] R.F. Gebhard. A Queueing Process with Bilevel Hysteretic Service Rate Control. *Naval Research Logistics Quarterly*, 14:55–68, 1967.
- [Gerla81] M. Gerla. Bandwidth Routing in X.25 Networks. In *Proc. Pacific Telcomm. Conf.*, pages A3.25–A3.32, Hawaii, January 1981.
- [Gerla83] M. Gerla. Packet, Circuit and Virtual Circuit Switching. In W. Chou, editor, *Computer Communications*, chapter 16. Prentice Hall, Englewood Cliffs, 1983. Vol. 2.
- [Gerla84a] M. Gerla. Controlling Routes, Traffic Rates and Buffer Allocations in Packet Networks. *IEEE Commun. Mag.*, 22(11):11–23, November 1984.
- [Gerla84b] M. Gerla and R.A. Pazos-Rangel. Bandwidth Allocation and Routing in ISDN's. *IEEE Commun. Mag.*, 22(2):16–26, February 1984.
- [Gersht88] A. Gersht and K.J. Lee. Virtual-Circuit Load Control in Fast Packet-Switched Broadband Networks. In *Proc. IEEE GLOBECOM'88*, pages 7.3.1–7.3.7, November 1988.
- [Gersht89] A. Gersht and K.J. Lee. A Congestion Control Framework for ATM Networks. In *Proc. IEEE INFOCOM'89*, pages 701–710, April 1989.
- [Gopal83] I.S. Gopal and T.E. Stern. Optimal Call Blocking Policies in an Integrated Services Environment. In *Proc. Conf. Infor. Sciences and Syst.*, pages 383–388, The John Hopkins University, March 1983.
- [Greaves90] D. Greaves, D. Lioupis, and A. Hopper. The Cambridge Backbone Ring. In *Proc. IEEE INFOCOM'90*, pages 8–15, June 1990.
- [Griffiths89] J.W. Griffiths et al. The Unison Project. In *Proceedings of the 1989 Singapore International Conference on Networks*, July 1989.
- [Gross85] D. Gross and C.M. Harris. *Fundamentals of Queueing Theory*. John Wiley and Sons, New York, 1985.
- [Gruber81] J.G. Gruber. Delay Related Issues in Integrated Voice and Data Networks. *IEEE Trans. Commun.*, COM-29(6):786–799, June 1981.
- [Halsall88] F. Halsall. *Data Communications, Computer Networks and OSI*. Addison Wesley, Wokingham, 1988.
- [Harita89] B.R. Harita and I.M. Leslie. Dynamic Bandwidth Management of Primary Rate ISDN to Support ATM Access. *SIGCOMM'89 Symp., Comp. Commun. Review*, 19(4):197–210, September 1989.

- [Hasleton83] E.F. Hasleton. A PCM Frame Switching Concept Leading to Burst Switching Network Architecture. *IEEE Commun. Mag.*, 21(6):13-19, September 1983.
- [Hearn85] A.C. Hearn. *Reduce Users Manual*. The Rand Corporation, California, version 3.2 Edition, 1985.
- [Heffes86] H. Heffes and D.M. Lucantoni. A Markov Modulated Characterization of Packetised Voice and Data Traffic and Related Statistical Multiplexer Performance. *IEEE J. Select. Areas in Commun.*, SAC-4(6):856-868, September 1986.
- [Hellstern89] K. S. M. Hellstern. The Analysis of a Queue Arising in Overflow Models. *IEEE Trans. Commun.*, 37(4):367-373, April 1989.
- [Hemrick88] C.F. Hemrick, R.W. Klessig, and J.M. McRoberts. Switched Multi-Megabit Data Service and Early Availability via MAN Technology. *IEEE Commun. Mag.*, 26(4), April 1988.
- [Heyman68] D.P. Heyman. Optimal Operating Policies for M/G/1 Queueing Systems. *Oper. Res.*, 16:362-382, 1968.
- [Hirano89] M. Hirano and N. Watanabe. Characteristics of a Cell Multiplexer for Bursty ATM Traffic. In *Proc. IEEE ICC'89*, pages 13.2.1-13.2.5, June 1989.
- [Hopper86] A. Hopper, S. Temple, and R. C. Williamson. *Local Area Network Design*. Addison Wesley, London, 1986.
