CONGESTION CONTROL IN ASYNCHRONOUS TRANSFER MODE NETWORKS

A THESIS

Submitted in fulfillment of the
requirements for the award of the degree
of
DOCTOR OF PHILOSOPHY

By

BRIJESH KUMAR GUPTA

DEPARTMENT OF PAPER TECHNOLOGY
(INDIAN INSTITUTE OF TECHNOLOGY ROORKEE)
SAHARANPUR CAMPUS,
SAHARANPUR-247 001 (INDIA)

JULY, 2002

INDIAN INSTITUTE OF TECHNOLOGY, ROORKEE
(Formerly University of Roorkee)

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled, **"CONGESTION CONTROL IN ASYNCHRONOUS TRANSFER MODE NETWORKS"**, in fulfilment of the requirement for the award of the degree of Doctor of Philosophy, and submitted in the Department of Paper Technology of the institute is an authentic record of my own work carried out during a period from January, 1999 to July, 2002 under the supervision of **Dr. S.C. Sharma,** and **Dr. Mohan Lal.**

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other Institute/University.

Dated: 22nd July, 2002

(Brijesh Kumar Gupta)
Signature of the Candidate

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

(S.C. Sharma)
Asstt. Prof., Instt. Sec.
Deptt. of Paper Technology
(I.I.T. Roorkee)
Saharanpur-247001
India

Dated: 25·7·2002

(Mohan Lal)
Asstt. Prof.,
Computer Center
I.I.T. Roorkee
Roorkee-247667
India

Dated: 25-07-02

The Ph.D. viva-voce examination of Brijesh Kumar Gupta, Research-Scholar, has been held on 25-1-03 at 12-1 PM.

(Signature of Supervisors)

25/1-2003
(Signature of H.O.D.)

(Signature of External Examiner)

Head of Department
Department of Paper Technology
(I.I.T. Roorkee) Saharanpur Campus
SAHARANPUR-247001

Abstract

Asynchronous Transfer Mode (ATM) is the target switching technique for the future public Broadband Integrated Services Digital Networks (B-ISDN). The switch fabric provides the essential routing and buffering functions, but ATM switching systems also require management and control functions necessary for the efficient operation of the network. Congestion is defined as the state of network elements in which network is not able to meet the negotiated performance objective for the established connection.

ATM connections can be multiplexed deterministically or statistically. Deterministic multiplexing is done by allocating resources in accordance to peak rate. Here congestion is eliminated and delay is also very small. While this approach is safe but average utilization of bandwidth is poor. In statistically multiplexing, the peak rate of all connections is above the link capacity. In this case if the large number of connections are very bursty, then all of them may be assigned to the same link in the hope that statistically they will not burst at the same time. If some of them do burst simultaneously, there is sufficient elasticity so that the burst can be buffered up at the input. This partly offsets the probability that at one time, offered input at the switch may become higher than the link capacity. However, a longer duration of congestion might result in cell loss and higher delay. ATM networks have an inherent risk of network congestion if they are based on statistical multiplexing. To take care of this problem various congestion control schemes are reported in the literature.

The success of ATM networks depends on the development of effective congestion control schemes. These schemes are responsible for maintaining an acceptable QoS level that is delivered by the network. In managing the wide range of traffic types and performance requirements, the key challenge is the equitable and efficient allocation of network resources. Since real-time VBR is inherently bursty, due to unpredictable fluctuations congestion can occur frequently. Therefore an

important issue in the congestion control of ATM networks is how to handle the conditions of a large number of cells being in transit between two ATM switching nodes.

In this thesis work efforts have been made to develop congestion control schemes in ATM networks. Various schemes, which are outcome of the work have been proposed / studied under the followings:

i. Congestion control for VBR service in non-blocking ATM switches

ii. Congestion control in ATM network using fuzzy approach

iii Congestion control for multicast bursty traffic in shared memory ATM switch

iv Congestion control in wireless ATM network

The performance analysis of schemes includes various Quality of Service (QoS) parameters such as throughput, average cell delay, and cell loss probability.

i. Congestion control for VBR service in non-blocking ATM switches:

In non-blocking ATM switches, when cells are served according to First-In-First-Out (FIFO) strategy, then due to Head-Of-Line (HOL) blocking the performance of the switch is degraded. HOL blocking limits the throughput of each input port to a maximum of 58.6% under uniform random traffic and much lower than that for bursty traffic.

Traffic streams in the real world are often characterized as bursty. Most of the Application level Data Units (ADUs), such as video frame, are too large to be encapsulated into a single 53-byte ATM cell and must be segmented into a sequence of cells in order to be transmitted over ATM networks. As a result, consecutive arriving cells in a burst are strongly correlated by having the same destination which addresses the same output of the switch. Keeping this point in mind we have simulated and analyzed performance of non-blocking multiple input ATM switches for VBR bursty traffic using Parallel Iterative Matching (PIM) technique as described in [5]. In this technique, each input port maintains a separate queue for each output port, during a single time slot a maximum of one cell per input port can be transferred, and maximum of one cell per out put port can be received. The switch operation is based on the PIM algorithm to find the maximal matching. Maximal matching is used to determine which inputs transmit cells over the switch to which outputs in the current time slot. It has been observed by simulation that the technique increases throughput and reduces not only mean cell delay but cell loss probability also. These results are very much suitable for providing better QoS for real time VBR bursty traffic applications.

ii. Congestion control in ATM network using fuzzy approach:

Congestion is a result of a mismatch between the network resources (buffer space, processing and transmission capacity) and the amount of traffic admitted for transmission. Consequently, congestion prevention can be interpreted as the problem of matching the admitted traffic to the network resources. This, in turn, could be viewed as a classical problem of feedback control i.e. matching the output to the input of dynamical systems. Fuzzy logic system have been successfully applied to deal with congestion control related problems in ATM networks and have provided a robust mathematical frame work for dealing with real world imprecision. A fuzzy logic based scheme has been proposed for dynamic feed-back threshold. In this scheme, out-put buffer is divided into various equal number of parts(N) viz. two, three, and four for the purpose, then the feedback had applied after 50%, 33%, and 25% of the buffer space i.e. when N=2, 3, and 4 respectively. Depending upon which threshold has been crossed, the network gets a mild warning, a stern warning or an ultimatum. A gradual change is more intuitive here, this has been incorporated with fuzzy logic. In applied fuzzy scheme, burst length as well as buffer occupancy are represented by triangular functions. After fuzzification and defuzzification process, the percentage blocking to offered at that particular buffer occupancy level and at given burst length has been determined. The performance of the proposed scheme has been compared with that of the constant threshold and dynamic feedback threshold schemes. It has been shown that the fuzzy scheme improves the performance with reference to major parameters of QoS namely - throughput, average delay and cell loss probability.

iii. Congestion control for multicast bursty traffic in shared memory ATM switch:

Shared-memory ATM switches play a leading role in practical, experimental implementation of ATM. Multicast capability is going to play a key role in the design of ATM broadband switch. The management of multicast traffic in shared memory ATM switches is of particular interest.

Congestion control is an important phenomena, because ATM is connection-oriented and support real-time service. The concept of threshold helps in maintaining a fine balance between the number of cells propagated and the average delay observed by a typical cell in a switch. We have studied and simulated the effect of various threshold schemes to control congestion for multicast bursty traffic in shared memory ATM switch, and proposed a fuzzy scheme for the purpose. In proposed Fuzzy Threshold Scheme, wherein line occupancy as well as buffer occupancy are represented by fuzzy sets, and represented by triangular functions. After fuzzification and defuzzification process, the percentage allocation space to offered at that particular line occupancy and at given buffer occupancy can be determined. The simulation results obtained for evaluation the performance of various congestion control schemes under unicast (90%) and multicast (10%) mixed bursty traffic load. The results have showed that there is a trade-offs exits between cell delay and cell loss rate parameters. The proposed scheme gives minimum cell delay (unicast and multicast) at the cost of higher cell loss rate.

iv. Congestion control in wireless ATM network:

In a Personal Communication Networks (PCN), the covered geographical area is typically partitioned into a set of microcells. Each microcell has a Base Station (BS) to exchange radio signals with wireless mobile terminals. Due to the limited range of wireless transceivers, mobile users can communicate only with BSs that reside within the same microcell at any instance. The number of handoffs/handovers during a call will increase as the cell radii decrease, thus affecting the QoS. Frequent handoff in wireless/mobile networks introduces a new paradigm in the area of network congestion and admission control. The increase in processing load due to demand for service and fast handoffs to mitigate the propagation effect, a high speed backbone network for the

PCN to connect BS is required. The ATM technology, which has recently emerged to be a predominant switching technology, is suited to be an infrastructure to interconnect the BSs of the PCN.

To support network-wide handoff, new and handoff call requests will compete for connection resources in both the mobile and backbone networks. Handoff calls require a higher congestion related performance, i.e., blocking probability, relative to new calls because forced terminations of ongoing calls due to hand-off call blocking are generally more objectionable then new call blocking from the subscriber's perspective.

A hybrid scheme has been proposed for handover in ATM-based PCN, which combines queuing and reservation schemes. In the scheme, FIFO and Measurement Based Priority Scheme (MBPS) queuing discipline [158] and Reserved Channel Scheme (RCS) [79] is used. This scheme gives handovers higher priority than queuing or reservation schemes. When reservation is applied on both radio and backbone channels, it leads to significant improvement in QoS. After applying proposed scheme there is a remarkable reduction in Forced Termination Probability (FTP) at the cost of tolerable Call Blocking Probability (CBP).

Acknowledgements

I feel expedient to express my profound indebtedness deep sense of gratitude and sincere thanks to Dr. S.C. Sharma, *Asstt. Prof.*, Instt. Sec., Department of Paper Technology (Indian Institute of Technology-Roorkee), Saharanpur and Dr. Mohan Lal, *Asstt. Prof.*, Computer Center, Indian Institute of Technology, Roorkee, for their valuable guidance, inspiration, encouragement and wholehearted co-operation in understanding the relevant field and help presenting this work. Their keen interest and efforts in planning the work in this form can not be expressed in words, as they devote their valuable time in discussions and in critical analysis of the work. I am highly obliged to them.

I am extremely grateful to Dr. J.S. Upadhyaya, Professor and Head, D.P.T.- Saharanpur, and Dr. R.C. Joshi, Professor and Co-ordinator, Computer Center, I.I.T.-Roorkee for providing necessary administrative help for carrying the work. The author is also thankful to Dr. M.C.Bansal, Professor, and former Head, I.P.T.- Saharanpur whose co-operation and encouragement helped a great deal in this endeavour.

I would like to express my sincere thanks and regards to Mrs. Shalini and her family, for providing me a homely environment during my stay at Roorkee. In a similar way my special thanks to Dr. Chhaya Sharma, Lecturer, D.P.T.-Saharanpur and her family for giving me morale support and inspirations time-to-time.

My thanks to Mr. S.P. Arora, Paper Lab., D.P.T.-Saharanpur, and Mr. Rakesh Arora, Chemistry Lab., D.P.T.-Saharanpur for sharing light spiritual moments during our leisure which give me more powers and enthusiasm to deal with the work in a better way.

My special thanks to Dr. P.K. Jain, Reader and Head, Deptt. of English, Govt. P.G. College, Deoband - Saharanpur for grammatically corrections in the thesis, and Mr. Hans Raj Sharma, I.I.T.-Roorkee for different types of co-operation and help.

It is my pleasant duty to acknowledge my thanks to all my well wishers, employees of D.P.T., Saharanpur, and Computer-Center I.I.T.-Roorkee (since it is not possible to mention all the names), who support, and inspired me through their love and regards directly or indirectly.

The financial support from University Grants Commissions(UGC), New-Delhi is gratefully acknowledged. The financial assistance in the form of Project-Fellowship provided by UGC proved indispensable to me.

No words, No language is ever adequate to express my heart felt veneration for my respected parents Sri O.N. Gupta, Smt. Kamala Devi, and my family members for their cheerful, enthusiastic and dedicated efforts to engage myself in higher pursuits. Above all, I praise the almighty 'God' for his blessings.

BKGupta

Dated: 22nd July, 2002. (Brijesh Kr. Gupta)

TABLE OF CONTENTS

List of Publications

In Journals:

1. Gupta, B.K., Sharma, S.C., and Lal, M. : "Evaluation of cell loss probability of shared-memory ATM switch," Journal of the CSI, vol. 30, no. 4, pp. 15-20, Dec. 2000, Published by Computer Society of India, Mumbai, INDIA.

2. Gupta, B.K., Lal, M., and Sharma, S.C. : "Fuzzy logic control for VBR traffic in ATM networks," Journal of the CSI, vol. 31, no. 2, pp. 36-40, June 2001, Published by Computer Society of India, Mumbai, INDIA.

3. Gupta, B.K., Sharma, S.C., and Lal, M. : "Analysis and simulation of non-blocking multiple input ATM switches based on input queuing," Accepted for IETE –Technical Review Journal, vol. 19, no. 6, Nov./ Dec. 2002, published by The Institute of Electronics and Telecommunications Engineers, New-Delhi, INDIA.

In International Conferences:

4. Sharma, S.C., Lal, M., Kumar, V., and Gupta, B.K. : "Fuzzy system based queue management scheme for multicast bursty traffic in shared-memory of ATM switches," Proc. 5th Int'l Conf. on Cognitive System, pp. 331-339, 15th-18th Dec. 1999, New-Delhi, INDIA.

5. Lal, M., Sharma, S.C., Gupta, B.K., and Kumar, V. : "Dynamic queue length threshold for multicast support in shared-memory-based ATM switch architecture," 7th Int'l Conf. on Advanced Computing and Communications (ADCOM-1999), Sponsored by IEEE Computer Society USA, 20th-22nd Dec. 1999, Roorkee, INDIA.

6. Lal, M., Sharma, S.C., and Gupta, B.K. : "Congestion control with dynamic feedback threshold in ATM networks - Fuzzy approach," 9th Int'l Conf. on Advance Computing & Communications(ADCOM-2001), Sponsored by IEEE Computer Society USA, pp. 277, 16-19th Dec.2001, Bhubaneswar, INDIA.

In National Conferences:

7. Sharma, S.C., and Gupta, B.K. : "Developments of congestion control in ATM networks: A comparative study," Proc. National Seminar on Role of Information Technology in 21st

Century Challenges and Strategies, pp. 16-25, 11th–12th Feb. 1999, Jabalpur (M.P.), INDIA.

8. Kumar, V., Gupta, B.K., Sharma, S.C., and Lal, M.: "Fuzzy traffic controller in Asynchronous Transfer Mode (ATM) networks," Proc. IE (I) Conf. on Info. –Tech., pp. II-1 to II- 8, 14th Nov. 1999, Palghat (Kerala), INDIA.

9. Sharma, S.C., Lal, M., Gupta, B.K., and Tyagi, N. : "Queue length management of multicast traffic for shared memory Asynchronous Transfer Mode(ATM) switch," Proc. National Conf. on Towards Excellence in Technical Education in 21st Century, pp. 295-299, 23rd - 24th Nov. 1999, Pusad (Maharastra), INDIA.

10. Sharma, S.C., Lal, M., and Gupta, B.K. : "Performance analysis of non-blocking multiple input ATM switches for VBR bursty traffic," National Conf. on Technology Convergence for Information, Communication and Entertainment, (NICE-2001), pp. 160-164, 23rd-24th Feb. 2001, Cochin(Kerala), INDIA.

List of Figures

List of Tables

List of Acronyms

AAL	ATM Adaptation Layer
ABR	Available Bit Rate
ATM	Asynchronous Transfer Mode
B-ISDN	Broadband Integrated Service Digital Network
BS	Base Station
CAC	Connection Admission Control
CBP	Connection / Call Blocking Probability
CBR	Constant Bit Rate
CCITT	Consultative Committee on International Telecommunication & Telegraph
CDV	Cell Delay Variation
CDVT	Cell Delay Variation Tolerance
CLP	Cell Loss Priority
COS	Cross Over Switch
CPCS	Common Part Convergence Sublayer
CPE	Customer's Premises Equipment
CRC	Cyclic Redundancy Code
CS	Convergence Sublayer
EPD	Early Packet Discard
FIFO	First-in First-out
FTP	Force Termination Probability
GCRA	Generic Cell Rate Algorithm
GFC	Generic Flow Control
HDTV	High-Definition Television
HEC	Header Error Control
HOL	Head of Line
IBP	Interrupted Bernoulli Process
ISDN	Integrated Service Digital Network
ITU	International Telecommunication Union
NNI	Network Parameter Control
MAS	Mobile ATM Switch
MT	Mobile Terminal
OAM	Operation And Maintenance
PCI	Protocol Control Information

PCR	Peak Cell Rate
PCN	Personal Communication Network
PDU	Protocol Data Unit
PIM	Parallel Iterative Matching
PMD	Physical Medium Dependent
PPD	Partial Packet Discard
PRM	Protocol Reference Model
PSTN	Public Switched Telephone Network
PTI	Payload Type Identifier
PVC	Permanent Virtual Connection
QoS	Quality of Service
SAP	Service Access Point
SAR	Segmentation And Reassembly
SCR	Sustainable Cell Rate
SCD	Selective Cell Discard
SDU	Service Data Unit
SONET	Synchronous Optical Network
SSCS	Service-Specific Convergence Sublayer
STM	Synchronous Transfer Mode
SVC	Signalling Virtual Channels
TC	Transmission Convergence
TDM	Time Division Multiplexing
UBR	Unspecified Bit Rate
UNI	User-Network Interface
UPC	Usage Parameter Control
VBR	Variable Bit Rate
VC	Virtual Channel / Circuit
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier
VP	Virtual Path
VPC	Virtual Path Connection
VPCI	Virtual Path Connection Identifier
VPI	Virtual Path Identifier
WATM	Wireless ATM

Chapter-1

Introduction

1.1 Introduction

1.1.1 Broadband - ISDN
1.1.2 Synchronous Transfer Mode
1.1.3 Asynchronous Transfer Mode

1.1.3.1 Advantages of ATM

1.1.4 Wireless ATM

1.2 Problem of Congestion

1.3 Why need Congestion Control?

1.4 Objective of the Research Work

1.5 Organization of the Thesis

1.1 Introduction:

Communication of information is a fundamental need in human society. In order to satisfy this need, a variety of telecommunications systems have been developed to support different types of traffic requirements. Before 1962, the transmission facilities for the telephone network were all analog [46]. Voice originates as an analog signal and, after multiplexing, is transmitted as a set of 64-bit frames. Data are transmitted as bursts of bits whose format depends on transmission protocols. Voice and data transmissions were traditionally separate business having separate networks. However, voice is transmitted as a digital signal and voice and data use the same physical networks, such as Public Switched Telephone Network (PSTN) [120]. The many advantages of digital systems - small cost and size, large transmission and switching capacities, high signal quality, flexibility, and ease of maintenance - have promoted the conversion of transmission and switching systems from analog to digital in many telephone administration all over the world.

The adoption of digital techniques in the public switched network makes it possible and cost-effective to integrate voice, data, and video services into a single Integrated Service Digital Network (ISDN). The advantages of an ISDN stem from the fact that the same access, transmission, and switching facilities can be shared by services that are provided by separate networks. Since telecommunications networks [80] are an integral part of the infrastructure of society, an objective of the ISDN is to synergistically responsive to the evolving needs of the information society, and in addition, to spur the demand for information services.

Even though ISDN is supporting many services, there are still some advance services like video telephony, video conferencing, high-speed digital information, telefax, video/document retrieval services, and television distribution which can not be provided by ISDN. For example, if somebody is trying to transfer a high resolution graphics image of 10^9 bits, one must wait for over 4 hours using a 64Kbps access line and 11 minutes on a 1.5 Mbps circuit. To come out all these limitations an advanced digital transmission system named Broadband Integrated Service Digital Network (B-ISDN) is introduced. This B-ISDN is an effort to provide, among other things, data rates that are high enough to comfortably handle image data in the future.

1.1.1 Broadband - ISDN:

Consultative Committee on International Telecommunication and Telegraph(CCITT) defines B-ISDN as "a service or system requiring transmission channels capable of supporting rates greater than the primary rate"[60]. The drive for B-ISDN standardization follows essentially the same philosophy as that of the ISDN to provide a public communication network that can support a range of integrated services over a unique customer interface, with customer control and management, rather than through separate, service specific networks with the multiple interfaces, control, and management. New service opportunity are a major driving force for future broadband ISDN. The different services that a B-ISDN should support are [114]:

- ***Interactive and Distributive Services:***

  Interactive services are those which involve a two way exchange of information between two subscribers or a subscriber and a service provider. Distributive services primarily involve a one way information transfer. These services will include voice, video and data.

- ***Broadband and Narrowband Rates:***

  These include high bit rate capabilities for applications like video conferencing etc. as well as low bit rate (64kbps) support for all transmission services offered by ISDN.

- ***Variety of Traffic Support:***

  Data communications like file or image transfer generate highly bursty traffic while real time services like voice calls generate fairly uniform traffic. Constraints on both also vary in the form of loss sensitivity for the data traffic and delay sensitivity of voice traffic.

- ***Connection-oriented and Connectionless Services:***

  Continuous-Bit-Rate (CBR) applications, are best served by connection-oriented services. These services are characterized by separation of the procedures for connection establishment procedures, which must precede information transfer, determine a route and set up a path between users. Other services, including mail and other data-oriented communications, may not warrant separate connection establishment and information transfer phases, with route-related and user information carried in the same message.

The multiplexing and switching technology to be applied within the B-ISDN is an important issue. The switching facility will have to be capable of handling a wide range of different bit rates and traffic parameters (e.g., burstiness). Several techniques have been proposed for the switching and multiplexing schemes (transfer modes) [35], [73], [99], [150], [162], [165], [168], and [180].

1.1.2 Synchronous Transfer Mode:

Synchronous Transfer Mode(STM), a circuit switching based technique, was initially considered an appropriate transfer mode for B-ISDN because of its compatibility with existing systems. In STM, bandwidth [9] is organized in a periodic frame, which consists of time slots. A framing slot indicates the start of each frame. As in traditional circuit switching, each slot in an STM frame is assigned to a particular call, and the call is identified by the position of the slot. In STM, slots are assigned based on the peak transfer rate of the call so that the required service quality can be guaranteed even at the peak load. Because of its circuit-like nature, STM is suitable for fixed-rate services; however, STM cannot support traffic efficiently since, in STM, bandwidth is wasted during the period in which information is transported below peak rate. Thus, the need for a universal transmission system arises, and this is why the Asynchronous Transfer Mode (ATM) system was devised.

1.1.3 Asynchronous Transfer Mode:

ATM, which is essentially a connection-oriented transmission system, conveys packets, or cells, through the network. "Connection-oriented," means that a path for the cell must be established before the actual transmission starts. The term asynchronous mode means that the time when a given cell transmission may start is not predetermined. The packet in an ATM is a fixed length cell containing 53 bytes. This unit of transmitted information carries a 48 bytes payload and a 5-bytes header, or overhead [157]. One of the original motivations for ATM to employ cells and implement virtual circuits was so that significant statistical multiplexing gain could be exploited by capitalizing on the inherent burstiness of many applications[130].

ATM eliminates the inflexibility and inefficiency found in STM. ATM's fundamental difference from STM is that slot assignments are not fixed; instead, the time slots are assigned in an asynchronous (demand-based) manner. In ATM, therefore, no bandwidth is consumed unless information is actually being transmitted. Because slots are allocated to service on demand, ATM can easily accommodate Variable Bit Rate(VBR) services. ATM can also gain bandwidth efficiency by statistically multiplexing bursty sources. Since bursty traffic does not require continuous allocation of the bandwidth at its peak rate, a large number of bursty traffic sources can share the bandwidth. ATM can also support circuit-oriented and CBR services by allocating bandwidth based on peak rate. Because of these advantages, ATM is considered more suitable for B-ISDN. The reasons for selection of ATM over STM for the transfer mode for B-ISDN can be summarized as follows:

- STM does not provide a flexible interface for meeting a variety of needs.
- Many data applications are bursty in nature and can more efficiently be handled with some sort of packet switching approach.
- Another inflexibility of STM is that it does not lend itself to rate adaptation.
- The use of multiple high data rates complicates the switching system. It requires switches that can handle data streams of multiple high data rates. This is in contrast to narrowband ISDN, which has just the 64kbps data stream to switch.

ATM has been intended for transmission over a Synchronous Optical Network (SONET), which is synchronous network, even though the system is called asynchronous transfer mode. Behind this apparent incongruity is the fact that the user's cells arrive at unpredictable times that is, asynchronously – while the data stream is transmitted synchronously. This is done by filling the idle intervals between the data with either timing bits or timing cells. Timing cells, of course, carry the headers that allows the network to distinguish them from user's cell, a feature that adds another degree of sophistication to an ATM network.

1.1.3.1 Advantages of ATM:

There are a number of advantages that makes ATM one of the most important transmission systems in use[120]:

- It can carry any type of information.(Indeed, a 48-byte payload can carry voice, data, and even video traffic)
- It supports both circuit switching and packet switching. In packet switching, each ATM switch is activated only when the cell is being transmitted. This enables effective use of the available network bandwidth but strictly regulates the speed of the switches. In circuit switching, the ATM transmits cell through an established line, just as any other circuit - switching does.
- It support transmission on any scale from LANs to MANs and to WANs without any additional transformation of the signal. This saves time, holds down the equipment cost, and makes Operation And Maintenance (OAM) much easier to carry out.
- It provides Quality-of-Service (QoS) guarantees, such as throughput, cell loss rate, bandwidth, delay, and error control. This is certainly one of ATM's key advantages.

ATM transmission speed varies widely, which is another advantage of the system because it allows the ATM to work with any application and leaves room for improvement as technology advances. Four basic techniques used in ATM system allow the network to achieve high-speed transmission:

- ATM switches operates with the fixed length of transmitting cells. This relegates the switching function more to the system's hardware than to its software. Since hardware-switching is faster than software-switching, ATM networks can transmit at very high speeds.
- Since a route is established at the beginning of each conversation, the time for setting up the route is minimized. This is done by established a Virtual Channel(VC) and Virtual Path(VP) using Virtual Channel Identifier (VCI) and the Virtual Path Identifier(VPI), which are the parts

of a cell header. Reading these identifiers, each ATM switch simply directs the cell to the appropriate predetermined port rather than spending time to find out to which transmission port the cell should next be directed.

- ATM cell travel along only one path and arrive in the same sequence in which they were transmitted. This save time for routing and for reassembling data from the arriving cell stream.
- ATM perform only some very basic error checking, which also saves transmission time. Through error monitoring is unnecessary because ATM transmission is done through presumably reliable networks, such as SONET, that have their own error-control capabilities.

1.1.4 Wireless ATM:

Wireless ATM (WATM) is considered as the wireless access network to interconnect the mobile users to the ATM network. "WATM" originates by applying the ATM-type transport mechanism through the radio link for B-ISDN services[27]. Advantages of the ATM-cell relay paradigm are the flexibility of the resource management of the bursty mixed traffic, and the QoS control and guarantee of each traffic. The mixed traffic is output of the B-ISDN services, which are characterized by integrating the highly diverse rate-bearer services, such as low-rate voice and data services, and high-rate data, image, and video services. Also, the B-ISDN services demand variable measure of the performance parameters, such as delay tolerance, delay jitter, and bit-error rate. Hence, WATM exploits the ATM-cell transport mechanism on the radio design for the B-ISDN services through the radio link.

1.2 Problem of Congestion:

A major problem encountered in the design of ATM based networks, is to handle bursty sources for which it is difficult to predict and control the rate at which cells are generated. Congestion is defined as the state of network elements in which network is not able to meet the negotiated performance objective for the established connection. ATM networks have an inherent risk of network congestion if they are based on statistical multiplexing. ATM provides flexibility to integrate the wide varieties of services. ATM connections can be multiplexed deterministically or statistically. Deterministic multiplexing is done by allocating resources in accordance to peak rate, here congestion is eliminated. While this approach is safe but average utilization of bandwidth is poor. In statistically multiplexing, the peak rate of all connections is above the link capacity [48]. In this case if the large number of connections are very bursty, then all of them may be assigned to the same link in the hope that statistically they will not burst at the same time. If some of them do burst simultaneously, there is sufficient elasticity so that the burst can be buffered up at the input. This partly offsets the probability that at one time, offered input at the switch may become higher than the link capacity. However, a longer duration of congestion might result in cell loss and higher delay. So congestion control in ATM is difficult[1] because of the high link speed, diverse service requirements, and diverse characteristics of the traffic ATM is expected to ~~support [61]. The high speed and multipoint~~ [illegible] broadcast packets networks place new demands on congestion control methods. To take care of this problem various congestion control schemes are reported in the literature [6], [18], [23], [33], [36], [40], [41], [44], [47], [49], [57], [67], [71], [74], [86], [92], [94], [98], [116], [129], [144], [163], [164], [166], [172], and [175].

1.3 Why Need Congestion Control?:

Congestion control is concerned with allocating the resources in a network such that the network can operate at an acceptable performance level, when the demand exceeds or is near the capacity of network resources. The resources include bandwidth of links, buffer space and processing capacity at internal nodes. Although resources allocation is necessary even at low load, the problem becomes more important as load increases because the issue of fairness and low overheads becomes increasingly important. With out proper congestion control mechanism the throughput of network may be reduced.

The assumption that statistical multiplexing can be used to improve the link utilization is that the

---

[1] *More details are given in Chapter-2*

users do not take their peak rate values simultaneously. But since the traffic demands are stochastic and can not be predicted, congestion is unavoidable. Whenever the total input rate is greater than the output link capacity, congestion happens. Under a congestion situation, the queue length may become very large in a short time, resulting in buffer overflow and cell loss. So congestion control is necessary to ensure that users get the negotiated QoS.

1.4 Objective of the Research-Work:

The success of ATM networks depends on the development of effective congestion control schemes. These schemes are responsible for maintaining an acceptable QoS level that is delivered by the network. In managing the wide range of traffic types and performance requirements, the key challenge is the equitable and efficient allocation of network resources. Since real-time VBR is inherently bursty, due to unpredictable fluctuations congestion can occur frequently. Therefore an important issue in the congestion control of ATM networks is how to handle the conditions of a large number of cells being in transit between two ATM switching nodes. The specific objectives of the present research-work are listed below:

i. Congestion control for VBR service in non-blocking ATM switches
ii. Congestion control in ATM network using fuzzy approach
iii. Congestion control for multicast bursty traffic in shared memory ATM switch
iv. Congestion control in wireless ATM network

1.5 Organization of the Thesis:

This thesis contains seven chapters. The organization of the thesis is as follows:

Chapter-1:

This chapter introduces in brief Broadband Integrated Service Digital Network (B-ISDN), Synchronous Transfer Mode (STM), Asynchronous Transfer Mode (ATM), and Wireless ATM (WATM). The importance of the ATM networks in future Broadband infrastructure is described in the chapter. The phenomena of congestion in ATM networks is a major problem. Chapter also deals with the significance of congestion control in ATM networks.

Chapter-2:

A review of the basic concepts of ATM and congestion control is presented in this chapter. Various causes of congestion like high link speed, multiple service requirement, and diverse traffic characteristics are briefly discussed in the chapter. Congestion control schemes are given in the Section 2.11 of this chapter, which can be classified as either preventive or reactive. In addition to the conventional schemes, fuzzy mechanism to control the congestion is also reviewed in the Section 2.12 of the chapter.

Chapter-3:

This chapter describes congestion control for VBR service in non-blocking ATM switches. The Head-of-Line (HOL) blocking problem has been discussed in the chapter. Due to this blocking problem congestion arises as a result the throughput of each input port gets reduce. In this chapter Parallel Iterative Matching (PIM) technique is used for VBR bursty traffic to control the congestion. Various QoS parameters like average cell delay, throughput, and cell loss probability are evaluated.

Chapter-4:

In this chapter a fuzzy scheme is proposed for congestion control in ATM network. A brief description of fuzzy theory and fuzzy traffic controller is also given in the chapter. Comparative studies have shown that the proposed fuzzy scheme significantly improved network performance compared with conventional schemes.

Chapter-5:

This chapter discusses congestion control for multicast bursty traffic in shared memory ATM switch. Different issues on multicasting are presented in the chapter. Various applied schemes are given in Section 5.7, which includes conventional as well as fuzzy schemes. The performance of these schemes studied under various QoS parameters like unicast cell delay, multicast cell delay, throughput, and cell loss probability in the chapter.

Chapter-6:

A hybrid scheme has been proposed for congestion control in wireless ATM network in this chapter. A review of various handover schemes is also given in the chapter. The aim of the proposed hybrid scheme is to give handover calls higher priority than new calls. Different QoS parameters like forced termination probability, call blocking probability, and total carried traffic are evaluated.

Chapter-7:

This chapter concludes the thesis by summarizing the important results of the present work. The scope for future work in this area is also include.

Chapter-2

Standardization & Congestion Control in ATM Networks: A Review

2.1 Introduction:

The ISDN was introduced in 1984 by the International Telecommunication Union (ITU) as a common access technology to support a range of digital services offered by telecommunication carriers from the customers' premises to the carriers' networks[130]. This technology has as its primary goal the integration of voice and non-voice services. ISDN supports both switched and non-switched connections. It is simultaneously carrying a wide range of voice and non-voice applications. All services are arranged compatible with the 64 kbps switched digital connection[16]. If we envisage a universal network, it must be able to transport a large number of service such as:

- Low speed (e.g., telemetry, telecontrol, telealarm, voice, telefax, low speed data)
- Medium speed (e.g., hi-fi sound, video telephony)
- High-speed (e.g., high speed data transfer for LAN inter connection)
- Very high speed (e.g., high quality video distribution, video library, video education)

Therefore, the transport mode for this universal network cannot be designed specifically for one server. Universal network supports a large range of services, with an estimated bit rate of a few bps to a few hundreds of Mbps as represented in Table 2.1. So the need for a faster, smarter, and more flexible network continues, this is the target of B-ISDN[133].

Service	Estimated Bit Rate	Burstiness
Voice	32 kbps	2
Interactive data	1-100 kbps	10
Bulk data	1-10 Mbps	1-10
Standard quality video	1.5-15 Mbps	2-3
High definition TV	15-150 Mbps	1-2
High quality video telephony	0.2-2 Mbps	5

Table 2.1: Bit Rate and Burstiness for Broadband Services

B-ISDN has received increased attention as a communication architecture capable of supporting multimedia applications. B-ISDN was introduced as the successor [67] to narrowband ISDN after the latter fell short of meeting the high demand for bandwidth required by emerging application such as real-time video and High Definition TV(HDTV). Among the techniques proposed to implement the B-ISDN, Asynchronous Transfer Mode (ATM) is considered to be the most suitable transfer technique because of its efficiency and flexibility. As the ATM is recommended by CCITT as an integrated transport technique for the future B–ISDN. CCITT recommend the definition of ATM is as a transfer mode in which information is organized into fixed length packets called cells and it is asynchronous in the scene that the occurrence of cells, which containing the information is not periodic. One of the primary attractions of the ATM networks is its statistical multiplexing capability. This allows VBR and bursty connections to share the same limited resources (e.g. bandwidth and buffers) stochastically. This, in turn, allows more connections to be supported simultaneously than in a circuit switched or a STM network.

2.2 Structure of ATM Cell and Basic Idea:

ATM stands for *Asynchronous Transfer Mode* of data transport in digital communication network. The ATM concepts results from the merging of two well-known concepts: packet switching and Time Division Multiplexing (TDM). Each of these techniques has been modified in the following manner related to packet switching[133]:

- No error control (on data field) and flow control on the links inside the ATM network.
- Connection oriented at the lowest level. All information is transferred in a Virtual Circuit (VC) assigned for the complete duration of the connection.

ATM transfers information in fixed-size units called cells. Each cell consists of 53 octets, or bytes. The first 5 bytes contain cell-header information, and the remaining 48 bytes contain the "payload" (user information)[134]. Small fixed-length cells are well suited to transferring voice and video traffic because such traffic is intolerant of delays that result from having to wait for large data packet to download, among other things. The structure of an ATM cell is shown in Fig. 2.1.

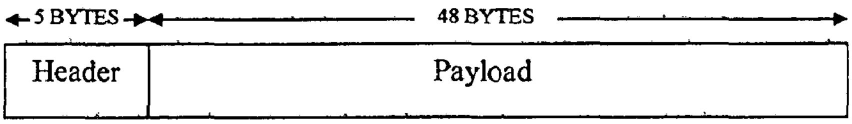


Fig. 2.1: ATM Cell

In the ATM layer, two interfaces are distinguished as: User-Network Interface (UNI) and the Network-Node Interface (NNI). The former defines the boundary between a host and an ATM network (in many cases between the customer and the carrier). The latter applies to the line between two ATM switches (the ATM term for routers).

In both cases the cells consist of a 5-byte header followed by a 48-byte payload, but the two headers are slightly different[60]. The only difference between the header at UNI and NNI is a four bit Generic Flow Control (GFC) field provided at the UNI. Both types of header are shown below in Fig. 2.2(a) and Fig. 2.2(b).

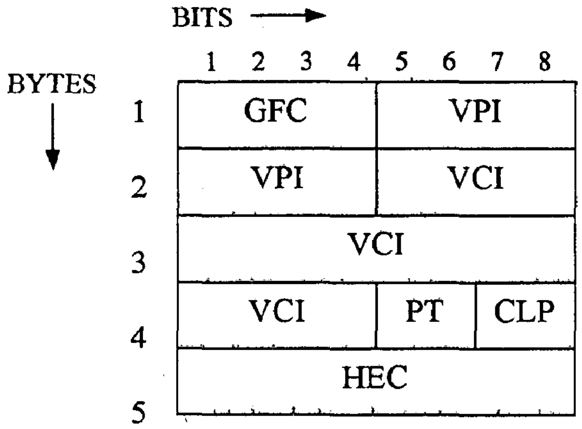


Fig. 2.2(a): ATM Cell Header at UNI

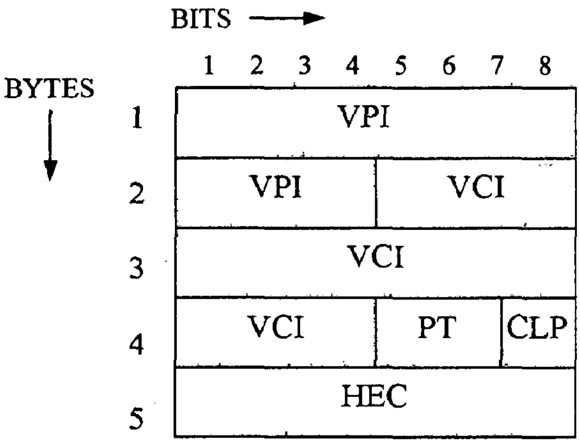


Fig.2.2(b): ATM Cell Header at NNI

The header fields are as follows:

2.2.1 Generic Flow Control:

Multiple terminals may share a single access link to the network. The Generic Flow Control(GFC) field is used to assist the customer premises in controlling the flow of traffic for different QoS.

2.2.2 Virtual Path Identifier:

The Virtual Path Identifier (VPI) field is a small integer selecting a particular Virtual Path(VP). A VP consists of a bundle of Virtual Channels(VC) that are carried end-to-end between VP terminators on the same physical facility.

2.2.3 Virtual Channel Identifier:

The Virtual Channel Identifier (VCI) field selects a particular VC within the chosen VP. The VCI identifies a particular connection, for a given connection the value of the VCI may change as the cell traverses the networks.

2.2.4 Payload Type Identifier:

Three header bits are used for the Payload Type(PT) identification. The following Table 2.2, describes the Payload Type Identifier (PTI) coding:

PTI	***Interpretation***
000	User data cell, no congestion, user-to-user indication=0
001	User data cell, no congestion, user-to-user indication=1
010	User data cell, congestion, user-to-user indication=0
011	User data cell, congestion, user-to user indication=1
100	OAM F5 segment associated cell
101	OAM F5 end-to-end associated cell
110	Resource management cell
111	Reserved for future functions

Table 2.2: Payload Type Identifier Values

2.2.5 Cell Loss Priority:

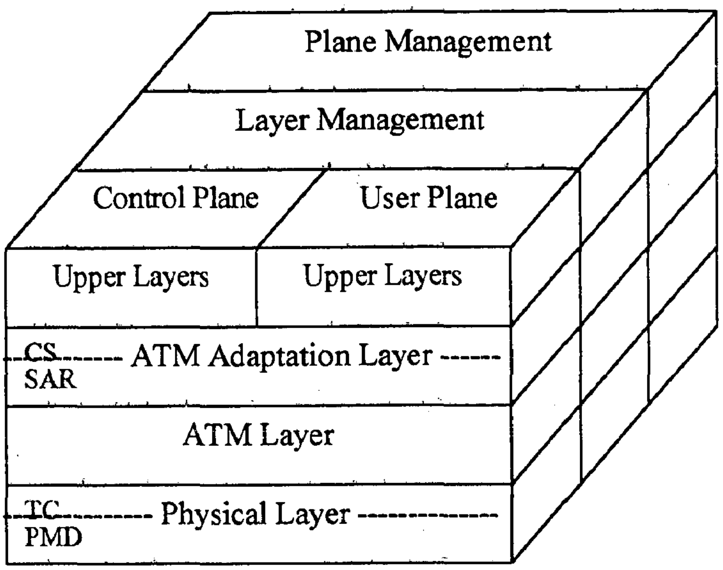
The Cell Loss Priority (CLP) field consists of one bit which is used explicitly to indicate the cell loss priority. It is used to prevent buffer overflow at the switching nodes. A marked cell will be discarded at the time of congestion.

2.2.6 Header Error Control:

Header Error Control (HEC) field is needed to prevent errors in the VCI, VPI, and PT fields from causing cells to be misdelivered or misinterpreted. This will be done by Cyclic Redundancy Code (CRC) of the header.

2.3 ATM Protocol Reference Model:

The ATM Protocol Reference Model is based on standards developed by the ITU. It consists of three layers, the physical, ATM and ATM adaptation layers, plus whatever the users want to put on top of that[157]. The ATM layered network architecture is shown in Fig. 2.3.



CS : Convergence sublayer
SAR : Segmentation & Reassembly sublayer
TC : Transmission Convergence sublayer
PMD : Physical Medium Dependent sublayer

Fig. 2.3: The B-ISDN ATM Reference Model

Various functions of different layers are described as follows:

2.3.1 Physical Layer :-

The Physical Layer is responsible for the electrical or optical transmission and receipt along the physical media between two devices. The physical layer is responsible for supporting different physical media and different media interface rates. As shown in the Fig. 2.3, Physical Layer is divided into 2 sub-layers: the Physical Medium Dependent (PMD) sub-layer and the Transmission Convergence (TC) sub-layer. The PMD sub-layer is responsible for the correct transmission and the reception of bits on the physical medium and is medium dependent (optical, electrical). The TC sub-layer's main function after bit reconstruction is the mapping of the ATM cells to the transmission system used. Header Error Control(HEC), cell delineation, transmission frame adaptation, transmission frame generation /recovery are also functions of this layer.

