M.Sc. Engg. Thesis

QoS Support for Multimedia Sessions in the Diffserv-Aware
ATM based MPLS Network

By
Mohammad Ali

Submitted to
Department of Computer Science & Engineering
In partial fulfillment of the requirements for the degree of
M.Sc. Engg. (Computer Science & Engineering)

Department of Computer Science & Engineering
Bangladesh University of Engineering & Technology
Dhaka-1000, Bangladesh

April 20, 2004
The thesis, "QoS Support for Multimedia Sessions in the DiffServ-Aware ATM based MPLS Network" submitted by Mohammad Ali, Roll No. 100105014P, Session October, 2001 to the Department of Computer Science and Engineering of Bangladesh University of Engineering and Technology has been accepted as satisfactory for partial fulfillment of the requirements for the degree of M.Sc. Engg. in Computer Science and Engineering and approved as to its style and contents. The examination has been held on April 20, 2004.

Board of Examiners

1
(Dr. Md. Mostofa Akbar)
(Supervisor)
Assistant Professor
Department of CSE
BUET, Dhaka-1000

2
(Dr. Md. Shamsul Alam)
Professor and Head
Department of CSE
BUET, Dhaka-1000

3
(Dr. M. Kaykobad)
Professor
Department of CSE
BUET, Dhaka-1000

4
(Dr. Md. Monirul Islam)
Assistant Professor
Department of CSE
BUET, Dhaka-1000

5
(Dr. Md. Saifur Rahman)
Professor
Department of EEE
BUET, Dhaka-1000
Declaration

I do, hereby, declare that the work presented in this thesis is done by me under the supervision of Dr. Md. Mostofa Akbar, Assistant Professor, Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000. I also declare that neither this thesis nor any part thereof has been submitted elsewhere for the award of any degree or diploma.

Countersigned

[Signature]

Dr. Md. Mostofa Akbar
(Supervisor)

[Signature]

Mohammad Ali
(Mohammad Ali)
To Hasan Arham Bin Mohammad, my only son

and

to my parents

Md. Fazlul Rahman & Mazeda Khatun,

Md. Eusuf Ali & Suraiya Khanum
Acknowledgement

I would like to express my sincere thanks to my supervisor Dr. Md. Mostofa Akbar for his guidance, thoughtful suggestions, encouragement, support and belief in me. I would like to thank Md. Kazi Rafiqul Alam, Executive Director, Dhaka Ahsania Mission, for his support. Without them this thesis would not have been possible.

I would like to thank Dr. Md. Saifur Rahman and Dr. M. Kaykobad for their suggested corrections in the thesis.

I also want to thank all the members of Ahsanullah Institute of Information & Communication Technology (AIICT) and the teachers of Ahsanullah University of Science & Technology (AUST). I really enjoyed the suggestions of Md. Shahriar, Md. Eunus Ali and Md. Abdul Hakim Newton.

It is impossible for me to name each and every person who contributed to this work, but I would like to thank my wife for her encouragement and everyone else who has participated in making my time enjoyable.
Abstract

Communication links with high data throughput capacity (bandwidth) and low latency are the prime requirements to support multimedia applications such as voice on demand services, video conferencing and virtual reality. These applications require a particular Quality of Service (QoS) defined by the delivery of voluminous continuous media (audio or video) data packets over the network with a certain delay and jitter bound. Best-effort service is unsuitable for all the above-mentioned requirements because all packets compete equally for network resources. The necessity to develop better QoS solutions to address these issues of multimedia transmission has led Internet Engineering Task Force (IETF) to develop Integrated Services (Intserv) architecture with Resource Reservation Protocol (RSVP) and Differentiated Services (Diffserv) architecture. In Intserv architecture, RSVP is used to set up paths and reserve resources towards receiver before sending data. But in large networks it enjoys the drawbacks of scalability.

In this thesis a new flow-based admission control algorithm has been proposed for routing the multimedia traffics with a predefined QoS through a Diffserv-Aware ATM based MPLS network. The routers in the networks work as distributed admission controllers to find a particular path satisfying a particular QoS. A scheme of rerouting multimedia flows during the congestion in a particular part of the network is also presented.

Admission controlling would be done by the router connected to the source of the multimedia traffic by observing the current delay and jitter of the calculated path for the multimedia request. Virtual Circuits (VC) in an ATM-based MPLS network are used to reduce end-to-end delay and router load and for higher throughput. In the proposed technique, QoS of a multimedia session can be violated if congestion occurs in a router on its path. Diverting traffic towards low congested path is a common technique used by the routers to avoid congestion in the network. Consideration has been given in one particular rerouting option during congestion, to guarantee the overall QoS. However, this rerouting decision would be made by the source of the multimedia traffic. Unlike the
traditional routing of data packets, the intermediate routers on the flow of a multimedia session have nothing to do with this routing scheme.