- [Hopper88] A. Hopper and R.M. Needham. The Cambridge Fast Ring Networking System. *IEEE Transactions on Computers*, 37(10):1214-1223, October 1988.
- [Hopper90] A. Hopper. Pandora - An Experimental System for Multimedia Applications. *Operating Systems Review*, 24(2):19-34, April 1990.
- [Hui88] J.Y. Hui. Resource Allocation for Broadband Networks. *IEEE J. Select. Areas in Commun.*, 6(9):1598-1608, December 1988.
- [Hui90] J.Y. Hui. *Switching and Traffic Theory for Integrated Broadband Networks*. Kluwer Academic Publishers, Boston, 1990.
- [IEE88] IEEE 802.6 Working Group. *Proposed IEEE Standard 802.6 - DQDB - Metropolitan Area Network*, July 1988.
- [Jacobsen90] S.B. Jacobsen, K.Moth, and L. Dittmann. Load Control in ATM Networks. In *Proc. Int. Switching Symp. (ISS'90)*, pages V.131-V.138, May 1990.
- [Jenq84] Y.C. Jenq. Approximations for Packetized Voice Traffic in Statistical Multiplexer. In *Proc. IEEE INFOCOM'84*, pages 256-259, April 1984.
- [Joos89] P. Joos and W. Verbiest. A Statistical Bandwidth Allocation and Usage Monitoring Algorithm for ATM Networks. In *Proc. IEEE ICC'89*, pages 13.5.1-13.5.8, June 1989.
- [Kamitake89] T. Kamitake and T. Suda. Evaluation of an Admission Control Scheme for an ATM Network Considering Fluctuations in Cell Loss Rate. In *Proc. IEEE GLOBECOM'89*, pages 49.4.1-49.4.7, November 1989.
- [Kawarazaki90] M. Kawarazaki, H. Saito, and H. Yamada. Traffic Analysis of an ATM Transport Network. In *Proc. IEEE ICC'90*, pages 308.1.1-308.1.5, April 1990.
- [Kawashima89] K. Kawashima and H. Saito. Teletraffic Issues in ATM Networks. In *ITC Specialist Seminar*, pages 17.5.1-17.5.8, Adelaide, Australia, 1989.

- [Kermani79] P. Kermani and L. Kleinrock. Virtual Cut-Through : A New Computer Communication Switching Technique. *Computer Networks*, 3:267-286, 1979.
- [King82] P.J.B. King and I. Mitrani. Modelling the Cambridge Ring. *Perf. Eval. Review*, 11(4):250-258, 1982.
- [Kleinrock75] L. Kleinrock. *Queueing Systems*, volume 1. John Wiley, New York, 1975.
- [Knight87] G.J. Knight, D.Z. Deniz, and J. Fan. Gateways to the ISDN. In *Proceedings of ISDN'87*, June 1987.
- [Koistpaiboon90] R. Koistpaiboon, G. M. Woodruff, and P. Richards. ATM Traffic Control for Guaranteed Performance. *To appear in International Journal of Digital and Analog Cabled Systems*, 1990.
- [Kraimeche84] B. Kraimeche and M. Schwartz. Circuit Access Control Strategies in Integrated Digital Networks. In *Proc. IEEE INFOCOM'84*, pages 230-235, April 1984.
- [Kraimeche85] B. Kraimeche and M. Schwartz. Analysis of Traffic Access Control Strategies in Integrated Services Networks. *IEEE Trans. Commun.*, COM-33(10):1085-1093, October 1985.
- [Kraimeche86] B. Kraimeche and M. Schwartz. Bandwidth Allocation Strategies in Wide-Band Integrated Networks. *IEEE J. Select. Areas in Commun.*, SAC-4(6):869-878, September 1986.
- [Kulzer84] J. J. Kulzer and W.A. Montgomery. Statistical Switching Architectures for Future Services. In *Proc. Int. Switching Symp. (ISS'84)*, pages 43.A.1.1-43.A.1.6, May 1984.
- [Kummerle74] K. Kummerle. Multiplexor Performance for Integrated Line and Packet-Switched Traffic. In *Proc. IEEE ICC'74*, pages 507-515, August 1974.
- [Law82] A.M. Law and W.D. Kelton. *Simulation, Modelling and Analysis*. McGraw-Hill Book Company, New York, 1982.
- [Lee89] D.S. Lee and S.Q. Li. Transient Analysis of a Switched Poisson Arrival Queue under Overload Control, 1989. Submitted to Performance Evaluation.