2.3.2 ATM Layer :

ATM layer is the layer above the physical layer. This layer is fully independent of the physical medium used to transport ATM cells and thus for the physical layer. Congestion control is also located here. It mainly performs following functions:

- ***Cell Multiplexing/Demultiplexing*** **:**

This function multiplexes cells from individual VPs and VCs into one resulting cell stream in the transmit direction, and divides the arriving cell stream into individual cell flows with respect to VC or VP in the receive direction.

- ***Cell Header Generation/Extraction:***

This function adds the appropriate ATM cell header (except for the HEC value) to the received cell information field from the AAL in the transmit direction. VPI/VCI values could be obtained by translation from the Service Access Point (SAP) identifier. It does opposite i.e. removes cell header in the receive direction. Only cell information field is passed to the AAL.

- ***VPI/VCI Translation:***

This function is performed at the switching and/or cross-connect nodes. With a VP node the value of the VPI field of each incoming cell is translated into a new VPI value for the outgoing cell. The values of VPI and VCI are translated into new values at a VC switch.

- ***Generic Flow Control:***

This function supports control of the ATM traffic flow in a customer network. This is only defined at the B-ISDN User-to-Network Interface (UNI).

2.3.3 ATM Adaptation Layer (AAL):

The ATM Adaptation Layer (AAL) enhances the services provided by the ATM layer according to the requirements of specific services. These services can be user services as well as control and management functions. The AAL maps the user, control or management protocol data units into the information field of ATM cell and vice versa. AAL layer is divided into two sub layers, Convergence Sub-layer (CS) and Segmentation And Reassembly (SAR) Sub-layer. There are many and varied requirements for applications using the B-ISDN and several AAL protocols defined, the classification of these protocols was guided by considering the following parameters:

- ***Timing Relationship:***

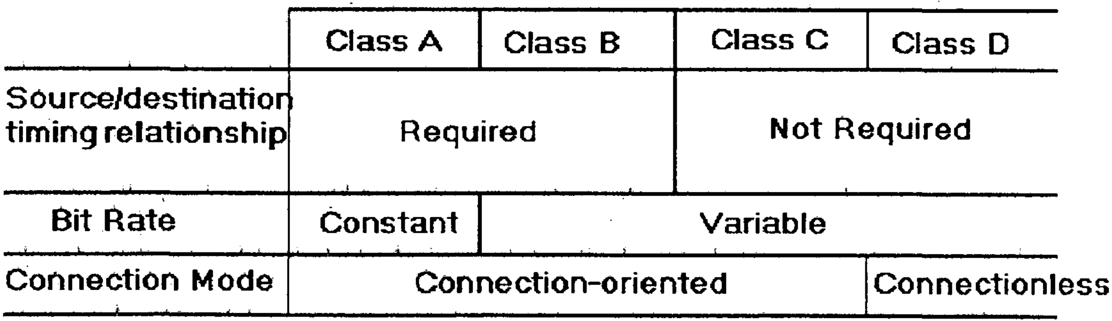
The (non-)requirement for synchronization between the receiver and sender

- ***Bit Rate:***

Constant or Variable

- ***Connection Mode:***

Connection oriented or Connectionless



	Class A	Class B	Class C	Class D
Source/destination timing relationship	Required		Not Required	
Bit Rate	Constant	Variable		
Connection Mode	Connection-oriented			Connectionless

Table 2.3: AAL Service Classification

A summary of the (sensible) combinations of these parameters is given in the above Table 2.3. CCITT has recommended five types of AAL to support four service classes A, B, C, and D. The AAL types are in consonance with the service classes. CCITT has left open the possibility to have an empty AAL for users who may find the ATM services sufficient for their requirements.

2.3.3.1 AAL Type 0:

This is effectively a NULL AAL. It is not really an official AAL type but is mentioned for completeness.

2.3.3.2 AAL Type 1:

This AAL type is normally used by Class A (CBR) services. The function performed by this AAL are:

- Segmentation and reassembly of user information
- Handling of Cell Delay Variation (jitter)
- Handling of lost and misinserted cells
- Source clock frequency recovery at the receiver
- Source data structure recovery at the receiver
- Monitoring and handling of AAL-PCI for bit errors
- Monitoring and (possibly) correcting the bit errors in the user information field
- For circuit emulation, monitoring and maintenance of end-to-end QoS

2.3.3.3 AAL Type 2:

This AAL type would be used with Class B (VBR). This type is not well defined and it seems possible that the it may be merged with AAL Type 1 in the future. Some of its functions are similar to ALL Type 1:

- Segmentation and reassembly of user information
- Handling of CDV(jitter)
- Handling of lost and misinserted cells
- Source clock frequency recovery at the receiver
- Monitoring and handling of AAL-PCI for bit errors
- Monitoring and (possibly) correcting the bit errors in the user information field

It also has the additional functions:

- Handle SDUs from a variable bit rate source

- Transfer timing information between source and destination
- Notify the higher layers of uncorrectable errors in AAL

2.3.3.4 AAL Type 3/4 :

There was once separate Type 3 and Type 4 AALs, but they have now been merged. This AAL is now intended to support both Class C (ABR) and Class D (UBR) services.

In this AAL, the Convergence Sublayer (CS) is split into two parts, the Service-Specific Convergence Sublayer (SSCS) and the Common Part Convergence Sublayer (CPCS) as shown in Fig. 2.4. The SSCS is - application dependent and may be null. The CPCS is responsible for constructing Protocol Data Units (PDUs) that can be sent to the other end user. There are two modes of operation of ALL Type 3/4; Message Mode (MM) and Streaming Mode (SM).

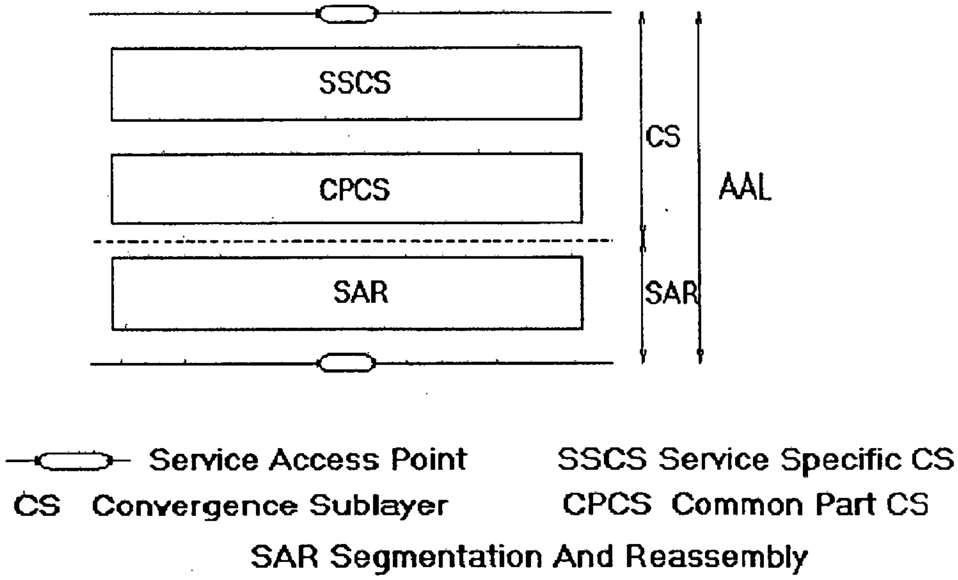


Fig. 2.4: AAL Type 3/4 Sub-layering

The Message Mode is intended for use framed data transfer where the AAL-Service Data unit (SDU) is a logical unit of data with respect to the B-ISDN user. It allows the transport of a single AAL-SDU in one or (optionally) more than one CS-PDU. The CS-PDU may be then further spilt into several SAR- PDUs. The AAL-SDU can be of an arbitrary size. Fig. 2.5 shows the operation of this service.

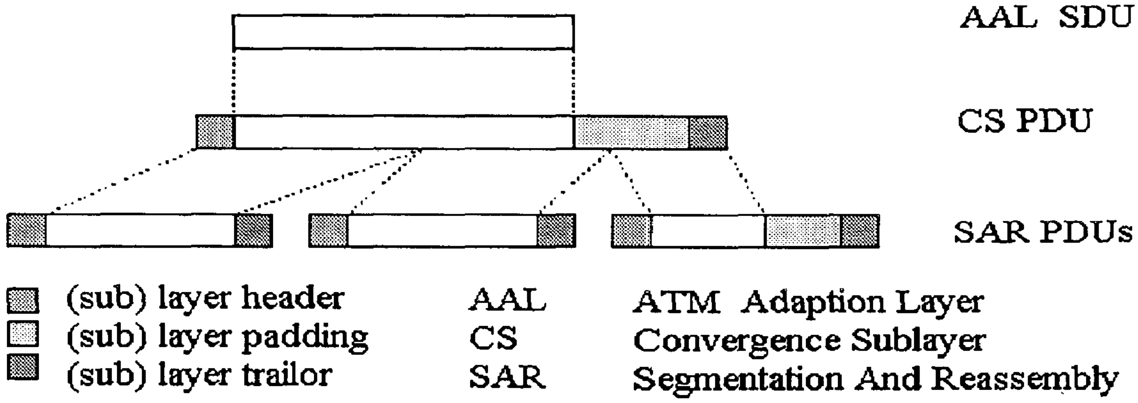


Fig.2.5:AAL Type 3/4 Message Mode Service

In Streaming Mode(SM), the AAL-SDUs are of fixed size and one or more of them may be transported in a single CS-PDU. Each AAL-SDU is delivered in a separate SAR-PDU. Fig. 2.6 illustrates the operation of the SM service.

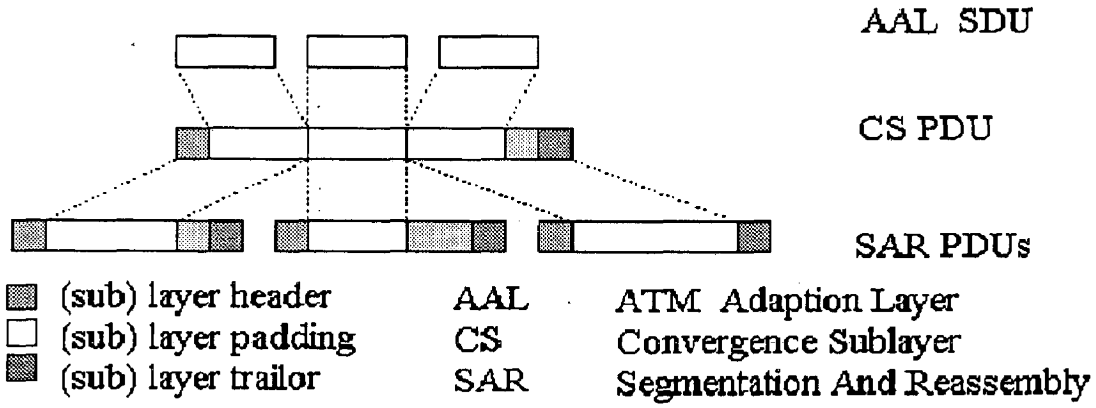


Fig.2.6:AAL Type 3/4 Streaming Mode Service

In both cases, the SAR sublayer provides error detection and both these modes can offer the following operational procedures:

- ***Assured Operation:***

Flow control and retransmission of missing or errored AAL-SDUs takes place under this operation. Flow control restricted to point-to-point connections at the ATM layer and point-to-multipoint flow control possible.

- **Non-assured Operation:**

No retransmission of missing or erroneous SAR-PDUs takes place. Optionally deliver erroneous PDUs to user. Allow flow control for point-to-point connections but not point-to-multipoint. This AAL type also provides multiplexing at the SAR sublayer.

2.3.3.5 AAL Type 5:

This AAL type provides similar services as AAL Type 3/4, but has a reduced overhead when compared to AAL Type 3/4. It is intended for use by VBR sources with timing relationship between source and destination. It has identical modes and operational procedures as AAL Type 3/4. The difference is that this AAL does not provide the AAL Type 3/4 multiplexing capability. The SAR sublayer accepts only AAL-SDUs that are an integer multiple of 48 octets. So, it would be possible for this AAL to offer an efficient cell based interface to the B-ISDN user.

AAL Types 1 and Type 2 are not used much.

The ATM model is defined as being three-dimensional. The User Plane deals with data transport, flow control, error correction and other user functions. The Control Plane is concerned with connection management. The layer and plane management function relate to resource management and inter layer coordination.

2.4 ATM Switching and Queueing Disciplines:

ATM Switching is also known as fast packet switching[162], and [169]. ATM switching node transports cells from the incoming links to outgoing links using the routing information contained in the cell header and information stored at each switching node using connection set-up procedure.

Two functions at each switching node are performed by a connection set up procedure:

- A unique connection identifier at the incoming link and the link identifier and a unique connection identifier at the outgoing link are defined for each connection.

- Routing tables at each switching node are set up to provide an association between the incoming and outgoing links for each connection. VPI and VCI are the two connection identifiers used in ATM cells.

Thus the basic functions of an ATM switch can be stated as follows:

- Routing (space switching) which indicates how the information is internally routed from the inlet to outlet.
- Queueing, which is used in solving contention problems if two or more logical channels, contend for the same output.
- Header translations all cells, which have a header equal to some value j on incoming link, are switched to outlet and their header is translated to a value k.

There are mainly three different queueing disciplines available determined by their physical location as follows.

2.4.1 Input Queueing:

In this queueing, the contention problem is solved at the input buffer of the inlet of the switching element. Each inlet contains a dedicated buffer, which is used to store the incoming cells until the arbitration logic decides to serve the buffer.

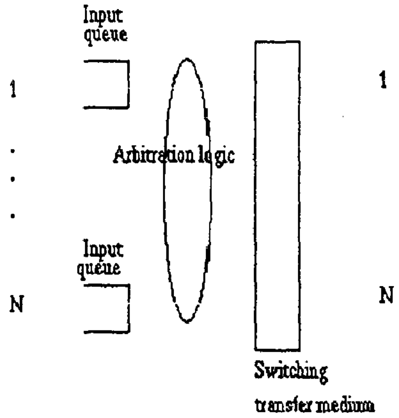


Fig. 2.7: Input Queueing

The switching transfer medium then switches the ATM cells from the input queues to the outlet avoiding an internal contention. The arbitration logic can be as simple as round robin or can be complex such as taking into account the input buffer filling levels. However, this scheme has Head of Line (HOL) blocking[2] problem i.e. if two cells of two different inlets contend for the same output, one of the cells is to be stopped and this cell blocks the other cells in the same inlet which are destined for different outlet. This queueing discipline can be shown by the Fig. 2.7.

2.4.2 Output Queueing:

In this queueing discipline, queues are located at each outlet of the switching element and the output contention problem is solved by these queues. The cells arriving simultaneously at all inlets destined for the same output are queued in the buffer of the outlet. The only restriction is that the system must be able to write N cells in the queues during one cell time to avoid the cell loss where N is the total number of inlets of the switch.

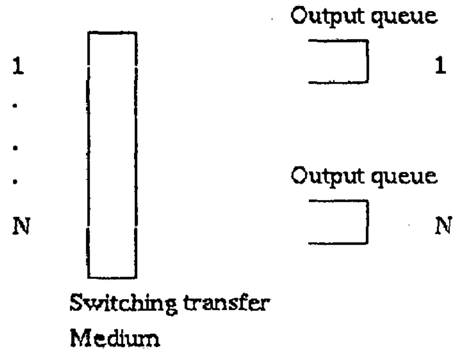


Fig. 2.8: Output Queueing

In this mechanism, no arbitration logic is required as all the cells can be switched to their respective output queue. The cells in the output queue are served on FIFO basis to maintain the integrity of the cell sequence. The Fig. 2.8 illustrates this mechanism.

2.4.3 Central Queueing:

In this scheme, the queueing buffers are shared between all inlets and outlets. All the incoming cells are stored in the central queue and each outlet chooses the cells, which are destined for it from this central memory. Since cells for different outlets

---

[2] *The term is defined in Chapter-3.*

are merged in this central queue, FIFO discipline is not followed in reading and writing of this queue. Cells can be written and read at random memory locations and this needs a complex memory management system for this scheme. The following Fig. 2.9 shows this mechanism.

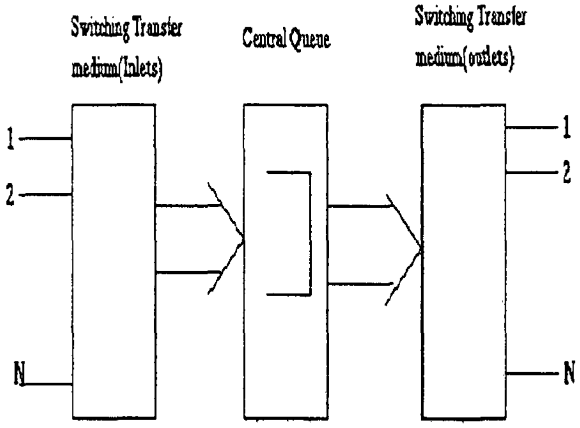


Fig. 2.9: Central Queueing

2.5 Multiplexing:

Multiplexing is the sharing of one physical transmission medium by more than one data stream. In ATM, cells containing different forms of data are multiplexed over the same bandwidth. Each time slot may contain cells from voice, video or data traffic types. The traffic type with the most throughput required will take up most of the bandwidth. For example, a multiplexed multimedia application may have five out of eight cells for video, two for sound and one for data. Multiplexing improves efficiency by maximizing resources. ATM connections can be allocated resources based on either statistical multiplexing or deterministic multiplexing. ATM uses statistical multiplex gain to improve efficiency. This process involves dynamically assigned time slots only to users who need them. Time slots are not reserved for individual users and they are not sent if no data needs to be transmitted as shown in Fig. 2.10(a). The statistical gain is the factor by which the sum of the peak bandwidths exceeds the output channel's capacity [48]. Statistical multiplexing, therefore, relies on the input channels being bursty due to variable information transfer rates. Hence, the statistical gain directly depends on the bandwidth utilization and traffic characteristics of the input channels. With deterministic multiplexing, the sum of the peak bandwidths for the constituent connections is less than the peak bandwidth capability of the channel into which they are multiplexed as represented in Fig. 2.10(b). An ATM network that uses only deterministic multiplexing can be designed with a limited set of traffic management controls. The issue of statistical multiplexing and deterministic multiplexing in ATM network is given in [112].

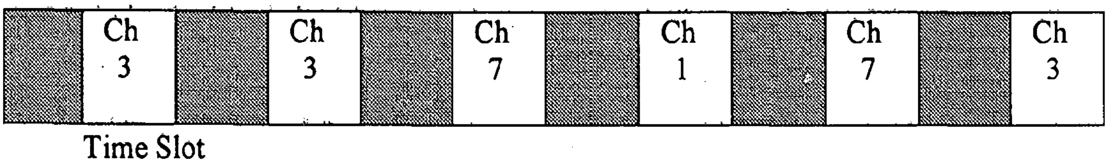


Fig.2.10 (a): Statistical Multiplexing

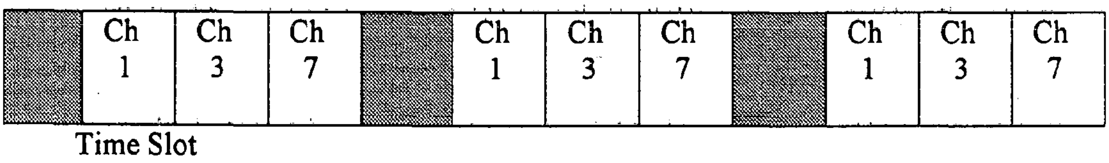


Fig. 2.10 (b): Deterministic Multiplexing

2.6 Signalling :

The Signalling capability for ATM networks has to satisfy the following functions:

- Set up, maintain and release ATM Virtual Channel Connections(VCCs) for information transfer.
- Negotiate the traffic characteristics of a connection (CAC algorithms are considered for these functions.)

Signalling functions may also support multi-connection calls and multi-party calls. Multi-connection call requires the establishment of several connections to set up a composite call comprising various types of traffic like voice, video, image and data. It will also have the capability of not only removing one or more connections from the call but also adding new connections to the existing ones. Thus the network has to correlate the connections of a call. A multi - party call contains of several connections between more than two end-users like conferencing calls.

Signalling messages are conveyed out-of band in dedicated Signalling Virtual Channels (SVCs) in broadband networks. There are different types of SVCs that can be defined at the B-ISDN User-to-Network Interface (UNI). They can be described as follows:

- Meta-Signalling Virtual Channel is used to establish, check and release point-to-point and selective broadcast signalling virtual channels. It is bi-directional and permanent.
- Point-to-point signalling channel is allocated to a signalling end point only while it is active.

These channels are also bi-directional and are used to establish, control and release VCCs to transport user information. In point-to-multipoint signalling access configuration, meta-signalling is needed for managing the SVCs.

2.7 ATM Service Categories:

The ATM Forum has defined the service/traffic categories based on different requirements of quality of service parameters. There are five categories of service, the QoS parameters for those categories are explained below[160]:

2.7.1 Constant Bit Rate (CBR):

This category is used for emulating circuit switching. The cell rate is constant. Cell loss ratio is specified for CLP =0 cells and may or may not be specified for CLP=1 cells. To handle this type of traffic, the ATM network must provide a constant amount of bandwidth and deliver low cell delay variation. Telephone, T.V. and video conferencing are the example of applications of CBR.

2.7.2 Variable Bit Rate (VBR):

This category allows user to send at a variable rate. Statistical multiplexing is used and so there may be a small non-zero random loss. Depending upon whether or not the application is sensitive to cell delay variation, this category is sub-divided in two categories: Real time VBR and Non real time VBR. For non-real time VBR, only mean delay is specified, while for real time VBR, maximum delay and peak to peak Cell Delay Variation (CDV) are specified. Because VBR traffic does not transmit permanently at its peak rate, the bandwidth available can be released to UBR and ABR services. Real time video conferencing and multimedia e-mail are the examples for Rt-VBR and Nrt-VBR respectively.

2.7.3 Available Bit Rate (ABR):

This category is designed for normal data traffic such as file transfer and email. Although, the standard does not require the cell transfer delay and cell loss ratio to be guaranteed or minimized, it is desirable for switches to minimize the delay and loss as much as possible.

S.N	Description	Class	Characteristics				Example
			Bandwidth Guarantee	Suitable for R-T Traffic	Suitable for Bursty Traffic	Feedback about Congestion	
1.	Constant Bit Rate	CBR	Yes	Yes	No	No	T1 Circuit
2-a	Variable Bit Rate: Real Time	VBR: RT	Yes	Yes	Yes	No	Real time Video Conferencing
2-b	Variable Bit Rate: non Real Time	VBR: nRT	Yes	No	Yes	No	Multimedia Email
3.	Available Bit Rate	ABR	Optional	No	Yes	Yes	Browsing the Web
4.	Unspecified Bit Rate	UBR	No	No	Yes	No	Background File Transfer
5.	Guaranteed Frame Rate	GFR	Yes	No	Yes	-	E-Mail, Frame Relay Networking

Table 2.4: ATM Service Categories and Their Characteristics

2.7.4 Unspecified Bit Rate (UBR):

This category is designed for those data applications that want to use any left over capacity and are not sensitive to cell loss or delay. Such connections are not rejected on the basis of bandwidth shortage and not policed for their usage behavior. UBR is sometimes called best efforts service, since the transport of cells in UBR service is not guaranteed by mechanisms operating at cell level. Examples of such type of service are email, file transfer, news feed etc.

2.7.5 Guaranteed Frame Rate (GFR):

This service category is intended to support non Real-Time applications requiring a minimum rate guarantee. It does not require adherence to a rate control protocol. An example application is frame relay interworking.

The ATM service categories and their characteristics are summarized in Table 2.4.

2.8 Basics of Traffic Management:

ATM is a technology that can support a wide variety of applications in several different network environments. Different applications require distinct degrees of transmission quality, in particular as relates to data loss, introduced delay, and delay variations. Applications with timing requirements must be treated differently from application without-for example, they impose strict data-loss requirements[15]. QoS classes are separated with buffering schemes in a network element. These schemes can vary in degree of sophistication, ranging from a few FIFO queues with internal priorities, to queues per connection with advanced scheduling mechanisms. Hence there is a need of proper traffic management to obtain best resource optimization. Similarly, when attaching a wireless ATM network to a fixed ATM network proper interaction of traffic and resource management functions throughout both networks is necessary to achieve stringent QoS objectives[151].

Traffic management is the art of providing users with the service they need and have paid for. With proper traffic management, the network is able, at the right level of quality, to deliver data to the appointed destination. Likewise, good traffic management protects the network from traffic that misbehaves. Traffic management has been widely reported in [24], [38], [50], [78], [86], [110], [113], [125], [135], and [170].

2.8.1 Traffic Contract:

A user and a network must agree on traffic characteristics and on the kind of service the network is to deliver. This agreement, called the traffic contract, consists of following three parts:

2.8.1.1 Traffic Descriptors:

Traffic descriptors are provided as part of a connection request to define the expected characteristics for the connection. These descriptors contain bandwidth parameters, plus other useful information such as the QoS requirements. The values for these traffic descriptor parameters must be derived from collected statistics, or from a network operator's estimate of what is required. For new connection requests the subscriber may not know in advance the traffic behavior. Therefore, the user should only have to specify the type of service to be provide in high level terms, such as voice, videophony, or image database browsing, as well as QoS parameter (e.g., image resolution-high, medium or low). The network then must be responsible for deriving appropriate traffic parameters and ensuring that the application level QoS is achieved.

To accurately perform connection admission and traffic monitoring functions, a deterministic estimate of the connection's bandwidth characteristic is required. This can be accomplished by specifying the maximum cell counts that are expected during given time intervals. For example, the connection's peak cell transfer rate could be specify as the minimum inter-cell arrival time between any two consecutive cells. The connection's burst transfer rate is specified as the minimum time allowed for the arrival of B cells, where B is the maximum expected burst length. Finally, the connection's average bandwidth can be specified by the number of cells expected to arrive during a fairly long time period.

2.8.1.2 Quality of Service:

Quality of Service (QoS) is an important issue for ATM networks, in part because they are used for real time traffic, such as audio and video. When a virtual circuit is established, both the transport layer and the ATM network must agree on a contract defining the service. In the case of a public network, this contract may have legal implication. To make it possible to have concrete traffic contracts, the ATM standard defines a number of QoS parameters whose values the customer and carrier can negotiate. While setting up a connection on ATM networks, users can specify the following parameters related to the input traffic characteristics and the desire QoS.

- **Peak Cell Rate (PCR):** The maximum instantaneous rate at which the user will transmit.
- **Sustained Cell Rate (SCR):** This is the average rate as measured over a long interval.
- **Cell Loss Ratio (CLR):** The percentage of cells that are lost in the network due to error and congestion and are not delivered to the destination or delivered so late as to be useless.
- **Cell Transfer Delay (CTD):** The delay experience by a cell between network entry and exit point is called the cell transfer delay.
- **Cell Delay Variation (CDV):** This is a measure of variance of CTD. High variation implies large buffering for delay sensitive traffic such as voice and video.
- **Maximum Burst Size (MBS):** The maximum number of back to back cells that can be sent to peak cell rate but without violating the sustainable rate is called maximum burst size.
- **Minimum Cell Rate (MCR):** This is the minimum rate desired by a user.

2.8.1.3 Cell Delay Variation Tolerance (CDVT):

A tolerance to accommodate cell delay variation introduced for making reliable connection acceptance/denial decisions. For example the Terminal Equipment (TE) or the Customer's Premises Equipment (CPE), which may alter the negotiated limits of the expected traffic volume. It tells how much variation will be present in cell transmission times. It is specified independently for PCR and SCR.

2.8.2 Traffic Control:

In order to fulfill the ATM objectives, the traffic contract procedure must be supported by control functions. Traffic control consists of a set of mechanisms that regulate the flow of traffic in order to prevent random fluctuations from adversely effecting the quality of offered services. Three essential traffic control functions are:

2.8.2.1 Connection Admission Control:

Connection Admission Control (CAC) represents the set of actions taken by the network at call set-up phase in order to accept or reject an ATM connection [152]. Traffic management is about giving users the services they require and admitting new users into the network without compromising the quality of existing connections. Excessive traffic violates quality of service guarantees, whereas excessive control results in poor network utilization. When the CAC function accepts a connection, the traffic contract is signed.

2.8.2.2 Usage/Network Parameter Control:

Usage/Network Parameter Control (UPC/NPC) are performed at the User-Network Interface (UNI) and Network-Node Interface (NNI), respectively, and represent the set of actions taken by the network to monitor and control traffic on an ATM connection in terms of cell traffic volume and cell routing validity. This function is sometimes also called "police function". The main purpose is to enforce the compliance of every ATM connection to its negotiated traffic contract. Without a UPC/NPC function, a terminal equipment failure, excessive cell delay variation in for example the CPE or even traffic abuse could seriously affect the QoS committed to other already established connections. To determine compliance with a traffic contract, the UPC function implements a policing algorithm such as 'leaky bucket' Generic Cell Rate Algorithm(GCRA).

2.8.2.3 Congestion Control:

QoS guarantees are statistical in nature. On rare occasions the switch may become congested, which implies increased delay, cell loss, or both. In the case of cell loss, a switch must have a cell-discard policy for controlling congestion. When congestion is imminent, the switch begins discarding the least important cells. Above a given threshold, only high priority cells (CLP=0) are allowed. This method of discarding cells is known as Selective Cell Discard (SCD). Data traffic is usually frame-based. Thus, the loss of a single cell implies that the entire frame is useless. Efficiency increases dramatically if the discard decision coincides with the arrival of the first cell of a frame. This policy is called Early Packet Discard (EPD). A related policy, called Partial

Packet Discard (PPD), discards the rest of the frame as soon as a cell is discarded. Congestion control complements Connection Admission Control(CAC).

2.8.3 Traffic-Pattern:

Traffic patterns are at the heart of any performance evaluation of communications networks. An accurate estimation of network performance is critical for the success of broadband networks. Such networks need to guaranteed an acceptable QoS level to the users. Therefore, traffic patterns need to be accurate and able to capture the statistical characteristics of the actual traffic[2]. Among the traffic patterns used in various simulations, the following three are most common:

2.8.3.1 Persistent Sources:

These sources, also known as "greedy" or "infinite" sources always have cells to send, thus the network is always congested.

2.8.3.2 Staggered Sources:

The sources start at different times. This allows us to study the ramp-up (or ramp-down) time of the schemes.

2.8.3.3 Bursty Sources:

These sources oscillate between active state and idle state. During active state, they generate a burst of cells. This is a more realistic source model than a persistent source.

2.9 Phenomena of Congestion and Its Causes:

Modern networks are typically high speed networks, deploying high capacity links, and integrating multiple services. Service integration means that the network has to support both services that need reserved resources, and those which can not reserve resources and must rely on the available resources whenever the need arises. High transmission speed and high capacity mean that large amount of data may be transit through the network, which implies inability of the source to react in time to feedback coming from the network. Modern networks differ substantially from the traditional networks in many aspects-the most important of which is flow control and congestion control. Flow control is a point-to-point traffic control mechanism between the server and client[3], [4], and [181]. It regulates the server into sending no faster than the client can handle. On the other hand congestion is a phenomena where the amount of traffic injected into the network is more than the capacity of the network. In ATM, bursty calls are statistically multiplexed. Each call is assigned some bandwidth that is lower than its peak bit-rate. Statistical multiplexing is more bandwidth efficient and allows more calls to enter the network. However it presents challenges in protecting network to and its users from congestion. Congestion control in ATM networks is difficult phenomena because of the high link speed, diverse service requirement and diverse characteristics of the traffic [61]. Each of the three issues are discussed briefly below:

2.9.1 High Link Speed:

Due to high speed link, the cell processing schemes in ATM must be performed at speeds comparable to the high switching speeds. Which requires simple protocol to avoid the excessive processing time of the software. Another problem caused by the high link rate is the increased propagation delay-bandwidth product, the amount of the traffic that can be in transit during propagation delay time.

Consider a 1Gb/s line with a propagation delay of 20ms for a cross-continental distance of 6,000km. In reactive control systems[3], when the network becomes congested, the destinations can send choke packets to the senders to stop or slow transmission. By the time the choke packets are transmitted, 20 Mb are already in transit. By the time the choke packet reaches the sender, 40Mb will have to be retransmitted. Large retransmission can cause severe buffer management problem.

2.9.2 Multiple Service Requirement:

ATM must provide proper QoS for different service classes. A service class is a set of services that have the same QoS requirements. These requirements are usually measured in terms of maximum delays and cell loss rates[171]. Some services such as voice, real-time video, and data for real

---

[3] *Which we have defined later in this Chapter*

time control have strict delay requirements. If packets are not delivered within their delay requirement, they are consider to be lost because of the real time nature of the service. Delay jitter, the variance of delays, should also be small so the output can be reconstructed in a continuous fashion.

In ATM, even if a call is admitted to the network, the network delay and cell loss rate may not be guaranteed due to ATM's packet switching nature. To satisfy diverse services requirements of synchronous and asynchronous traffic, favorable treatment of some classes may be necessary.

2.9.3 Diverse Traffic Characteristics:

B-ISDN is required to support traffic such as interactive data and video which are highly bursty, while some traffic such as large file is continuous. Burstiness is generally characterized by the ratio of call's peak and average bit rates of transmission. This is not an accurate measure since two calls with similar peak and average rates may actually have dissimilar traffic characteristics. This implies other factors such as burst length, amount of time spent at peak rate or some equivalent measure must be considered.

2.10 Goals of Congestion Control:

The high-level goals of congestion control in ATM networks can be summarized as follows[134]:

2.10.1 Simplicity:

Simple control algorithms are more likely to prove implementable. A simple control architecture improves understandability and the ability to predict control behavior and performance, and is more likely to achieve buy-in from the multiple parties involved in the development, deployment, and eventual success of B-ISDN/ATM. Often it is claimed that resource efficiency must be sacrificed to achieve control simplicity. The challenge is to achieve high resource efficiency while maintaining relatively simple control structures.

2.10.2 Robustness:

While it is natural to expect that the control architecture would be "tuned" with respect to certain assumptions, the controls should remain effective even if some of the assumptions are only partially valid. That is, robustness equates to relative insensitivity to imperfect assumptions.

2.10.3 Flexibility:

It has been noted that B-ISDN is targeted to support a rich mixture of services and applications. An ATM congestion control architecture that has the flexibility to adapt as needed to new situations that arise is thus of paramount importance.

2.10.4 Controllability:

Finally, through congestion control architectures, congestion can be adequately controlled so that efficient network resources utilization is achieved with out paying a penalty in performance.

2.11 Congestion Control in ATM Networks:

The basic problem of ATM networks is the statistical behavior of the cell arrival process. Congestion control in ATM networks can be described at different levels with respect to time scales[48]. For example, the call level has a typical time scale of seconds up to hours, the burst level a millisecond range up to seconds, and the cell level a microsecond range. These levels have different impacts on an ATM network. Congestion control schemes can be divided into two categories: Reactive Control and Preventive Control Schemes.

Reactive control reacts to the congestion after it happens and tries to bring the degree of network congestion to an acceptable level. On the other hand preventive control does not wait until congestion actually occurs, but rather tries to prevent the network from reaching an unacceptable level of congestion [9]. The classification of control options and their time scales is shown in Fig. 2.11. To achieve effective congestion control, the ATM network control function is divided into the following control levels:

2.11.1 Call-Level Control:

At Call Level Control long term congestion is avoided through Call Admission Control (CAC) and traffic load is maintained at a manageable level. Admission control is based on allocation of resources. Admission control decides whether to accept or reject new calls based on whether the required performance can be maintained. A criterion for this must exist in order to preserve the network from getting into a congestion state[154]. When a new call is requested, the network examines its service requirements and traffic characteristics. The network then examines the current load and decide whether or not to accept the new calls . Finally, call-routing is an essential component of the call-level controls, and may be a primary factor in the acceptance or denial of the call, and in determining the call service parameters that the network can support[41]. There are various Call Admission Control schemes for ATM networks as follows:

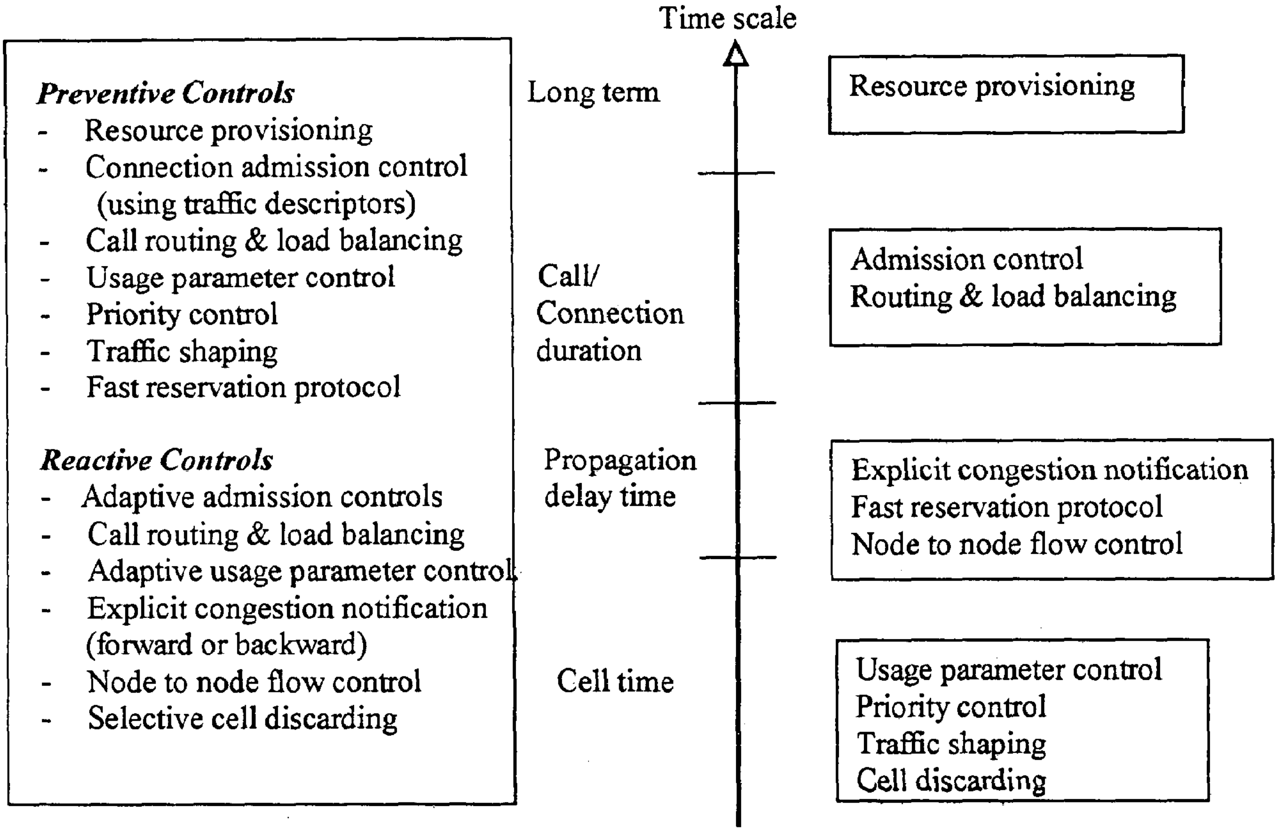


Fig. 2.11: Classification of Control Options with Time Scales

2.11.1.1 Equivalent Bandwidth Method:

Equivalent bandwidth of a call is the amount of bandwidth that call is expected to require. The factors that determines equivalent bandwidth are ratio of the peak bit rate of the call to the link rate, burst length, mean bit rate and loss delay requirement of the call. In this method a new call is admitted if the new bandwidth of new call does not push the expected bandwidth usage for a link over the threshold of available bandwidth. This scheme may not fairly admit calls. Accepting few calls requiring a large amount of bandwidth can block several smaller bandwidth calls so the call blocking probability becomes unacceptably high. Above problem can be overcome by using partial allocation scheme.

In Partial Allocation Scheme[66], a new call is admitted if its bandwidth requirement does not exceed some percentage of available bandwidth. By using this method the arrival rate of high bandwidth calls will not affect the blocking probability of lower bandwidth calls. However, this is at the expense of higher blocking probability for high bandwidth calls. Dynamic bandwidth allocation is necessary for ATM streams that carry VBR traffic. In [13] different allocation policies have been presented and analyzed to understand their dynamics in the presence of link congestion. A Dynamic Bandwidth Controller is proposed for ABR service by A. Smith et. al. in [153].

2.11.1.2 Scout Packet Method:

The scout packet is a method of call admission and set-up in which intermediate nodes do not have to perform costly operations during call admission. In this method, a short stream of pseudo-packets, called scout, with similar traffic characteristics as the desired connections (such as average bit rate and burstiness), is transmitted on the desired route[102]. If the scout finds congestion along the path, the scout will either return to their source in excess of a specified maximum delay or be dropped at the congested link(s). In either case the call is rejected. If no congestion is detected, then call is admitted.

The statistical performance of this method is unknown it can falsely admit a call if a large number of calls were idle when scout were sent. Alternatively, it can falsely reject a call if there was temporary congestion in the network caused by a large number of active calls, even when there is generally enough capacity in the network to accommodate a new call. It performance may depend upon the length of the pseudo-packet stream.

2.11.1.3 Dynamic Call Admission Control Method:

Dynamic call admission control method uses the distribution of the number of cells arriving during fixed interval. This distribution is estimated from the measured number of cells arriving at the output buffer during the fixed interval and traffic parameters specified by the users. Call acceptance is decided on the basis of on-line evaluation of the upper bound of cell loss probability, derived from the estimated distribution of the number of cells arriving. This control mechanism is effective when the number of calls classes are larger.

2.11.2 Cell Level Control:

This level can be controlled by mechanisms such as policing or traffic shaping and priority control. At multiplexing points the traffic pattern at cell level determines the buffer size required for real-time connections [14], and [60]. Cell arrival variations are caused by the principles of asynchronous multiplexing of cells. At cell control level short term congestion is avoided by using two mechanism :

2.11.2.1 Policing Mechanism (Traffic Enforcement):

In cell level control the policing function controls the cell stream during the entire active phase of all and restrict the behavior of traffic source to the characteristics negotiated in the contract. The policing parameter proposed by the CCITT are mean cell rate, peak cell rate and duration of peak (burst length). A leaky bucket method is one of the typical bandwidth enforcement mechanism used for ATM. Leaky bucket appears as the most efficient for mean rate policing [30], and [164].

2.11.2.2 Priority Schemes:

To provide multiple grade of service with ATM, one can use priorities between and within services classes. Having determined priority level for various services. Prioritized cell must handle in appropriate manner during cell discarding and cell scheduling. Connection level priority and cell level priority are two priority control mechanism with different roles and effects on the QoS of ATM services and performance of the network. At any network node, cell belonging to connection supporting reserved connection bandwidth service shall be handled with higher priority than cell belonging to connections supporting the no-reserved-bandwidth services. It is shown that a prioritized system is capable of achieving better performance than non-prioritized system [82], and [142].