Based upon the knowledge of the customers' Service Level Agreements (SLAs), simulation results are also presented that show the improvement of session success rate by rerouting the multimedia traffic to relatively under-utilized links during congestion.
# Table of Contents

Title Page ................................................................. i  
Declaration .................................................................. iii  
Acknowledgement .............................................................. v  
Abstract ....................................................................... vi  
Table of Contents ............................................................. viii  
List of Figures .................................................................. xi  
List of Tables .................................................................... xiv  
Glossary of Terms .............................................................. xv  

1. Introduction and Literature Review ......................................... 1  
   1.1 Overview ............................................................... 1  
   1.2 Motivation .............................................................. 1  
   1.3 Related Works .......................................................... 2  
   1.4 Scope and Focus ....................................................... 4  
   1.5 Outline of the Thesis ................................................... 4  

2. QoS Requirements and Issues to Transmit Multimedia Traffic ........................................... 6  
   2.1 Introduction .............................................................. 6  
   2.2 Factors Affecting the Quality of Service (QoS) ......................... 6  
      2.2.1 Delay ............................................................... 6  
      2.2.2 Jitter (Delay Variability) ............................................ 7  
      2.2.3 Packet Loss and Out of Order Packets ......................... 7  
   2.3 Efforts to ensure QoS .................................................... 8  
   2.4 Integrated Services (Intserv) .......................................... 8  
   2.5 Differentiated Services (Diffserv) ...................................... 9  
      2.5.1 Per-Hop Behaviors (PHBs) ....................................... 10  
      2.5.2 Differentiated Services Code Points (DSCPs) .................. 10  
      2.5.3 Services provided .................................................. 11  
      2.5.4 Network Resource Allocation .................................... 13  
   2.6 Multiprotocol Label Switching (MPLS) .............................. 14  
      2.6.1 Classification of Packets using Labels in MPLS Networks ...... 15  
      2.6.2 Mechanism of Label Swapping .................................. 16  
      2.6.3 MPLS Traffic Engineering Mechanisms ....................... 18  
      2.6.4 Resource Reservation Protocol (RSVP) ...................... 20  
      2.6.5 CR-LDP ........................................................... 21
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5</td>
<td>Background Source Sink</td>
<td>69</td>
</tr>
<tr>
<td>5.6</td>
<td>LER and LSR Node Models</td>
<td>69</td>
</tr>
<tr>
<td>5.6.1</td>
<td>Root Module</td>
<td>70</td>
</tr>
<tr>
<td>5.6.2</td>
<td>QoS Buffer Module</td>
<td>72</td>
</tr>
<tr>
<td>5.6.3</td>
<td>Scheduler Module</td>
<td>74</td>
</tr>
<tr>
<td>5.6.4</td>
<td>Sources</td>
<td>75</td>
</tr>
<tr>
<td>5.7</td>
<td>Packets Forwarding using MPLS</td>
<td>76</td>
</tr>
<tr>
<td>5.8</td>
<td>MPLS Service Mapping</td>
<td>78</td>
</tr>
<tr>
<td>5.9</td>
<td>Random Number Generation</td>
<td>79</td>
</tr>
<tr>
<td>5.10</td>
<td>Mean Calculation</td>
<td>79</td>
</tr>
<tr>
<td>5.11</td>
<td>Variance Calculation</td>
<td>80</td>
</tr>
<tr>
<td>5.12</td>
<td>Experimental Results and Discussion</td>
<td>80</td>
</tr>
<tr>
<td>5.13</td>
<td>Summary</td>
<td>93</td>
</tr>
<tr>
<td>6.</td>
<td>Conclusion</td>
<td>94</td>
</tr>
<tr>
<td>6.1</td>
<td>Major Contributions</td>
<td>94</td>
</tr>
<tr>
<td>6.2</td>
<td>Future Scope of Work</td>
<td>95</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>97</td>
</tr>
<tr>
<td>Curriculum Vitae</td>
<td></td>
<td>101</td>
</tr>
</tbody>
</table>
List of Figures