- [Leslie84] I.M. Leslie et al. The Architecture of the UNIVERSE Network. *SIGCOMM'84 Symp., Comp. Commun. Review*, 14(2):2-9, 1984.
- [Li85] S.Q. Li and J.W. Mark. Performance of Voice/Data Integration on a TDM System. *IEEE Trans. Commun.*, COM-33(12):1265-1273, December 1985.
- [Li89a] S.Q. Li, August 1989. Personal Communication.
- [Li89b] S.Q. Li. Overload Control in a Finite Message Storage Buffer. *IEEE Trans. Commun.*, 37(12):1330-1338, December 1989.
- [Li90] S.Q. Li. A General Solution Technique for Discrete Queueing Analysis of Multi-Media Traffic on ATM. In *Proc. IEEE INFOCOM'90*, pages 1144-1155, June 1990.
- [Lu84] F.V. Lu and R.F. Serfozo. M/M/1 Queueing Decision Processes with Monotone Hysteretic Optimal Policies. *Oper. Res.*, 32(5):1116-1132, 1984.
- [Maglaris87] B. Maglaris et al. Performance Analysis of Statistical Multiplexing for Packet Video Sources. In *Proc. IEEE GLOBECOM'87*, pages 1890-1899, November 1987.

- [McAuley87] D.R. McAuley. Unity : An RPC Mechanism. Technical Report UC027, Project Unison, March 1987.
- [McAuley90] D.R. McAuley. Protocol Design for High Speed Networks. Technical Report 186, University of Cambridge Computer Laboratory, January 1990.
- [Minzer89] S.E. Minzer. Broadband ISDN and Asynchronous Transfer Mode (ATM). *IEEE Commun. Mag.*, 27(9):17-24, September 1989.
- [Mitrany68] I.L. Mitrany and B.Avi-Itzhak. A Many Server Queue with Service Interruptions. *Oper. Res.*, 16:628-638, 1968.
- [Nagle87] J.B. Nagle. On Packet Switches with Infinite Storage. *IEEE Trans. Commun.*, COM-35(4):435-438, April 1987.
- [Neuts81] M. F. Neuts. *Matrix Geometric Solutions in Stochastic Models: An Algorithmic Approach*. The John Hopkins University Press, Baltimore, 1981.
- [Neuts84] M.F. Neuts. Problems of Reducibility in Structured Markov Chains of M/G/1 type and Related Queueing Models in Communication Engineering. In G. Iazellola, P.J. Courtois, and A. Hordijk, editors, *Mathematical Computer Performance and Reliability*, pages 139-151. North Holland, 1984.
- [Neuts85] M. F. Neuts. A Queueing Model for a Storage Buffer in Which the Arrival Rate is Controlled by a Switch with a Random Delay. *Performance Evaluation*, 5:243-256, 1985.
- [Newman86] R. Newman. Introduction of ISDN into the British Telecom Network. *IEEE J. Select. Areas in Commun.*, SAC-4(3):385-389, May 1986.
- [Newman88] P. Newman. A Fast Packet Switch for the Integrated Services Backbone Network. *IEEE J. Select. Areas in Commun.*, SAC-6(9):1468-1479, December 1988.
- [Ohta88a] M. Ohta. Effects of Control Delay and Transmission Delay on Congestion Controls. In *Proc. IEEE ICC'88*, pages 127-131, 1988.
- [Ohta88b] S. Ohta, K. Sato, and I. Tokizawa. A Dynamically Controllable ATM Transport Network Based on the Virtual Path Concept. In *Proc. IEEE GLOBECOM'88*, pages 39.2.1-39.2.5, November 1988.
- [Oie90] Y. Oie et al. Survey of Switching Techniques in High-Speed Networks and their Performance. In *Proc. IEEE INFOCOM'90*, pages 1242-1250, June 1990.
- [O'Reilly86] P. O'Reilly. Performance Analysis of Data in Burst Switching. *IEEE Trans. Commun.*, COM-34(12):1259-1263, December 1986.
- [Pazos-Rangel82] R.A. Pazos-Rangel and M. Gerla. Express Pipe Networks. In *Proc. IEEE GLOBECOM'82*, pages B2.3.1-B2.3.5, Miami, December 1982.