2.11.3 Burst Level Control:

Due to the unpredictable nature of applications using the reserved burst bandwidth service, accurate traffic characteristics of the connections are difficult to obtain, and therefore the necessary bandwidth cannot be pre-allocated, with sufficient confidence, at connection setup time. This level can be controlled by fast resource reservation mechanisms and adaptive flow control protocols like ABR. It determines the large buffers needed for non-real-time connections. Under normal circumstances, connection admission control will accept all connections supporting this services category without allocating any bandwidth for them. The bandwidth will be allocated only to the bursts, according to their peak bandwidth requirements, by the burst admission control prior to their transmissions. If there is not sufficient bandwidth available from the pool of bandwidth reserved for this category, the transmission of the burst will be blocked or delayed [170]. Furthermore, specific CAC algorithms may allow connections to be statistically multiplexed at burst level. Burst periods are generated by end applications which transmit more or less information during the connection's lifetime. A burst-level priority scheme for bursty traffic in ATM networks is proposed by Fernandez, R.J. et. al. in [42].

2.11.4 Network-Wide ATM-Level Control:

With service parameters having been established at the call level, there must be a set of mechanisms for observing service agreements during the transport of ATM cells. This is one of the functions of the network-wide ATM–level controls; another function is to provide congestion status information to end terminals. Capabilities in this control domain need to be supported by ATM cell header functionality. The ATM–level control structure has three basic control capabilities[40]:

First, it has a capability for selectively shedding load under congestion conditions, and thus providing network resiliency to traffic uncertainties, implemented via a single indicator in the ATM cell header, termed the "Cell Loss Priority",(CLP) indicator [21], and [22]. If this indicator is set {CLP=1}, it signifies that the cell may be discarded in any network element along the VC/VP path if local congestion above a threshold is encountered in that network element by that cell.

Second, there is a capabilities for forward conveyance of encountered congestion conditions along the VC/VP to the ATM destination terminal, implemented via an Explicit Forward Congestion Indicator (EFCI), borne within the header of ATM cells and set to {EFCI =1} by any congested network element through which that cell passes when congestion exceeds certain defined thresholds.

A third capability, which is equally essential in achieving congestion control objectives, is the ability for a network element to perform service scheduling and resource allocation consistent with the service classes supported by that element. This could involve buffer allocation between service classes, and real-time allocation of link transmission bandwidth so as to provide at least a minimum bandwidth for each service class from "locking out" another from transmission.

2.11.5 Internal Control:

There are potentially many network-element-internal implementations of actions like traffic monitoring, selective cell discard, setting of the Forward Congestion Indicator(FCI), and triggering the backward notification of congestion. However, an additional capability, which is equally essential in achieving overall congestion objectives, is the ability for a network element to perform service scheduling for band width and buffers consistent with the service classes supported by the element. This could involve buffer allocation between service classes, and real time allocation of link transmission bandwidth so as to provide at least a minimal bandwidth for each service class, and thus prevent one service class from "locking out" another from transmission. This established a system of relative priorities between classes.

2.12 Fuzzy Mechanism for Congestion Control:

In modern communication networks, the increasing demand for bandwidth, high reliability and high availability increase demand for network design and control. To be able to provide the QoS guaranteed to users and maintain high network utilization at the same time, all network parameters must be optimized. A major problem of congestion is raised by the statistical multiplexing of cells associated to non-homogeneous traffic flows. It requires a number of congestion control mechanisms for ATM to be applied in order to have a stable network while maximizing network utilization[57]. A system becomes a fuzzy system when its operation are entirely or partially governed by fuzzy logic or are based fuzzy sets. Operations of systems are defined by several basic problems such as control, estimation(prediction, forecasting), modeling, pattern recognition (classification, clustering), optimization, and data compression[12].

Fuzzy set theory, compared to other mathematical theories, is perhaps the most easily adaptable theory to practice. The main reason is that a fuzzy set has the property of relativity, variability, and inexactness in the definition of its elements. Instead of defining an entity in calculus by assuming that its role is exactly known, we can use fuzzy sets to define the same entity by allowing possible deviations and inexactness in its role. This representation suits well the uncertainties encountered in practical life, which make fuzzy sets a valuable mathematical tool. Fuzzy system based traffic controllers are considered as one of the most promising application area in ATM networks. Two complementary, related traffic management mechanisms have been described previously: CAC and UPC. The CAC decides, during the call set up phase (or the call re-negotiation phase), to accept a new connection on a link and, consequently, to allocate a certain portion of bandwidth to it, if the required QoS can be guaranteed for both the connections already established and the new one. As the CAC relies on the negotiated parameter values, it can only perform correctly if all bandwidth contracts are respected. So, a UPC is required in order to ensure that each source conforms to its negotiated parameters. The UPC function can be defined as the set of actions taken by the network, during the entire phase of the call, to monitor and control the offered traffic, with the purpose of protecting network resources from malicious, as well as, unintentional misbehavior which can affect the QoS of already established connections, by detecting violations of negotiated parameters and taking appropriate action. This action can either be cell dropping, cell-marking, or shaping the source rate.

In literature, the UPC function has also been referred to as traffic enforcement or policing. The policing function should fulfill the basic requirements:

- High selectivity with respect to the traffic monitored (that is, the capability of detecting any illegal traffic situation and transparency for connections that respect the parameter values negotiated, on whose cells no policing action need be taken)
- High responsiveness, that is, low response time to parameter violations.
- Simplicity of implementation and cost effectiveness.

Several problems arise in defining an efficient policing mechanism. One is identifying the traffic parameters which best characterize the behavior of a source. The difficulty lies in the fact that the sources to be characterized have different statistical properties as they range from video to data services and it is necessary to define parameters that can be monitored during the call. A traffic parameter contributing to a source traffic descriptors should be significant use in resource allocation, enforceable by the network operator, and understandable by the user. The latter requirement is specifically necessary to allow the user to estimate the value of the parameter in relation to the type of traffic that will be generated. This is still an open issue as, in the case of both average parameters such as long-term average cell rate, average burst duration, average inter-burst time, and in the case of upper-bound parameters such as the sustainable cell rate [161], it is difficult for the user to estimate their value accurately.

The difficulty of characterizing a policer accurately if traditional methods and models are used led us to explore alternative solutions based on artificial intelligence techniques, specifically, in the field of fuzzy system. In this kind of network one of the critical functions is 'policing' which has the task of ensuring that each user complies with the traffic parameters negotiated during cell setup in order to avoid network congestion. Thomas D. Ndouse et. al. [122] presented a fuzzy logic implementation of Virtual Leaky Bucket (VLB) called Fuzzy Leaky Bucket (FLB), which is based on fuzzy control theory. A.R.Bonde et. al. proposed a fuzzy system based Queue–Management scheme in which binary [17] state of buffer occupancy is replaced by fuzzy sets. The ATM forum has accepted rate based framework or family of rate based close loop scheme for congestion control. Within this framework Andreas Pitsillides et. al. [131] designed an effective explicit rate based traffic control strategy, Fuzzy Explicit Rate Marking (FERM) for ABR traffic. In order to prevent the QoS from severely degrading during short-term congestion an appropriate congestion control must also be provided. One approach to congestion is via traffic smoothing. In [119], Masayuki et. al. investigated how network performance depended on the degree of burstiness of the input traffic and observed that smoothing input traffic could reduce network congestion. More deep discussion can be found in [11], [28], [29], [85], [89], and [167].

2.13 Conclusions:

Among the techniques proposed for B-ISDN transfer mode, the ATM concept is considered to be the most promising transfer technique because of its flexibility, and efficiency in bandwidth. In this chapter, a number of topics related to ATM networks are surveyed and reviewed. Congestion control plays an important role providing differentiated QoS and supporting the integration of variety of broadband services. Congestion control schemes in ATM networks can improve overall loss performance, as well as fairness. A great deal of work has already been done on ATM networks although there is more ahead.

A review of the key ATM congestion control mechanisms is presented in the Section 2.11 of this chapter, which can be classified as either preventive or reactive. Preventive mechanisms limit the total amount of traffic admitted to the network in order to virtually eliminate probability of cell-level congestion. In conjunction, reactive controls assist the network and individual connections to avoid the onset of congestion and minimize its severity when it does occur. In addition, ATM congestion control functions must ensure that all connections receive their required QoS. Therefore, ATM connections with more stringent performance requirements must receive some degree of priority over those connections which can tolerate longer cell delay or potential cell losses.

In addition to the above described conventional mechanisms, in section 2.12 fuzzy mechanism to control the congestion is also reviewed. Fuzzy system based traffic controllers are considered as one of the most promising application area in ATM networks. In this kind of network one of the critical functions is 'policing' which has the task of ensuring that each user complies with the traffic parameters negotiated during cell setup in order to avoid network congestion.

Chapter-3

Congestion Control for VBR Service in Non-Blocking ATM Switches

3.1 Introduction:

The introduction of the Internet to the public has increased the amount of traffic passing through today's networks exponentially and this situation is expected to continue for several more years in the future. To cope with sudden changes in network environments, such activities as research and development for relevant technologies are continuing all over the world. The technologies include high-speed packet switching, high-speed transmission, access, and so on. Among them, it is the high-speed packet switching technology that is crucial and indispensable for future broadband digital networks.

The recent emergence of ATM technology makes available a unique driving technology for high-speed communication platforms. The progress in the field of ATM technology has brought up new design principles of high performance, high capacity of switching fabric to be used in integrated network. Virtually all proposals of switching fabric architectures for ATM networks are based on self-routing structures in which each packet autonomously find its path through the interconnection network to the desired output port [58], and [59]. This feature enables the design of the switching fabrics characterized by a very high throughput, on the order of gigabits per second. Statistical multiplexing capabilities is one of the main attraction of ATM networks. This allows VBR and bursty connections to share the same limited resources (e.g. bandwidth and buffers) stochastically[42]. This, in turn, allows more connections to be supported simultaneously than in a circuit switched or STM.

ATM switches can be broadly classified into two main categories, namely internally blocking and non-blocking[45]. A non-blocking switch is a switch that does not suffer from internal blocking since output contention occurs in every switch. For the reason, the non-blocking switch is also called the internally non-blocking switch. Congestion can occur if the switch is a blocking network, that is, if there are not enough switch points to provide simultaneous, independent paths between arbitrary pairs of inputs and outputs as shown in Fig. 3.1.

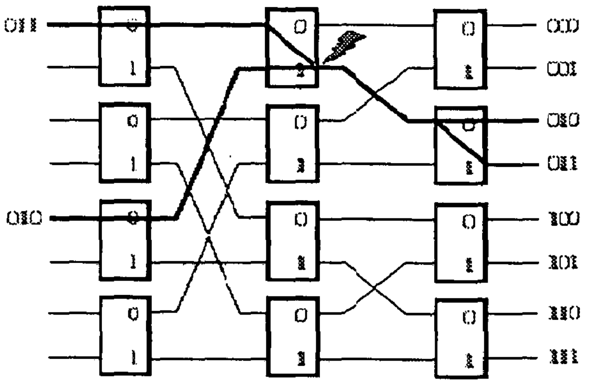


Fig. 3.1: Internal Blocking

But this can also be happen in the case of a non-blocking switch fabric, when two or more packets arrive simultaneously on different inputs requiring to be routed to the same output as represented in Fig. 3.2.

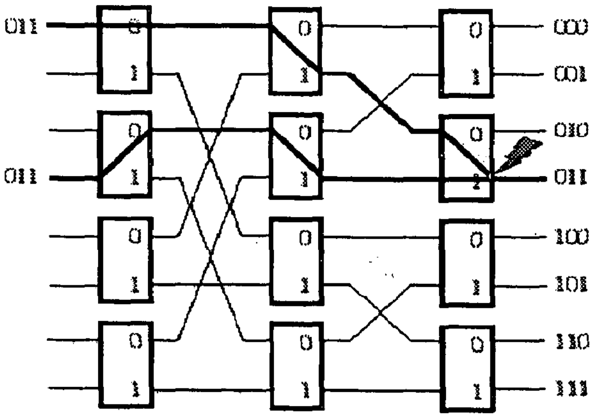


Fig. 3.2: Output Contention

One of these contending packets attains switching to the output. Queueing is required for the others to wait for a later route to the output[139]. It is evident that this form of congestion is unavoidable in fast packet switching systems and it often leads to an increased complexity in the switch architecture. To improve the performance of the switches design proposals were based on various type of queuing strategies: input queuing, dedicated internal queuing, shared internal queuing, or output queuing[128].

In this chapter, we consider only non-blocking ATM switch. The studies [76], and [100] show that a non-blocking switch with input/output queueing achieves the best delay/throughput performance.

The simplest approach, however, is input queuing. In this architecture the approach is taken to solve the possible contention problem at the input[134]. Even with an internally non-blocking switch, when several cells destined for the same output arrive in a time slot, at most one can actually leave the switch; the other must be buffered at the inputs. They are FIFO or allow random access. The simplest approach is to maintain a FIFO queue of cells at each input; only the first cell in the queue is eligible to transmit during the next time slot. This approach has a difficulty, when the cell at the head of an input queue is blocked all the cells behind it in the queue are prevented from being transmitted, even when the output link they need is idle. This is called Head of Line (HOL) blocking. It was shown through mathematical analysis and computer simulation that HOL blocking limits the throughput of each input port to a maximum of 58.6% under uniform random traffic, and much lower than that for bursty traffic[76]. Various approaches have been proposed to overcome the problems associated with FIFO input queuing [72], [76], [117], and [128].

Traffic streams in the real world are often characterized as bursty. Most of the Applications level Data Units (ADUs), such as video frame, are too large to be encapsulated into a single 53-byte ATM cell and must be segmented into a sequence of cells in order to be transmitted over ATM networks. As a result, consecutive arriving cells in a burst are strongly co-related by having the same destination which addresses the same output of the switch. Keeping this point in mind we have simulated and analyzed performance of non-blocking multiple input ATM switches for VBR bursty traffic [55], and[145] using Parallel Iterative Matching (PIM) technique as described in [5].

3.2 Review of the Related Work:

In 1987, Hui and Arthurs [64] and Karol et. al. [76] published their study on the performance analysis of the single input-queued switch under different selection policies, respectively. Their common result was that the switch equipped with a FIFO memory at each input port has a limited throughput performance of $2 - 2^{1/2} = 0.586$. To improve the limited throughput of the input-queued switch, many alternatives have been proposed. If incoming traffic is periodic, Li [103] showed that the aggregate switch throughput can be as small as the throughput of a single link, even for very large switches; this is called stationary blocking. In [19] Cao et. al. investigated the maximum throughput of an ATM switch in the presence of an offered load of multi-cell packets. For the case of input queueing coupled with a round-robin policy for transferring cells from inputs to output, the system is approximated by a product form queueing network. According to San-Qi Li[101], if the source access rate is substantially lower than the link transmission rate, the effect of input traffic correlation on the output contentions can generally be ignored. A concept of input buffer controller is given in [72]. More details can be found in [69], [96], [111], and [174].

3.3 The Switch Model:

The ATM switch under the consideration is an N×N non-blocking switch, i.e., the N inputs are connected to the N outputs via a non-blocking interconnection network. A cell may be sent from any input to the any output, provided that no more than one cell is sent from the same input and no more than one cell is received by the same output. Each input queue of the switch is a random access buffer. This random access buffer can be used to construct the N FIFO queues, each of which is used to store the cells that are destined for one of the N output ports. The architecture of this switch is shown in the Fig. 3.3. A key parameter is the size of each switch, here we have considered the switch size with three values of N viz. 4, 8, and 16 i.e. 4×4, 8×8 and 16×16 switch.

3.3.1 Internal Interconnection:

Once the switch size has been decided, internal data path is needed to transport cells from the switch is organized internally as multistage network of smaller switches. A cell destined for one output can be delayed because of contention at the internal switches with cells destined for other outputs. The following two criteria must be considered for this purpose:

- The switch must route as many cells among the ones that arrived at its inputs as possible to the appropriate outputs to maximize the throughput.
- The switch must solve the output contention problem.

3.3.2 Fixed Length Vs Variable Length Packets:

In this work data are transmitted in fixed length cells rather than variable length packets. We support standard 53 byte ATM cells with 5 byte cell headers and 48 bytes are user data. The use of fixed length cells has number of advantages for frequent switch operation. Such as switch make more frequent scheduling decisions and that a greater portion of link bandwidth is consumed by the overhead of cells header and fragmentation. Second using the cells packet latency can be improved for both short and long packets. Short packets do better because they can be interleaved over a link with long packets. A long packet can't monopolize a connection for its entire duration. For long packets, cells simulate the performance of cut-through[77], while permitting a simpler store-and-forward implementation.

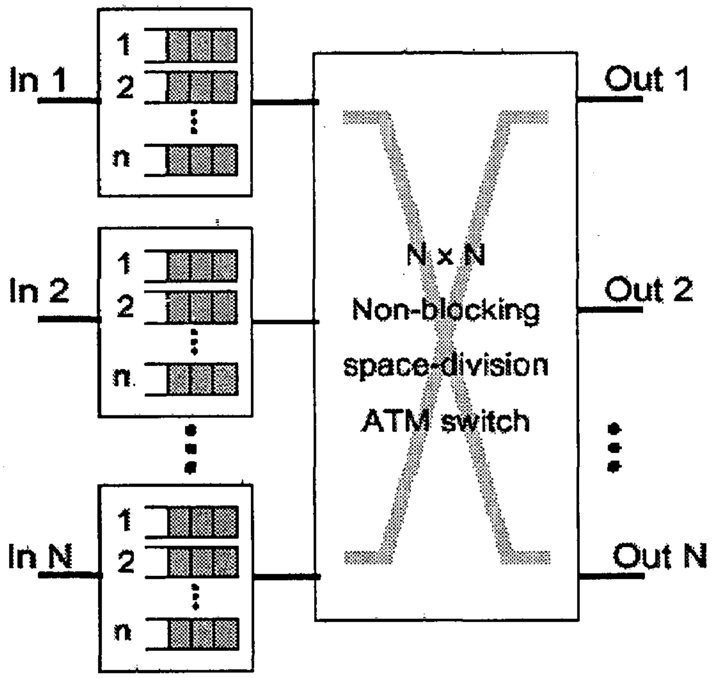


Fig. 3.3: Architecture of a Multiple-Input Queues ATM Switch (N ×N)

3.3.3 Buffer Organization:

Even with an internally non-blocking switch, when several cells destined for the same output arrive in a time slot, at most one can actually leave the switch; the others must be buffered. We put the buffered at the inputs, they are FIFO or allow random access. The simplest approach is to maintain a FIFO queue of cells at each input; only the first cell in the queue is eligible to transmit during the next time slot. This approach has a difficulty is when the cell at the head of an input queue is blocked all the cells behind it in the queue are prevented from being transmitted, even when the output link they need is idle. This is called Head of Line (HOL) blocking [76]. It was shown through mathematical analysis and computer simulations that HOL blocking limits the throughput of each input port to a maximum of 58.6% under uniform random traffic, and much lower than that for bursty traffic.

3.4 Parallel Iterative Matching :

The Parallel Iterative Matching (PIM) is an effective technique[5] used to reduce the HOL blocking. In this technique, each input port maintains a separate queue for each output port, during a single time slot a maximum of one cell per input port can be transferred, and maximum of one cell per out put port can be received. The switch operation is based on the parallel iterative matching algorithm to find the maximal matching between the inputs and outputs of the switch. Here we are considered the architecture of switch with random access buffer and cast the switch scheduling for finding the conflict free pairing of inputs to the outputs. The high throughput and low latency of an ATM switch dictates that the scheduling algorithm must be able to find the matching of as many as conflict free pairings as possible.

3.4.1 Maximal Matching:

A maximal match is one in which pairing can not be trivially added; each node either is matched or has no edge to an unmatched node or in other words it is a matching in which no unmatched input has a queued cell destined for an unmatched output.

3.4.2 Maximum Matching:

Maximum matching is a matching with maximum number of paired input and outputs. There is no other pairing that matches with inputs and outputs. A maximum matching must be a maximal, but reverse is not true.

3.4.3 Parallel Iterative Matching Algorithm:

For switches used in high-performance ATM networks, the switch scheduling algorithm must be able to provide high throughput, low latency, and graceful degradation under heavy traffic loads. As considered in [5] the architecture of a non-blocking switch with random access buffers, the algorithm uses parallelism, randomness, and iteration to find a maximal matching between the inputs that have queued cells for transmission and the outputs that have queued cells (at the inputs) destined for them. Maximal matching is used to determine which inputs transmit cells over the non-blocking switch to which outputs in the current time slot. Their matching algorithm iterates the following three steps until a maximal matching is found or until a fixed number of iterations is performed.

- ***Request*:** Each unmatched input sends a request to every output for which it has a queued cell.
- ***Grant:*** if an unmatched output receives any request, it grants to one by randomly selecting a request uniformly over all requests.
- ***Accept:*** if an input receives grants, it accepts one by selecting an output among those that granted to this input.

Each of these steps occurs independently and in parallel at each input /output port. At the end of one iteration of the protocol, we have a legal matching of inputs to the outputs. More than one input can request the same output; the grant phase chooses among them, ensuring that each output is paired with at most one input. More than one output can grant to the same input (if the input made more than one request); the accept phase choose among them, ensuring that each input is paired with at most one output. An output whose grant is not accepted may be able to paired with an input, none of whose request were granted. We repeat this process of request grant and accept protocol, retaining the matches made in previous iterations.

3.5 Throughput Analysis of Single Iteration:

A potential problem in the ATM networks may be the contention in the switch caused by either long-term or short-term congestion. The first is caused by more incoming traffic that the network can handle, and the second is caused by burstiness in the traffic. Congestion control contributes greatly to a stable and efficient operation of an ATM network[173].

Congestion control through adequate buffering, is becoming particularly significant in minimizing the probability of cell loss and cell delay. As described in [76] and [128], the authors are able to obtain maximum throughput with input queuing up-to 58.6% when N is large. In the following expression we analyzed the ATM switch with one iteration for bursty traffic and find out a maximum throughput up-to approximately 63.2% can be obtained for large value of N.

The throughput of ATM switch with one iteration PIM scheduling $\rho(1)$ is equal to the probability that an out put O_j gets matched after the first iteration. Since each output selects uniformly from all of the input requests, the probability of an input request being accepted by an output $p = \frac{1}{N}$

then, $\rho(1)=\sum_{k=0}^{N-1}\binom{N-1}{k}p^{k}\frac{(1-p)^{N-1-K}}{(k+1)}=1-\left(1-\frac{1}{N}\right)^{N}$

and $\lim_{N\to\infty}\rho(1)=1-\frac{1}{e}=.632$ ------------------------------------(1)

Let Pr {m(1)} be the probability that m(1) inputs (outputs) gets matched and output O_j remains unmatched after the first iteration, then

$\Pr\{m(1)\}=\binom{N-1}{m(1)}m(1)!\,S_N^{(m(1))}\;{}^{1}\!\Big/_{N^{N}}$ ------------------------------(2)

Where $S_n^{(m)}$ is the striling number of the second kind, which gives the number of ways of partitioning a set of *n* elements into *m* nonempty subsets.

$$S_n^{(m)}=\frac{1}{m!}\sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}k^{n}$$

In Equ.(2) $m(1)!\,S_N^{(m(1))}$ accounts for the number of ways in which N inputs contend for the given *m(1)* outputs with condition that each output is requested by at least one input.

Let $\Pr\{n(1)\}$ denote the probability that *n(1)* inputs (outputs) remains unmatched after the first iteration with the condition that output O_j remains unmatched, then

$$\Pr\{n(1)\} = \Pr\{m(1)\} = N - n(1)\}$$

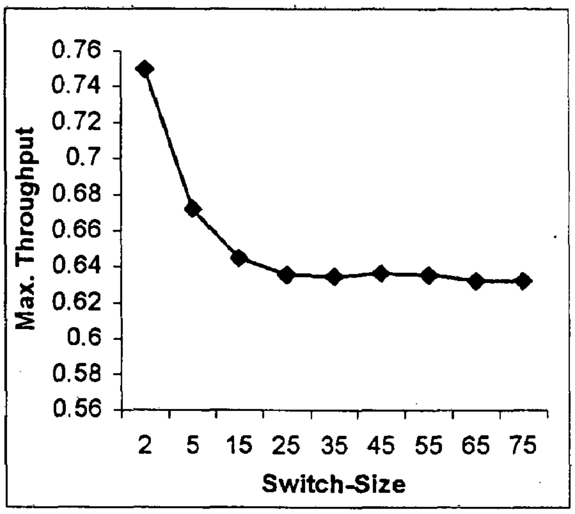


Fig. 3.4: Maximum Throughput Vs Switch Size

The results are verified by computer simulation and are presented in the Fig.3.4.

3.5.1 Throughput with Multiple Iterations:

The throughput of two iterations PIM scheduling is equal to the sum of $\rho(1)$ and the probability that output O_j gets matched in the second iteration, that is

$$\rho(2) = \rho(1) + \Pr\{\text{ output } O_j \text{ gets matched in the second iteration }\}$$

$$= \rho(1) + \sum_{n(1)=1}^{N-1}\left(1-\left(1-\frac{1}{n(1)}\right)^{n(1)}\right)\Pr\{n(1)\}.$$

Similarly, we can calculate the throughput of three iteration scheduling as follows:

$$\rho(3) = \rho(2) + \Pr\{\text{output } O_j \text{ gets matched in the third iteration}\}$$

$$=\rho(2)+\sum_{n(2)=1}^{N-1}\left(1-\left(1-\frac{1}{n(2)}\right)^{n(2)}\right)\Pr\{n(2)\}.$$

To derive the throughput of i iterations PIM scheduling $\rho(i)$, can be written as

$$\rho(i)=\rho(i\text{-}1)+\sum_{n(i-1)=1}^{N-(i-1)}\left(1-\left(1-\frac{1}{(n(i-1))}\right)^{n(i-1)}\right).\Pr\{n(i-1)\}. \quad\text{-------------(3)}$$

and

$$\Pr\{n(i)\}=\sum_{n(i-1)=n(i)+1}^{N-(i-1)}\Pr\{m(i)=n(i-1)-n(i)\}\Pr\{n(i-1)\}$$

Where

$$\Pr\{m(i)\}=\binom{n(i-1)-1}{m(i)}m(i)!S_{n(i-1)}^{(m(i))}\Big/(n(i-1))^{n(i-1)}$$

The maximum throughput of an ATM switch with one iteration PIM scheduling converges to 0.632. When the switch size increase the throughput increase significantly after each iteration of PIM scheduling.

3.6 Simulation and Interrupted Bernoulli Process (IBP) Traffic Model:

For developing simulation program for communication systems we considered the following points:

- Modeling of random user demands for the network.
- Simulation of network resources needed for processing those demands.
- Estimating performance base on output data generated by simulation.

In the present case, VBR traffic (bursty source) is modeled by an Interrupted Bernoulli Process (IBP) with parameter α_i and β_i as shown in Fig. 3.5. The IBP can be described as follows:

- The ON and OFF states appear in turn in IBP.

- The duration of each state can be approximated by a geometric distribution whereas cells arrival during ON state is approximated by Bernoulli Distribution.
- An IBP changes from ON to OFF state with probability α_i per slot and from OFF to ON state with probability β_i per slot. It may stay in an ON or OFF state a geometrically distributed length of time with the average $1/\alpha_i$ & $1/\beta_i$ respectively.
- An IBP generates a cell with the rate per slot when it is in the ON state and OFF state does not generate any cell[9].

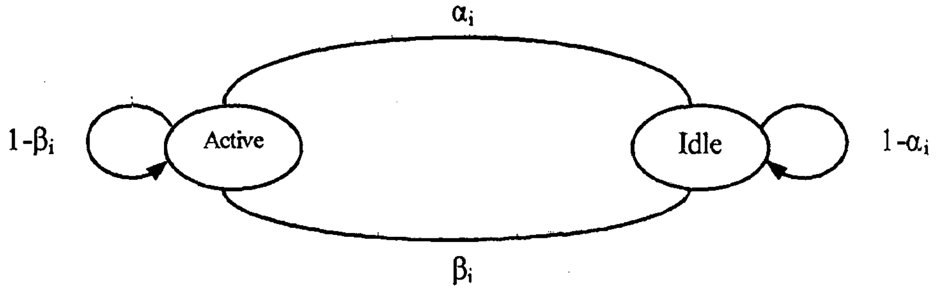


Fig. 3.5: A Two State ON-OFF Model

For IBP source modelling, three parameters are required to described,

1. Peak Cell Rate (P)
2. Average Cell Rate (A)
3. Average Burst Length (B)

The steady state probability (ρ) that any slot contains a cell is given by:

$$\rho = \frac{B}{(B+I)},$$

from the given Fig. 3.6,

$$(B+I)\,A = PB$$

therefore $I = \frac{(PB - BA)}{A} = B\left(\frac{P}{A} - 1\right)$

$$\alpha_i = \frac{1}{B}, \text{ and } \beta_i = \frac{1}{I}$$

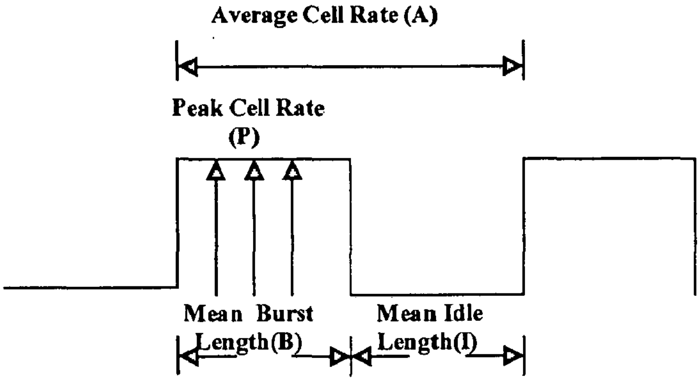


Fig. 3.6: Bursty Traffic Characteristics

Here α_i and β_i characterizes the duration of the active and idle periods respectively, so probability that the active period lasts for a duration of j time slots is given by :

$P(j) = \alpha_i\,(1-\alpha_i)^{j-1}$ for $j \geq 1$

and the corresponding average burst period is given by :

$$E_B[j] = \sum_{j=1}^{\infty} jP(j) = \frac{1}{\alpha_i}$$

Similarly, the probability that an idle period lasts for k time slots is:

$R(k) = \beta_i(1-\beta_i)^k$ for $k \geq 0$

and the corresponding mean idle period is given by:

$$E_I(k) = \sum_{i=0}^{\infty} k\,R(k) = \frac{(1-\beta_i)}{\beta_i}$$

So offered load is given by

$$L = \frac{E_B[j]}{(E_B[j] + E_I[k])} = \frac{\beta_i}{(\alpha_i + \beta_i - \alpha_i\beta_i)}$$

3.7 Results:

The performance of non-blocking ATM switches have been evaluated for VBR service. In this work, link bandwidth is taken as 155.5 Mbps. So minimum delay suffered by a cell is 2.827 μs. Each input VBR source i, i=1,2,.......,N is modeled by two state ON-OFF Interrupted Bernoulli Process (IBP). We considered three non-blocking ATM switch configurations with dimensions 4×4, 8×8 and 16×16, with input buffer size (Siq)=4. Simulation results are taken for four iterations after applying different load conditions. The results are obtained for the three important performance indices i.e. throughput, average cell delay and cell loss probability Vs load, which are presented in the form of Fig. 3.7-3.15 and their respective Tables 3.1-3.9.

Figs. 3.7-3.9 (or Tables 3.1-3.3) show that the switch throughput as a function of offered load for switch sizes 4,8 and 16, with various PIM scheduling iteration numbers 1,2,3 and 4 respectively. It can be seen that when the switch size increases, the throughput of the switch decreases under high offered load. At lower load or initial load conditions the value of throughput is 100% to 90%, but at higher load due to sudden burst the throughput decreases. Also from the figs./tables, we observe that the saturation throughput will increase when the PIM scheduling iteration increases. It is expected that with more iterations, more HOL cells get matched during a scheduling iteration. The results show that three iterations are sufficient to get the maximum throughput.

Figs. 3.10 - 3.12 (or Tables 3.4 - 3.6) show that mean cell delay as a function of offered load for switch sizes 4,8 and 16 with various PIM iteration numbers. These figs./tables indicate that when the switch size increases, the mean cell delay also increased. The mean cell delay decreased for a particular switch size with the traffic load as number of iterations increased.

In a similar fashion like average cell delay, the mean cell loss probability of PIM switches varies with switch size and number of iterations as cleared from Figs. 3.13 – 3.15 (or Tables 3.7 –3.9). It is not difficult to understand that, under low traffic load, the opportunity that more than one HOL cell contends for a common input/output is small. That is, single iteration PIM scheduling is typically enough to find a maximal matching. When the traffic load is increases, the chances of conflicts increases and more iterations are needed using PIM scheduling to achieve a maximal matching.

3.8 Conclusions:

This chapter has surveyed the causes of congestion and effect of HOL blocking on the performance of non-blocking ATM switches. In order to improve the performance of non-blocking ATM switches and to control the congestion for VBR traffic, we need to deal with the HOL blocking problem. An effective technique called Parallel Iterative Matching (PIM) is used to reduced the HOL blocking. For the switch with PIM scheduling, the contention resolution process consists of two stages. According to PIM algorithm a HOL cell in an input queue will contend for transmission not only with the HOL cells of the same input, but also the HOL cells destined for the same output. According to the inventors of PIM switches it is difficult to mathematically analyze the performance of a PIM switch even for the simplest traffic model. The problem lies in the evolution and interdependence of the state of each arbiter and their dependence on arriving traffic[22].

In this chapter the throughput of an ATM switch with single iteration and multiple iteration is analyzed. Various performance parameters including throughput, average cell delay, and cell loss probability of a non-blocking ATM using a PIM scheduling scheme are presented. It was shown that PIM technique increases throughput and reduces not only mean cell delay but cell loss probability also. These results are very much suitable for providing better QoS for real time VBR bursty traffic applications. Finally we observed that various graphs almost clustered to-gather under different iterations, which indicates that three iterations are enough to find maximal matching under any traffic load.

Table 3.1: Avg.Throughput Vs. Load For 4X4 PIM Switch, Siq = 4

Load	*Iteration-I*	*Iteration-II*	*Iteration-III*	*Iteration-IV*
0.1	1	1	1	1
0.3	0.9888	0.9962	0.9963	0.9963
0.5	0.9438	0.9589	0.9619	0.9619
0.7	0.8135	0.8763	0.8747	0.8792
0.9	0.6594	0.7319	0.7334	0.7334

Fig. 3.7: Avg. Throughput Vs. Load
for 4X4 PIM Switch, Siq=4

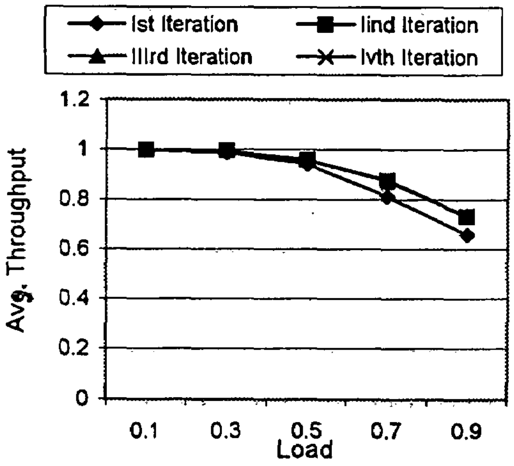


Table 3.2: Avg.Throughput Vs. Load For 8X8 PIM Switch, Siq = 4

Load	*I teration-I*	*Iteration-II*	*Iteration-III*	*Iteration-IV*
0.1	0.9995	0.9995	0.9995	0.9995
0.3	0.9947	0.9958	0.9957	0.9957
0.5	0.9536	0.9768	0.9784	0.9785
0.7	0.7892	0.9118	0.917	0.9177
0.9	0.6143	0.7289	0.7405	0.7405

Fig. 3.8: Avg. Throughput Vs. Load
for 8X8 PIM Switch, Siq=4

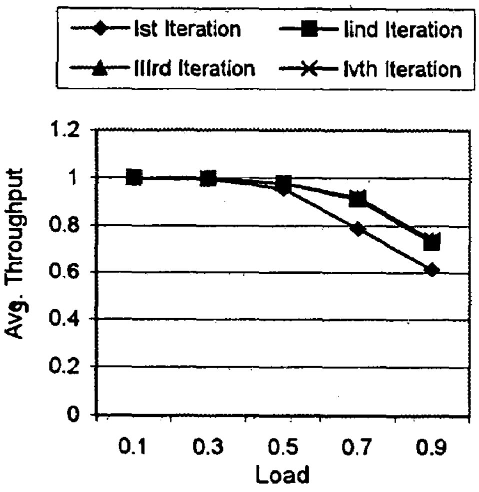


Table 3.3: Avg.Throughput Vs. Load For 16X16 PIM Switch, Siq = 4

Load	*Iteration-I*	*Iteration-II*	*Iteration-III*	*Iteration-IV*
0.1	0.9998	0.9993	0.9993	0.9993
0.3	0.9966	0.9986	0.9978	0.9979
0.5	0.9706	0.9903	0.9868	0.9868
0.7	0.7841	0.9301	0.9425	0.943
0.9	0.6107	0.7319	0.7452	0.7456

Fig. 3.9: Avg. Throughput Vs. Load
for 16 X16 PIM Switch, Siq=4

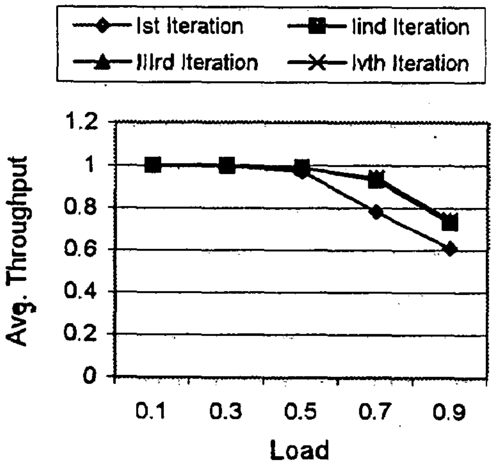


Table 3.4: Avg. Cell Delay Vs. Load For 4X4 PIM Switch, Siq = 4

Load	*Iteration-I*	*Iteration-II*	*Iteration-III*	*Iteration-IV*
0.1	4.2719	4.0231	4.0231	4.0231
0.3	8.2191	6.821	6.821	6.821
0.5	15.1226	11.5056	11.4785	11.4785
0.7	25.1602	18.09	17.9015	17.943
0.9	35.2001	27.6378	27.4835	27.4835

Fig. 3.10: Avg. Cell Delay Vs. Load for 4X4 PIM Switch, Siq=4

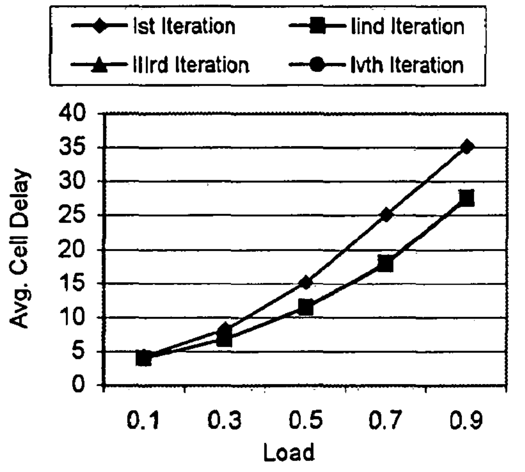


Table 3.5: Avg. Cell Delay Vs. Load For 8 X8 PIM Switch, Siq = 4

Load	*Iteration-I*	*Iteration-II*	*Iteration-III*	*Iteration-IV*
0.1	5.0434	4.3301	4.3314	4.3314
0.3	9.1337	7.7031	7.679	7.6794
0.5	23.56	12.736	11.6119	11.6119
0.7	56.372	29.1401	27.3865	27.011
0.9	79.9761	57.4801	55.1096	54.3539

Fig. 3.11: Avg. Cell Delay Vs. Load for 8X8 PIM Switch Size, Siq=4

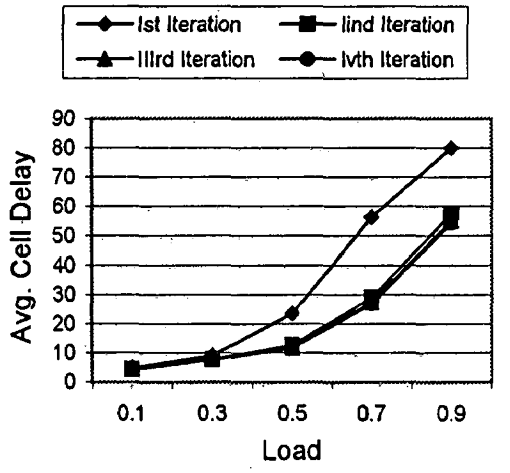


Table 3.6: Avg. Cell Delay Vs. Load For 16X16 PIM Switch, Siq = 4

Load	*Iteration-I*	*Iteration-II*	*Iteration-III*	*Iteration-IV*
0.1	4.7935	4.4624	4.4624	4.4624
0.3	9.7694	7.4813	7.7699	7.7699
0.5	34.056	12.8907	12.628	12.628
0.7	116.94	48.264	42.5546	41.3169
0.9	161.7086	111.5177	106.675	106.94

Fig. 3.12: Avg. Cell Delay Vs. Load for 16X16 PIM Switch, Siq=4

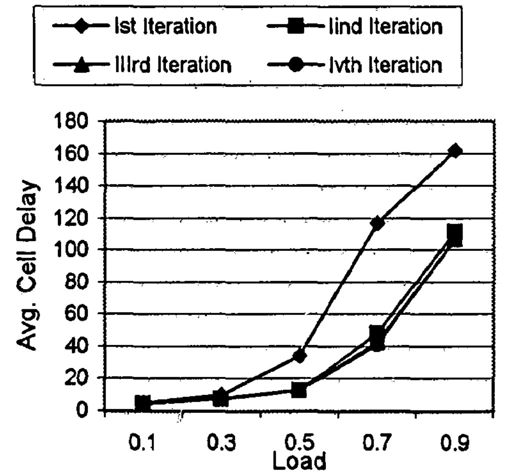


Table 3.7: Cell Loss Prob. Vs. Load
For 4X4 PIM Switch, Siq = 4

Load	*Iteration-I*	*Iteration-II*	*Iteration-III*	*Iteration-IV*
0.1	0	0	0	0
0.3	0.0112	0.0037	0.0037	0.0037
0.5	0.0562	0.0411	0.0381	0.0381
0.7	0.1865	0.1236	0.1253	0.1208
0.9	0.3406	0.2681	0.2665	0.2665

Fig. 3.13: Cell Loss Prob. Vs. Load
for 4X4 PIM Switch, for Siq=4

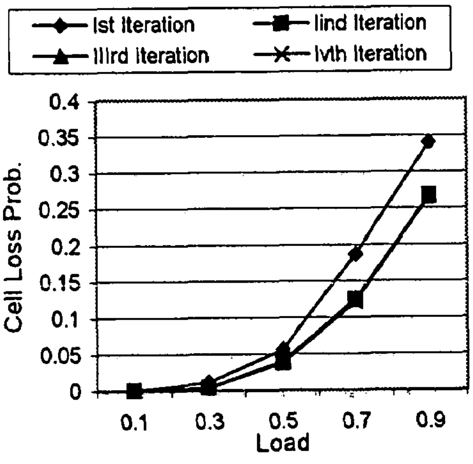


Table 3.8: Cell Loss Prob. Vs. Load
For 8X8 PIM Switch, Siq = 4

Load	*Iteration-I*	*Iteration-II*	*Iteration-III*	*Iteration-IV*
0.1	0.00047	0.00047	0.00047	0.00047
0.3	0.0053	0.00423	0.00424	0.00424
0.5	0.04636	0.02313	0.02154	0.02154
0.7	0.2108	0.0882	0.083	0.0823
0.9	0.3856	0.2711	0.2595	0.2595

Fig. 3.14: Cell Loss Prob. Vs. Load
for 8X8 PIM Switch, Siq=4

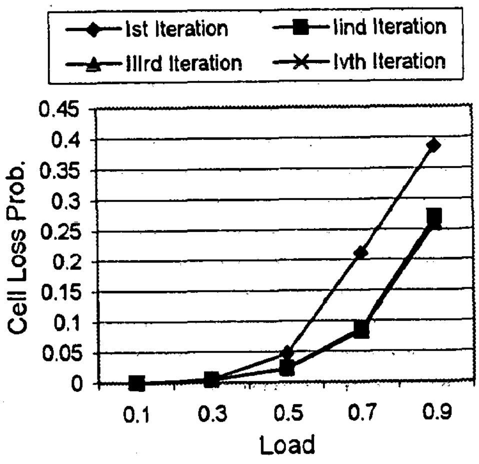
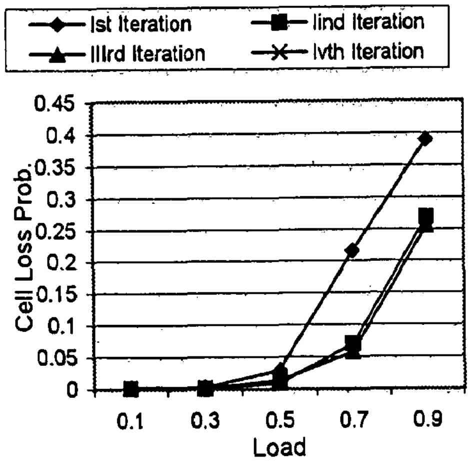


Table 3.9: Cell Loss Prob. Vs. Load
For 16X16 PIM Switch, Siq = 4

Load	*Iteration-I*	*Iteration-II*	*Iteration-III*	*Iteration-IV*
0.1	0.0022	0.00066	0.00066	0.00066
0.3	0.0034	0.0014	0.0021	0.0021
0.5	0.02938	0.0097	0.0132	0.0132
0.7	0.2159	0.0698	0.0575	0.0575
0.9	0.3893	0.2683	0.2548	0.2544

Graph 3.15: Cell Loss Prob. Vs Load
for 16X16 PIM Switch, Siq=4



4.1 Introduction:

A major development in high-speed networking is the emergence of B-ISDN's and ATM. ATM has been designed to support various classes of multimedia traffic with different bit rates and QoS requirements. Due to the unpredictable fluctuations and burstiness of traffic flow within multimedia networks, congestion can occur frequently. Therefore, it is necessary to design appropriate congestion control mechanisms to ensure the promised QoS is met. Meanwhile, due to the high speed transfer rate and rather short cell length, the ratio of propagation delay to cell transmission time and the ratio of processing time to cell transmission time of ATM networks are significantly higher than that of existing networks. This leads to a shift in the network's performance bottleneck from channel transmission speed to propagation delay of the channel and the processing speed at the network switching nodes[9].