Figure 2-1 Logical view of a packet classifier and traffic conditioner ........................................... 12
Figure 2-2 An MPLS Domain ........................................................................................................... 14
Figure 2-3 Forwarding Equivalence Class ....................................................................................... 15
Figure 2-4 MPLS 'Shim' header ...................................................................................................... 16
Figure 2-5 Forwarding packets in an MPLS domain ..................................................................... 17
Figure 2-6 Extensions to RSVP to establish an ERLSP ................................................................. 20
Figure 2-7 CR-LDP Path Establishment ........................................................................................ 21
Figure 2-8 Fundamental difference between RSVP and CR-LDP ............................................... 23
Figure 2-9 Basic format of an ATM Cell .......................................................................................... 24
Figure 2-10 ATM cell header (UNI Format) .................................................................................. 25
Figure 2-11 ATM cell header (NNI Format) .................................................................................. 25
Figure 2-12 ATM Virtual Path and Virtual Connection ................................................................. 27
Figure 3-1 The proposed ATM-based MPLS network architecture ............................................ 35
Figure 3-2 Next Hop Resolution Protocol .................................................................................... 37
Figure 3-3 The Diffserv model in the proposed MPLS router system ......................................... 40
Figure 3-4 Queuing architecture ..................................................................................................... 42
Figure 3-5 Diffserv policy types and their deployment ................................................................. 44
Figure 4-1 Proposed algorithm for admission control ................................................................. 51
Figure 4-2 Congestion detected in the best effort traffic ............................................................... 52
Figure 4-3 Packet loss calculated at the input buffers ................................................................. 52
Figure 4-4 Average time spent by a packet in the System ............................................................ 60
Figure 4-5 Packet loss for a given utilization & buffer size ......................................................... 60
Figure 4-6 P_balking vs Offered utilizations for different hop number ....................................... 61
Figure 5-1 The network used for simulation ............................................................................... 64
Figure 5-2 Parent process source module processing states ....................................................... 65
Figure 5-3 Child process tagged source processing states ......................................................... 66
Figure 5-4 Source module process relationship with Tagged sources' child process...66
Figure 5-5 Buffer stream ........................................................................................................67
Figure 5-6 Parent process background source state machine .................................................67
Figure 5-7 Child process on off source state machine ..........................................................68
Figure 5-8 Parent background source module relationship with Child on off processes...68
Figure 5-9 Simulation model of LER/LSR ...........................................................................69
Figure 5-10 Finite state machine of the root module ............................................................72
Figure 5-11 Finite state machine for QoS buffer module .......................................................74
Figure 5-12 Simulation model of the scheduler .................................................................75
Figure 5-13 Finite state machine for the scheduler ............................................................75
Figure 5-14 Source .................................................................................................................76
Figure 5-15 Queuing delay (simulation) for EF1 and average delay (theory) with different utilization ..........................................................................................................................81
Figure 5-16 Queuing delay (simulation) for EF2 and average delay (theory) with different utilization ..........................................................................................................................82
Figure 5-17 Queuing delay (simulation) for AF1 and average delay (theory) with different utilization ..........................................................................................................................82
Figure 5-18 Queuing delay (simulation) for AF2 and average delay (theory) with different utilization ..........................................................................................................................83
Figure 5-19 Queuing delay (simulation) for AF3 and average delay (theory) with different utilization ..........................................................................................................................83
Figure 5-20 Queuing delay (simulation) for AF4 and average delay (theory) with different utilization ..........................................................................................................................84
Figure 5-21 Queuing delay (simulation) for BE and average delay (theory) with different utilization ..........................................................................................................................84
Figure 5-22 Packet loss (simulation) for EF1 and average packet loss (theory) with different utilization ..........................................................................................................................85
Figure 5-23 Packet loss (simulation) for EF2 and average packet loss (theory) with different utilization ..........................................................................................................................86
Figure 5-24 Packet loss (simulation) for AF1 and average packet loss (theory) with different utilization

Figure 5-25 Packet loss (simulation) for AF2 and average packet loss (theory) with different utilization

Figure 5-26 Packet loss (simulation) for AF3 and average packet loss (theory) with different utilization

Figure 5-27 Packet loss (simulation) for AF4 and average packet loss (theory) with different utilization

Figure 5-28 Packet loss (simulation) for BE and average packet loss (theory) with different utilization

Figure 5-29 Data served vs. Utilization

Figure 5-30 Average end to end delay for various utilizations vs. rerouting policy

Figure 5-31 Packet loss for various utilizations

Figure 5-32 Session success rate vs. % Reroute

Figure 5-33 Average waiting time with different loads and routing probability

Figure 5-34 Session success rate vs. Offered load in different class of services
List of Tables