- [Porter90] J.D. Porter. *MAC layer Bridging of Homogeneous Local Area Networks*. PhD Thesis, University of Cambridge Computer Laboratory, 1990. In preparation.
- [Prycker88] M.D. Prycker. Data Communication in an ATM Network. *Electrical Communication*, 62(3/4):333-337, 1988.
- [Rathgeb90] E.R. Rathgeb and T.H. Theimer. The Policing Function in ATM Networks. In *Proc. Int. Switching Symp. (ISS'90)*, pages V.127-V.130, May 1990.

- [Richards79] M. Richards et al. TRIPOS : A Portable Operating System for Mini-Computers. *Software - Practice and Experience*, 9:513-526, 1979.
- [Rider89] M. J. Rider. Protocols for ATM Access Networks. *IEEE Commun. Mag.*, 27(1):17-22, January 1989.
- [Saito89] H. Saito. A Simplified Dimensioning Method of ATM Networks. Technical Report SSE89-112, IEICE, Japan, November 1989.
- [Saito90a] H. Saito. Call Admission Control in an ATM Network without Using Traffic Measurement. Technical Report SSE89-155, IEICE, Japan, February 1990.
- [Saito90b] H. Saito and K. Kawashima. Performance Modelling of Integrated Voice and Data Communication Networks. In H. Takagi, editor, *Stochastic Analysis of Computer and Communication Systems*, pages 807-849. North Holland, 1990.
- [Sallberg90] K. Sallberg, B. Stavenow, and I. Andersen. A Resource Allocation Framework in B-ISDN. In *Proc. Int. Switching Symp. (ISS'90)*, pages I.111-I.118, May 1990.
- [Sato90] K. Sato, S. Ohta, and I. Tokizawa. Broad-Band ATM Network Architecture Based on Virtual Paths. *IEEE Trans. Commun.*, 38(8):1212-1222, August 1990.
- [Schwartz82] M. Schwartz and B. Kraimeche. Comparison of Channel Assignment Techniques for Hybrid Switching. In *Proc. IEEE ICC'85*, pages 2F.3.1-2F.3.5, June 1982.
- [Schwartz87] M. Schwartz. *Telecommunication Networks : Protocols, Modeling and Analysis*. Addison Wesley, Reading, Mass., 1987.
- [Serres88] Y.D. Serres and L.G. Mason. A Multiserver Queue with Narrow- and Wide-Band Customers and Wide-Band Restricted Access. *IEEE Trans. Commun.*, 36(6):675-684, June 1988.
- [Sotelo87] W. Sotelo and A. Fukuda. On Multiserver Queues with Synchronous Fluctuation of Traffic Intensity. *Trans. of the IEICE*, E 70(10):951-959, October 1987.
- [Sotelo88] W. Sotelo, K. Mukumoto, and A. Fukuda. M/M/K Systems with M-Phase Fluctuations of Traffic Intensity. In *Proc. IEEE INFOCOM'88*, pages 7A.3.1-7A.3.9, April 1988.
- [Sriram83] K. Sriram, P.K. Varshney, and J.G. Shantikumar. Discrete-Time Analysis of Integrated Voice/Data Multiplexers With and Without Speech Activity Detectors. *IEEE J. Select. Areas in Commun.*, SAC-1(6):1124-1132, December 1983.
- [Sriram86] K. Sriram and W. Whitt. Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data. *IEEE J. Select. Areas in Commun.*, SAC-4(6):833-846, September 1986.
- [Sriram89] K. Sriram and D.M. Lucantoni. Traffic Smoothing Effects of Bit Dropping in a Packet Voice Multiplexer. *IEEE Trans. Commun.*, 37(7):703-712, July 1989.
- [Tennenhouse86] D.L. Tennenhouse. The CFR Unison Data-Link Protocol Specification. Technical Report UC021, Project Unison, October 1986.
- [Tennenhouse87] D.L. Tennenhouse et al. Exploiting Wideband ISDN: The Unison Exchange. In *Proc. IEEE INFOCOM'87*, pages 1018-1026, April 1987.

- [Tennenhouse89a] D.L. Tennenhouse. Site Interconnection and the Exchange Architecture. Technical Report 184, University of Cambridge Computer Laboratory, October 1989.
- [Tennenhouse89b] D.L. Tennenhouse and I.M. Leslie. A Testbed for Wide Area ATM Research. *SIGCOMM'89 Symp., Comp. Commun. Review*, 19(4):182-190, September 1989.