The congestion is a result of a mismatch between the network resources (buffer space, processing and transmission capacity) and the amount of traffic admitted for transmission. Consequently, congestion prevention can be interpreted as the problem of matching the admitted traffic to the network resources. This, in turn, could be viewed as a classical problem of feedback control i.e. matching the output to the input of dynamical systems[10]. In feedback controls, when possible traffic congestion is detected at any network element, feedback signals are sent back to all sources. ATM layer congestion control refers to the set of actions taken by the network to minimize the intensity, spread, and duration of congestion[68]. Feedback flow control is one of the solutions which has been reported in the literature [7], [104], [105], [115], [123], and [155].

The growing success of fuzzy logic in various fields of applications, such as control, decision support, knowledge base systems, data base information retrieval and pattern recognition, is due to its inherent capacity to formalize control algorithms that can tolerate imprecision and uncertainty, emulating the cognitive processes that human beings use every day [12], [118], [149], [178], and [179]. Fuzzy logic system have been successfully applied to deal with congestion control related problems in ATM networks and have provided a robust mathematical frame work for dealing with real world imprecision [11], [28], [29], [85], [89], [122], [131], and [167]. The fuzzy approach exhibits a soft behavior, which means a greater ability to adapt itself to dynamic, imprecise and bursty environments. Comparative studies[17], [20], and [85] have shown that the fuzzy approaches significantly improve system performance compared with conventional approaches.

In conventional schemes, a binary threshold divides the buffer space in two parts: below or equal to the threshold level, for every arriving cell is given entry to the network and above the threshold every cell is rejected. In fixed threshold case as described by Bonde et. al. [17], two states of buffer - block and admit can be replaced by fuzzy sets. At each of the permissible level of buffer occupancy, a fixed level of percentage blocking is defined. Thus, selective blocking is used, here a sigmoid shaped function is chosen to represent degree of blocking offered to the incoming cells. However, in the scheme no weightage is given to the burst length in connection to congestion control. The concept of using burst length for bandwidth reservation have been utilized in some existing protocol such as in Fast Reservation Protocol (FRP). The problem in this approach is that, even when the buffer is almost empty, a long burst of cells may overflow the buffer. Similarly, even when the buffer is almost full, buffer may accommodate a very short burst of cells. So without considering the burst length, congestion control is less sensitive to the incoming traffic changes. We have proposed the use of fuzzy logic for dynamic feed-back threshold in our scheme [53], and [90], in which burst length as well as buffer occupancy are represented by triangular functions .

4.2 Fuzzy Expert System:

Fuzzy logic provides a general concept for description and measurement. Most fuzzy logic systems encode human reasoning into a program to make decisions or control a system. Fuzzy logic comprises fuzzy sets, which are a way of representing non-statistical uncertainty and approximate reasoning, which includes the operations used to make inferences in fuzzy logic.

Chapter-4

Congestion Control in ATM Network using Fuzzy Approach

The goal of the defuzzifier is to map the output fuzzy sets to a crisp output value. It combines the different fuzzy sets [140] with different degrees of possibility to produce a single numerical value.

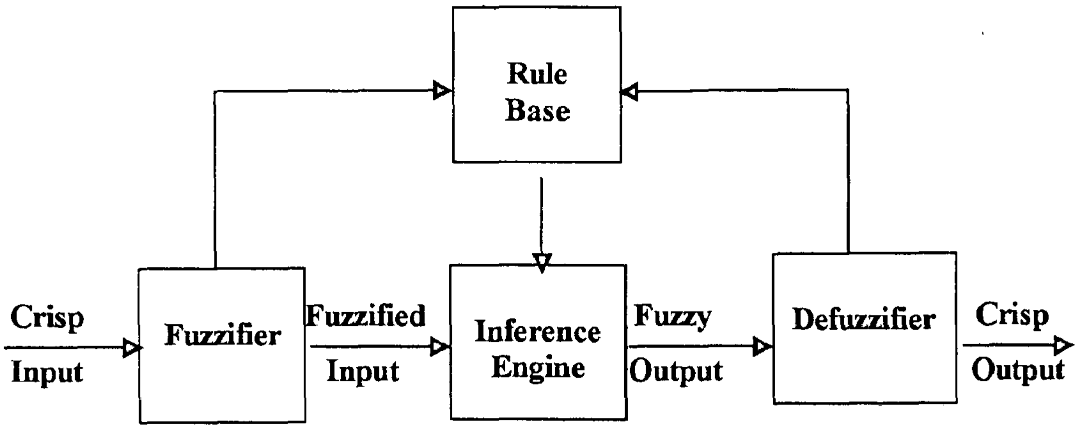
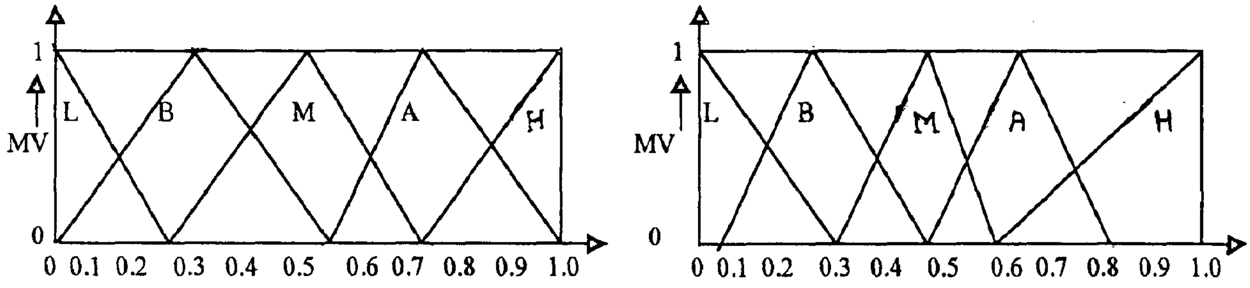


Fig. 4.1: Model of Fuzzy Controller

The fuzzy inference engine defines how the system should infer through the rules in the rules base to determine the output fuzzy sets. Ray-Guang Cheng et. al. [28] developed a model of a fuzzy traffic controller in which inputs linguistic variables are chosen so that the controller is a closed-loop system with the stable and robust operation.

The heart of a Fuzzy system is a rule base which consists of a set of If-Then rules. The rules are statements in which some words are characterized by continuous membership functions. For example, IF the link is close to congestion THEN reduce the input rate, the words close to congestion are characterized by a membership function as shown in the Figs.4.2(a) and 4.2(b), where congestion is considered happening when the link utilization is above 0.8.



(a) (b)

Fig. 4.2: A Typical Representation of Buffer Occupancy as well as Burst Length by Fuzzy Sets

L, B, M, A and H represent Low, Below medium, Medium, Above medium and High membership sets respectively. M.V. represents membership values.

The fuzzy system encodes expert knowledge about the system to be implemented rather than modeling the actual system, therefore it resembles a rule based expert system. However, unlike expert system fuzzy system does not fail when faced with a control situation in which no rule is defined. Instead, controls are inferred using the membership function to generate approximate control actions.

4.4 Applied Fuzzy Approach:

To tackle congestion, schemes are needed which can keep the resource utilization under check. These schemes should be simple, flexible, and should not degrade the performance. For applied scheme out-put buffer divided into various number of equal parts viz. two, three, and four for the purpose, then the feedback had applied after 50%, 33%, and 25% completion of the buffer space i.e. when N=2, 3, and 4 respectively. Depending upon which threshold has been crossed, the network gets a mild warning, a stem warning or an ultimatum. A gradual change is more intuitive here, this has been incorporated with fuzzy logic. In applied fuzzy scheme, burst length as well as buffer occupancy are represented by triangular functions as shown in Fig. 4.2. The degree of membership of a particular set, associated with each valid buffer occupancy can be read from this figure. This quantification of membership is called fuzzification. From these membership values and corresponding sets, blocking to be offered, again in fuzzy terms can be find out. This process is called rule-based inference. As an example, a typical rule is when buffer occupancy is high and burst length is high, number of blocked cells is also high as shown in Lookup Tables 4.1(a), and 4.1(b). Then, by applying suitable defuzzification method, the percentage blocking to offered at that particular buffer occupancy level and at given burst length can be determined. For defuzzification, with the set such as shown in the Table 4.2, weighted average is used.

BL BO	L	B	M	A	H
L	L	L	L	B	B
B	B	B	M	M	M
M	B	M	M	A	H
A	M	M	A	H	H
H	H	H	H	H	H

Blocking

Table 4.1(a) : Lookup Table

BL BO	L	B	M	A	H
L	L	L	B	M	M
B	L	B	B	M	A
M	B	B	M	A	A
A	M	M	A	H	H
H	M	A	H	H	H

Blocking

Table 4.1(b): Lookup Table

SET	L	B	M	A	H
I	0.05	0.25	0.50	0.75	0.95
II	0.05	0.20	0.40	0.60	0.80

Table 4.2: Defuzzification Table

L: Low Set, B: Below Medium Set, M: Medium Set, A: Above Medium Set, H: High Set

A typical example is explained as follows : Let us assume that buffer occupancy as well as burst length both are characterized by the fuzzy set described in Fig. 4.2(a). Also, maximum buffer size is kept at 8 and maximum burst length is assumed to be 8. Suppose, at the time of the new arriving cell burst, buffer occupancy = 5 and arriving burst length = 6. When normalized with respect to maximum value of 8, these variables are mapped as buffer occupancy =0.625 and burst 0.75. Using fuzzy set of Fig. 4.2(a) for fuzzification it is seen that, buffer occupancy is a member of set M with associated value 0.5 and a member of set A with associated value 0.5. Burst length is a member of set A with associated value 1.0 and a member of set H with associated value 0.0. Using Table 4.1(a) and min-max method of evaluation, we get:

Buffer occupancy M(0.5) and burst length A(1.0)=>Blocking of A(0.5)

Buffer occupancy M(0.5) and burst length H(0.0)=>Blocking of H(0.0)

Buffer occupancy A(0.5) and burst length A(1.0)=>Blocking of H(0.5)

Buffer occupancy A(0.5) and burst length H(0.0)=>Blocking of H(0.0)

Thus, taking maximum of the four values associated with H, blocking has membership of set A with value (0.5) and membership of set H with value (0.5). Using these sets with weighted average of membership values, percentage blocking offered can be found. For defuzzification, set 1 of Table 4.2 is used.

$$\text{Blocking} = \frac{[(0.5\times0.75)+(0.5\times0.95)]}{(0.5+0.5)} = 0.85$$

Thus the percentage blocking to be offered, as per the proposed scheme is 85%. Based on this method of determining percentage blocking for the incoming cells, an ATM node is simulated, and performance of the scheme has been compared with static and dynamic feed-back schemes.

4.5 Results:

The simulation results are shown in the Tables 4.3-4.11 and Figs. 4.3-4.11 indicated that the over all performance of the ATM node improved when we applied fuzzy logic to Dynamic Feed-back Threshold scheme. In this work, link bandwidth is taken as 155.5 Mbps. So minimum delay suffered by a cell is 2.827 μs. Each input VBR source i, i=1,2,.......,N is modeled by two state ON-OFF Interrupted Bernoulli Process (IBP)$^{\Psi}$. We first considered switch of size 10×10, with input burst length = 4 and output burst length = 8. We had applied a Constant Threshold (C.Th.) = 4, the size of the output buffer (Bop) is kept 10.

Out-put buffer had divided into equal number of parts viz. two, three, and four. The feedback had applied after 50%, 33%, and 25% of the buffer space gets filled i.e. when N =2, 3, and 4 respectively under Dynamic Threshold Feed-back (D.Th.Fb.) scheme. Simulation results are taken for these values of N, after applying different load conditions. A gradual change is more intuitive here, this has been done with fuzzy logic in our proposed Fuzzy Feed-back (F.Fb.) scheme.

The results are obtained for the three important performance indices i.e. throughput, average cell delay and cell loss probability Vs load. The performance of the new proposed scheme has been compared with Constant Threshold and Dynamic feed-back Threshold based schemes. From the

---

$^{\Psi}$ The IBP Traffic Model is described in Section 3.6 of chapter –3.

results we observed that all the QoS parameters as described above are the function of offered load and number of buffer parts(N). (But for Constant Threshold Scheme these parameters don't depend on the value of N).

Figs. 4.3, 4.6, and 4.9 (Respective Tables 4.3, 4.6, and 4.9) indicate that, for low loads ($L \leq 0.5$) all the schemes provide about the same throughput (100% - 99%). Which shows that all the incoming cells are served by the switch, so we will limit our discussion to higher loads. For moderate loads ($0.5 \leq L \leq 0.7$), due to rigidness of the Constant Threshold the throughput decreases from 99% to 97%, but the remaining schemes again have same results. The reason is that since after the completion of every 50%, 33% and 25% of the buffer space the network gets a proper signal to control the incoming burst of cells. At higher loads ($0.7 \leq L \leq 0.9$), the throughput decreases up-to 96% for Constant Threshold Scheme. At these load conditions the value of throughput increases gradually for Dynamic Feed-back Scheme with respect to N, while remains constant for proposed Fuzzy Scheme.

Figs. 4.4, 4.7, and 4.10 (Respective Tables 4.4, 4.7, and 4.10) show plots of the average cell delay against load for all the schemes taking various value of N. The value of average cell delay for Constant Threshold Scheme increases rapidly as offered load changes from lower to higher. This parameter is again doesn't depend upon value of N. But the value of average cell delay is very low for Dynamic Feedback Threshold Scheme. For proposed Fuzzy Scheme it is minimum.

The results for average cell loss probability against load for the applied schemes are shown in Figs. 4.5, 4.8, and 4.11 (Tables 4. 5, 4.8, and 4.11). Like average cell delay, the results show that Proposed Fuzzy Scheme has minimum value of average cell loss probability too. The reason can be explained as follows:

The threshold function determines, for each cell-burst, how many of the arriving cells to admit into the buffer. This function bears significant influence on the performance of the network including the fraction of cells lost due to dropping or excessive delays and the delay distribution of

Table 4.3: Avg.Throughput Vs Load,When N=2

Load	*C-Th*	*D-Th-Fb.*	*F-Fb.*
0.1	1	1	1
0.3	1	1	1
0.5	0.9939	0.9992	0.9992
0.7	0.9733	0.9969	0.9985
0.9	0.9608	0.9766	0.9962

C.Th. = Constant Threshold
D.Th.Fb.=Dynamic Threshold Feed-back
F.Fb.=Fuzzy Feed -back

Fig. 4.3: Avg. Throughput Vs Load for Switch Size 10 X10, N=2

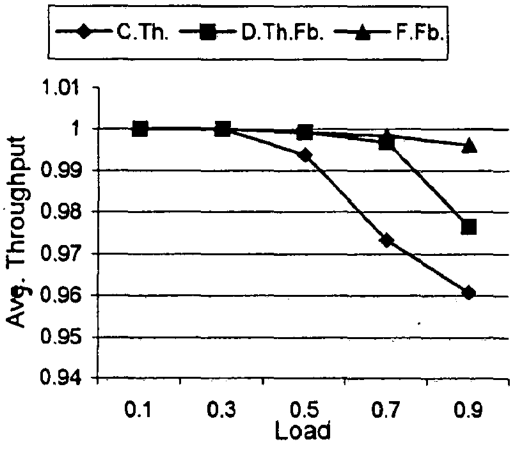


Table 4.4: Avg. Cell Delay Vs Load, When N=2

Load	*C-Th*	*D-Th-FB*	*F-FB*
0.1	0.0062	0.0062	0.0062
0.3	0.5469	0.0195	0.01628
0.5	0.5673	0.0401	0.0433
0.7	0.7422	0.1223	0.0878
0.9	0.6494	0.3419	0.1269

Fig. 4.4: Avg. Cell Delay Vs. Load for Switch Size 10X10, N=2

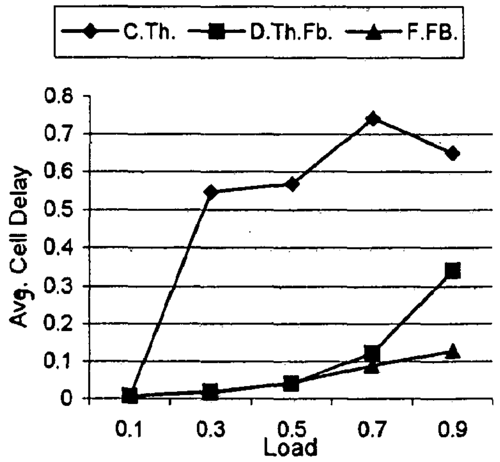


Table4.5:AvgCell Loss ProbVs Load,WhenN=2

Load	*C-Th*	*D-Th-FB*	*F-FB*
0.1	0	0	0
0.3	0	0	0
0.5	0.0061	0.0008	0.0007
0.7	0.0268	0.0031	0.0015
0.9	0.0393	0.0234	0.0038

Fig. 4.5: Avg. Cell Loss Prob. Vs Load for Switch Size 10X10, N=2

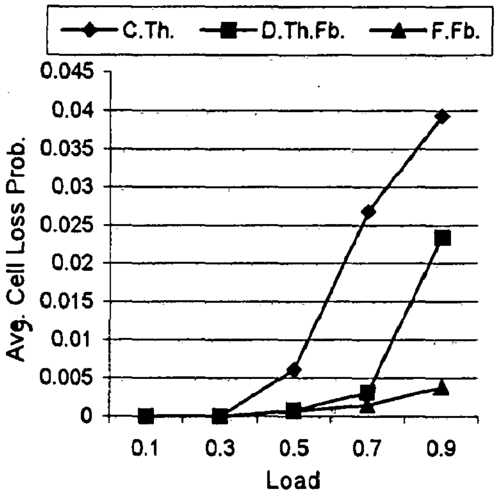


Table 4.6: Avg.Throughput Vs Load, When N=3

Load	*C-Th*	*D-Th-FB*	*F-FB*
0.1	1	1	1
0.3	0.9977	1	1
0.5	0.9861	0.9991	0.9991
0.7	0.9361	0.9968	0.9986
0.9	0.9028	0.9809	0.9962

C.Th.= Constant Threshold
D.Th.Fb.=Dynamic Threshold Feed -back
F.Fb=Fuzzy Feed-back

Fig. 4.6: Avg. Throughput Vs Load for Switch Size 10X10, N=3

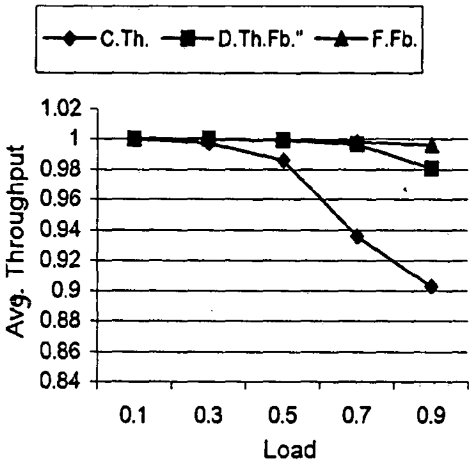


Table 4.7 : Avg. Cell Delay Vs Load, When N=3

Load	*C-Th*	*D-Th-FB*	*F-FB*
0.1	0.0062	0.0119	0.0085
0.3	0.5469	0.0474	0.0311
0.5	0.5673	0.0667	0.0413
0.7	0.7422	0.0987	0.0845
0.9	0.6494	0.3707	0.1124

Fig. 4.7: Avg. Cell Delay Vs. Load for Switch Size10X10, N=3

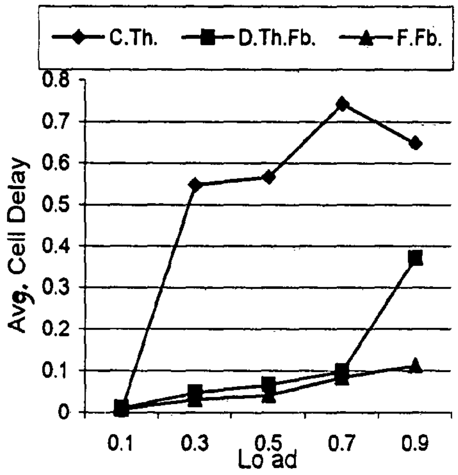


Table4.8:Avg.Cell Loss Prob.Vs Load,WhenN=3

Load	*C-Th*	*D-Th-FB*	*F-FB*
0.1	0	0	0
0.3	0.0023	0	0
0.5	0.0139	0.00086	0.0008
0.7	0.0639	0.00315	0.0014
0.9	0.0972	0.01901	0.0038

Fig. 4.8: Cell Loss Prob. Vs. Load for Switch Size 10 X10, N=3

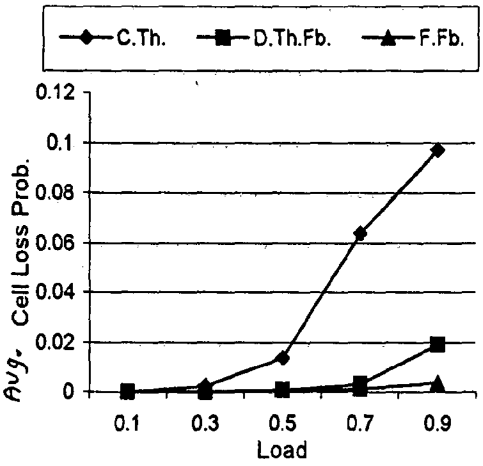


Table 4.9: Avg.Throughput Vs Load, When N=4

Load	*C-Th*	*D-Th-Fb*	*F-Fb*
0.1	1	1	1
0.3	1	1	1
0.5	0.9939	0.9997	0.9997
0.7	0.9733	0.9954	0.9991
0.9	0.9608	0.9973	0.999

C.Th.= Constant Threshold
D.Th.Fb.=Dynamic Threshold Feed -back
F.Fb.=Fuzzy Feed-back

Fig. 4.9: Avg.Throughput Vs Load for Switch Size 10X10, N=4

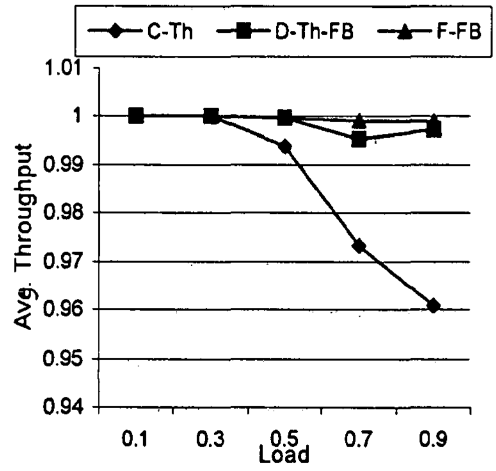


Table 4.10: Avg.Cell Delay Vs Load, When N=4

Load	*C-Th*	*D-Th-Fb*	*F-Fb*
0.1	0.0062	0.0062	0.0045
0.3	0.547	0.01563	0.0142
0.5	0.5673	0.0262	0.0215
0.7	0.7422	0.0763	0.0195
0.9	0.6494	0.0617	0.0188

Fig. 4.10: Avg. Cell Delay Vs. Load for Switch Size 10X10, N=4

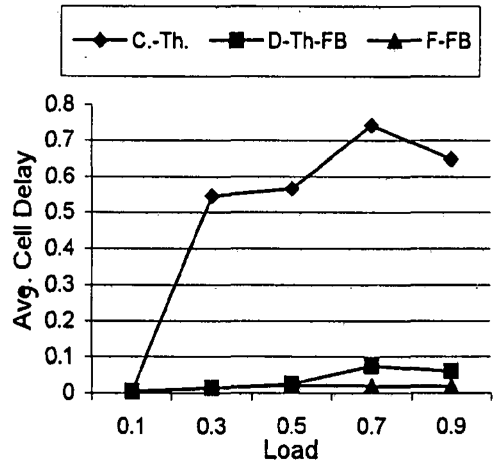
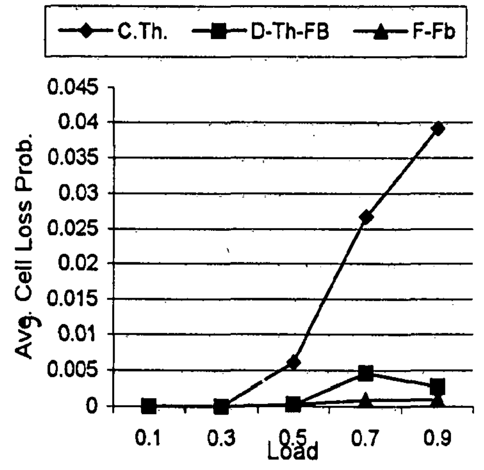


Table4.11:AvgCell Loss ProbVs Load,WhenN=4

Load	*C-Th*	*D-Th-Fb*	*F-Fb*
0.1	0	0	0
0.3	0	0	0
0.5	0.0061	0.0003	0.0003
0.7	0.02674	0.0046	0.0009
0.9	0.03923	0.0027	0.00094

Fig. 4.11: Avg. Cell Loss Prob. Vs Load for Switch Size 10X10, N=4



the cells. The traditional 'fixed' scheme utilizes a binary threshold: admit or no-admit, depending on the occupancy of the buffer. In the proposed Fuzzy Scheme, blocking decision is based on triangular membership function.

4.6 Time Complexity Analysis:

As shown in the previous Section 4.5, the new fuzzy logic based scheme performs very well in comparison to the 'constant' or 'fixed' threshold scheme. However, the price paid, is in terms of increased time complexity. The constant/fixed threshold case has a time complexity of the order of O(l) because it has to be decided only once if incoming burst can be accommodated or not. In this Section an analysis is carried out to determine worst case time complexity of the suggested scheme.

4.6.1 Step 1: Fuzzification

If sample space of buffer occupancy is represented by *n* sets and if burst length sample space is represented by *m* sets, in the worst case, this is bounded by time complexity of O(n) and O(m) respectively.

4.6.2 Step 2: Look-up Table

At the end of fuzzification process, member functions along with membership values are obtained. The set of membership functions can be at the most *m* in case of burst length and at most *n* in the case of buffer occupancy. Now, for each membership function of buffer occupancy, every membership function of burst length is taken and the look-up table is referred and corresponding entry is noted down. Whole of this process takes constant time. As this process is referred for a total *m.n* of times, the time complexity of this step is O(*m.n*).

4.6.3 Step 3: Inference Engine

Assuming that, these *m.n* data elements are clustered in *m* groups, each group having *n* data elements, it can be easily seen that finding out the minimum element in *n* data elements is of time complexity O(n). As this procedure is repeated for *m* times, the total time complexity of finding out minimum is O(m.n). It can be safely assumed that de-fuzzification sample space is represented by not more than l element where $l \le m.n$. This assumption is valid as at the most *m.n* look-up table entries are referred only. So finding out maximum for these entries takes time of the order of O(m.n) again. Total time complexity of this step is thus O(m.n)+O(m.n).

4.6.4 Step 4: Defuzzification

Here, the fuzzy sets passed on along with the membership values are defuzzified to determine the crisp output value. Without losing generality, it can be assumed $n \ge m$. In this case. we can say that at the most *m.n* membership functions along with the values are passed on from step 3. For each function, respective de-fuzzification value is retried and multiplying with associated weight takes a constant time. So worst case time complexity is O(*m.n*). Addition of all the elements after this has time complexity of O(*m.n*) and divided by addition of respective weight (complexity of O(*m.n*) again). Division takes constant time out of time.

So the total time complexity of this step (O(m.n) +O(m.n) +O(m.n)).

4.6.5 Total Time Complexity

Total time complexity of fuzzy logic is obtained by adding the individual time complexities as:

O(m) +O(n) +O(m.n) +O(m.n) +O(m.n) +O(m.n) +O(m.n) +O(m.n)

The value turns out to be of time complexity of O(m.n), again for n>=m the time complexity of the new proposed scheme can be expressed as $O(n^2)$.

Thus proposed scheme can be easily implemented for small value of n. For large 'n' time complexity may become a liability for the scheme.

4.7 Conclusions:

In this chapter, we have introduced fuzzy approach to control congestion in ATM networks. When a number of bursty traffic sources add cells, the network is inevitably subject to congestion. Various traditional approaches to congestion management reported in the literature, utilize 'fixed' threshold, i.e., either binary or a limited number of predetermined values based on the cell priorities, to determine when to permit or refuse entry of cells into the buffer. The aim is to achieve a desired tradeoff between the number of cells carried through the network, propagation delay of the cells, and the number of discarded cells. Conventional thresholds suffer from some fundamental limitations. One of the limitations is the difficulty of obtaining complete statistics on input traffic to a network. As a result, it is not easy to accurately determine the equivalent capacity or effective thresholds for multimedia high-speed networks in various bursty traffic flow conditions. Besides, these approaches/schemes provide optimal solutions only under a steady state. A control scheme that dynamically regulates traffic flow according to changing network conditions, however requires understanding of network dynamics. To minimize congestion, for a gradual change we proposed fuzzy approach. In our scheme, burst length as well as buffer occupancy are represented by triangular membership functions of fuzzy sets. From these functions; the degree of membership, associated with burst length and buffer occupancy for a particular set, can be read. From these membership values and corresponding sets, blocking to be offered, again in fuzzy terms can be find out. Then, by applying suitable defuzzification method, the percentage blocking to offered at that particular buffer occupancy level and at given burst length can be determined. A comparative study has revealed the proposed scheme is able to achieve lower average delay and higher throughput than the constant as well as dynamic feed-back threshold schemes and that too with lower cell loss probability. However, these improvements are achieved at the cost of higher time complexity.

Chapter-5

Congestion Control for Multicast Bursty Traffic in Shared Memory ATM Switch

5.1 Introduction:

In recent years, large technological progress has occurred both in the field of electronics as well as in the field of optics. This progress will allow the economical development of new communication networks running at very high speeds. An explosion in demand for enterprise–wide connectivity is now taking place and certain to accelerate in coming years. It is becoming increasingly difficult for network managers to maintain cost efficiency and respond fully to customer requirements. No single technology can reserve inter working integration and network simplification issues. It seems certain that ATM will play central strategic role in the network of future.

Recently, a large number of switching architectures has been proposed [81], [95], [107], and [162]. All the approaches point to the need of a very high speed hardware switches because of the high transfer rates involved; on the other hand, due to the statistical multiplexing, buffering is also required in order to avoid packet loss whenever there are multiple input packets arriving simultaneously on different inputs ports and destined for the same output. Only one packet at a time can be transmitted over an output link, the rest must be temporarily stored in a buffer for later transmission.

The management of multicast traffic in shared memory ATM switches is of particular interest. Multicast mechanisms provide group communications by reducing the amount of duplicate traffic in the network to conserve bandwidth and switch resources [88], [124], [148], and [177]. Many envisioned applications in ATM Networks are multicast in nature and are expected to generate a significant portion of the total traffic. Examples of such application are broadcast video, video conferencing, multi party telephony and work group applications. While using a shared memory switch, a copy network[97] can also be used for replication of multicast cells, and then the copies of multi-cast cells can be stored in their respective logical queues in a shared memory ATM switch for point-to-point switching. The proper placement and arrangement of the buffering system have a dramatic impact on the switch performance[45]. Buffer management refers to the discarding policy for the input of cells into the buffers and scheduling policy for the output of cells from the buffers. These functions are a component of the traffic control functions handled by the switch management. Inside the switch fabric, the queue need to be monitored for signs of congestion to alter the switch management and attempt to control the congestion. Congestion control is an important phenomena, because ATM is connection-oriented and support real-time service[25]. The concept of threshold helps in maintaining a fine balance between the number of cells propagated and the average delay observed by a typical cell in a switch [17], [31], [32], [34], and [87]. We have studied and simulated [54], [91], and [146] the effect of various threshold schemes to control congestion for multicast bursty traffic in shared memory ATM switch, and proposed a fuzzy scheme for the purpose [147].

5.2 ATM Switching System Architecture:

ATM has distinct advantages over other switching architecture because of its switched structure and scalability. In ATM every slice of bandwidth can be used for creating point to point connections. The overall structure of the ATM switch shown in Fig. 5.1.

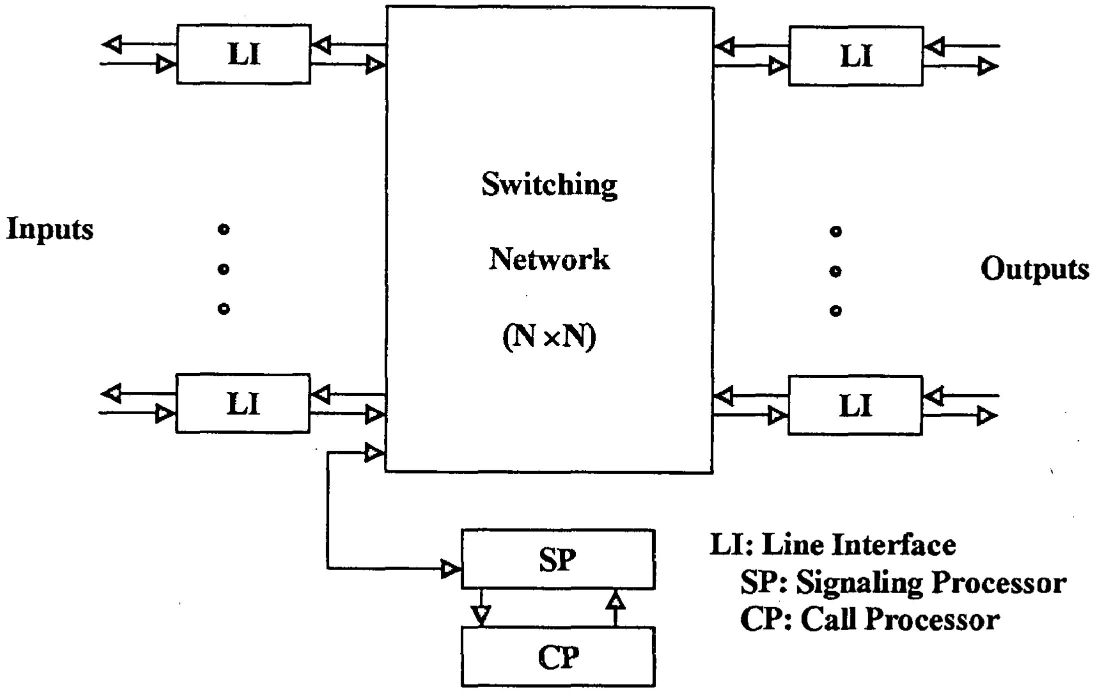


Fig. 5.1: Overall Structure of ATM Switch

The Line Interface(LI) performs optical-electrical signal conversion cell synchronization, header Translation, and insertion and extraction of routing information. The Call Processor (CP) and the Signaling Processor (SP) are concern with the ATM connection setup and release. The switch module routes cells using routing information of the cells.

In this chapter, only transport port of switch is modified for supporting multicast traffic and control part is assumed to have a multicast routing table with the normal VCI/VPI translation table. Transport is defined as all physical means which are responsible for the correct transportation of the cell from an inlet to an ATM outlet, within the QoS of ATM. The control part of the switch is that controls the transport network. It decides which inlet to connect which outlet and accept or discard connection. The decision is based on incoming signaling cell or semi-permanent connection. More details can be found in [26], [75], [93], and [97].

5.2.1 Switching Network:

In an ATM switch, the ATM cells have to be transported from an inlet to one (point-to-point) or more (multicast) outlets. This switching can be combined with concentrator, expansion, multiplexing and demultiplexing of ATM traffic. In switching element there is no co-ordination

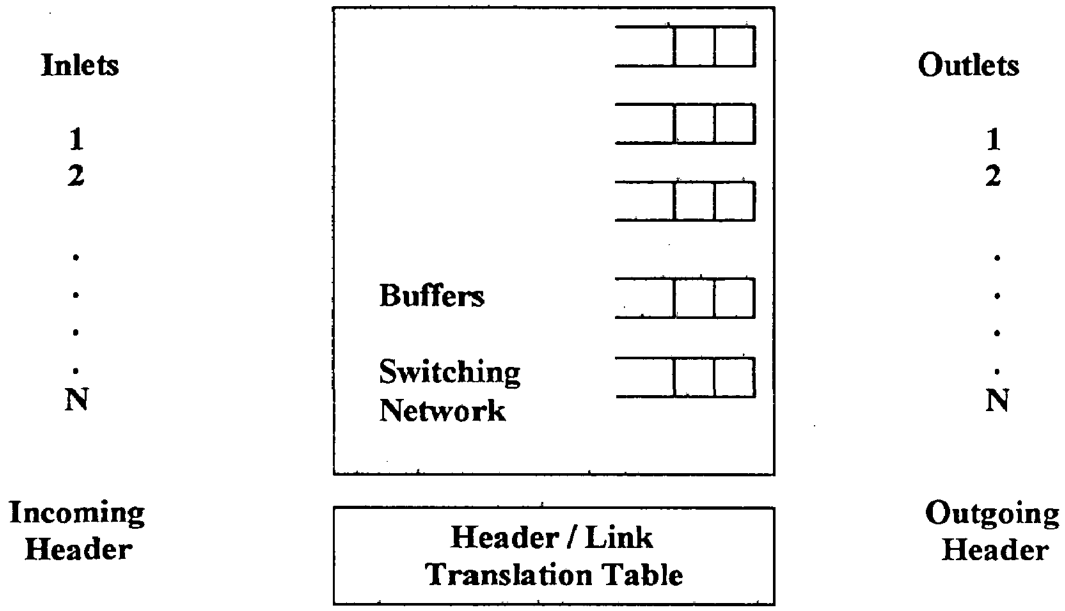


Fig. 5.2: Switching Network of ATM Switch

among arrival cells as far as their destinations request are concerned. Thus more then one cell arriving in the same time slot may be destined to go to the same output port. This event raises cell conflict so buffering of cells within the switch is provided as shown in Fig. 5.2.

5.3 Shared Memory ATM Switch:

One of the most promising solutions of ATM switches is based on the shared-memory principle. Such switches play a leading role in practical, experimental implementations of ATM. A shared-memory ATM switch provides sharing of memory space among its switches ports and superior cell loss rate performance compared to input-buffer-based and output-buffer-based ATM switches

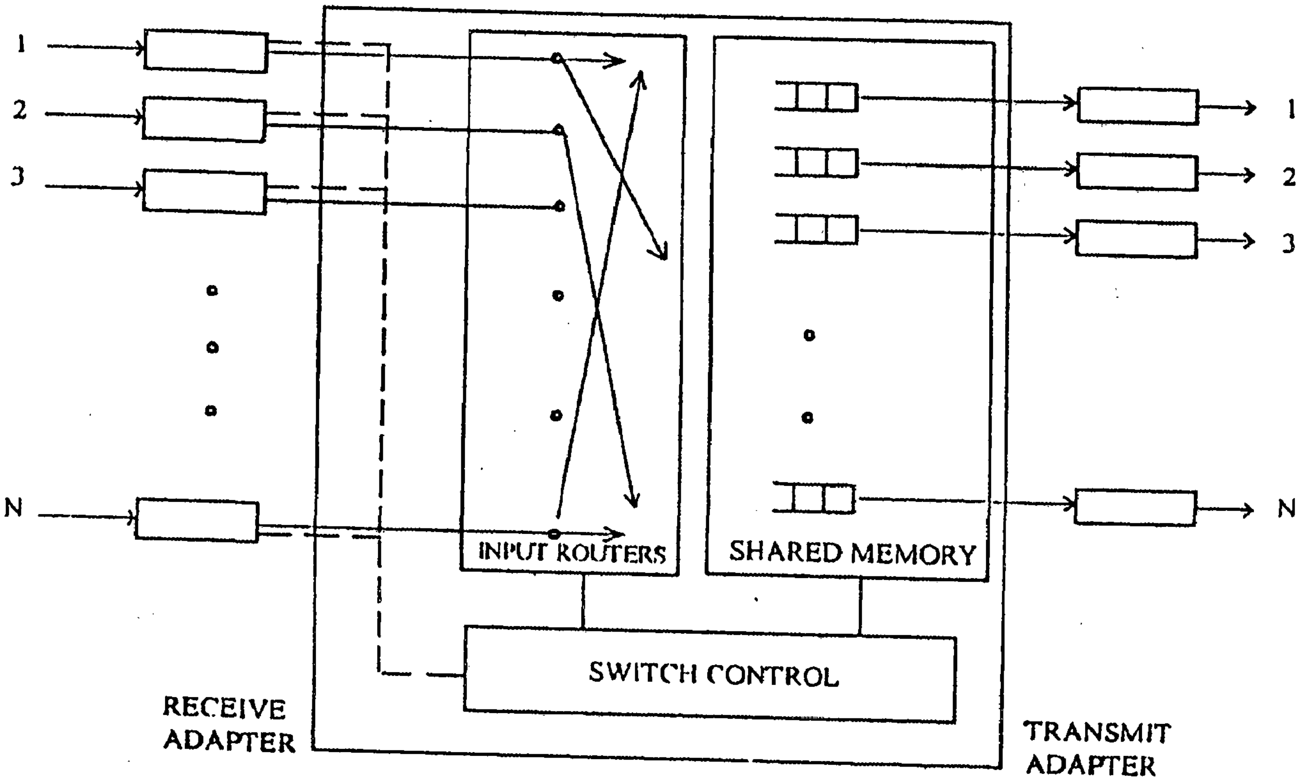


Fig. 5.3: Shared Memory Switching Element

under conditions of identical memory size. Shared memory ATM switch consist of a single dual-port memory shared by all input and output line as shown in Fig. 5.3. Cells arriving on all input lines are multiplexed into a single stream which is fed to the common storage. Internally to the memory, cells are organized into separate output queues. Simultaneously an output stream of cells is performed by retrieving packets from the output queues sequentially.

The N input ports are connected to the receive adapters and the N output ports are connected to the transmit adapters. A 1 to N router for each input provides full contention-free connectivity to all output ports. The routing is determined by the destination address in the cell header. Each output is a logical queue with N inputs and one output. All output queues are located in a block of shared memory where the space allocated to a specific output varies dynamically with the load of the switch i.e. the memory is full shared among all queues. In-fact switch output queues are implemented through queues of pointers, with each pointer addressing a cell in the shared memory. This allows better memory utilization then fixed memory allocated to each output queue.

For reasons of fairness, the maximum share of memory that can be allocated to each output must be limited to prevent a temporarily overloaded connections from using the entire memory space and degrading the throughput of other connection to other outputs. This policy is implemented in the control section of the switch. In shared memory ATM switches two main constraints must be satisfied:

- The processing time required for determining where to queue up the cells and for issuing the proper signals for the purpose should be sufficiently small to keep up with the flow of incoming cells.
- There is limitation on the memory access speed. Thus the size of the switch is determined by the available memory speeds and achievable processing speeds.

5.4 Multicast and its Requirements for ATM :

Multicast can be simply defined as the ability to send one message to one or more in a single operation. This is different than using replicated unicast which sends messages from one node to a group of nodes by sending to each node individually. This will incur one operation for each destination node and is non-atomic. A 1 : N multicast allows one source to reach N destinations. An M:N multicast allows M sources to reach N destinations.

ATM imposes specific requirements that must be considered for the design of the multicast services. The short, fixed length of the ATM cell requires an Adaptation Layer at the end points to transfer complete message and guarantee that cells from different sources are not interleaved by intermediate switches. The overhead of call setup and tear-down for the connection-based ATM protocol require an efficient mechanism for adding and removing users to a multicast group. The limited size of the VPI/VCI field prevents the use of source-based routing, as used by the existing IP multicast routing programs[56].

Connection oriented protocols can incur a high overhead when setting up and destroying links, therefore connectionless links are preferable. Some applications that require multicasting such as audio would prefer a CBR, where as video is less susceptible to the jitters. This is due to a human's high sensitivity to audio jitters.

5.5 Issues on ATM Multicasting:

Firstly the issue of the fan-out (1 to N) copy of cells will have some associated state data. This data has to be put in each cell or in the switch itself. The cell structure is incapable of storing this information. The most efficient place to store the fan-out is at each switch, associated with the VCI/VPI information. Thus cells can be replicated on a per connection basis and under the control of the connection setup mechanism provided in ATM network.

Secondly, some means of identifying copies of a cell has to be provided to unable subsequent switches and the final destinations to be de-multiplexed. This brings up the issue of separating the replication function from the routing function. The following issues must also be taken into consideration when designing an extended ATM switch.

Delay- important for interactive applications

Jitter - critical for audio and video

Capacity- contention for bandwidth will cause delays

Reliability- of both switch and media will affect error recovery

Priority- CLP bit indicates primitive priority in ATM

5.6 Multicast Support for Shared Memory ATM Switch:

Multicast operations are important in supporting point-to-multipoint communications. There are various ways of supporting multicast operation with shared memory architecture. Replication-at-Receiving (RAR) is one of the scheme to support such operation, which has been used for the purpose. In RAR scheme, a multicast cell arriving at the switch and destined to 'm' destinations is first copied 'm' times as represented in Fig. 5.4. A copy of the cell is linked to each output queue to which the multicast cell is destined[84]. All copies are stored in the buffer then each copy is served independently. So this scheme is also known as Multiple Write Multiple Read (MWMR) scheme.

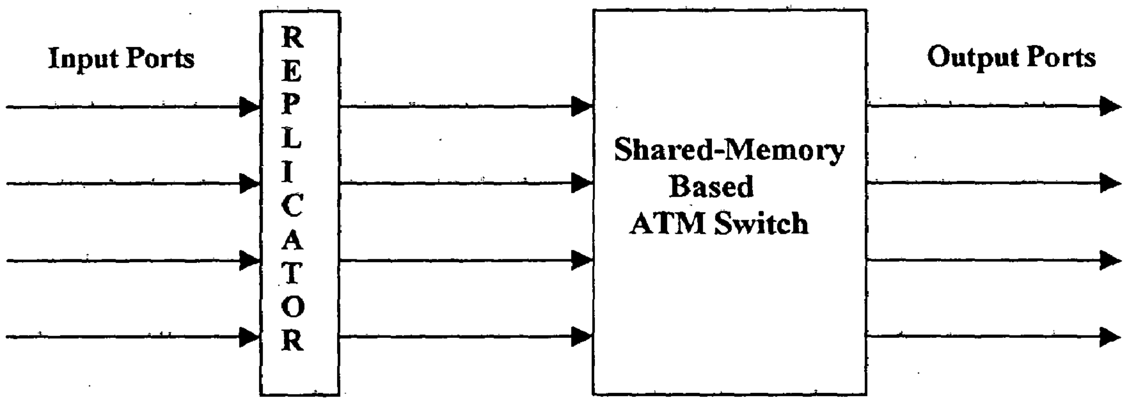


Fig. 5.4: RAR Multicast Scheme

The scheme has been used in several existing shared memory switches because it is relatively simple to implement. While using a shared memory switch, a copy network [97] can also be used for replication of multicast cells, and then the copies of multi-cast cells can be stored in their respective logical queues in a shared memory ATM switch for point-to-point switching. In fact once a cell has been replicated each copy of the cell can be treated in a same way as a uni-cast cell. Consequently, both the control and structure of the linked-lists are basically the same as those used in a uni-cast switch.

The advantage of this scheme is that, this is a straightforward solution to providing multicast support. This scheme is a fair scheme in the sense that copies of a multicast cell destined to a loaded output are more likely to be dropped, and an idle output port would always get its cell if the copies are present in the memory.

5.6.1 Multicast Traffic Model :

Let X be the random variable denoting the number of copies associated with an incoming multicast cell. Let P(X= r) be the probability that the number of copies associated with an incoming multicast cell is r. If f is a fraction of incoming bursts consisting of multicast cells with a geometric fanout, then in such a case,

Effective offered load = [(1-f).L] + [f.L.E(X)]

If X has a truncated geometric distribution g(X) with the parameter q, then:

$$g(k) = P_r[X = k] = \frac{(1-q)q^{k-1}}{1-q^n}; 1 \le k \le n$$

and the average number of copies per multicast cell

$$E(X) = \left[\frac{1}{1-q} - \frac{nq^n}{1-q^n}\right]$$

Thus, the effective offered load can be given as follows:

$$[1-f]\left[\frac{\beta_i}{\alpha_i+\beta_i-\alpha_i\beta_i}\right] + f\left[\frac{\beta_i}{\alpha_i+\beta_i-\alpha_i\beta_i}\right] \cdot \left[\frac{1}{1-q} - \frac{nq^n}{1-q^n}\right]$$

where n is the switch size.

5.7 Various Applied Threshold Schemes:

Most of the ATM switch architectures that have been proposed in the literature use some buffering to accommodate packets whose service has been delayed due to contention for some resource within the switch. The location of these buffers and the threshold scheme directly affect the performance of such a switch. Various threshold schemes, which are applied to control congestion in multicast bursty traffic in shared memory ATM switch are as follows:

5.7.1 Static Threshold (ST) Scheme:

The maximum or minimum amount of buffering that should be available to any individual queue, called Static Threshold (ST). In this method[32], an arriving cell is admitted only if the queue length at its destination output port is smaller than a given threshold. It requires only queue length counters, which is likely to be required for network management purposes anyway, and a comparator.

ST scheme can be extended for multiple services classes-one can simply use different thresholds for different classes. The performance of this method suffers from the fact that it is not adaptive. When so many queues are active at once that the sum of their thresholds exceeds the buffer capacity, then it is possible for the buffer to fill up completely even though all queues are obeying their threshold constraints. This allows some queues to become starved for space, which can lead to under utilization of the switch. At other times, when very few output queues are active, these queues are needlessly denied access to the idle buffer space beyond the sum of their thresholds. This creates higher cell loss rate and lower throughputs for these active queues than they would experience if they had access to extra buffer space.

5.7.2 Dynamic Threshold (DT) Scheme:

An efficient queue management-strategy is of prime importance for an output-queued shared memory ATM switch in order to maintain a high throughput under a variety of load condition[32]. Dynamic Threshold is a scheme to fairly regulate the sharing of memory among different output queue for traffic of a single loss priority. This scheme is conceptually similar to bandwidth balancing in Distribution Queue Dual Bus (DQDB) network and to bottleneck flow control [70]. The DT scheme deliberately wastes a small amount of buffer space, but attempt to equally share the remaining buffer space among the active output queues.

Each output queue attempts to limit its length to some function *f* of the unused buffer space; output queues with less demand than this can have all the space they wish. At time *t*, let *T(t)* be the control threshold and let $Q^i(t)$ be the length of queue *i*. Let $Q(t)$ be the sum of all of queue lengths, i.e., the total occupancy of the shared memory. If B is the total buffer space, then

$$T(t) = f(B - Q(t)) = f(B - \sum_i Q^i(t)) \qquad (1)$$

An arriving cell for queue *i* will be blocked at time *t* if $Q^i(t) \geq T(t)$. All cells going to this queue will be blocked until the queue length drains below the control threshold and/or the threshold rises above the queue length[7]. The simplest method is to set the control threshold to a multiple α of the unused buffer space

$$T(t) = \alpha . (\mathrm{B} - \mathrm{Q(t)}) = \alpha .(B - \sum_i Q^i(t)). \qquad (2)$$

If α is a power of two (either positive or negative), then the threshold computation is extremely easy to implement.

The DT scheme adapts to changes in traffic conditions. Whenever the load changes, the system will go through a transient. For example, when a lightly loaded output port suddenly becomes very active, its queue will grow, the total buffer occupancy will increase, the control threshold will decrease, and queues exceeding the threshold will have their arrivals blocked temporarily while they drain, freeing up more cell buffers for the newly active queue.

When referring to steady-state values, we simply drop the time parameter *t* from the variable names. For instance, Q denotes the steady-state value of $Q(t)$ and T denotes the steady-state value of $T(t)$. If there are S very active queues, then the total buffer occupancy Q in steady- state will be

$$\mathrm{Q} = \mathrm{S.T} + \Omega \qquad (3)$$

Where Ω is the space occupied by the uncontrolled queues, i.e., those with lengths below the control threshold. The steady-state length of each controlled queue can be found by substituting (3) into (2) and solving for *T*.

$$Q^i = T = \alpha(\mathrm{B} - \Omega) / (1+\alpha \mathrm{S}) . \qquad (4)$$

The amount of memory held in reserve by the algorithm will be $(\mathrm{B} - \Omega) / (1+\alpha\mathrm{S})$. If $\alpha=2$ and S=10, for instance, then each of the ten queues take 2/21 of the buffer, and 1/21 is left unallocated.

5.8 Proposed Fuzzy Threshold (FT) Scheme:

In ATM shared memory switch, it has the total buffer space B. This buffer space is divided into the number of output line. We use 64 X 64 ATM switch, i.e. buffer space distributed 64 line space. It is to be calculated buffer line occupancies and total buffer occupancy. A binary threshold is considered in each line space. This binary threshold divides the line space into two parts. Below or equal to the threshold level, every arriving cell is given entry to the ATM switch, threshold levels in each line will be change upto maximum line capacity. If total arriving cells in each line is large to maximum line capacity then cells will be lost.

In proposed Fuzzy Threshold Scheme[147], wherein line occupancy as well as buffer occupancy are represented by fuzzy sets, and these are represented by triangular functions as shown in Fig.5.5 (a),5.5(b).

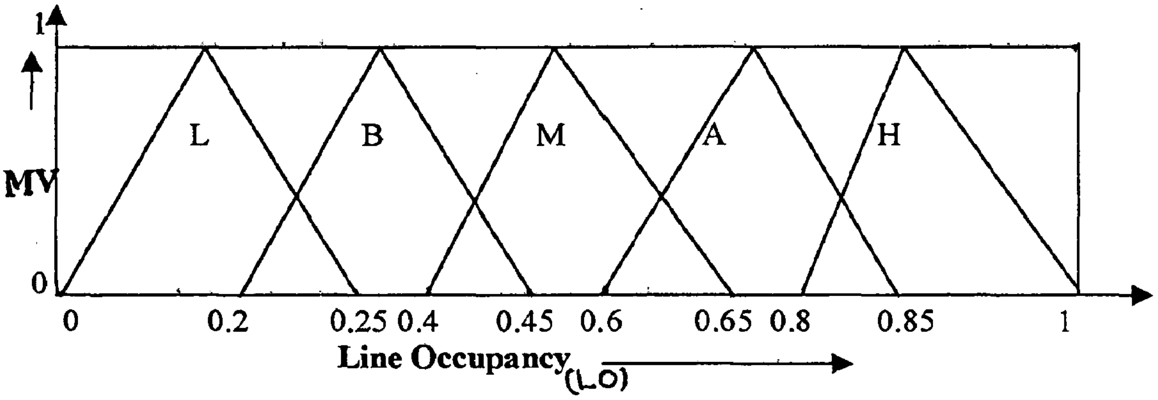


Fig. 5.5(a): Triangular Functions for Line Occupancy

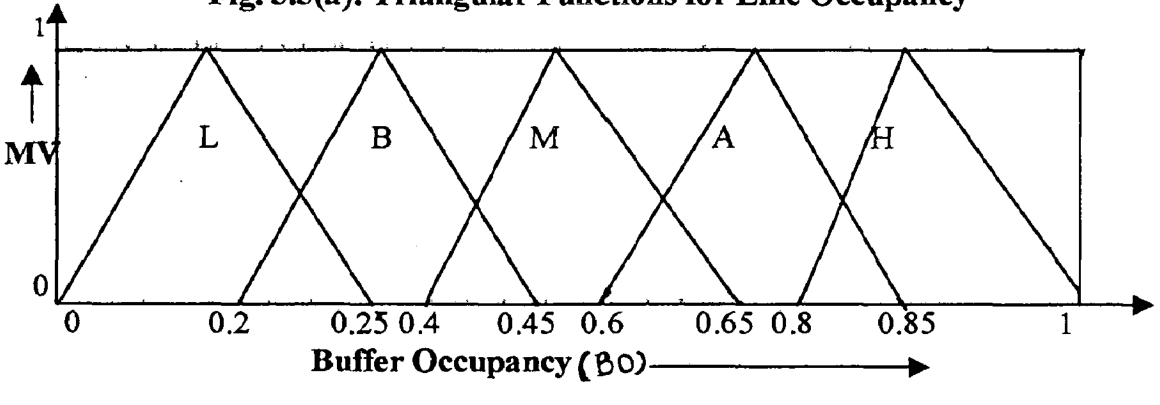


Fig. 5.5(b): Triangular Functions for Buffer Occupancy

In the above figures L, B, M, A and H represent Low, Below medium, Medium, Above medium and High membership sets respectively, M.V. represents membership value. From these, the degree of membership, to particular set, associated with each valid buffer occupancy and line occupancy, can be read. This quantification of membership is called fuzzification.

From these membership values and corresponding sets allocating space to be offered again in fuzzy term can be find out. This process is called rule based inference. As an example, a typical rule is when line occupancy is high and buffer occupancy is high, threshold value is high, i.e., allocation space will be high as shown in Lookup Tables5.1(a), and (b).

Then, by applying suitable defuzzification method, such as min-max or max-min defuzzification, the percentage allocation space to offered at that particular line occupancy and at given buffer occupancy can be determined. For defuzzification, with the set such as shown in the following Table 5.2 weighted average is used.

BO LO	**L**	**B**	**M**	**A**	**H**
L	L	L	L	B	M
B	L	B	B	M	M
M	B	M	M	A	A
A	M	A	A	A	H
H	A	A	H	H	H

Space Allocated

Table 5.1(a) : Lookup Table

BO LO	**L**	**B**	**M**	**A**	**H**
L	L	L	B	B	B
B	B	B	B	M	M
M	M	M	A	A	A
A	M	M	A	A	H
H	A	A	H	H	H

Space Allocated

Table 5.1(b): Lookup Table

Set	L	B	M	A	H
I	0.2	0.4	0.6	0.8	1
II	0.4	0.5	0.7	0.9	1
III	0.5	0.6	0.8	0.9	1
IV	0.3	0.7	0.8	0.9	1

Table 5.2: Defuzzification Table

L: Low Set, B: Below Medium Set, M: Medium Set, A: Above Medium Set, H: High Set.

For example, it is assumed that line occupancy as well as the buffer occupancy both are characterized by the fuzzy sets. Assume maximum line capacity is kept at 40 and total buffer space is assumed 200. Suppose, at the time of a new arriving cells at particular queue of buffer, line occupancy =20 and buffer occupancy =48. When normalized with respect to maximum value, these variables are mapped as line occupancy =0.5 and buffer occupancy = 0.24. Using Fig. 5.5(a) and 5.5(b) for fuzzification it is seen that line occupancy is a member of set M with associated value 0.8. Buffer occupancy of set B with associated value 0.32 and a member of set L with associate value 0.08. Using Lookup Table 5.1(a) and min-max method of evaluation, we get :

Line occupancy M(0.8) and buffer occupancy B(0.32) => Allocated value for threshold M(0.32)

Line occupancy M(0.8) and buffer occupancy L(0.08) => Allocated value for threshold B(0.08)

Thus, taking maximum of the two values associated with same membership of set, allocating value for threshold has membership of set M with value 0.32 and membership of set B with value 0.08 using the set with weighted average of membership values. Percentage of allocated value for each line in ATM switch for defuzzification, set I of Table 5.2 is used.

Allocated Space (in %) = [(0.32*0.6) + (0.08*0.4)]/ (0.32 + 0.08) = 0.56%

and we multiply with maximum value of line capacity and calculate the threshold level 0.56 × 40 = 22.4 (i.e. 22), where 40 is the maximum line capacity of each queue. This fuzzy scheme is applied for total number of line N where we used N × N ATM shared memory switch.

5.9 Simulation Process:

For ATM multicast, the parameters of interest are queuing delay, throughput and cell loss ratio. If there is no multicasting, all the cells are switched to the buffer corresponding to their destination field. But in multicasting case, one cell may be duplicated (not necessarily) and therefore can occupy more than one buffer space. The general simulation process for multicasting in ATM switching can be described as below and represented in the form of flow-chart in Fig. 5.6:

5.9.1 Generate Input:

The first step is to generate the traffic for the switch and generate the output ports for the cells randomly. Cells are marked either unicast or multicast type according to combination required.

5.9.2 Setup Load and Buffer:

The performance analysis is to be done with varying loads and multicast combinations. Load is varied by varying OFF periods of the traffic. The buffer size can be varied for performance but in this case, it is taken as 2048 cells space for the switch.

5.9.3 Initialize Parameters:

All the out put parameters i.e. cell loss, unicast and multicast cell delay and throughput are initialized for all the ports and all simulation runs.

5.9.4 Initialize Sources and Buffer:

All the VBR sources are initialized. For VBR sources, ON/OFF state are taken according to load and random number generator. The buffers are also initialized. At each input port for each time slot do the following:

- **Switching the Cells Arriving:**

All the cells arriving at the input port are switched to corresponding output buffer for unicast cells. For multicast cells duplication is done (not in all the cases) before forwarding to destinations. If there is no cell, wait till next time slot.

- **Manage Buffer:**

If there is any cell in the output register of switching network, place it in buffer. If buffer is full, the cell will be lost. If buffer is not full, then apply various threshold schemes i.e. ST, DT, and proposed Fuzzy Threshold scheme as shown in Fig. 5.7(a), (b), and (c) respectively.

- **Increment Delay:**

The minimum service time is one slot-time, so after each slot, number of simulation time slot counter is increasing by one. Hence delay is automatically increase by one for all the cells (delay = simulation counter-cell generation time).

- **Process Buffer**

At the end of each time slot, a cell is taken from each buffer and put into the sink for calculating the performance parameters and then free and number of cells field for buffer is updated.

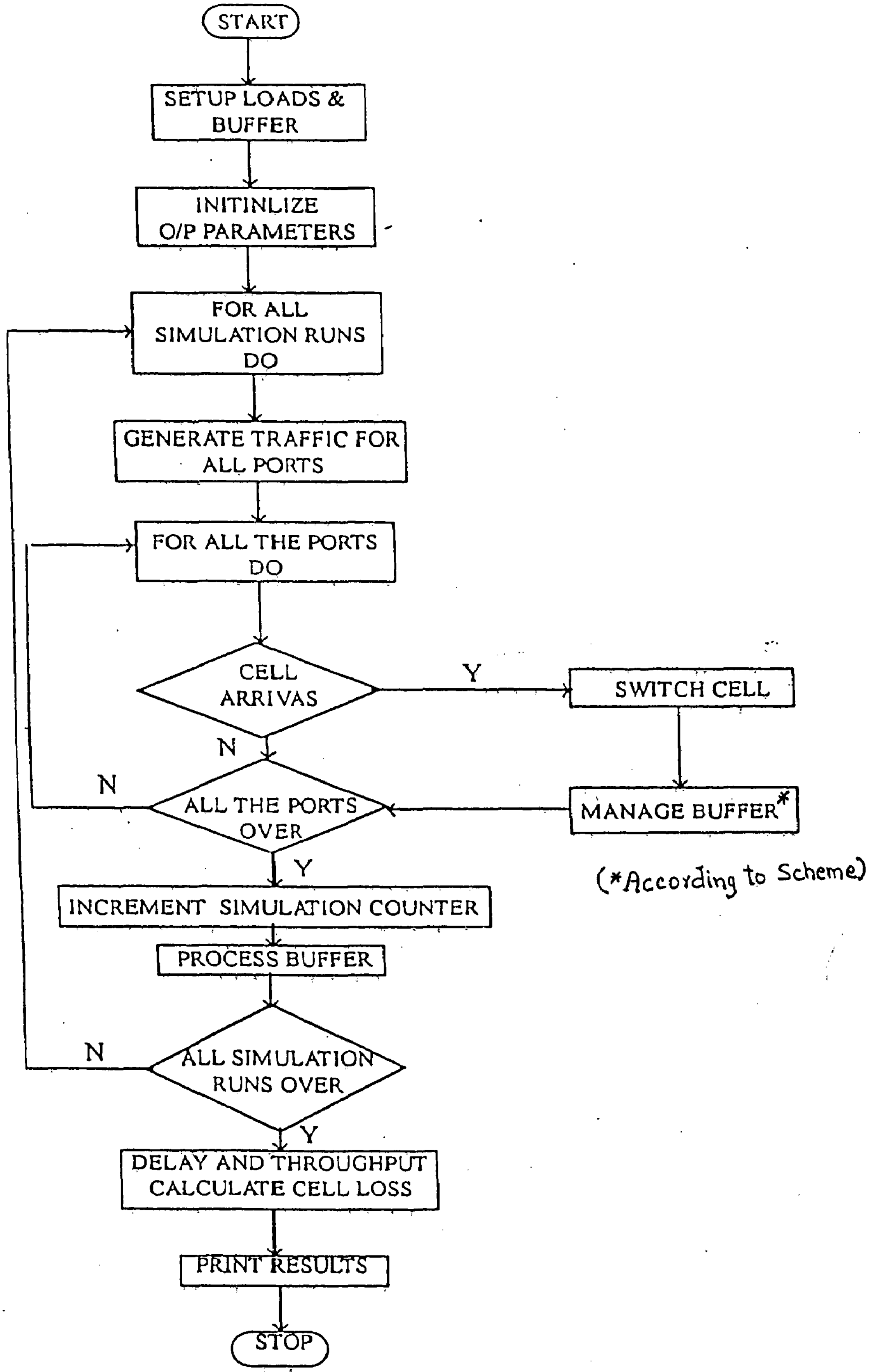


Fig. 5.6: Flow Chart for Simulation Process

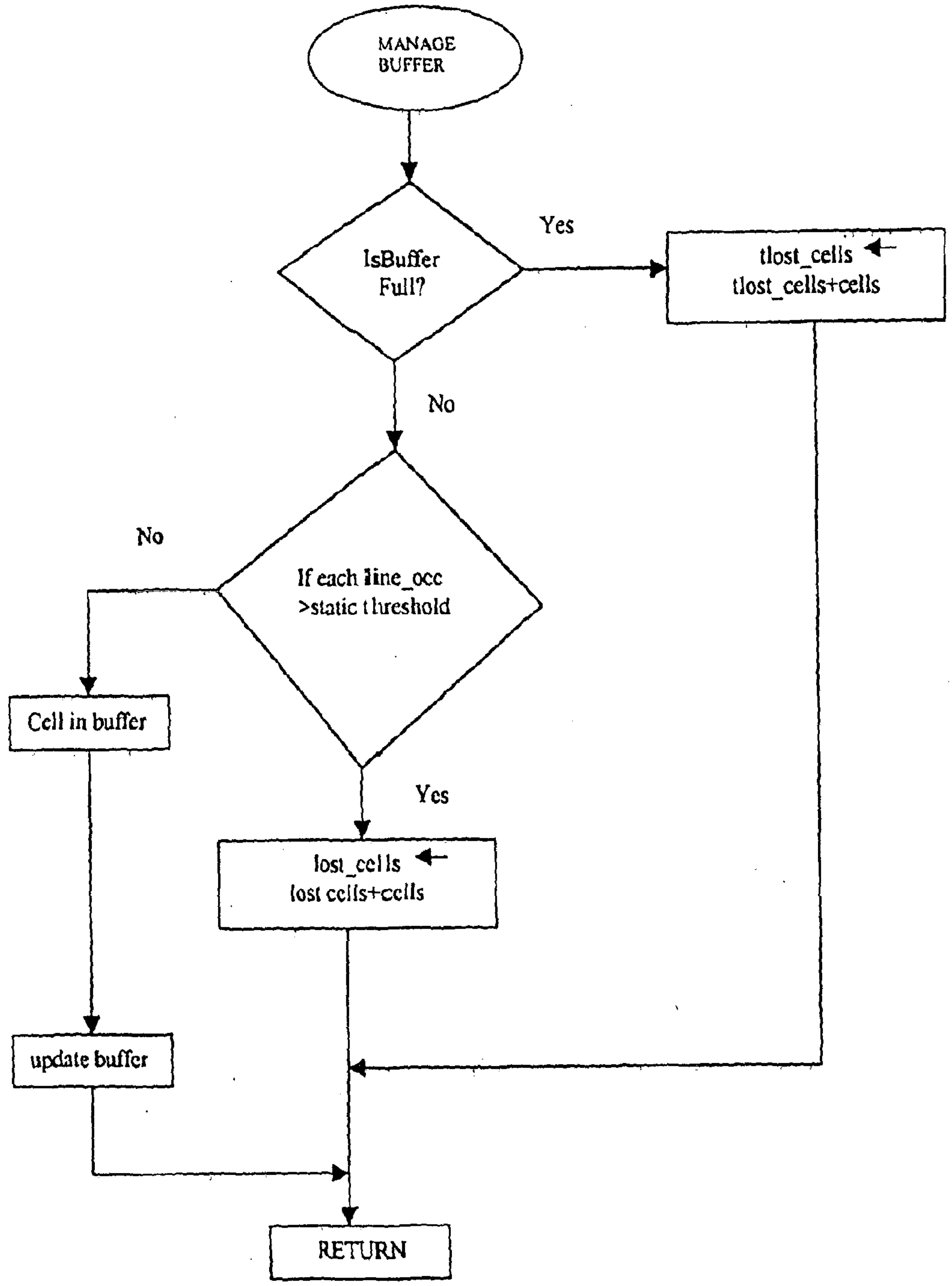


Fig. 5.7(a): Flow-Chart for Manage Buffer of Static Threshold Scheme

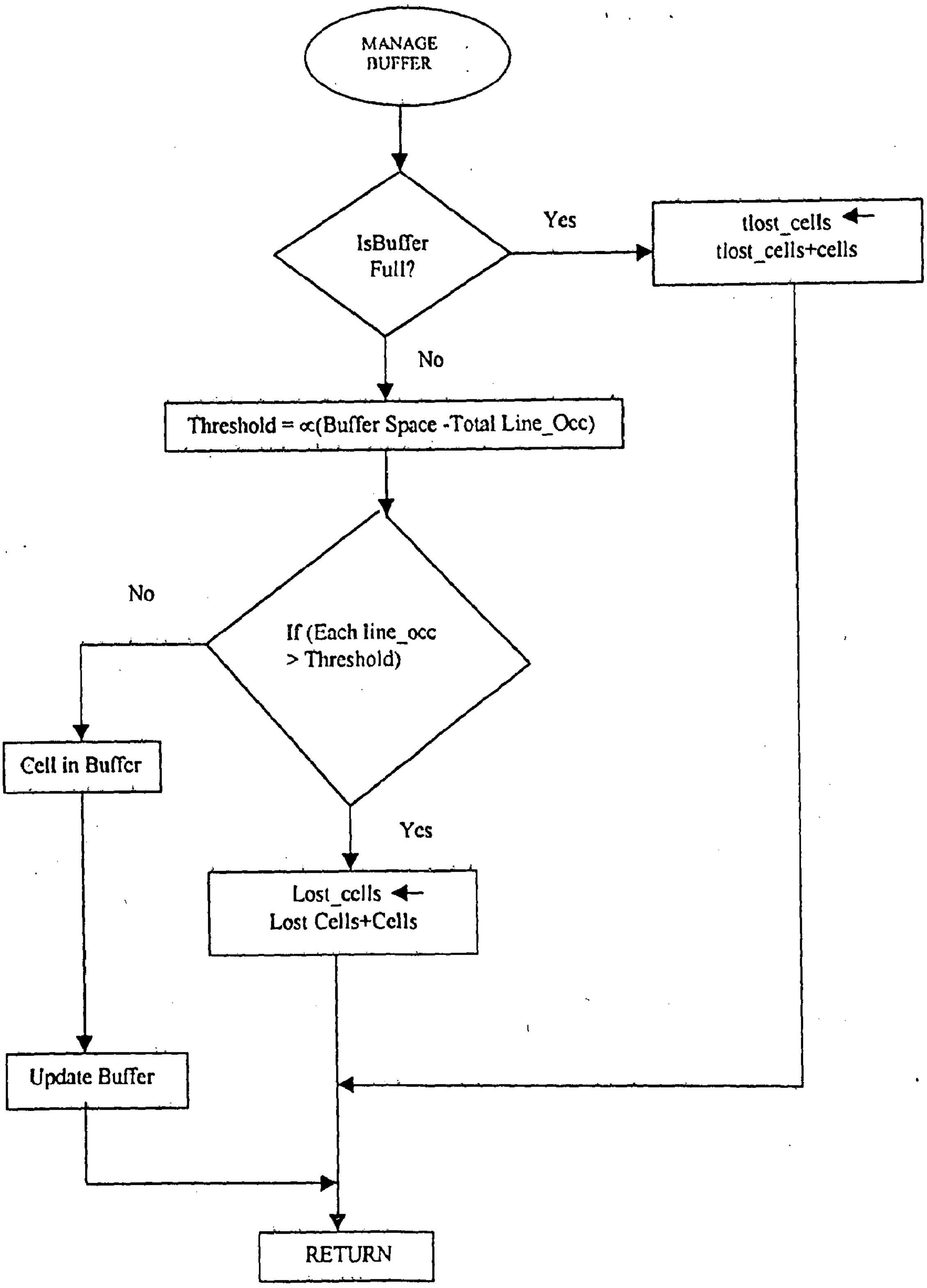


Fig. 5.7(b): Flow-Chart for Manage Buffer of Dynamic Threshold Scheme

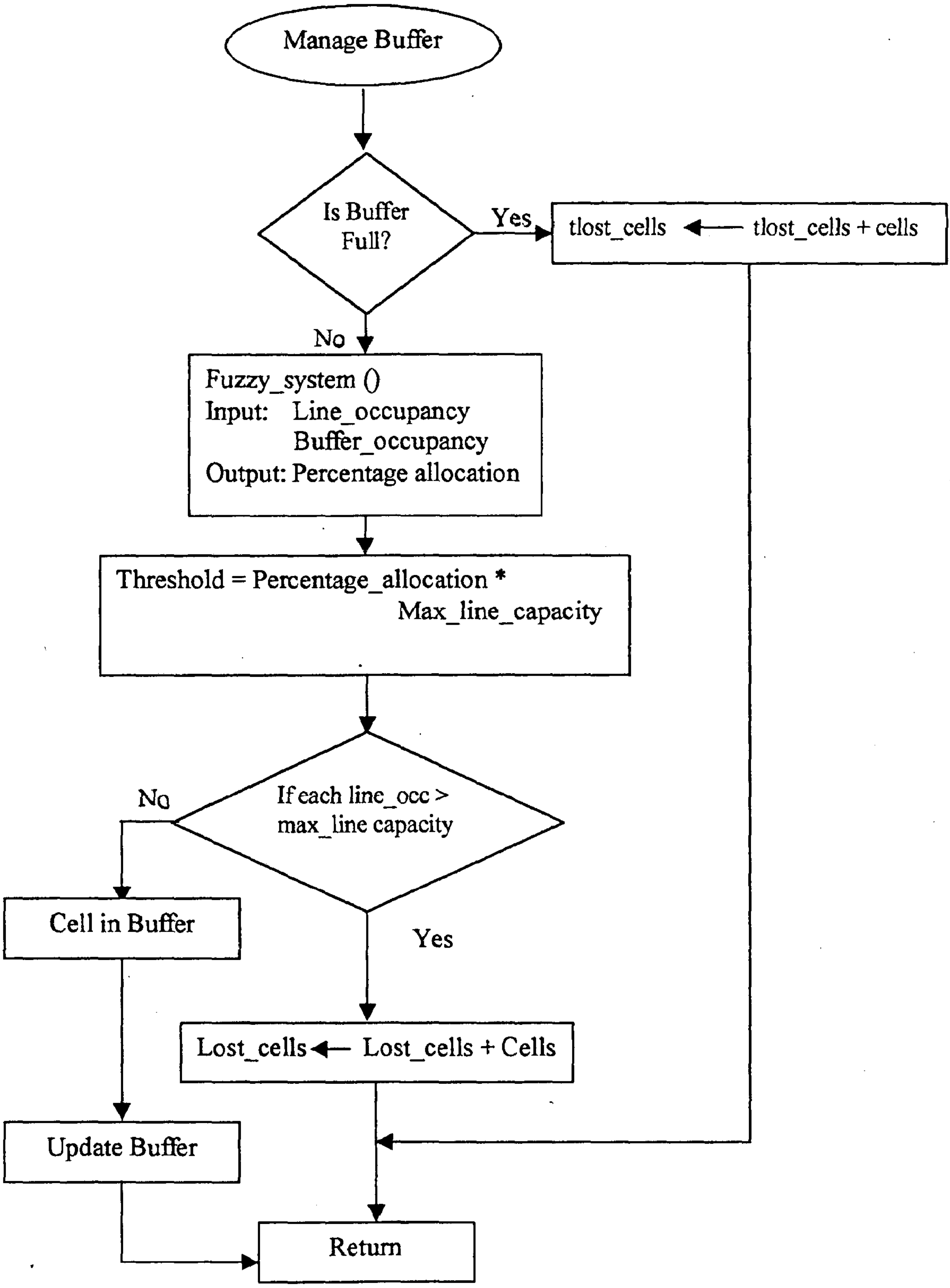


Fig. 5.7(c): Flow-Chart for Manage Buffer of Fuzzy Threshold Scheme

5.10 Results:

For modeling of an ATM switch, discrete-time model is used. Switch is N X N ATM switch, which uses output queueing and shared memory, i.e., each output port of the fabric has a logical queue, but these queues all share the same fabric memory. Time axis is slotted with slot size equal to a cell transmission time on the output link. In onetime slot, only one cell can be switched and other cells are put in input queue. Cells arrivals and transmission are considered to occur at slot boundaries. Hence the minimum time required to serve a cell is one time slot. The standardized ATM cell size of 53 bytes leads to a transmission time 2.827 μs, at a transmission rate of 155.5Mbps. The switch fabric randomly selects a maximum of $1 \leq L \leq N$ cells with the same output address and places them at their output buffer in one slot. A cell leaves its queue after receiving service of one time slot duration. Cells are served according to FIFO strategy. A cell arriving at a queue which is full will be lost.

We first consider a nominal switch configuration with dimensions 64 X 64, buffer size (B) = 2048 cells. Simulation results are taken for three burst lengths i.e. for BL = 20, BL = 30 and BL = 40. We compared the performance of proposed Fuzzy Threshold Scheme with Static Threshold (S.T.), Dynamic Threshold (D.T.) schemes. The performance of switch also evaluated when the multicast bursty traffic is under Without Threshold (W.T.). The value of constant Static Threshold is taken equal to 30, like Static Threshold the Dynamic Threshold has a single parameter, i.e., proportionality constant $\propto$. We have taken two values for this constant in our simulation i.e. $\propto = 1.015625$ for D.T.(1) and $\propto = 2.015625$ for D.T.(2). In present work VBR traffic is used for the purpose. Each input VBR source i, i= 1, 2 ,........ N, is modeled by two state ON-OFF Interrupted Bernoulli Process (IBP)[#].

The simulation results obtained for evaluation the performance of various congestion control schemes under unicast (90%) and multicast (10%) mixed traffic load are chosen as follows and

---

[#] The IBP Traffic Model used is presented in Section 3.6 of Chapter-3.

presented in the form of Figs. 5.8 – 5.19 and their respective Tables 5.3 – 5.14 :

- Average cell delay as experienced by unicast cells in the incoming mixed traffic
- Average cell delay as experienced by multicast cells in the incoming mixed traffic
- Average Throughput
- Cell Loss Rate

From the results we observed that all the QoS parameters as described above are the function of burst length and offered load. It has been shown from the Figs. 5.8 - 5.13 (or Tables 5.3 - 5.8), that for proposed Fuzzy Threshold scheme, average unicast as well as multicast cell delay for a given burst length is minimum under different load conditions. The ST scheme shows the better performance in comparison to DT scheme for delay parameter. The reason for increase in average cell delay for unicast and multicast cells under DT scheme, the shared memory space is utilized more efficiently than that under ST scheme; and, as a result of efficient memory utilization, more cells are able to stay for a longer duration within the memory without being lost. This phenomena decreases the cell loss rate at the cost of increasing the average queuing delay. The average unicast as well as multicast cell delay is maximum when there is no threshold applied on the incoming mixed bursty traffic.

From Figs. 5.14 – 5.16 (or Tables 5.9 – 5.11), we observed that, at the initial stage of the offered load (20% – 40%) all the schemes provide same throughput. Which shows that all the incoming cells are served by the switch. In a similar way when offered load increases from 40% to 60%, all the schemes show almost same behavior with a little difference in throughput. But at the higher loads (70% - 100%) the throughput for proposed Fuzzy Scheme is minimum, and it is maximum for DT Scheme. The problem relates to the fact that the memory of the switch fabric is shared across all the ports or lines of that switching element. Under most circumstances, this sharing is good for performance. There is a risk, however, that a small number of queues could become overloaded, take over all the shared memory in their switching element, and thereby block traffic to other queues.