- [Tirtaatmajda90] E. Tirtaatmajda and R.A. Palmer. The Application of Virtual Paths to the Interconnection of IEEE 802.6 Metropolitan Area Networks. In *Proc. Int. Switching Symp. (ISS'90)*, pages II.133-II.146, May 1990.
- [Turner86a] J.S. Turner. Design of an Integrated Services Packet Network. *IEEE J. Select. Areas in Commun.*, SAC-4(8):1373-1379, November 1986.
- [Turner86b] J.S. Turner. New Directions in Communications (or Which Way to the Information Stage). *IEEE Commun. Mag.*, 24(10):8-15, October 1986.
- [Verbiest88a] W. Verbiest, L. Pinoo, and B. Voeten. The Impact of ATM Concept on Video Coding. *IEEE J. Select. Areas in Commun.*, 6(9):1623-1632, December 1988.
- [Verbiest88b] W. Verbiest, L. Pinoo, and B. Voeten. Statistical Multiplexing of Variable Bit Rate Video Sources in Asynchronous Transfer Mode Networks. In *Proc. IEEE GLOBECOM'88*, pages 7.2.1-7.2.6, November 1988.
- [Vukotic88] M. Z. Vukotic, I.G. Niemegeers, and D.S. Valk. Performance Analysis of Slotted Ring Protocols in HSLAN's. *IEEE J. Select. Areas in Commun.*, 6(6):1011-1024, July 1988.
- [Weinstein80] C. J. Weinstein, M. L. Malpass, and M.J. Fisher. Data Traffic Performance of an Integrated Circuit- and Packet-Switched Multiplex Structure. *IEEE Trans. Commun.*, COM-28(6):873-878, June 1980.
- [Wernik88] M. R. Wernik. Architecture and Technology Considerations for Multimedia Broadband Communications. In *Proc. IEEE GLOBECOM'88*, pages 21.2.1-21.2.5, November 1988.
- [Wilkes79] M.V. Wilkes and D.J. Wheeler. The Cambridge Digital Communication Ring. In *Proceedings of Local-Area Communications Network Symposium*, May 1979.
- [Williams84] G.F. Williams and A.L. Garcia. Performance Analysis of Integrated Voice and Data Hybrid-Switched Links. *IEEE Trans. Commun.*, COM-32(6):695-706, June 1984.
- [Woodruff88] G. M. Woodruff, R. G. H Rogers, and P.S. Richards. A Congestion Control Framework for High-Speed Integrated Packetized Transport. In *Proc. IEEE GLOBECOM'88*, pages 7.1.1-7.1.5, November 1988.
- [Woodruff90] G. M. Woodruff and R. Kositpaiboon. Multimedia Traffic Management Principles for Guaranteed ATM Network Performance. *IEEE J. Select. Areas in Commun.*, 8(3):437-446, April 1990.
- [Yadin67] M. Yadin and P. Naor. On Queueing Systems with Variable Service Capacitors. *Naval Research Logistics Quarterly*, 14:43-54, 1967.
- [Yechiali71] U. Yechiali and P. Naor. Queueing Problems with Heterogenous Arrivals and Service. *Oper. Res.*, 19:722-734, 1971.

- [Yechiali73] U. Yechiali. A Queuing-Type Birth and Death Process Defined on a Continuous-Time Markov Chain. *Oper. Res.*, 21:604–609, 1973.
- [Zukerman85] M. Zukerman. *Performance of Queuing and Telecommunication Systems under Mode and Time Dependent Traffic Models*. PhD Thesis, UCLA, Los Angeles, Calif., 1985.
- [Zukerman86a] M. Zukerman and I. Rubin. On Multi-Channel Queuing Systems with Fluctuating Parameters. In *Proc. IEEE INFOCOM'86*, pages 600–608, April 1986.
- [Zukerman86b] M. Zukerman and I. Rubin. Queue Size and Delay Analysis for a Communication System Subject to Traffic Activity Mode Changes. *IEEE Trans. Commun.*, COM-34(6):622–628, June 1986.
- [Zukerman88] M. Zukerman and P. Kirton. Applications of Matrix-Geometric Solutions to the Analysis of the Bursty Data Queue in a B-ISDN Switching System. In *Proc. IEEE GLOBECOM'88*, pages 49.6.1–49.6.5, November 1988.