TABLE 5.3 : AVERAGE UNICAST CELL DELAY Vs LOAD

For BL=20

Load	*W.T.*	*S.T.*	*D.T.(1)*	*D.T.(2)*	*F.T.*
0.2	5.1202	3.9858	4.5761	4.7799	2.43179
0.4	10.2762	7.473	9.8565	9.9771	3.403296
0.6	21.7384	11.452	22.226	25.9126	4.19243
0.8	41.8648	14.968	38.2071	37.5061	4.795557
1	42.7273	18.1302	36.1636	36.4893	5.233615

W.T. = Without - Threshold
S. T. = Static -Threshold
D.T.(1) = Dynamic - Threshold, α=1.015625
D.T.(2) = Dynamic - Threshold, α=2.015625
F.T. = Fuzzy Threshold

Fig. 5.8: Avg. Unicast Cell Delay Vs. Load, for BL=20

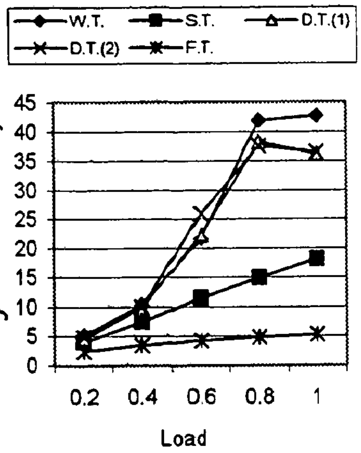


TABLE 5.4: AVERAGE UNICAST CELL DELAY Vs LOAD

For BL=30

Load	*W.T.*	*S.T.*	*D.T.(1)*	*D.T.(2)*	*F.T.*
0.2	6.8178	4.9001	6.5285	6.3425	2.62314
0.4	15.0431	8.9495	14.6465	15.6604	3.6553
0.6	35.9809	12.6886	30.4512	32.8104	4.421818
0.8	47.432	16.2619	41.2532	41.6802	5.007696
1	46.5201	19.3768	38.4722	39.0235	5.46504

Fig. 5.9: Avg. Unicast Cell Delay Vs. Load, for BL=30

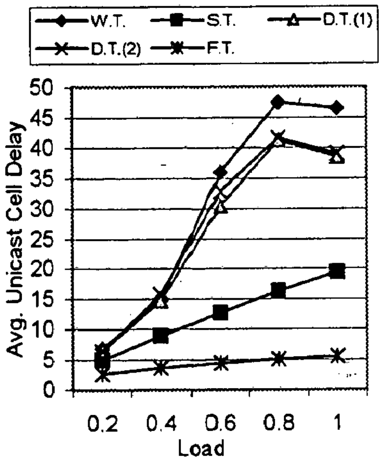


TABLE 5.5 : AVERAGE UNICAST CELL DELAY Vs LOAD

For BL=40

Load	*W.T.*	*S.T.*	*D.T.(1)*	*D.T.(2)*	*F.T.*
0.2	8.1214	5.6736	8.8847	8.9417	2.558143
0.4	21.1664	9.9561	21.8897	18.6585	3.871553
0.6	38.8708	13.8296	39.5961	39.0121	4.588369
0.8	51.1744	17.1669	43.5229	44.9101	5.104234
1	51.2696	19.9293	39.5668	41.0749	5.834389

Fig. 5.10: Avg. Unicast Cell Delay Vs. Load, for BL=40

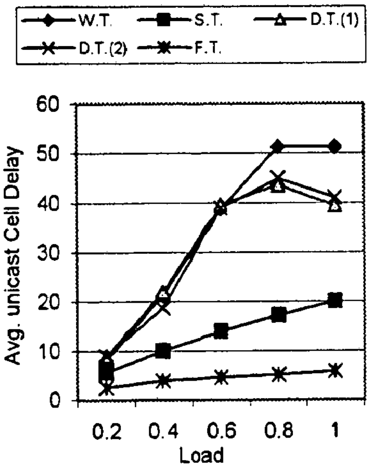


TABLE 5.6: AVERAGE MULTICAST CELL DELAY Vs LOAD

For BL=20

Load	*W.T.*	*S.T.*	*D.T.(1)*	*D.T.(2)*	*F.T.*
0.2	2.0916	1.9484	1.9799	2.1183	1.419878
0.4	4.9645	3.778	4.7311	5.1982	1.892389
0.6	16.0351	6.466	15.6829	22.3165	2.437752
0.8	43.0886	10.807	35.4606	41.0813	2.997471
1	42.8291	12.8095	36.1834	38.7862	3.517804

W.T. = Without - Threshold
S. T. = Static -Threshold
D.T.(1) = Dynamic - Threshold, α=1.015625
D.T.(2) = Dynamic - Threshold, α=2.015625
F.T. = Fuzzy Threshold

Fig. 5.11: Avg. Multicast Cell Delay Vs. Load, For BL= 20

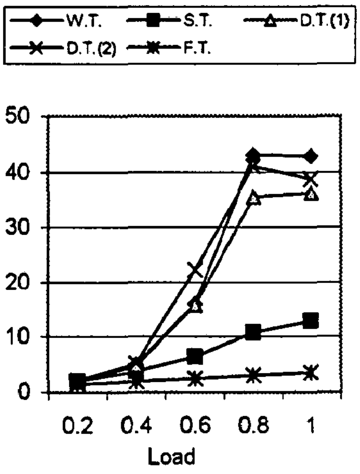


TABLE 5.7: AVERAGE MULTICAST CELL DELAY Vs LOAD

For BL=30

Load	*W.T.*	*S.T.*	*D.T.(1)*	*D.T.(2)*	*F.T.*
0.2	2.7383	2.02	2.6901	2.4459	1.500079
0.4	7.5354	3.9686	6.6537	7.1891	1.939566
0.6	32.5093	6.8513	20.1056	26.0012	2.405141
0.8	40.0977	9.7818	36.0706	38.9817	2.921138
1	44.6321	13.5026	35.8796	38.1975	3.41156

Fig. 5.12 : Avg. Multicast Cell Delay Vs. Load, for BL=30

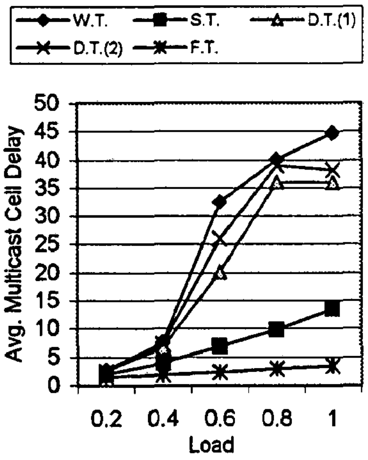


TABLE 5.8 :AVERAGE MULTICAST CELL DELAY Vs LOAD

For BL=40

Load	*W.T.*	*S.T.*	*D.T.(1)*	*D.T.(2)*	*F.T.*
0.2	2.8321	2.2156	3.0213	2.9627	1.459351
0.4	9.1184	4.1497	9.2986	8.7203	1.934768
0.6	24.7594	6.5323	25.1045	27.2037	2.523848
0.8	37.6435	9.5956	34.8049	36.8287	2.94433
1	38.0424	12.8998	37.0402	37.2728	3.53744

Fig. 5.13: Avg. Multicast Cell Delay Vs. Load for BL=40

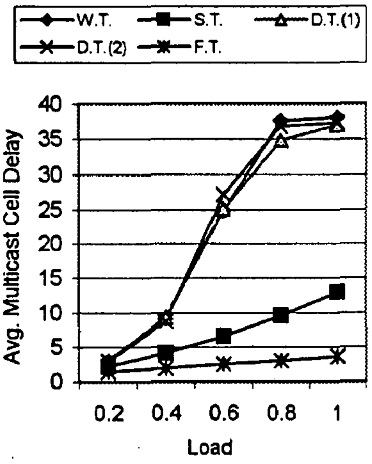


TABLE 5.9: AVERAGE THROUGHPUT Vs LOAD

For BL=20

Load	W.T.	S.T.	D.T.(1)	D.T.(2)	F.T.
0.2	0.2078	0.2037	0.2088	0.2065	0.198092
0.4	0.3941	0.3911	0.3941	0.3964	0.362559
0.6	0.6011	0.5664	0.5982	0.6035	0.510169
0.8	0.7299	0.7101	0.7732	0.7673	0.63575
1	0.7378	0.7284	0.8552	0.8444	0.72593

W.T. = Without - Threshold
S. T. = Static -Threshold
D.T.(1) = Dynamic - Threshold, α=1.015625
D.T.(2) = Dynamic - Threshold, α=2.015625
F.T. = Fuzzy Threshold

Fig. 5.14: Avg. Throughput Vs. Load, for BL=20

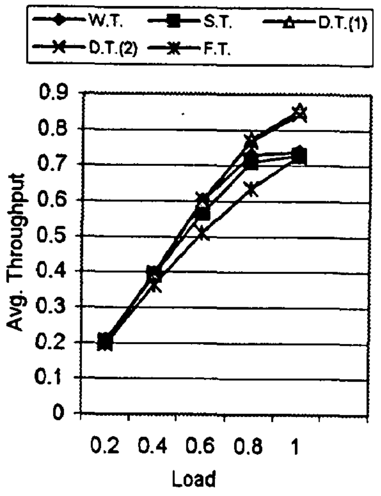


TABLE 5.10: AVERAGE THROUGHPUT Vs LOAD

For BL=30

Load	W.T.	S.T.	D.T.(1)	D.T.(2)	F.T.
0.2	0.2036	0.2039	0.2062	0.2039	0.199072
0.4	0.3992	0.3856	0.3968	0.3961	0.361766
0.6	0.5963	0.5511	0.6011	0.5952	0.501013
0.8	0.6961	0.6904	0.7472	0.74451	0.619644
1	0.6894	0.7982	0.82034	0.8122	0.716061

Fig. 5.15: Avg. Throughput Vs. Load for BL=30

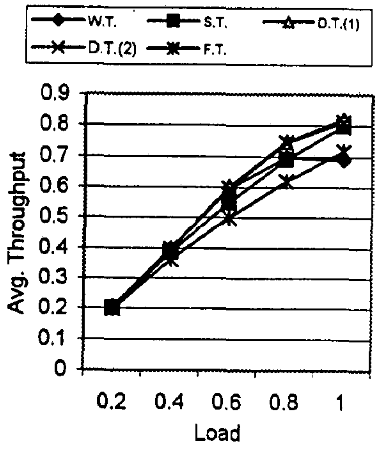


TABLE 5.11 : AVERAGE THROUGHPUT Vs LOAD

For BL=40

Load	W.T.	S.T.	D.T.(1)	D.T.(2)	F.T.
0.2	0.2021	0.1976	0.2053	0.2027	0.195539
0.4	0.4063	0.3863	0.4091	0.4067	0.362128
0.6	0.5903	0.5425	0.5926	0.5883	0.506728
0.8	0.6711	0.6701	0.7339	0.7264	0.608211
1	0.6732	0.7763	0.7991	0.788	0.665727

Fig. 5.16 : Avg. Throughput Vs Load for BL=40

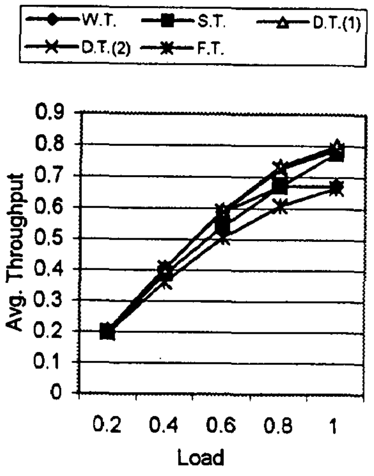


TABLE 5.12: AVERAGE CELL LOSS RATE Vs LOAD

For BL=20

Load	W.T.	S.T.	D.T.(1)	D.T.(2)	F.T.
0.2	0	0.0055	0	0	0.038859
0.4	0	0.0216	0	0	0.089962
0.6	0	0.0596	0	0	0.153736
0.8	0.0965	0.117	0.042	0.0455	0.225391
1	0.1935	0.2019	0.15	0.1646	0.293828

W.T. = Without - Threshold
S. T. = Static -Threshold
D.T.(1) = Dynamic - Threshold, α=1.015625
D.T.(2) = Dynamic - Threshold, α=2.015625
F.T. = Fuzzy Threshold

Fig. 5.17: Avg. Cell Loss Rate Vs. Load, for BL=20

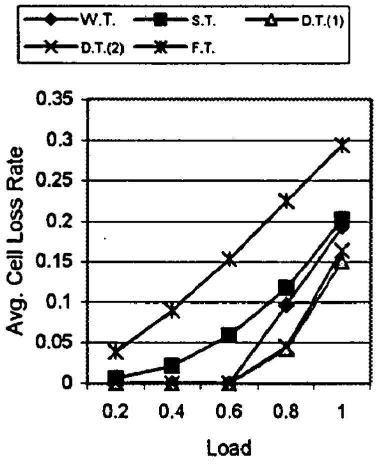


TABLE 5.13 : AVERAGE CELL LOSS RATE Vs LOAD

For BL=30

Load	W.T.	S.T.	D.T.(1)	D.T.(2)	F.T.
0.2	0	0.0102	0	0	0.043474
0.4	0	0.0376	0	0	0.098412
0.6	0.0009	0.079	0.00008	0.0006	0.162117
0.8	0.133	0.1385	0.0737	0.0719	0.233068
1	0.3148	0.2126	0.188	0.1998	0.304457

Fig. 5.18: Avg. Cell Loss Rate Vs. Load, for BL=30

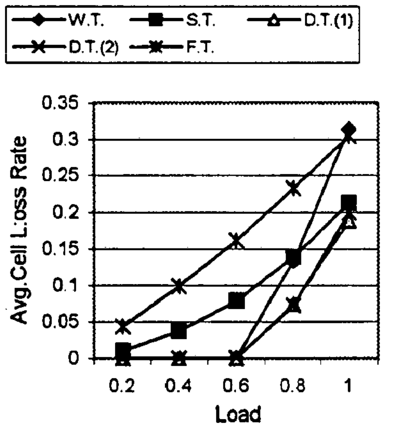
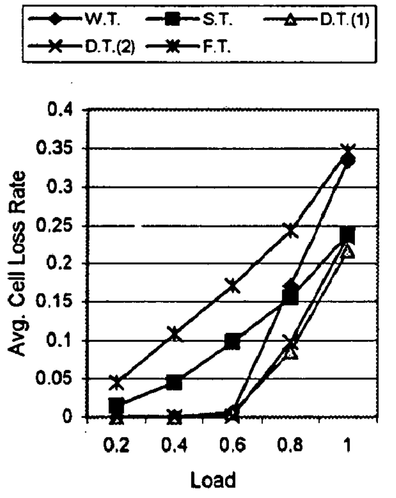


TABLE 5.14 : AVERAGE CELL LOSS RATE Vs LOAD

For BL=40

Load	W.T.	S.T.	D.T.(1)	D.T.(2)	F.T.
0.2	0	0.0149	0	0	0.044769
0.4	0	0.0449	0	0	0.108491
0.6	0.0057	0.0975	0.0054	0.0012	0.170792
0.8	0.1705	0.1551	0.0849	0.0976	0.244046
1	0.334	0.2375	0.2166	0.234	0.346473

Fig. 5.19 : Avg. Cell Loss Rate Vs. Load for BL=40



It is indicated from Figs. 5.17 – 5.19 (or Tables 5.12 – 5.14) that average cell loss rate is minimum for DT threshold. As stated previously this is achieved at the cost of increasing in average cell delay. The second reason is that, the DT deliberately wastes a small amount of buffer space. This 'wasted' buffer space actually serves useful functions. The advantage of maintaining some spare space all time is that this provides a cushion during transient periods when all output queue first become active. This reduces cell loss for newly active queue during such transients.

5.11 Conclusions:

In shared-memory ATM switches, queue length threshold/control can promote fair and efficient use of the packet buffer memory. The simplest control, Static Threshold (ST), works well as long as the threshold is tuned properly for the load condition. However, the ST is not robust to many sorts of load changes unless the threshold setting is retuned. To get good performance, one has to be especially careful not to set the threshold too low or too high. The attraction of ST is its ease of implementation.

The results of the performance evaluations indicated that the throughput of the shared-memory switch improved when Dynamic Threshold (DT) is applied. As stated earlier the reason behind this is that, memory space is utilized more efficiently in this scheme. So cells are able to stay for a longer duration within the memory without being lost. Naturally it reduces cell loss rate or increases throughput. Secondly, deliberately wasted buffer space by the scheme also reduces cell loss rate for newly active queue during transients periods. DT is nearly as simple to implement as ST and, DT is adaptive.

At last, proposed fuzzy logic based scheme in which we consider two fuzzy variables namely line occupancy and buffer occupancy, according to fuzzification and defuzzification, we find the allocation space for each line queue up-to max-line capacity in shared memory switch. It means it is dynamically for every entering cell in ATM switch.

We used the Lookup-Tables 5.1(a) and 5.1(b) for fuzzification, and Sets of Table 5.2 for defuzzification, and find the trade-offs between average cell delay (unicast and multicast both) and throughput. Different Lookup-Tables based on experience and different defuzzification sets can be used, so that parameters can get fine tuned to achieve a desired performance.

Finally the RAR scheme, which is used for multicast support for shared memory ATM switch, has its limitations and shortcomings, which are to be overcome. One of its demerits is inefficient memory utilization, as a result at higher load conditions cell loss rate becomes high.

Chapter-6

Congestion Control in Wireless ATM Network

6.1 Introduction:

Personal Communication Network (PCN) is an emerging wireless network that promises many new services[1], [37], [51], [63], and [141]. With the availability of interface cards, mobile users are no longer required to be confined within a static network premise to get network access. Mobile users may move from one place to another and yet maintain transparent network access through wireless links. Information exchanged between users, may be bi-directional, which includes but not limited to voice, data, and image, irrespective of location and time while permitting users to be mobile.

In a PCN, the covered geographical area is typically partitioned into a set of microcells[143]. Each microcell has a Base Station (BS) to exchange radio signals with wireless mobile terminals. Due to the limited range of wireless transceivers, mobile users can communicate only with BSs that reside within the same microcell at any instance[106]. When a mobile terminal, engaged in a call or data transfer, moves out of the coverage area of a BS it is communicating with, the call must be transferred to another BS of the new cell; otherwise connection is lost. This transfer process, known as handoff, is transparent to the mobile user. The number of handoffs[Π] during a call will increase as the cell radii decrease, thus affecting the QoS. Frequent handoff in wireless/mobile networks introduces a new paradigm in the area of network congestion and admission control [121], and [127]. The increase in processing load due to demand for service and fast handoffs to mitigate the propagation effect, a high speed backbone network for the PCN to connect BS is required. The ATM technology, which has recently emerged to be a predominant switching technology, is suited to be an infrastructure to interconnect the BSs of the PCN [8], [43], [137], and [138].

In wireless network congestion can occur when a number of BSs can simultaneously send packets to the same switch in the network. Bursty communication requires dynamic bandwidth allocation,

---

[Π] Here the terms 'handoff' and 'handover' are used interchangeably.

which may be difficult to allocate in practice. Bandwidth management is crucial for maintaining communication in the wireless networks. It is helpful to define certain quantities which are used in the next sections of this chapter. Two types of probabilities may be defined as QoS parameters in wireless networks as follows:

➢ ***Forced Termination Probability (FTP):***

In wireless networks, when a Mobile User (MU) travels from one cell to another, the connection handover takes place between the new and previous cell. The Forced Termination Probability (FTP) is the probability that an original call is eventually not completed because of an unsuccessful handover attempt.

➢ ***Connection/Call Blocking Probability (CBP):***

The Connection/Call Blocking Probability (CBP) is the probability that a new connection/call request is rejected. The reason for rejection is generally the unavailability of the sufficient resources which are required to meet the demands made by the connection/call.

One of the most obvious merits of a wireless network is the total traffic it carries, which can be define as:

➢ ***Total Carried Traffic:***

It is the amount of traffic admitted to the wireless/cellular network as opposed to the offered load. In light traffic conditions, the carried traffic can be taken to be equal to the offered traffic. However, in general, the carried traffic is less than the offered load because of blocking of calls and handover failures.

To support network-wide handoff, new and handoff call requests will compete for connection resources in both the mobile and backbone networks. Handoff calls require a higher congestion related performance, i.e., blocking probability, relative to new calls because forced terminations of ongoing calls due to hand-off call blocking are generally more objectionable then new call blocking from the subscriber's perspective[127].

Handover initiated in PCN, a new channel has to be granted to handover request for successful handover. To keep FTP to desired minimum values, handover algorithm should avoid blocking handover request due to lack of resources i.e. radio and wired links. In our proposed Hybrid Handover Scheme, this has been achieved by giving handover high priority over initiating calls. After applying proposed scheme there is a remarkable reduction in FTP at the cost of tolerable CBP.

6.2 Various Handover Schemes : A Review

Many handover schemes have been proposed in literature [62], [109], [126], [132], [136], [159], and [176] in order to reduce handover failure and call forced termination. The reduction of handover failure, results an increase of new call blocking probability. It becomes a tradeoff between forced termination of ongoing call, and blocking a new originating call. However, the forced termination of ongoing calls is a less desirable event in the performance evaluation of a PCN network than blocking new calls. When a call is in progress in a cell, efforts are made to provide continuity to the current call when the user moves from one cell coverage area to another. To deal with initial access and handoff problem, several strategies are reported in the literature[52], and [83] as:

Non-prioritized; Reserved Channel; FIFO Priority; Measurement-Based Priority; and Subrating Schemes.

Queuing of handover requests is made possible by the existence of the time interval the Mobile Terminal (MT) spends in the handover area, where it is physically capable of communicating with both the current and next Base Terminal Stations (BTSs). In queuing handover, if all channels of a cell are occupied, calls originating within that cell are blocked and the handover requests to that cell are queued. FIFO is queuing discipline, in which, the call first queued, will be first served. A Measurement-Based Prioritized Scheme is proposed in [158], according to the scheme a queued MT gains a higher priority as its power ratio decreases from the handover threshold to the receiver threshold. The MTs waiting for a channel in the handover queue are sorted continuously according to their priorities. Signal Prediction Priority Queuing (SPPQ)[39], which uses both Received Signal Strength (RSS) and the change in RSS to determine the priority ordering of an MT. In order to optimize the system for the minimum number of dropped handovers, the handover that would be terminated next should be the first to be handed over.

Another scheme called Reserved Channel Scheme (RCS)[79], gives handover calls a higher priority than new calls. In RCS, a number of wireless channels, called guard channels, are exclusively reserved for handover calls, and the remaining channels, called normal channels, can be shared equally between handover and new calls. In [127] the objectives of Dynamic Reserved Channel Scheme (DRCS) are to satisfy a desired dropping of the probability of handover calls, to reduce the blocking probability of new calls, and to improve the channel utilization. Similarly a flexible channel assignment scheme is proposed in [156]. In the Subrating Scheme(SRS) certain channels are allowed to be temporarily divided into two channels at half the original rate to accommodate handover calls. This subrating occurs when all the channels occupied at the moment of handover arrival. When subrating channel is released, it forms into an original full-rate channel by combining with another subrated channel [83].

6.3 Proposed Hybrid Handover Scheme:

In this chapter we proposed a hybrid handover scheme, which aims to give handover calls higher priority than new calls. The Hybrid Scheme combines two priority schemes namely Handover Queuing Scheme and Reserved Channel Scheme (bandwidth reservation). In this work, FIFO and Measurement Based Priority Scheme (MBPS) queuing discipline [158] and Reserved Channel Scheme (RCS) [79] is used, and achieves a remarkable reduction in FTP. The network resources are limited due to physical limitation of wired link and frequency interference in radio link. Consequently, as FTP decrease, the blocking probability of new calls increases. Careful implementation of handover algorithm leads to minimum FTP and keeps blocking probability to the objective value. The various steps used in the proposed Hybrid Scheme are as follows:

6.3.1 Admission Control:

In general, given total resources (channel or bandwidth) that may be allocated to the new and handover calls, blocking occurs during Call Admission Control, when the call requires bandwidth over the radio channel in a cell or the link traversed over the backbone network in excess of what is available. Without prioritized allocation scheme, handover and new calls would have the same blocking probability [127].

In the proposed scheme, a new call is admitted only if number of free channels is more than number of guard channels, otherwise, the new call is blocked. Handover calls are admitted if any channel is free. If all the channels are occupied, then the handover is queued using queuing discipline like FIFO and MBPS queuing schemes. Handover requests are blocked only if it is waiting in the queue for free resource, and the tolerance time period elapsed before granting a free resource. This reflects the natural boundary of the queue size. Fig. 6.1 shows a call flow diagram of hybrid of queuing scheme and RCS. Queuing scheme gives the priority to handover calls by keeping them waiting for resources to be freed, and give them priority over the new calls, while, RCS gives priority to the handover calls by preventing new call to use certain number of channels, which are reserved exclusively for handover calls. The hybrid scheme, combines the priority from the both schemes, and gains a higher priority for handover calls.

6.3.2 Extension Over ATM-Based Network:

The fixed or dynamic guard channel methods can be extended to the nodes in the backbone network for link bandwidth allocation to enable prioritized handover admission control. Different bandwidth assignment schemes may be employed depending on the traffic characterization and multiplexing scheme. Cellular mobile systems employ channel assignment, whereas ATM-based backbone networks may employ statistical multiplexing and statistical bandwidth assignment. Since statistical bandwidth assignment can be mapped into per call equivalent bandwidth assignment, the concept of guard channel, i.e. RCS can be directly applied to set aside reserved bandwidth for handover calls [61], and [65].

6.3.2.1 Admission Control Over Backbone Network:

New calls may need to use the backbone network to communicate with other call parties, served by different Mobile ATM Switch (MAS). A new call is admitted only if radio and backbone resources are available, otherwise, the new call is blocked. Figs.6.2(a) and 6.2(b) show flow diagram for Call Admission Control and Handover Admission Control over radio and backbone links respectively. The handover calls, may also need to use a backbone link, in this case, RCS may be applied for both radio links and backbone links, as stated in [127].

Each cell is served by BTS, which has N_R radio channels to serve MT in the cell. Number N_{RG} of radio channels, can be reserved to serve handover requests. On the other hand, each backbone link has N_L channels, number in channels in a backbone link, is relatively larger than the number of channels in a BTS. As at the radio channels, N_{LG} guard channels can be reserved to serve handover calls, that request backbone link. The number of guard channels should be determined carefully in both radio and backbone links; this number depends on the traffic patterns [158] and network topology.

Queuing the handover request, which is used in considering radio link only, can be used in the extension to ATM-based network. The idea is that when resources are needed to serve handover are not available, is to queue handover request instead of blocking the handover call, the queuing is limited by a time interval, during which some resources are expected to be freed, so handover request can be served. If handover request needs a backbone link, then it is queued if radio resources are not available or backbone link resources are not available.

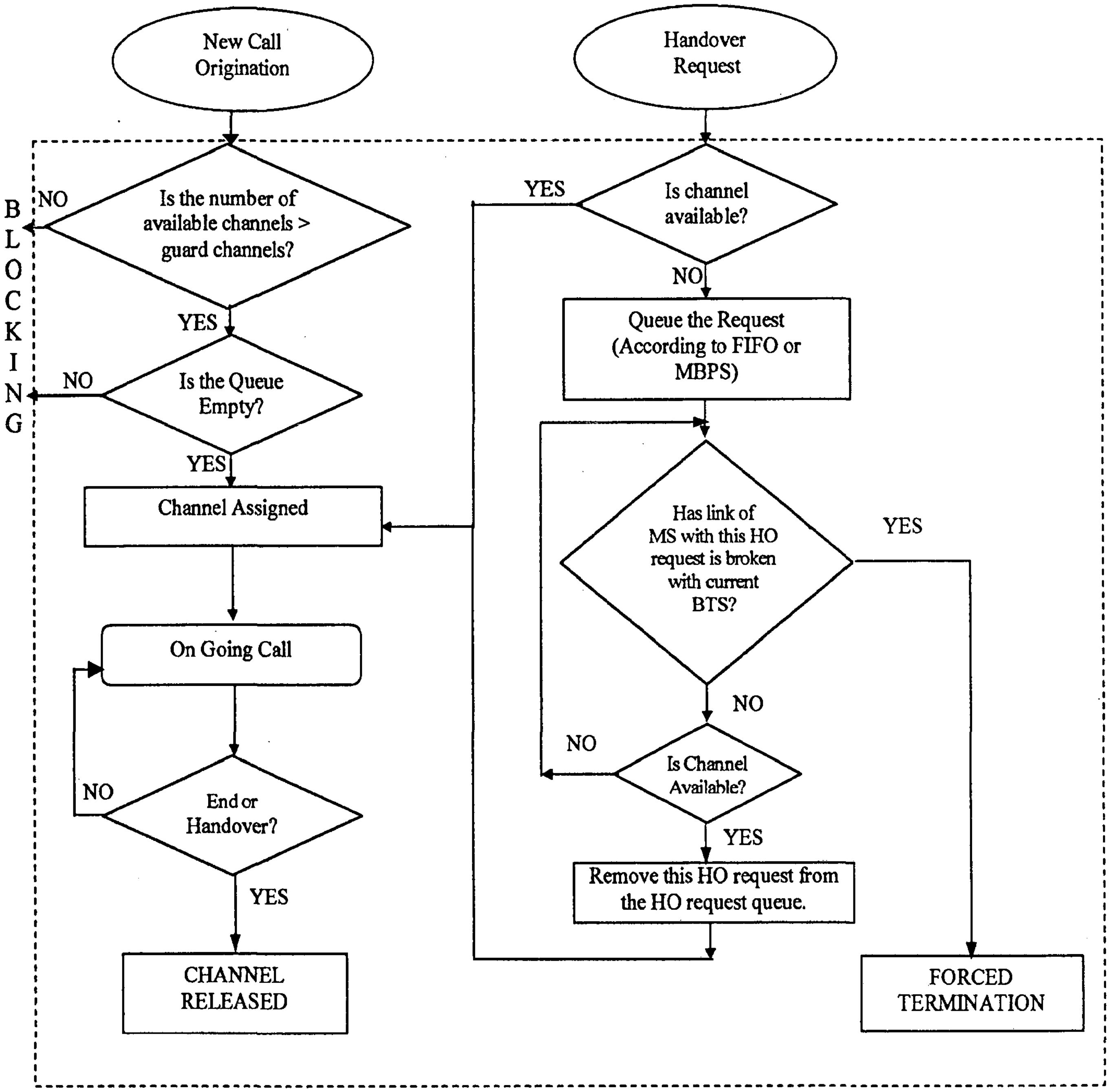


Fig. 6.1 Calls Flow Diagram for the Hybrid of Queuing Scheme with the Reserved Channel Scheme

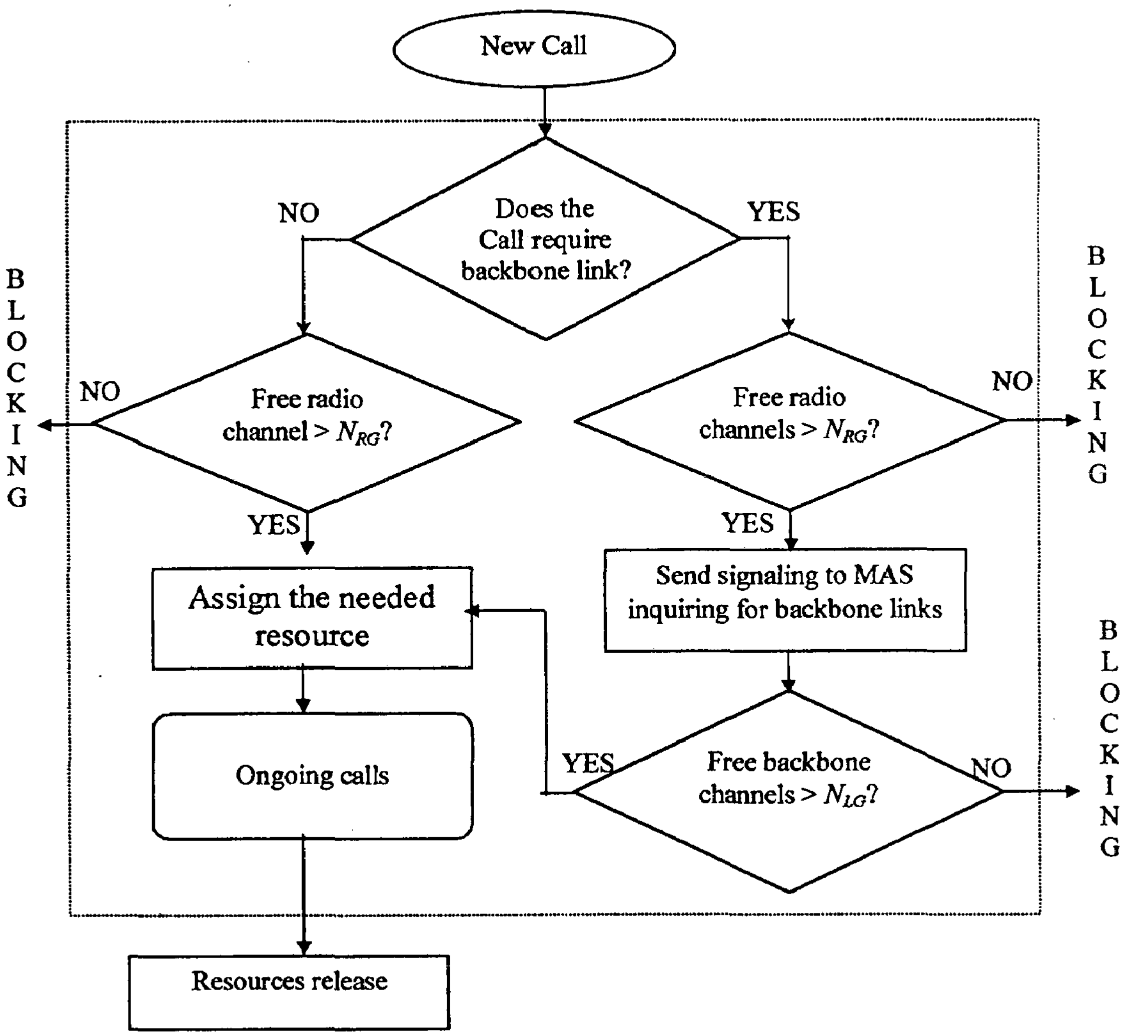
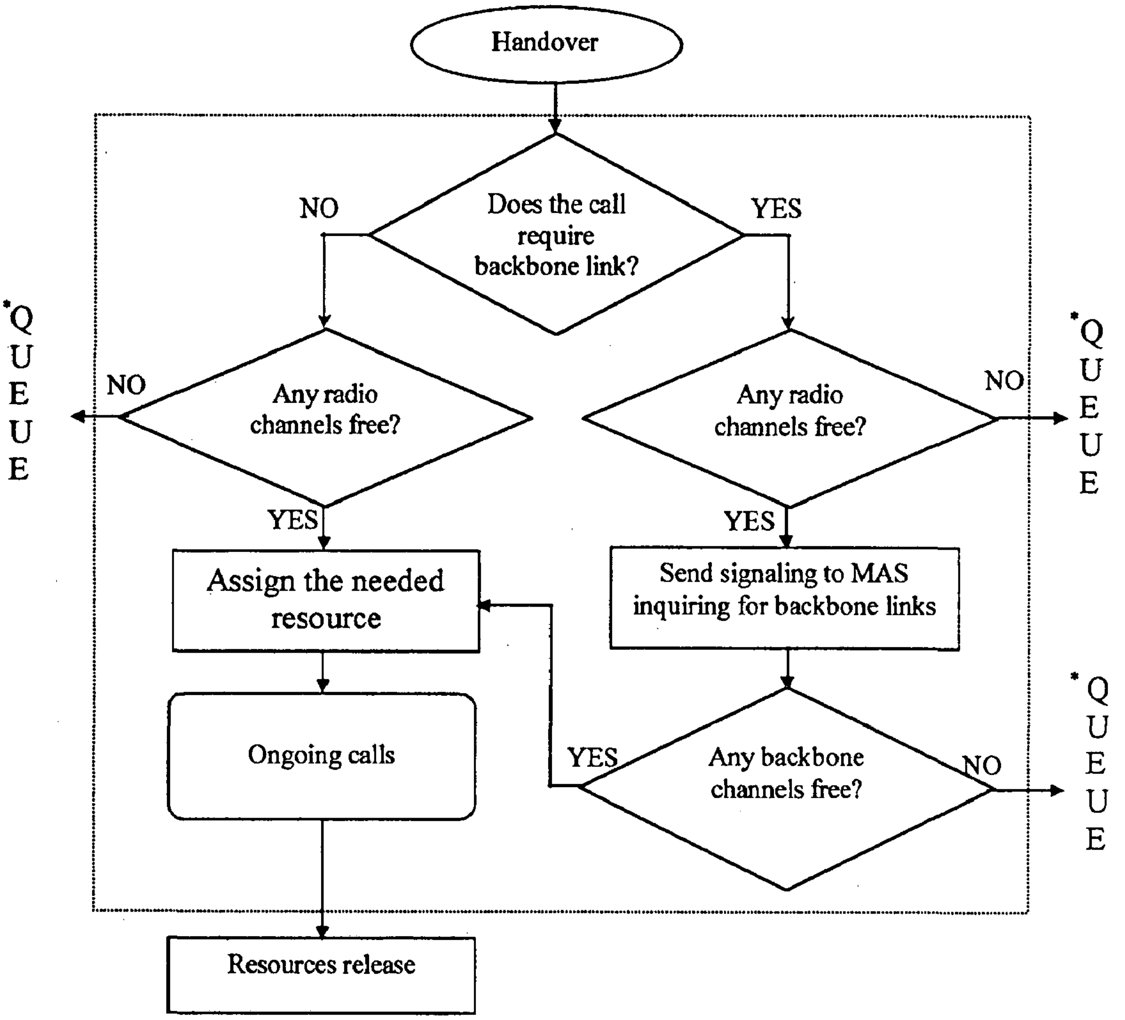


Fig. 6.2(a): New Calls Admission Control Flow Diagram for RCS Considering Both Radio and Backbone Links



* Queuing the handover request will be according to queue scheme as shown in Fig. 6.1.

Fig. 6.2 (b):Handovers Admission Control Flow Diagram for RCS Considering Both Radio and Backbone Links

6.3.2.2 Backbone Handover:

The physical network underlying the assumed connection architecture is shown in Fig. 6.3. The Radio Access Points (RAP) providing the physical radio coverage each are connected to a local MAS, thus forming one elementary mobility zone per MAS. Wide area mobility and consequently inter–zone handover and fixed network inter-working is facilitated through the concept of an Cross Over Switch (COS). While the MAS takes care of intra-zone handover, the COS is responsible for processing inter-zone handover requests [108].

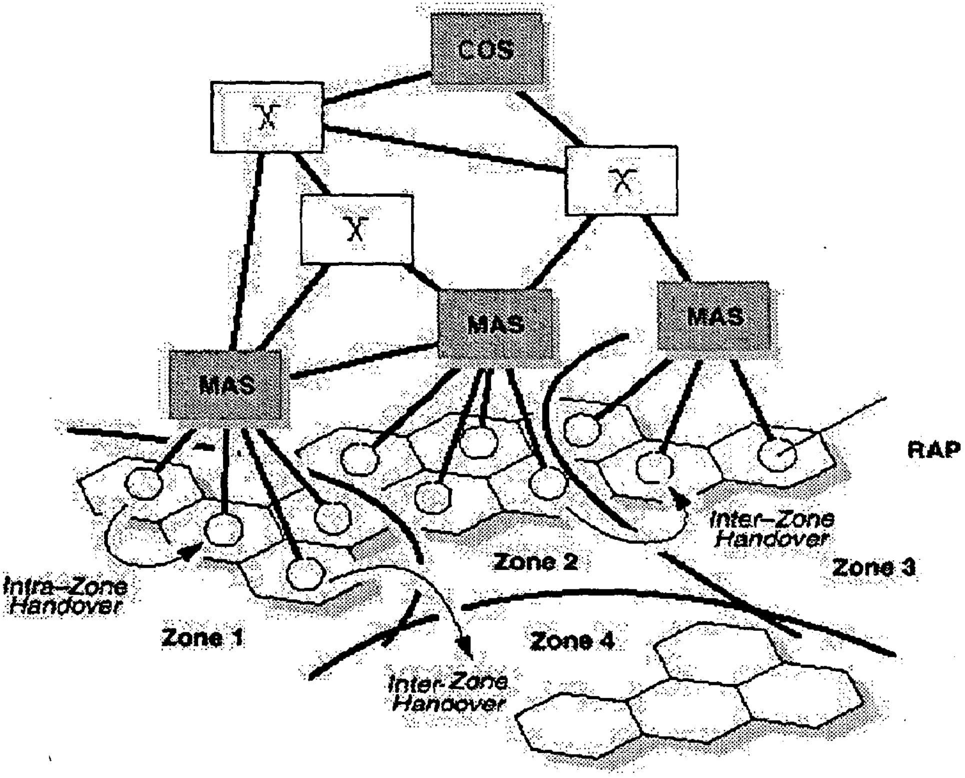


Fig. 6.3 Handover requires Backbone Link

A call may need one or more backbone link, to communicate with its destination. The number of links depends on the network topology and the destination of the call. The cost of a call is associated with the number of links, so efforts are made to optimize the number of links to the minimum. A partition methodology is proposed in [43], in which the set of cells are divided into subsets so that the base stations of the cells belonging to the same subset are connected to the same ATM switch. The optimum partition, which gives the minimum number of handover requires a change in the ATM switch. This methodology is based on handover rate between cells, which allows the choice of the best partition by taking mobility, radio, and fixed network aspect into account.

6.4 Simulation Model:

In this section, a simulation model is proposed in which multiple cells connected by MAS. The cells are connected by fixed backbone network. PCN architecture based ATM switches proposed in [106]. Simulation model can work with any channel assignment strategy. However, the results are obtained using Fixed Channel Assignment (FCA).

6.4.1 Traffic Model:

Traffic in a cell consists of new calls initiated inside the cell and handovers arriving to the cell from the neighboring cells. New calls and handovers follow Poisson distribution. The offered load (i.e. traffic) is variable, to obtain different points, while the fraction of total traffic due to the handover is kept fixed.

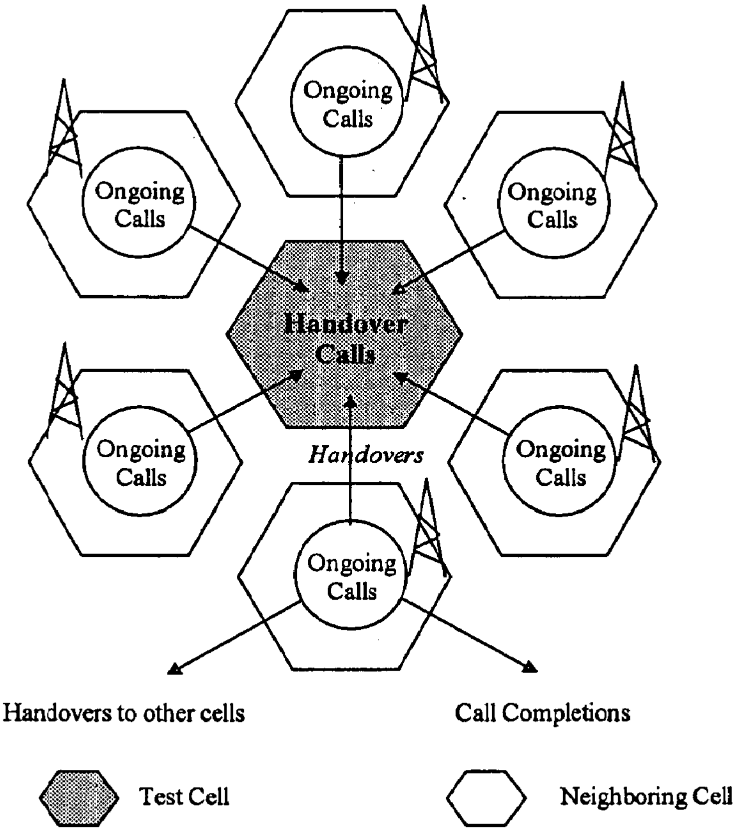


Fig. 6.4: Simulation Model for Single Cell

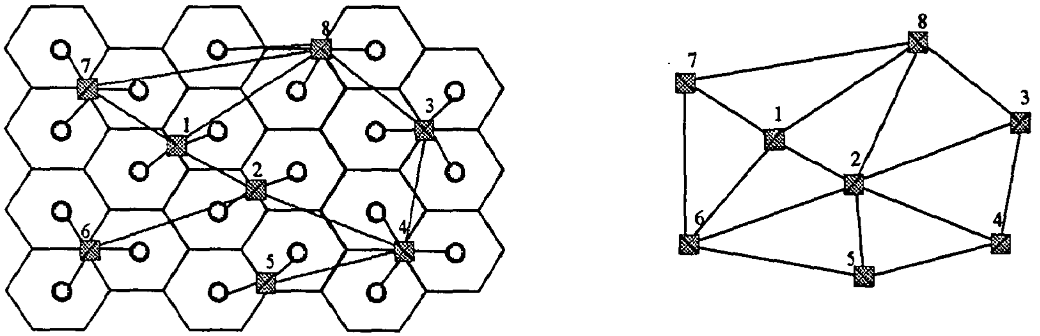
Call duration is assumed to be exponential. When a new call is originated in a cell and assigned a channel, the call holds until it is completed in the cell or handed over to another cell as the mobile moves out of the cell, Fig. 6.4 shows traffic flow into a cell. MT stays in the coverage area of a cell for a period of time (dwell time) that is exponentially distributed, and then it moves to one of surrounding cells.

The probability of requiring a handover depends on the cell coverage area, the MT movement, and the call duration. A call handover must be directed to one of the neighboring cells. The probability of each neighboring cell receiving the call depends on the amount of common boundary area and mobile direction [127]. In simulation model, we consider typical hexagonal cell, and we assume that the neighboring cells receive the handover with an equal probability of 1/6 for each.

6.4.2 Environment Description:

Each microcell has a BS to serve the MT within the cell. The geographical area is partitioned into a set $C=\{C_1, C_2, \dots, C_n\}$ of n disjoint clusters, each cluster consists of a set of microcells. An ATM switch is allocated within each cluster and each BS in this cluster is connected to one of the ports of this switch. The ATM switch offers the services of establishing / releasing channels for the MTs in the cluster, also this switch should have routing/rerouting capabilities. Two neighboring clusters can be interconnected via the associated ATM switches. The links between ATM switches are called backbone links, and the links between ATM switch and BS are called local links.

An ATM-based topology could be represented by an undirected graph $H = (V,F)$; where each vertex v_i in V stand for a cluster C_i (or an ATM switch) and an edge $e_{i,j}$ is in F if clusters C_i and C_i are adjacent in the given network. Fig. 6.5 shows an ATM based PCN topology, which consist of 21 cells, attached to 8 ATM switches, which connected by 9 backbone links. In [106], they have given PCN with different number of cells and ATM switches configuration. Corresponding H graph of Fig. 6.5(a) is shown in Fig.6.5(b).



(a) (b)

Fig. 6.5: An ATM Based Cellular Network (a) An ATM Based Architecture (b) Corresponding *H* graph of the PCN Architecture

Constructing a backbone network between MASs could be done in different ways. Depending on the geographical area, the cost of the backbone link, and traffic patterns. Figs. 6.6(a) and 6.6(b) show different possible backbone network for graph 6.5(b).

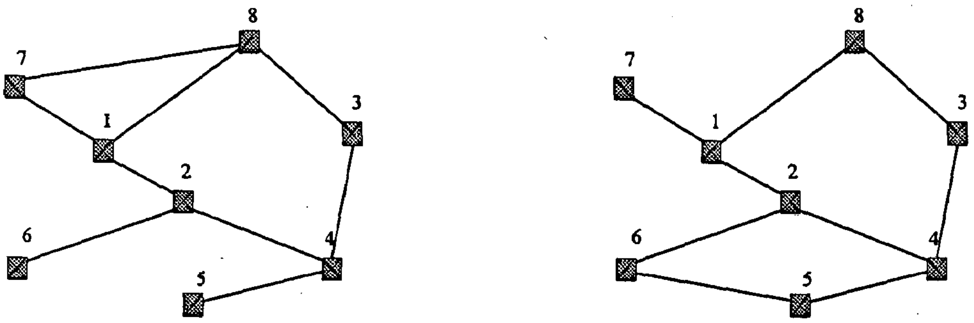


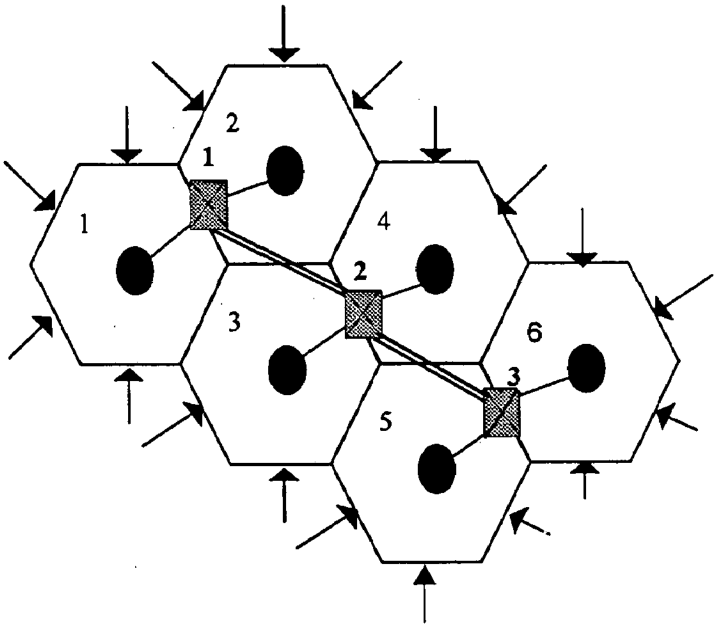
Fig. 6.6(a) **Fig. 6.6(b)**

Fig. 6.6: Different Possible Backbone Network Connections

MT engaged in a call or data transfer within the same cluster will consume two local links, one for each local link between base station and the associated switch. For intercluster communication, backbone links will be allocated in addition to local links. The channel occupied will depend on the communication path being assigned.

6.4.3 Simulation Environment:

In the simulation, we have simulated the traffic in six cells as a part of full network. From Fig. 6.7, we consider the ATM switches 1, 2 and 3. BTS 1 and BTS 2 form a cluster, and connected to ATM switch 1, BTS 3 and BTS 4 form a cluster and connected to ATM switch 2, similarly for ATM switch 3. ATM switches 1 with 2, and 2 with 3 are connected by backbone links; this configuration is illustrated in Fig. 6.7. To eliminate the boundary effect, wraparound topology is used. Traffic in the backbone link is from: calls between (BTS1 or BTS2), (BTS3 or BTS4) and (BTS5 or BTS6), and vice versa. Load from other parts of the network that may use this link in its communication. The number of channels available in this backbone network is relatively larger than that of each BTS radio channels. The initiated and handover calls in the cells, have Poisson rate as described above.



- Arrows indicate handover from surrounding cells

▨ Mobile ATM Switch

Fig. 6.7: Simulation Environment:ATM-Based Cellular PCN

6.4.4 Mobility Model:

MT may handover to cell which is inside or outside the six cells under simulation, this makes a radio channel release in that BTS, and may or may not effect the load in the backbone link, depending on the network topology and the destination. In our work we'll consider this has no effect on the backbone link.

Each cell has a BTS, which acts as a RAP, between MT and the core ATM network. BTS has a limited number of radio channels. Local link has a number of wired channels, which is equal to radio channels. Backbone link has a limited number of wired channels. For simplicity, the roaming of mobile terminal in the test cells only is considered.

The destination of the call is important to determine the need for local and backbone links, we define three call types according its destination, with probability of occurrence in the simulated environment, as follows:

- In cell call: in which the call source and destination at the same cell, the probability of this call is P_{cell}= 1/7, MT consumes only radio channel.
- In cluster call: in which the call source and destination at the same cluster, and different cells, probability of this call is $P_{cluster}$=1/7, MT consumes radio channel and local link channel.
- Out cluster call: in which the call sources and destination at different clusters, probability of this call is $P_{backbone}$=5/7, MT consumes radio channel, local link channel and backbone link channel.

Note that, there is no competition for local links, if the radio channel is available, then local link is granted, because only the user in the cell may use the local link channel. The call will occupy resources according to the probability above. If there are not enough resources available then, the call will be blocked. The new call may handover to the neighboring cell releasing the radio channel and local link channel, but may or may not release the backbone link depending on the destination and rerouting algorithm used.

6.4.5 Simulation Parameters:

The simulation parameters used for the purpose are as follows:

N_R: Number of radio channels in each cell.

N_{RG}: Number of radio guard channels in each cell.

λ_o: New call arrival rate.

λ_{hi}: Handover call arrival rate.

ρ : The offered load which is $\lambda_o + \lambda_{hi}$.

t_c: New call holding time.

t_h: Handover call holding time.

t_q: Maximum tolerable time in the queue.

N_L : Number of backbone channels in each backbone link.

N_{LG} : Number of backbone guard channels in each backbone link.

P_{cell} : Probability of in cell call.

$P_{cluster}$: Probability of in cluster call.

$P_{bacbone}$: Probability of out cluster call.

6.5 Results:

In this section, simulation results are obtained to evaluate the proposed Hybrid Scheme. Simulation program was run using default values of simulation parameters to obtain the results. 10000 calls were sampled in one arbitrary cell of the simulation environment. Calls may require a fixed part of the network to complete their connections. The default values for the simulation parameter are defined as follows[39]:

$N_R = 30$ Radio channels in each cell,

$N_{RG} = 3$ Reserved Radio channels in each cell,

$t_c = 60$ seconds average of new call holding time,

$t_h = 30$ seconds average of handover call holding time,

$t_q = 10$ seconds average time in the handover queue.

Handover has 50% of the total traffic. The offered load varies from 4 calls/min to 60 calls/min, which is considered as an overload traffic to the system.

6.5.1 Forced Termination Probability:

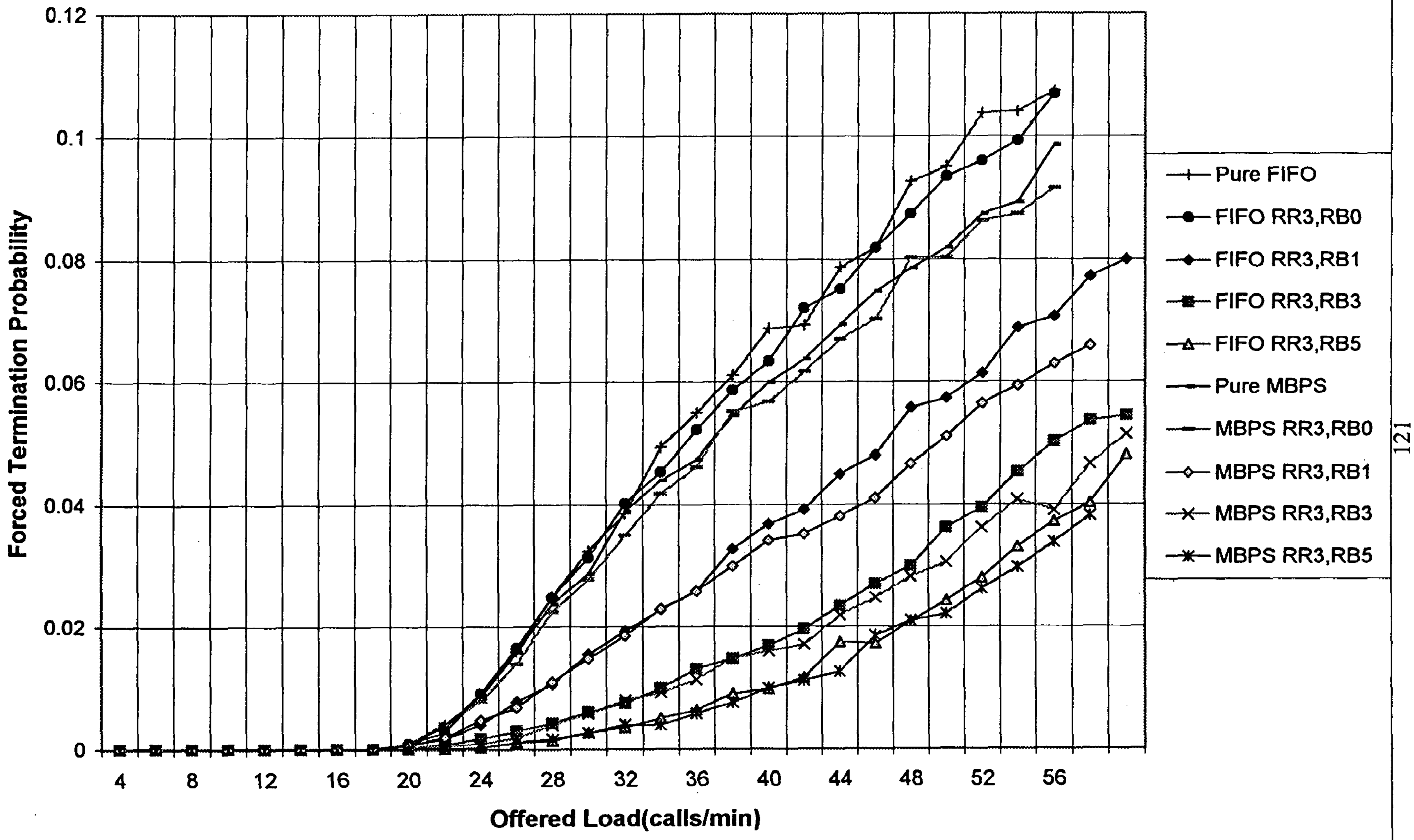
In our simulation model channels could be reserved at the backbone link as well as the radio link. Using reserved channels in both radio and backbone link lead to less Forced Termination Probability (FTP) as shown in Fig. 6.8. Channel reservation at ATM-based backbone link is valid as described in [127].

From the Fig. 6.8 it is clear that pure FIFO(i.e. for RR 0, RB 0) and pure MBPS(i.e. for RR 0, RB 0) schemes have maximum FTP with no reserved channel at the radio link and back bone link. Reserved channels at radio link (FIFO-RR 3, RB 0, and MBPS-RR 3, RB 0) improve the performance a little bit. MBPS with 3 reserved radio channels and 5 reserved backbone channels (MBPS, RR 3, RB 5) has the least FTP. FIFO scheme with 3 reserved radio channels and 5 reserved backbone channels (FIFO, RR 3, RB 5), has little improvement over (MBPS, RR 3, RB 3). FTP for (FIFO, RR 3, RB 3) is significantly higher than (FIFO, RR 3, RB 5). As mentioned before, MBPS and FIFO with 3 reserved radio channels and zero reserved backbone channels have significant higher FTP than the other described schemes, this shows the importance of using reserved channels on the backbone network links. There is significant improvement when MBPS queuing discipline is used over FIFO queuing discipline, this is clear when we use the same number of reserved channels on backbone link and radio link for both schemes.

6.5.2 Call Blocking Probability:

The Call Blocking Probability (CBP) is the probability that the new calls finds all the channels busy, and blocked. It is important to keep track of the blocking probability, to see how much various scheme yields blocking probabilities. Fig. 6.9 gives the CBP behavior of hybrid and non hybrid (FIFO and MBPS) schemes. It is clear that all the hybrid schemes have more blocking probability at higher offered load. Increase in Call Blocking Probability is always the price we have to pay for decrease in Forced Termination Probability. All hybrid schemes, which is

Fig. 6.8: Forced Termination Probabilty Vs. Load



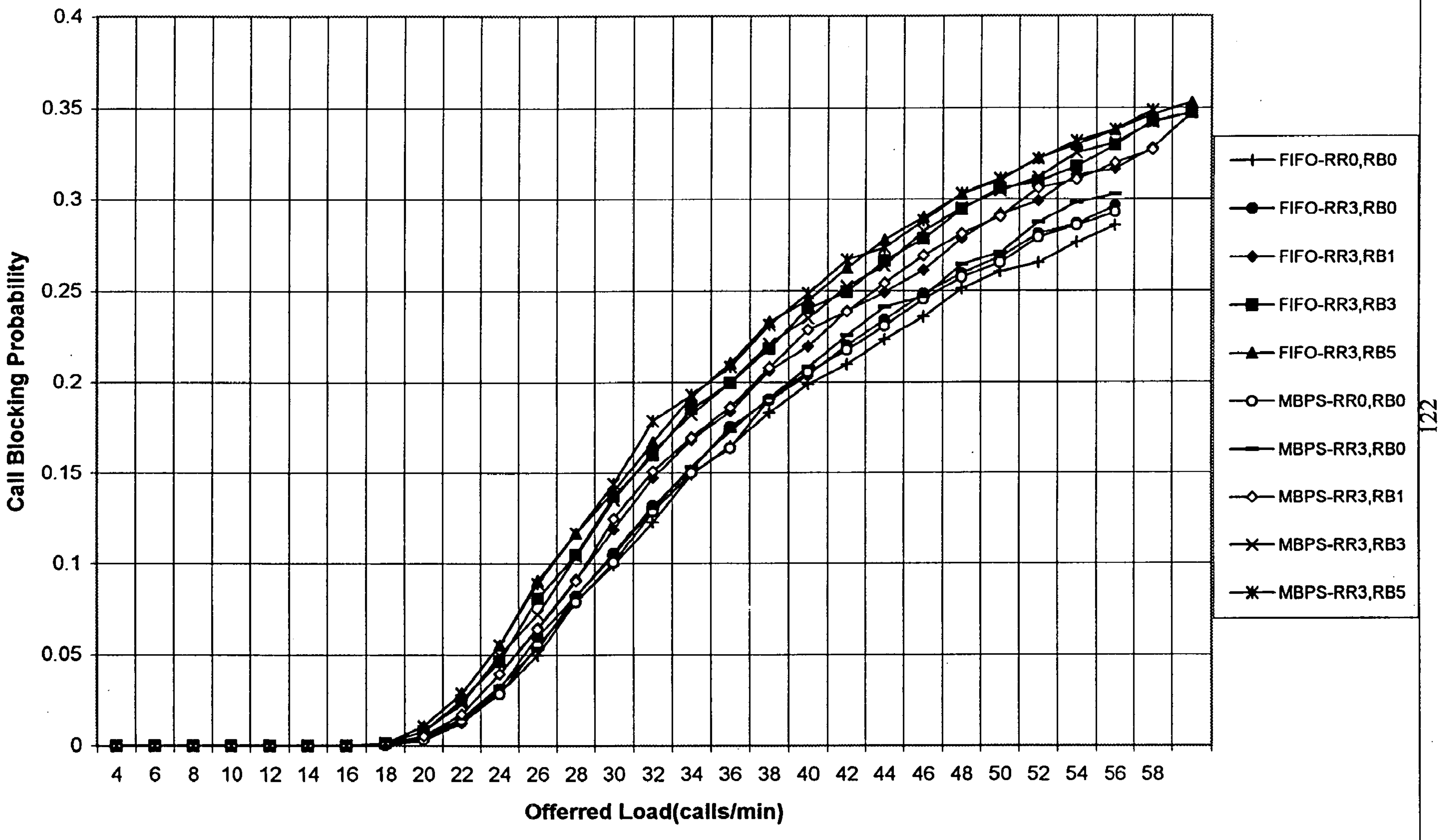


Fig. 6.9:Call Blocking Probability Vs Load

basically queuing and reservation, approximately have the same blocking probability with minor differences. This shows a trade-offs between the handover forced termination and new call blocking.

6.5.3 Total Carried Traffic:

An important performance measure is total carried traffic, which is the ratio of number of calls the system can handle to the total traffic in the system. Carried Traffic versus Offered Load is shown in Fig. 6.10. When the offered load is low, all calls are handled in all the schemes, as load increases the carried traffic decreases. The simple cause is that at higher loads FTP and CBP both increases, as a result the amount of traffic admitted to the network decreased. It is noticed that approximately all the schemes have the same performance, at moderate and higher loads.

6.5.4 Improvement in FTP due to Reserving Channels at Backbone Link:

The improvement study carried out to show how much improvement is achieved by using the hybrid scheme in the comparison to the other schemes. The improvement reflects the reduced percentage of FTP due to handover failure. The improvement of scheme S_1 over S_2 is calculated as follows:

$$Improvement\ (S_1, S_2)\ \text{in}\ \% = \frac{\{f(S_2) - f(S_1)}{S_2}\ X100$$

Where $f(S)$ is the FTP by using scheme S, substitute S_1 and S_2 for various schemes.

Fig. 6.11 shows improvement of Hybrid Scheme (i.e. using reserved channels with MBPS and FIFO), with respect to pure FIFO(RR 0, RB 0). The figure shows a significant improvement when 3 or 5 channels are reserved in backbone link. The maximum improvement is achieved by the schemes FIFO (RR 3, RB 5) and MBPS (RR 3, RB 5), which is 100% - 80% at moderate offered load and 70% - 50% at higher load. At initial stage of the environment i.e. at low offered load the improvement in FTP is zero, which indicates that all the schemes have same performance. Some improvement values are negative, the explanation is that for these points/offered load hybrid scheme has more FTP. This occurred when the load is low and the FTP is quite small, but as the load increases the improvement is clear. The general tendency of the graph is towards reducing the FTP.

Fig.6.12 explores the improvement of Hybrid Scheme with respect to FIFO (RR 3, RB 0). The figure shows a significant improvement when 3 or 5 channels are reserved in backbone link. The maximum improvements are 95 % - 80 % for moderate loads and 80% - 60% at higher loads. The improvement percentage is positive for all points, which means applying reserved channels on backbone link always behaves better than non-reserved channels on backbone link. Applying reservation scheme on backbone link leads to less FTP. When only one backbone channel is kept reserved for FIFO and MBPS Hybrid Schemes, than maximum improvement is 60% at moderate load and 50% - 40% at higher load.

6.5.5 Increase in CBP due to Reserving Channels at Backbone Link:

As consequence of reduction in FTP, an increase in new CBP is introduced. The number of resources is limited (i.e. radio channels) as more channels are assigned to serve handover request, blocking probability will increase. We have studied how the Hybrid Scheme introduces increase in CBP, in comparison to other schemes. The increased blocking reflects the percentage increase in CBP due to non-availability of resources of scheme S_1 over scheme S_2, the increase in CBP is calculated as follows:

$$\textit{Blocking Increase}\ (S_1, S_2)\ \text{in}\ \% = \frac{\{b(S_1) - b(S_2)\}}{S_1}\ X\ 100$$

Where $b(S)$ is the CBP by using scheme S, substitute S_1 and S_2 for various schemes.

Reserving channels on backbone link leads to slight increase in CBP. Fig. 6.13 explores the increase in CBP in Hybrid Scheme w.r.t. Pure FIFO(RR 0, RB 0). At low offered load (i.e. in initial stage) the increment in CBP is zero, which indicates that all the schemes have same performance. As the moderate load is offered to the network the increase in CBP suddenly raises up-to 100% for some duration, and than reduces to 30% gradually. The reason is that in the beginning of this duration no new calls get a channel all the channels are utilized by on going

Fig. 6.10: Total Carried Traffic Vs Load Offered

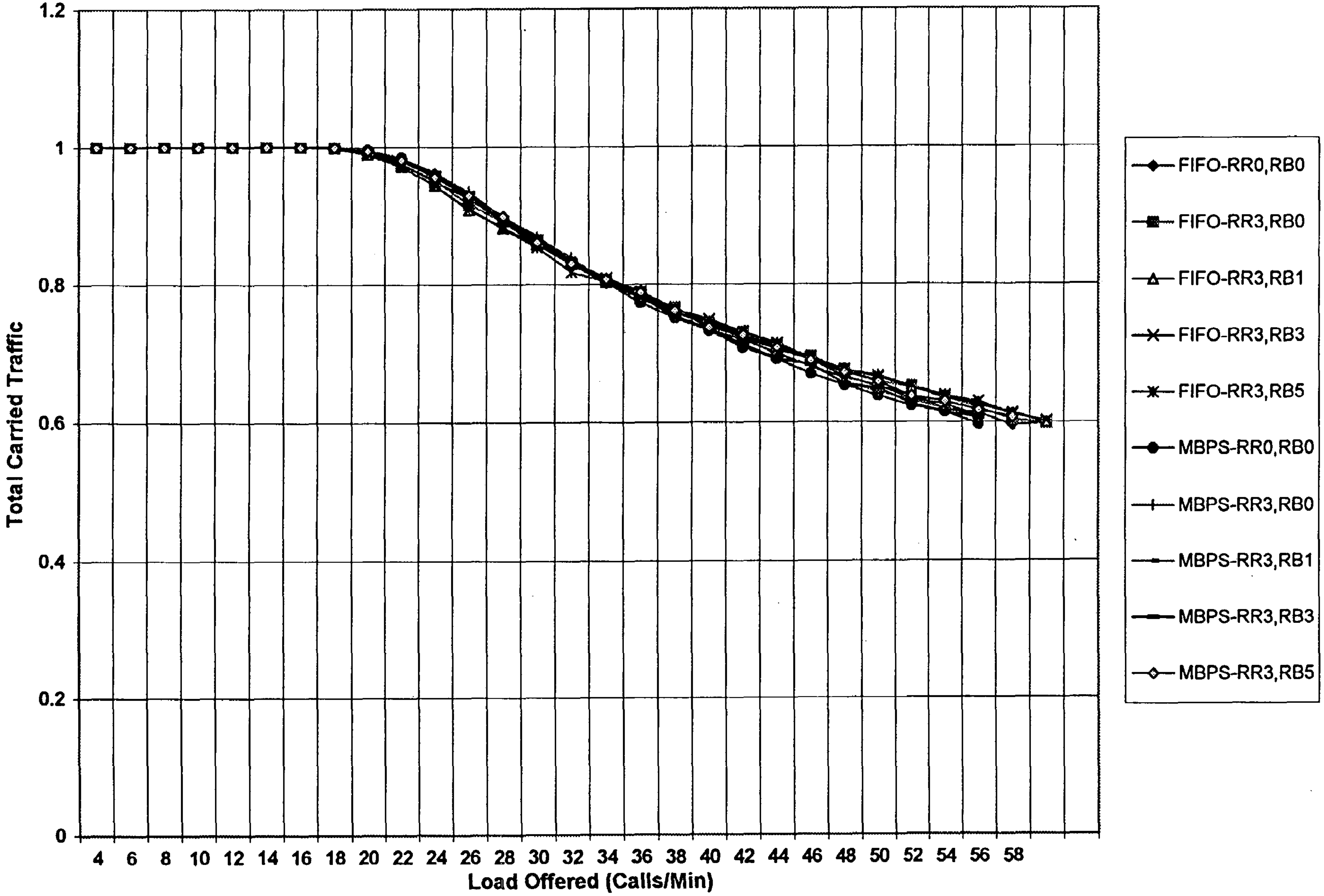


Fig. 6.11: Decrease in Forced Termination Prob. w.r.t. Pure FIFO(RR-0,RB-0)

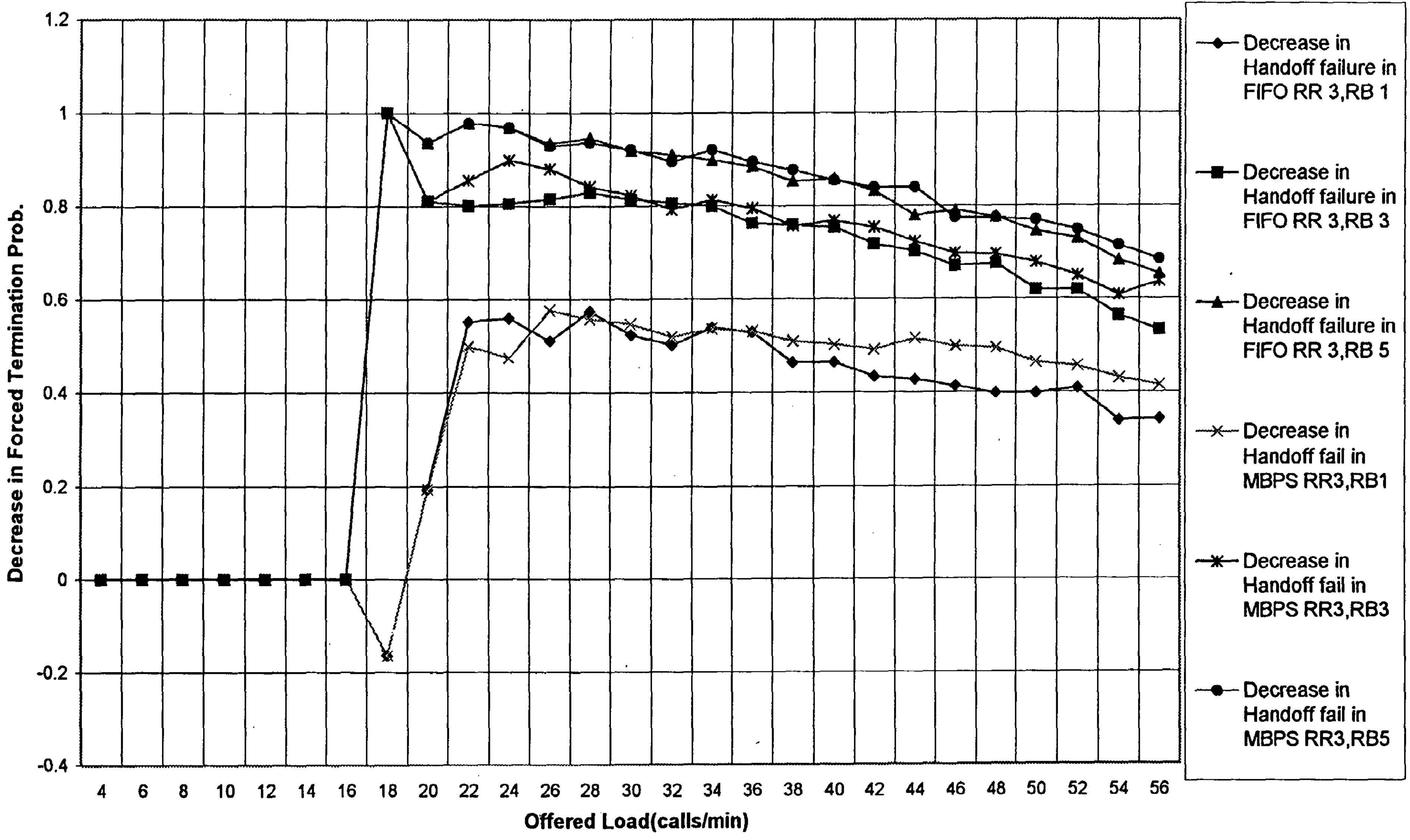


Fig. 6.12: Decrease in Forced Termination Probability w.r.t. FIFO-RR3,RB0

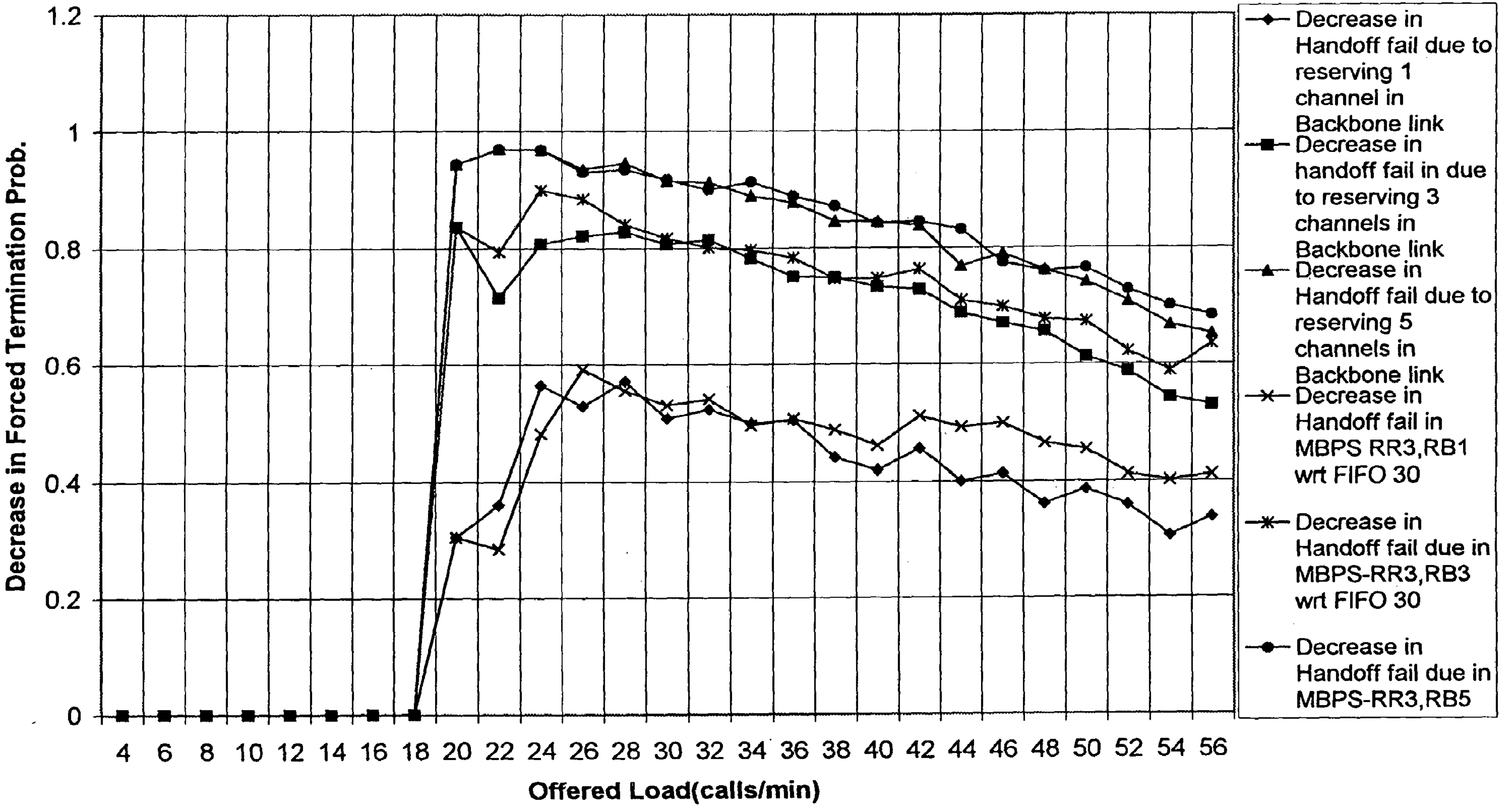


Fig. 6.13: Increase in Call Blocking Prob. w.r.t. Pure FIFO (RR-0, RB-0)

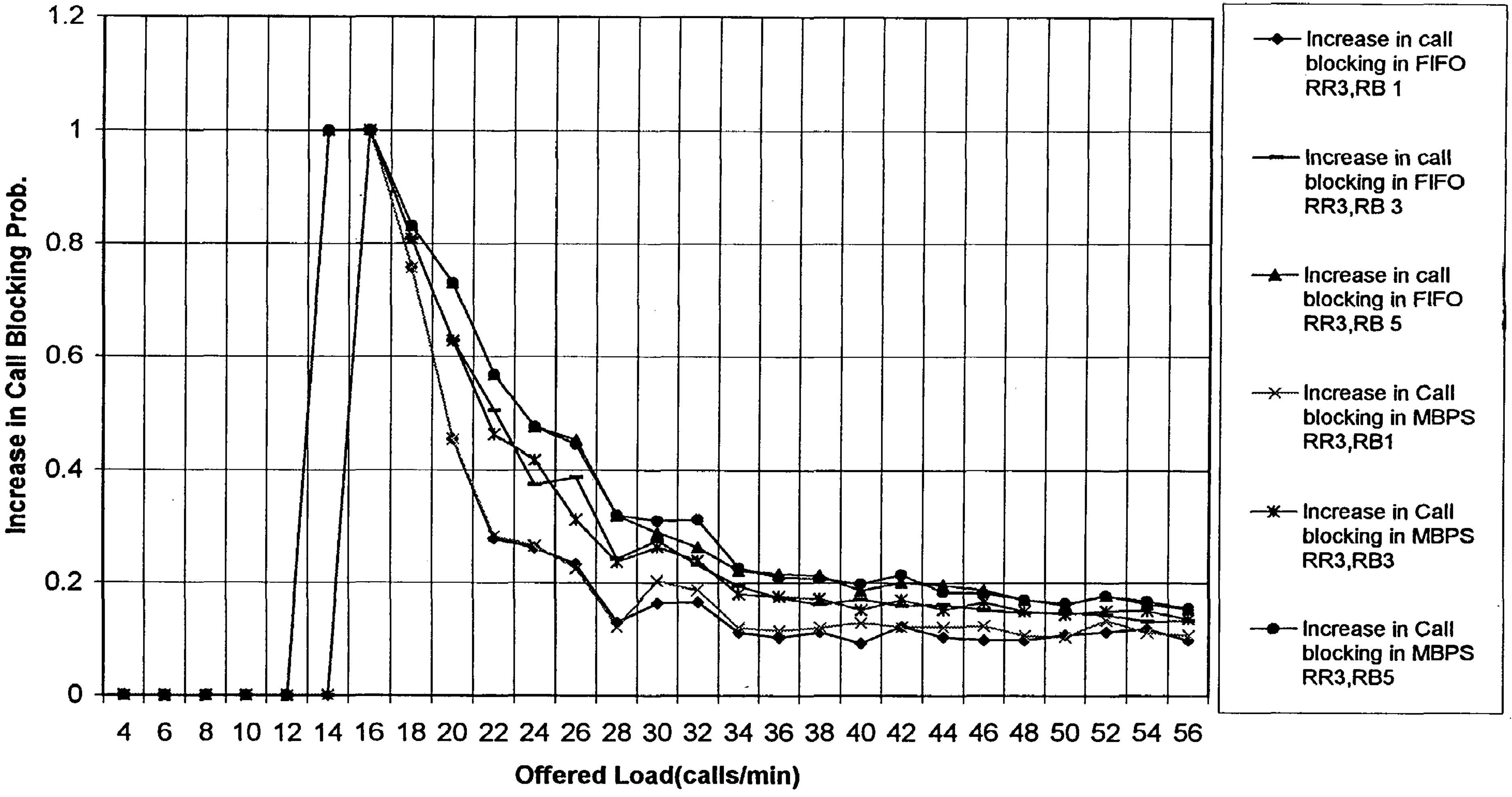
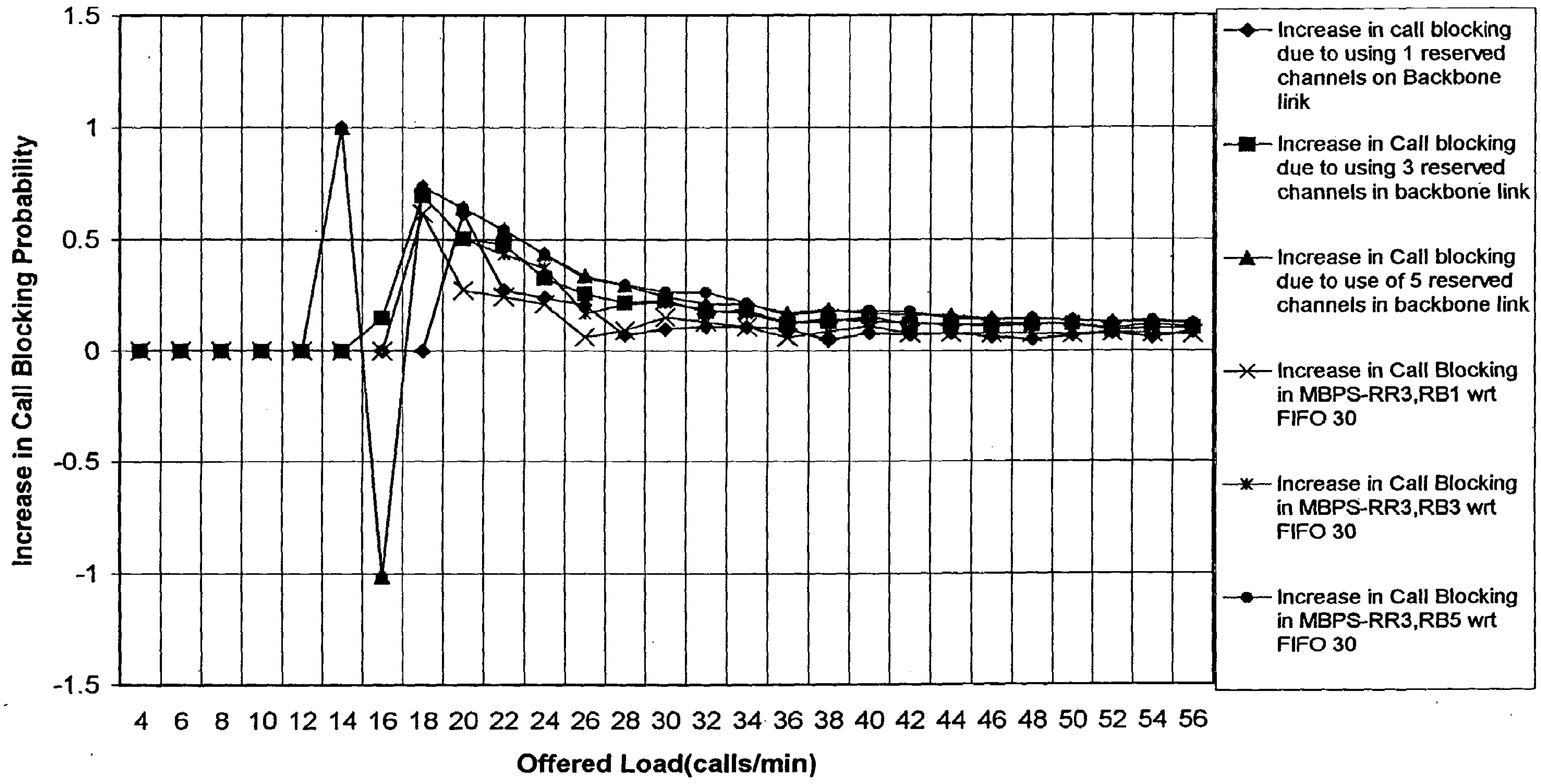


Fig. 6.14: Increase in Call Blocking Probability w.r.t. FIFO-RR3,RB0



calls. As some channels get freed for new calls the increment in CBP starts reduces till 30% as mentioned above. For higher loads this range is 30% - 10%.

Fig. 6.14 represents the increase in CBP of Hybrid Scheme w.r.t FIFO (RR 3, RB 0). The figure shows no increment in CBP at the initial stage. But in the beginning of moderate offered load for schemes FIFO (RR 3, RB 5) and MBPS (RR 3, RB 5) the graph oscillates suddenly between positive maximum and negative maximum values. The justification is that as the calls are generated randomly in our simulation, so 100% increment in CBP means no channel is available for new generated calls at that moment. But within the next few minutes the phenomena changes so rapidly that all the new generated calls get channels. So in-spite of increment in CBP there is significant improvement of 100%, when 5 channels are reserved in backbone link. As this situation passed all the schemes shows similar characteristics. For moderate loads this range varies from 70% to 25% and 25% to 15% at higher loads.

6.6 Conclusions:

This chapter focuses on the problem of congestion control in wireless ATM network. This is a practical fact that frequent handoff in PCN introduces the phenomena of congestion. After studying the currently used schemes, it is clear that there is some room for improvement for conventional handoff ordering schemes.

A simple new Hybrid Scheme is proposed to solve the hand off/ hand over problem in ATM-based PCN, which aims to give handover calls high priority over new calls. From simulation results, we find that using reserved channels at radio and backbone links, there is remarkable reduction in Forced Termination Probability (100% to 80% at moderate offered load and 70% to 50% at higher load). The price paid for using reserved channels is increase in Call Blocking Probability 75% to 25% approximately for moderate loads and 25% to 15% (approx.) for higher loads. This occurs because there is finite capacity for the network, and keeping more handoffs calls from being lost will result in more originating calls being lost because there are insufficient

This chapter concludes the thesis by summarizing the important results of the present work. The scope for future work in the area is also included. The summary of various chapter is as follows:

i. Congestion control for VBR service in non-blocking ATM switches:

To improve the performance of non-blocking ATM switch and to control the congestion for VBR traffic, we need to deal with the Head Of Line (HOL) blocking problem. Parallel Iterative Matching (PIM), a fast algorithm for choosing a conflict-free set of cells to forward across the switch during each time slot is used to reduced the HOL blocking. On the basis of simulation study for evaluating the performance of such switches, the following are conclusions:

- The results prove that PIM technique increases throughput and reduces not only mean cell delay but cell loss probability also.
- These results are very much suitable for providing better QoS for real time VBR bursty traffic applications.
- Finally we observed that various graphs almost clustered to-gather under different iterations, which indicates that three iterations are enough to find maximal matching under any traffic load.

Scope for future work:

- In this work, the performance of the non-blocking ATM switches has been evaluated for VBR service. Other service categories can be used for the purpose.
- The PIM algorithm can be extended to allocate resources fairly, when some part of the network is overloaded or congested.
- The consideration of queueing strategy and buffer size is also important, one can not ignore this while evaluating the performance of such switches. Keeping this point in mind, simulation study can be done for various queueing strategies and buffer size.

ii. Congestion control in ATM network using fuzzy approach:

The aim of various traditional approaches to congestion control is to achieve a desired trade-offs between the number of cells carried through the network, propagation delay of the cells, and the number of discarded cells. Fuzzy approaches are very much useful to deal with congestion control related problems in ATM networks. We have proposed the use of fuzzy logic for dynamic feed-back threshold in our Fuzzy Scheme. In this work, burst length as well as buffer occupancy are represented by triangular membership functions of fuzzy sets. By applying fuzzification, and defuzzification method, the percentage blocking to offered at that particular buffer occupancy level and at given burst length can be determined. After comparing the performance of proposed scheme with Constant and Dynamic Feed-back Threshold schemes, the following are conclusions:

- QoS parameters (Avg. throughput, cell delay, and cell loss probability) are the functions of offered load and number of buffer parts (N). However, in the case of Constant Threshold scheme QoS parameters don't depend upon the value of N.
- A comparative study has revealed the proposed scheme is able to achieve lower average cell delay and higher throughput than the Constant as well as Dynamic Feed-back Threshold schemes and that too with lower cell loss probability.
- However, these improvements are achieved at the cost of higher time complexity.

Scope for future work:

- The work presented in this chapter is based on many limitations and assumptions which must be eliminated to analyze the system behavior correctly.
- We have used On-Off IBP traffic model for simulation, other traffic model can be used for the purpose.
- In proposed Fuzzy Scheme Off Line Look-up Tables can be replaced by On Line Look-up Tables.

iii. Congestion control for multicast bursty traffic in shared memory ATM switch:

We have studied and simulated the effect of various threshold schemes to control congestion for multicast bursty traffic in shared memory ATM switch, and proposed a Fuzzy Scheme for the purpose. We considered two fuzzy variables namely line occupancy and buffer occupancy in our Fuzzy Scheme. After fuzzification and defuzzification process, we find the allocation space for each line queue up-to max-line capacity in shared memory switch . After applying the scheme, the conclusions based on comparative performance are as follows:

- The proposed Fuzzy Scheme provides minimum cell delay (unicast as well as multicast) at the cost of high cell loss rate.
- The Dynamic Threshold Scheme provides better throughput at the cost of very high cell delay.
- The performance of Static Threshold is in between Dynamic Threshold and Fuzzy Scheme.
- The results have showed that there is a trade-offs exits between cell delay and cell loss rate parameters.

Scope for future work:

- We used Lookup-Tables 5.1(a), (b) for fuzzification and Table 5.2 for defuzzification, and find the trade-offs between average cell delay and cell loss rate. To get better results one can use different Lookup-Tables and defuzzification sets based on experience. So that parameters can be get fine tuned to achieve a desired performance.
- Artificial Neural Network (ANN) for generating the inference engine can be used to get the optimal Lookup-Table.
- For multicast support in shared-memory ATM switch other schemes like Single Write, Multiple Read (SWMR) or Single Write, Single Read (SWSR) can be used.

iv. Congestion control in wireless ATM network:

In wireless network congestion can occur when a number of Base Stations can simultaneously send packets to the same switch in the network. Since handoff / handover calls require a higher congestion related performance, i.e., blocking probability, relative to new calls. Keeping this point in mind, a Hybrid Scheme has been proposed for handover in ATM-based Personal Communication Network. This scheme combines FIFO, and Measurement Based Priority Scheme (MBPS) queuing discipline, and Reserved Channel Scheme (RCS). The scheme gives handovers higher priority over initiating calls. The following are conclusions after simulating the scheme:

- There is significant improvement when MBPS queuing discipline is used over FIFO queuing discipline, when we use the same number of reserved channels on backbone link and radio link for both schemes.
- When reservation is applied on both radio and backbone channels, it leads to significant decrease in Forced Termination Probability with acceptable increase in Call Blocking Probability. Since the number of resources is limited (i.e. radio channels) as more channels are assigned to serve handover request, blocking probability will increase, which is obvious. This reflects the important of applying reservation scheme on both radio and backbone links and establishes a trade-offs also.
- Suddenly changes in the values of probabilities shows there is a need of careful determination/ implementation of reserved channels at radio as well as backbone levels. Consideration of traffic pattern and offered load is also important.

Scope for future work:

- Hybrid scheme can be implemented after applying dynamic reservation channels.
- Combination of the other schemes like channel borrowing and channel carrying may lead to more improvement in QoS.
- Various parameters used in the schemes can be calculated Off-Line and load into Look-up Tables to facilitate the dynamic allocation.

References:

1. Acampora, A., : "Wireless ATM : A perspective on issues and prospects," IEEE Personal Communications, vol. 3, no. 4, pp. 8-17, Aug. 1996.
2. Adas, A., : "Traffic models in broadband networks," IEEE Communications Magazine, vol . 35, no. 7, pp. 82-89, July 1997.
3. Agrawal, A. K., and Sanghi, D., : "An end-to-end send time based flow control scheme," Journal of Internetworking : Research & Experience, vol. 6, pp. 41-58, 1995.
4. Agrawal, A.K., Sanghi, D., and Leyuan Shi, : "Flow control of a virtual circuit," Proc. 11th Int'l Conference on Analysis and Optimization of Systems Discrete Event Systems, June 1994, Ecole des Mines, Sophia-Antipolis, France.
5. Anderson, T.E., Owicki, S.S., Saxe, J.B., and Thacker, C.P., : "High-speed switch scheduling for local area networks," ACM Trans, Comput. System, vol. 11, no. 4, pp. 319-352, Nov. 1993.
6. Anurag, Muppala, J.K., and Meer, H. De., : "Study of expedited forwarding in differentiated services and its performance characteristics using CBQ implementation," Proc. 7th Int'l Conf. on Advanced Computing and Communications (ADCOM), pp. 238-243, 20th–22nd Dec.1999, Roorkee, INDIA.
7. Atai, A., and Hui, J., : "A rate-based feedback traffic controller for ATM networks," Proc. IEEE ICC 94, New Orleans, pp. 1605-1615, 1994.
8. Ayanoglu, E., Eng, K.Y., and Karol, M.J., : "Wireless ATM : Limits, challenges, and proposals," IEEE Personal Communs., vol. 3, no.4, pp. 18-25, Aug. 1996.
9. Bae, J.J., and Suda, T., : "Survey of traffic control schemes and protocols in ATM networks," Proc. of the IEEE, vol. 79, no. 2, pp. 170 - 189, Feb. 1991.

10. Benmohamed, L., and Meerkov, S.M., : "Feedback control of congestion in packet switching networks : The case of a single congested node," IEEE / ACM Trans. on Networking, vol. 1, no. 6, pp. 693-708, Dec. 1993.

11. Bensaou, B., et. al., : "Estimation of the cell loss ratio in ATM networks with a fuzzy system and application to measurement–based call admission control," IEEE/ACM Trans. on Networking, vol. 5, no. 4, pp. 572-583, August1997.

12. Berkan, R. C., and Trubatch, S.L., : "Fuzzy system design principles : Building fuzzy If-Than rule Bases," Published by IEEE Press New Yark, pp. 7, 41-44, 227-99, 1997.

13. Biswas, S.K., and Izmailov, R., : " Design of a fair bandwidth allocation policy for VBR traffic in ATM networks," IEEE/ACM Transactions on Networking, vol. 8, no. 2, pp. 212-223, April 2000.

14. Biswas, S.K., Izmailov, R., and Sengupta, B., : "Connection splitting : An efficient way of reducing call blocking in ATM," IEEE/ACM Transactions on Networking, vol. 8, no.5, pp. 655-666, Oct. 2000.

15. Blau, S., et al, : "AXD 301 : A new generation ATM switching system," Computer Networks, vol. 31, no. 6, pp. 559-582, 1999.

16. Bocker, P., : "ISDN The Integrated Services Digital Network, Concepts, Methods, Systems," IInd Edition, pp. 19-27, Springer-Verlag, 1987.

17. Bonde, A.R., and Ghose, S., : "A comparative study of fuzzy versus 'fixed' threshold for robust queue management in cell switching networks," IEEE/ACM Trans. Networking, vol. 2, no. 4, pp. 337-344, Aug. 1994.

18. Bonomi, F., and Fendick, K.W., : "The rate–based flow control framework for the available bit rate ATM service," IEEE Network, vol. 9, no.2, pp. 25-39, March/April 1995.

19. Cao, Xi-Ren, and Towsley, D., : "A performance model for ATM switches with general packet length distributions," IEEE/ACM Trans. Networking, vol.3, no. 3, pp. 299-309, June 1995.

20. Catania, V., et. al., : “A comparative analysis of fuzzy versus conventional policing mechanism for ATM networks,” IEEE / ACM Trans. on Networking, vol. 4, no. 3, pp. 449-459, June 1996.

21. CCITT Draft Recommendation I. 150, Sec. 3.4.2.3.2, Study Group XVIII, Geneva, Jan. 1990.

22. CCITT Draft Recommendation I. 361, “ ATM Layer Specification for B-ISDN,” Study Group XVIII, Geneva, Jan. 1990.

23. Chan, S., et. al., : “Fair packet discarding for controlling ABR traffic in ATM networks,” IEEE Trans. on Communications, vol.45, no.8, pp. 913-916, Aug. 1997.

24. Chander, A., Akhila, et. al., : “The IIT-M Intranet: A Case Study of Cost – Effective Networking,” Presented at SECURENET-1998 National Seminar on Security Aspects of Intranet/ Internet, DOE/ITI/IETE, Banglore, 7-8th Jan. 1998.

25. Chen, T.M., and Liu, S.S., : “Management and control function in ATM switching systems,” IEEE Networks, vol. 8, no. 4, pp. 27-40, July/ Aug. 1994.

26. Chen, W.T., et. al., : “An efficient cell-scheduling algorithm for multicast ATM switching systems,” IEEE/ACM Trans. on Networking, vol. 8, no. 4, pp. 517- 525, Aug. 2000.

27. Cheng, F.C., and Holtzman, J.M., : “Wireless intelligent ATM network and protocol design for future personal communication systems,” IEEE JSAC, vol. 15, no. 7, pp. 1289-1307, Sept. 1997.

28. Cheng, R.G., and Chang, C.J., : “Design of a fuzzy traffic controller for ATM networks,” IEEE / ACM Trans. on Networking, vol. 4, no. 3, pp. 460-469 June 1996.

29. Cheng, R.G., Chang, C.J., and Lin, L.F., : “A QoS–provisioning neural fuzzy connection admission controller for multimedia high-speed networks,” IEEE/ACM Trans. on Networking, vol. 7, no. 1, pp. 111-121, Feb.1999.

30. Choi, B.D., Park, C.G., and Sung, D.K., : "Performance analysis of a Leaky Bucket scheme with a threshold in the data buffer," Computer Networks and ISDN Systems, vol. 29, no. 7, Pp.781-795, 1997.

31. Choudhury, A.K., and Hahne, E.L., : "A new buffer management scheme for hierarchical shared memory switches," IEEE / ACM Trans. on Networking, vol. 5, no. 5, pp. 728-738, Oct. 1997.

32. Choudhury, A.K., and Hahne, E.L., : "Dynamic queue length thresholds for shared-memory packet switches," IEEE / ACM Trans. on Networking, vol. 6, no.2, pp. 130-140, April 1998.

33. Chowdhury, S., and Sengupta, B., : "Design and performance of a threshold-based load balancing scheme in an ATM switch," Computer Networks and ISDN Systems, vol. 29, no. 2, pp. 157-164, 1997.

34. Cidon, I., et. al., : "Optimal buffer sharing," IEEE JSAC, vol. 13, no. 7, pp. 1229-1239, Sept. 1995.

35. Cidon, I., et. al. : "Real time packet switching : A performance analysis," IEEE JSAC, vol. 6, no. 9, pp. 1576-1586, Dec. 1988.

36. Cidon, I., Gopal, I., and Guerin, R., : "Bandwidth management and congestion control in planet," IEEE Communications Magazine, vol. 29, no. 10, pp.54-64, Oct.1991.

37. Cox, D.C., : "Personal communications - A viewpoint," IEEE Commun. Mag., vol. 28, no. 11, pp. 8-20, Nov.1990.

38. Drury, D.M., : "ATM traffic management and the impact of ATM switch design," Computer Networks and ISDN Systems, vol. 28, pp. 471-479, 1996.

39. Ebersman, H.G., and Tonguz, O.K., : "Handoff ordering using signal prediction priority queuing in personal communication systems," IEEE Transactions on Vehicular Technol., vol. 48, no.1, pp. 20-35, Jan 1999.

40. Eckberg, A.E., : "B-ISDN/ATM traffic and congestion control," IEEE Network, vol. 6, no. 5, pp. 28-37, Sept. 1992.

41. Eckberg, A.E., Doshi, B.T., and Zoccolillo, R., : "Controlling congestion in B-ISDN/ ATM issues and strategies," IEEE Commun. Magazine, vol. 29, no. 9, pp. 64-70, Sept. 1991.

42. Fernandez, R.J., and Mutka, M.W., : "A burst-level priority scheme for bursty traffic in ATM networks," Computer Networks, vol. 31, no. 1, pp. 33-45, 1999.

43. Ferracioli, M., and Verdone, R., : "A general methodology based on the handover rate for network planning of cellular radio networks based on ATM," IEEE JSAC, vol. 18, no. 7, pp. 1313-1325, July 2000.

44. Floyd, S., : "A Report on recent developments in TCP congestion control," IEEE Commun. Magazine, vol. 39, no.4, pp. 84 -90, April 2001.

45. Garcia-Haro, J., and Jajszczyk, A., : "ATM shared-memory switching architectures," IEEE Network, vol. 8, no. 4, pp.18- 26, July/Aug. 1994.

46. Gerla, M., and Pazos-Rangel, R.A., : "Bandwidth allocation and routing in ISDN's," IEEE Communications Magazine, vol. 22, no. 2, pp. 16-26, Feb. 1984.

47. Ghani, N., and Mark, J.W., : "Enhanced distributed explicit rate allocation for ABR services in ATM networks," IEEE/ACM Trans. on Networking, vol. 8, no. 1, pp. 71-86, Feb. 2000.

48. Gilbert, H., Aboul-Magd, O., and Phung, V., : "Developing a cohesive traffic management strategy for ATM networks," IEEE Communications Magazine, vol. 29, no. 10, pp. 36-45, Oct. 1991.

49. Golmie, N. et. al. : "ABR switch mechanisms : Design issues and performance evaluation," Computer Networks and ISDN Systems, vol. 30, no. 19, pp. 1749-1761, 1998.

50. Gonsalves, T.A., et. al., : "CygNet : Integrated management for modern telecom networks," Proceedings 6th National Conf. on Communications, Published by Viva Books Private Limited, 28th - 30th January 2000.

51. Goodman, D.J., : "Cellular packet communications," IEEE Trans. on Commun., vol. 38, no. 8, pp. 1272-1280, Aug. 1990.

52. Guerrero, L.O., and Aghvami, A.H., : "A prioritized handoff dynamic channel allocation strategy for PCS," IEEE Trans. on Veh. Technol., vol. 48, no. 4, pp. 1203-1215, July 1999.

53. Gupta, B.K., Lal, M., and Sharma, S.C., : "Fuzzy logic control for VBR traffic in ATM networks," Journal of the CSI, vol. 31, no. 2, pp. 36-40, June 2001, Published by Computer Society of India, Mumbai, INDIA.

54. Gupta, B.K., Sharma, S.C. and Lal, M., : "Evaluation of cell loss probability of shared-memory ATM switch," The Journal of CSI, vol. 30, no. 4, pp. 15-20, Published by Computer Society of India, Dec. 2000, Mumbai, INDIA.

55. Gupta, B.K., Sharma, S.C., and Lal, M., : "Analysis and simulation of non-blocking multiple input ATM switches based on input queuing," Accepted for IETE –Technical Review Journal, vol. 19, no. 6, Nov./ Dec. 2002, published by The Institute of Electronics and Telecommunications Engineers, New-Delhi, INDIA.

56. Habib, I., and Saadawi, T., : "Multimedia traffic characteristic in broadband networks," IEEE Communication Magazine, vol. 30, no.7, pp. 48-54, July 1992.

57. Habib, I.W., and Saadawi, T.N., : "Controlling flow and avoiding congestion in broadband networks," IEEE Commun. Magazine, vol. 30, no. 10, pp. 46-53, Oct. 1991.

58. Hamdi, M., and Muppala, J., : "Performance evaluation of non-blocking ATM switches under various traffic and buffering schemes," Int'l J. Commun. Syst., vol. 9, pp. 59-79, 1996.

59. Hamid, A., and Denzel, W.E., : "A survey of modern high - performance switching techniques," IEEE JSAC, vol. 7, no. 7, pp. 1091-1102, Sept. 1989.

60. Handel, R., Huber, M.N., and Schroder, S., : "ATM Networks-Concepts, Protocols, Applications," IIIrd Edition, pp. 86-87, 127-133, Published by Addison Wesley Longman, Inc.First Indian reprint, 1999.

61. Hong, D., and Suda, T., : "Congestion control and prevention in ATM Networks," IEEE Network Magazine, vol. 5, no. 4, pp. 10 -15, July 1991.

62. Hong, D., and Rappaport, S.S., : "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures," IEEE Trans. on Veh. Technol., vol. VT-35, no. 3, pp. 77-92, August 1996.

63. Huang, Nen-Fu, and Wang, Rui-Chi, : "An efficient link allocation algorithm for ATM-based personal communication network," Wireless Personal Communications, vol.4, pp.257-275, 1997.

64. Hui, J., and Arthurs, E., : "A broadband packet switch for integrated transport," IEEE JSAC, vol. 5, no. 8, pp. 1264-1273, Oct. 1987.

65. Hui, J.Y., et. al, : "A layered broadband switching architecture with physical or virtual path configurations," IEEE JSAC, vol. 9, no. 9, pp. 1416-1426, Dec. 1991.

66. Ilyas ,M., and Mouftah, H.T., : "Performance evaluation of congestion avoidance in broadband ISDNs," IEEE ICC, pp. 727-731, 1990.

67. Imer, O .C., et. al., : "Available bit rate congestion control in ATM networks," IEEE Control System Magazine, pp. 38-56, Feb. 2001.

68. ITU-T Recommendation I.371, "Traffic control and congestion control in B-ISDN," Helsinki, 1993.

69. Jacob, L. and Kumar, A., : "Comparative performance of scheduling strategies for switching and multiplexing in a hub based ATM network : A simulation study," Computer Networks and ISDN Systems, vol. 30, no. 14, pp. 1341-1354, Aug. 1998.

70. Jaffe, J.M., : "Bottleneck flow control," IEEE Trans. on Comm., vol. COM-29, no.7, pp. 954-961, July 1981.

71. Jain, R., : "Congestion control and traffic management in ATM networks : Recent advances and a survey," Computer Networks and ISDN Systems, vol. 28, pp. 1723-1738, 1996.

72. Jenq, Y.C., : "Performance analysis of a packet switch based single-buffered banyan network," IEEE JSAC, vol. sac-1, no. 6, pp. 1014-1021, Dec. 1983.

73. Jhunjhunwala, A., Ramamurthi, B., and Gonsalves, T.A., : "The role of technology in telecom expansion in India," IEEE Commun. Magazine, vol. 36, no. 11, pp.88-94, Nov. 1998.

74. Kalyanaraman, S., et. al., : "Performance and buffering requirements of internet protocols over ATM ABR and UBR Services," IEEE Commun. Magazine, vol. 36, no. 6, pp. 152-157, June 1998.

75. Kannan, R., and Ray, S., : "MSXmin : A modular multicast ATM packet switch with low delay and hardware complexity," IEEE / ACM Trans. on Networking, vol. 8, no. 3, pp. 407-418, June 2000.

76. Karol, M.J., Hluchyj, M.G., and Morgan, S.P., : "Input versus output queuing on a space division packet switch," IEEE Trans. Commun., vol. com-35, no. 12, pp.1347-1356, Dec. 1987.

77. Kermani, P., and Kleinrock, L., "Virtual cut through : A new computer communication switching technique," Computer Networks, vol. 3, pp. 267-286, Sept. 1979.

78. Kim, J. G., and Krunz, M.M., : "Bandwidth allocation in wireless networks with guaranteed packet-loss performance," IEEE/ACM Trans. on Networking , vol. 8, no. 3, pp. 337-349, June 2000.

79. Kim, Y.C., Lee, B.J., Kim, Y.S., and Mukherjee, B., : "Dynamic channel reservation based on mobility in wireless ATM networks," IEEE Commun. Mag., vol. 37, no. 11, pp. 47-51, Nov. 1999.

80. Kostas, D.J., : "Transition to ISDN - An overview," IEEE Communications Magazine vol..22, no.1, pp. 11-17, Jan.1984.

81. Kozaki,, T., et. al., : 32X32 Shared buffer type ATM switch VLSI's for B-ISDN's," IEEE JSAC, vol. 9, no. 8, pp. 1239-1247, Oct. 1991.

82. Kroner, H., Hebuterne, G., Boyer, P., and Gravey, A., : "Priority management in ATM switching nodes," IEEE JSAC, vol. 9, no. 3, pp. 418-427, Apr. 1991.

83. Kulavaratharasah, M.D., and Aghvami, A.H., : "Teletraffic performance evaluation of microcellular personal communication networks (PCNs) with prioritized handoff procedures," IEEE Transactions on Vehicular Technol., vol. 37, no.2, pp. 92-103, May 1998.

84. Kumar, S. and Agrawal, D.P. : "On multicast support for shared-memory based ATM switch architecture," IEEE Network, vol. 10, no. 1, pp. 34-39, Jan./Feb. 1996.

85. Kumar, V., Gupta, B.K., Sharma, S.C., and Lal, M., : "Fuzzy traffic controller in Asynchronous Transfer Mode (ATM) Networks," Proc. IE (I) Conf. on Information - Tech., pp. II-1 to II- 8, 14th Nov.1999, Palghat (Kerala), INDIA.

86. Kuri, J. and Gagnaire, M., : "ATM traffic management in an LMDS wireless access Network," IEEE Commun. Magazine, vol. 39, no. 9, pp.128-133, Sept. 2001.

87. Kwong, S.D., Park, H.M., Choi, B.S., and Park, J.H. : "Performance of a new cell loss mode of ATM switch with input and output buffers," Proc. 7th International Conf. on Advanced Computing and Communications (ADCOM), pp. 225-229, 20th–22nd Dec.1999, Roorkee, INDIA.

88. Lal, M., Sarje, A.K., and Yadav, B.K., : "Management of multicast traffic in shared memory ATM switches," Proc. All India Seminar on Information-Technology and Applications: Emerging Trends, pp. 56-58, 23rd –24th Dec. 1998, Roorkee, India.

89. Lal, M., Sarje, A.K., and Bohra, P., : "Fuzzy system based queue management scheme for congestion control in ATM networks," in Proc. International Conf. on Cognitive System, New Delhi, India, pp.408-420, Dec.1998.

90. Lal, M., Sharma, S.C., and Gupta, B.K., : "Congestion control with dynamic feedback threshold in ATM networks - Fuzzy approach," 9th Int'l Conf. on Advance Computing and Communications(ADCOM), pp. 277, 16-19th Dec.2001, Bhubaneswar, INDIA.

91. Lal, M., Sharma, S.C., Gupta, B.K. and Kumar, V., : "Dynamic queue length threshold for multicast support in shared-memory-based ATM switch architecture," Paper presented in 7th Int'l Conf. on Advance Computing and Communications (ADCOM), 20th-22nd Dec. 1999, Roorkee, INDIA.

92. Lapid, Y., et. al., : "Analysis of discarding policies in high-speed networks," IEEE JSAC, vol. 16, no. 5, pp. 764-777, June 1998.

93. Lea, C.T., : "A multicast broadband packet switch," IEEE Trans. on Communications, vol. 41, no. 4, pp. 621- 630, April 1993.

94. Lea, C.T., : "What should be the goal for ATM," IEEE Network, vol.6, no. 5, pp. 60-66, Sept. 1992.

95. Lee, J.Y., and Un, C.K., : "Performance analysis of an input and output packet switch with a priority packet discarding scheme," IEE Proc.– Communications, vol. 142, no. 2, pp. 67-74, April 1995.

96. Lee, M.J., and Ahn, D.S., : "Cell loss analysis and design trade-offs of non-blocking ATM switches with non-uniform traffic," IEEE/ACM Trans. Networking, vol.3, no. 2, pp. 199-209, April 1995.

97. Lee, T.T. : "Nonblocking copy networks for multicast packet switching," IEEE JSAC, vol. 6, no. 9, pp. 1455-67, Dec. 1998.

98. Lefelhocz, C., et. al., : "Congestion control for best-effort service : Why we need a new paradigm," IEEE Network, vol. 10, no. 1, pp. 10-19, Feb.1996.

99. Leslie, I.M., : "Shrinking network development time-scales : Flexible virtual networks," Electronics and Communication Engineering Journal, pp. 149-154, June 1999.

100. Li, S.Q.,: " Non-uniform traffic analysis on a non-blocking space division packet switch," IEEE Trans. Commun., vol. 38, no. 7, pp.1085-1096, July 1990.

101. Li, S.Q., : "Performance of a non-blocking space division packet switch with correlated traffic," IEEE Trans. Commun., vol. 40, no. 1, pp. 97-108, Jan. 1992.

102. Li, S.R., : "Algorithms for flow control and call set-up in multihop B-ISDN," IEEE INFOCOM, pp. 727-731, 1990.

103. Li, S.Y., : "Theory of periodic contention and its application to packet switching," Proceedings of INFOCOM, pp. 320-325, Mar. 1988.

104. Liu, Y.C., and Douligeris, C., : "Rate regulation with feedback controller in ATM networks-a neural network approach," IEEE JSAC, vol. 15, no. 2, pp. 200-210, Feb.1997.

105. Liu,Y., and Douligeris, C., : "Adaptive vs. static feedback congestion controller," Proc. IEEE GLOBECOM '95, Singapore, pp. 291-295, Nov. 1995.

106. Low, C.P., : "An efficient algorithm for link allocation problem on ATM-based personal communication network," IEEE JSAC, vol. 18, no. 7, pp. 1279-88, July 2000.

107. Markhamreh, I.I., et. al., : " Analysis of an output-buffered ATM switch with speed-up constraints under correlated and imbalanced bursty traffic," IEE Proc. Communications, vol. 142, no. 2, pp. 61-66, April 1995.

108. Marsan, M. A., Chiasserini, C.F., Cigno, R.L., and Maurizio, M., : "Local and global handovers for mobility management in wireless ATM network", IEEE Personal Communication, vol. 4, no. 5, pp.16-24 October 1997.

109. Marsan, M.A., et. al., : " Efficient estimation of call blocking probabilities in cellular mobile telephony networks with customer retrials," IEEE JSAC, vol. 19, no. 2, pp.332-345, Feb. 2001.

110. McGuire, A.,: "Management of optical transport networks," Electronics and Communication Engineering Journal, pp. 155-163, June 1999.

111. McKeown, N., et. al., : "Tiny Tera: A packet switch core," IEEE Micro, vol. 17, pp. 26-33, Jan./Feb. 1997.

112. Medhi, D., : " Models for network design, servicing and monitoring of ATM networks based on the virtual path concept," Computer Networks and ISDN System, vol.29, no. 3, pp. 373-386, 1997.

113. Medhi, D., : "Multi-hour, multi-traffic class network design for virtual path-based wide area dynamically reconfigurable ATM networks," IEEE/ACM Trans. on Networking , vol. 3, pp. 809-18, 1995.

114. Minzer, S.E., : "Broadband ISDN and Asynchronous Transfer Mode(ATM)," IEEE Communications Magazine, vol. 27, no. 9, pp. 7-24, Sept. 1989.

115. Mitra, D., : "Asymptotically optimal design of congestion control for high speed data networks," IEEE Trans. Communications, vol.40, no. 2, pp. 301-311, Feb. 1992.

116. Morris, D., and Pronk, V. : "Charging for ATM services," IEEE Commun. Magazine, vol. 37, no. 5, pp. 133-139, May 1999.

117. Motoyama, S., et. al., : "An iterative cell scheduling algorithm for ATM input-queued switch with service class priority," IEEE Commun. Letters, vol. 3, no. 12, pp. 323-325, Dec. 1999.

118. Munakata, T., and Jani, Y., : "Fuzzy systems : An overview," Commun. ACM., vol. 37, no. 3, pp. 69-76, Mar. 1994.

119. Murata, M., et al., : "Analysis of a discrete-time single server queue with bursty input for traffic control in ATM networks," Proc. IEEE, pp. 170-189, Feb. 1991.

120. Mynbaev, D.K., and Scheiner, L.L., : "Fiber-Optic Communications Technology," pp. 668, 683, Published by Addison Wesley Longman (Singapore) Pte. Ltd. Ist Indian Reprint, 2001.

121. Naghshineh, M., and Schwartz, M., : “Distributed call admission control in mobile /wireless network,” IEEE JSAC, vol. 14, no. 4, pp 711-716, May 1996.

122. Ndousse, T.D., : “Fuzzy neural control of voice cells in ATM networks,” IEEE JSAC, vol. 12, no. 9, pp. 1488-1494, Dec. 1994.

123. Newman, P., : “Backward explicit congestion notification for ATM local area network,” Proc. IEEE GLOBECOM '93, Houston, TX, pp. 719-723, 1995.

124. Ofek, Y., and Y., Bulent, : “Reliable concurrent multicast from bursty sources,” IEEE JSAC, vol. 15, no. 3, pp. 432-444, April 1997.

125. Okada, T., Ohnishi, H., and Morita, N., : Traffic control in Asynchronous Transfer Mode,” IEEE Communication Magazine, vol. 29, no. 9, pp. 58-62, Sept. 1991.

126. Oliveira, C., et al., : “An adaptive bandwidth reservation scheme for high-speed multimedia wireless networks,” IEEE JSAC, vol. 16, no. 6, pp. 858-873, Aug. 1998.

127. Oliver, T. W., and Victor, C.M., : “Adaptive resource allocation for prioritized call admission over an ATM-based wireless PCN,” IEEE JSAC, vol.15, no. 7, pp. 1208-1225, Sept. 1997.

128. Pattavina, A., and Bruzzi, G., : “Analysis of input and output queuing for non-blocking ATM switches,” IEEE/ACM Trans. Networking, vol. 1, no. 3, pp. 314-328, June 1993.

129. Peter, D.W., et. al., : “Priority discarding of speech in integrated packet networks,” IEEE JSAC, vol. 7, no. 5, pp. 644-656, June 1989.

130. Phillips, C., and Kirkby, P., : “ATM and the future of telecommunication networking,” Electronics and Communication Engineering Journal, vol. 11, no. 3, pp. 108-124, June 1999.

131. Pitsillides, A., Sekercioglu, Y. A., and Ramamurthy, G., : “Effective control of traffic flow in ATM networks using fuzzy explicit rate marking (FERM),” in IEEE JSAC, vol. 15, no. 2, pp.209-225, Feb.1997.

132. Pollini, G.P., : "Trends in handover design," IEEE Commun. Magazine, vol. 34, no. 3, pp. 82–90, March 1996.

133. Prycker, M. de, Peschi, R., and Landegem, T. Van, : "B-ISDN and the OSI protocol reference model," IEEE Network Magazine, vol. 7, no. 2. pp. 10-18, March 1993.

134. Prycker, M. de, : "Asynchrounous Transfer Mode solution for Broadband ISDN", IInd Edition. Chichester England: Ellis Horwood. pp. 112-18, 157, 159-68, 289, 323, 1993.

135. Ramaswamy, A., Murthy, H.A., and Gonsalves, T.A., : "Linear Prediction for Network Management," Proceedings 6th National Conf. on Communications, Published by Viva Books Private Limited, 28-30 January 2000.

136. Ramjee, R., et al., : "Performance evaluation of connection rerouting scheme for ATM-based wireless networks," IEEE/ACM Trans. on Networking vol. 6, no. 3, pp. 249-260, June 1998.

137. Raychaudhari, D., : "Wireless ATM networks : Architecture, system design and prototyping," IEEE Personal Communs., vol. 3, no. 4, pp. 42-49, Aug. 1996.

138. Raychaudhari, D., and Wilson, N.D., : "ATM-based transport architecture for multiservices wireless personal communication networks," IEEE JSAC, vol. 12, no. 8, pp. 1401-1414, October 1992.

139. Re, E.D., and Fantacci, R., : "Performance evaluation of input and output queueing techniques in ATM switching systems," IEEE Trans. on Commun., vol. 41, no. 10, pp. 1565-1575, Oct. 1993.

140. Ross, T.J., : "Fuzzy logic with engineering applications," McGraw Hill Inc., International Edition., pp. 87-147, 1997.

141. Ross, I.M., : "Wireless network directions," IEEE Communication Magazine, vol. 29, no. 2, pp. 40-42, Feb. 1991.

142. Saito, H., : "Optimal queueing discipline for real-time traffic at ATM switching nodes," IEEE Trans. Commun., vol. 38, no. 12, pp. 2131-2136, Dec. 1990.

143. Samecki, J., et. al, : "Microcell design principles," IEEE Communication Magazine, vol. 31, no. 4, pp. 76-82, April 1993.

144. Sharma, S.C., and Gupta, B.K., : "Developments of congestion control in ATM networks : A comparative study," Proc. of National Seminar on Role of Information Technology in 21st Century Challenges and Strategies, pp.16-25, Jabalpur (M.P), INDIA, 11th-12th Feb.1999.

145. Sharma, S.C., Lal, M., and Gupta, B.K., : "Performance analysis of non-blocking multiple input ATM switches for VBR bursty traffic," Proc. of National Conf. on Technology Convergence for Information, Communication and Entertainment, (NICE-2001), pp. 160-164, 23rd-24th Feb. 2001, Cochin, INDIA.

146. Sharma, S.C., Lal, M., Gupta, B.K., and Tyagi, N., : "Queue length management of multicast traffic for shared memory Asynchronous Transfer Mode(ATM) switch," Proc. of National Conf. on Towards Excellence in Technical Education in 21st Century, pp. 295-299, 23rd - 24th Nov. 1999, Pusad (Maharastra), INDIA.

147. Sharma, S.C., Lal, M., Kumar, V., and Gupta, B.K., : "Fuzzy system based queue management scheme for multicast bursty traffic in shared memory of ATM switches," Proc. 5th International Conference on Cognitive System, pp. 331-339, 15th-18th Dec. 1999, New-Delhi, INDIA.

148. Sestini, F., : "Recursive copy generation for multicast ATM switching," IEEE / ACM Trans. on Networking, vol. 5, no. 3, pp. 329-335, June 1997.

149. Shi, Y., et al., : "Implementation of evolutionary fuzzy systems," IEEE / ACM Trans. on Networking, vol. 7, no. 2, pp. 109-119, April 1999.

150. Shiomoto, K., and Yamanaka, N., : "Dynamic burst transfer time-slot-base network," IEEE Commun. Magazine, vol. 37, no. 10, pp. 88-96, Oct. 1999.

151. Sigle, R., Renger, T., : "Impact of wireless access on traffic management in ATM networks," Computer Networks, vol. 31, no. 9/10, pp. 985-997, 1999.

152. Sivabalan, M. and Mouftah, H.T., : "QUARTS-II : A routing simulator for ATM networks," IEEE Communications Magazine, vol. 36, no. 5, pp. 80-87, May 1998.

153. Smith, A., et. al., : "Available Bit Rate–A new service for ATM," Computer Networks and ISDN Systems, vol. 28, pp. 635-640, 1996.

154. Steenhaut, K., et. al., : "Scheduling and admission control policies : A case study for ATM ," Computer Networks and ISDN Systems, vol. 29, no. 5, pp. 539-554, 1997.

155. Su, C.F., et. al., : "Explicit rate flow control for ABR services in ATM networks," IEEE / ACM Trans. on Networking, vol. 8, no. 2, pp. 350 - 361, June 2000.

156. Tajima, J., and Imamura, K., : " A strategy for flexible channel assignment in mobile communication systems," IEEE Transactions on Vehicular Technol., vol. 48, no.1, pp. 137-152, Jan 1999.

157. Tanenbaum, A.S., : "Computer Networks," IIIrd Edition, Published by Prentice Hall of India Pte. Limited, New - Delhi, pp. 61-64, 1997.

158. Tekinay, S., and Jabbari, B., : "A measurement-based prioritization scheme for handovers in mobile cellular network," IEEE JSAC, vol. 10 no. 8, pp. 1343-1350, October 1992.

159. Tekinay, S., and Jabbari, B., : "Handover and channel assignment in mobile cellular network," IEEE Communications Magazine, vol. 29, no. 11, pp. 42-46, Nov. 1991.

160. The ATM Forum, Technical Committee, : "Traffic Management, Version 4.1," pp. 43-45, af-tm-0121.000, March 1999.

161. The ATM Forum, : "Traffic Management Specification, Version 4.0," Draft Version Apr. 1995.

162. Tobagi, F. A., : "Fast packet switch architectures for broadband integrated services digital networks," Proc. IEEE, vol. 78, no. 1. pp.133-167, Jan. 1990.

163. Trajkovic, L., and Golestani, S.J., : "Congestion control for multimedia services," IEEE Network, vol. 6, no. 5, pp. 20-26, Sept. 1992.

164. Turer, J. S., : "New directions in communications (or which way to the information age)," IEEE Communications Magazine, vol. 24, no. 10, pp. 8-15, Oct. 1986.

165. Turner, J.S., : "Design of a broadcast packet switching network," IEEE Trans. Communications, vol. 36, no. 6, pp. 734-743, June 1988.

166. Turner, J.S., : "Managing bandwidth in ATM networks with bursty traffic," IEEE Network, vol. 6, no. 5, pp. 50-58, Sept. 1992.

167. Uehara, K., and Hirota, K., : "Fuzzy connection admission control for ATM networks based on possibility distribution of cell loss ratio," IEEE JSAC, vol. 15, no. 2, pp. 179-190, Feb. 1997.

168. Venkatesulu, D., and Gonsalves, T.A., : "Efficient fault-tolerant reliable broadcast in a multi-switch extended LAN," Computer Communications, vol. 22, pp. 266-278, 1999.

169. Walrand, J., and Varaiya, P., : "High Performance Communication Networks," IInd Ed. Published by Harcourt Asia Pte. Ltd. Singapore, pp. 588-93, Ist Indian Reprint 2000.

170. Wernik, M., Aboul-Magd, O., and Gilbert, H., : "Traffic management for B-ISDN services," IEEE Network, vol. 6, no. 5, pp. 10-19, Sept. 1992.

171. Woodruff, G.M., et. al., : "Multimedia traffic management principles for guaranteed ATM network performance," IEEE JSAC, vol. 8, no. 3, pp. 437-446, Apr. 1990.

172. Yang, C. Q., and Reddy, A.V.S., : " A taxonomy for congestion control algorithms in packet switching networks," IEEE Network Magazine, vol. 9, no. 5, July/ August 1995.

173. Yapark, E, Chronopoulos, A.T., Psarris, K., and Xiao, Y., : "Dynamic buffer allocation in an ATM switch," Computer Networks, vol. 31, pp.1927-1933, 1999.

174. Yeh, Yu-S. et. al., : "The knockout switch : A simple, modular architecture for high-performance packet switching," IEEE JSAC, vol. sac-5, no. 8, pp. 1274 - 1284, Oct. 1987.

175. Yin, N., Li, S.Q., and Stern, T.E., : "Congestion control for packet voice by selective packet discarding," IEEE Trans. Communications, vol. 38, no. 5, pp. 674-683, May 1990.

176. Yoon, C.H., and Un., C.K., : "Performance of personal portable radio telephone system with and without guard channels," IEEE JSAC, vol. 11, no. 6, pp. 911-917, Aug. 1993.

177. Yurcik, W., Tipper, D., and Medhi, D., : "The use of hop-limits to provide survivable ATM group communications," 2nd Annual Workshop on Networked Group Communications, 8th – 10th Nov. 2000, Palo Alto, CA, USA.

178. Zadeh, L.A., : "Fuzzy logic, neural networks and soft computing," Communications ACM., vol. 37, no. 3, pp. 69-76, Mar. 1994.

179. Zadeh, L.A., : "Fuzzy Sets," Information Control, vol. 8, pp. 338-353, 1965.

180. Zhang, L., : "Designing a new architecture for packet switching communication networks," IEEE Communications Magazine, vol. 25, no. 9, pp. 5-12, Sept. 1987.

181. Zheng, B., and Atiquzzaman, M., : "Traffic management of multimedia over ATM networks," IEEE Communications Magazine, vol. 37, no.1, pp. 33-38, Jan. 1999.