Table 2-1 AF code points ........................................................................................................11
Table 2-2 Service categories and characteristics ...................................................................28
Table 4-1 Congestion indication table ....................................................................................53
Table 4-2 Average time spent in system per packets ..............................................................58
Table 4-3 Packet loss for different utilizations & hop numbers using buffer size of 100 packets .....................................................................................................................60
Table 5-1 On Off source simulation parameters ....................................................................76
Table 5-2 Proposed QoS mapping ..........................................................................................78
### Glossary of Terms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAL</td>
<td>ATM Adaptation Layer</td>
</tr>
<tr>
<td>AF</td>
<td>Assured Forwarding</td>
</tr>
<tr>
<td>ATM</td>
<td>Asynchronous Transfer Mode</td>
</tr>
<tr>
<td>ATM-LSR</td>
<td>Asynchronous Transfer Mode- Label Switching Routers</td>
</tr>
<tr>
<td>BA</td>
<td>Behavior Aggregate</td>
</tr>
<tr>
<td>BE</td>
<td>Best Effort</td>
</tr>
<tr>
<td>BGP</td>
<td>Border Gateway Protocol</td>
</tr>
<tr>
<td>CBS</td>
<td>Committed Burst Size</td>
</tr>
<tr>
<td>CBR</td>
<td>Committed Bit Rate</td>
</tr>
<tr>
<td>CDR</td>
<td>Committed Data Rate</td>
</tr>
<tr>
<td>CDVT</td>
<td>Cell Delay Variation Tolerance</td>
</tr>
<tr>
<td>CIR</td>
<td>Committed Information Rate</td>
</tr>
<tr>
<td>CLP</td>
<td>Cell Loss Priority</td>
</tr>
<tr>
<td>CLR</td>
<td>Cell Loss Ratio</td>
</tr>
<tr>
<td>CoS</td>
<td>Class of Service</td>
</tr>
<tr>
<td>CR-LDP</td>
<td>Constraint-based Routing Label Distribution Protocol</td>
</tr>
<tr>
<td>Diffserv</td>
<td>Differentiated Services</td>
</tr>
<tr>
<td>DS</td>
<td>Differentiated Services</td>
</tr>
<tr>
<td>DSCP</td>
<td>Differentiated Services Code Point</td>
</tr>
<tr>
<td>ECN</td>
<td>Explicit Congestion Notification</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>EF</td>
<td>Expedited Forwarding</td>
</tr>
<tr>
<td>ER</td>
<td>Explicit Route</td>
</tr>
<tr>
<td>E-LSP</td>
<td>EXP-Inferred-PSC LSPs</td>
</tr>
<tr>
<td>EXP</td>
<td>Experimental</td>
</tr>
<tr>
<td>FEC</td>
<td>Forwarding Equivalent Class</td>
</tr>
<tr>
<td>FIFO</td>
<td>First In First Out</td>
</tr>
<tr>
<td>FR</td>
<td>Frame Relay</td>
</tr>
<tr>
<td>FTN</td>
<td>FEC-to-NHLFE</td>
</tr>
<tr>
<td>GFC</td>
<td>Generic flow control</td>
</tr>
<tr>
<td>HEC</td>
<td>Header Error Control</td>
</tr>
<tr>
<td>IETF</td>
<td>Internet Engineering Task Force</td>
</tr>
<tr>
<td>ILM</td>
<td>Incoming Label Map</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>Intserv</td>
<td>Integrated Services</td>
</tr>
<tr>
<td>ITU</td>
<td>International Telecommunication Union</td>
</tr>
<tr>
<td>LC-ATM</td>
<td>Label Switching Controlled ATM</td>
</tr>
<tr>
<td>ISP</td>
<td>Internet Service Providers</td>
</tr>
<tr>
<td>LDP</td>
<td>Label Distribution Protocol</td>
</tr>
<tr>
<td>LER</td>
<td>Label Edge Router</td>
</tr>
<tr>
<td>LIS</td>
<td>Logical IP Subnetwork</td>
</tr>
<tr>
<td>LSP</td>
<td>Level Switch Path</td>
</tr>
<tr>
<td>LSR</td>
<td>Label Switching Routers</td>
</tr>
<tr>
<td>L-LSP</td>
<td>Label-Only-Inferred-PSC LSPs</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>MBS</td>
<td>Maximum Burst Size</td>
</tr>
<tr>
<td>MCR</td>
<td>Minimum Cell Rate</td>
</tr>
<tr>
<td>MCTD</td>
<td>Maximum Cell Transfer Delay</td>
</tr>
<tr>
<td>MF</td>
<td>Multi-Field</td>
</tr>
<tr>
<td>MPLS</td>
<td>Multi Protocol Label Switching</td>
</tr>
<tr>
<td>MPLS WG</td>
<td>Multi Protocol Label Switching Working Group</td>
</tr>
<tr>
<td>MOPA</td>
<td>Multi Protocol Over ATM</td>
</tr>
<tr>
<td>NHLFE</td>
<td>Next Hop Label Forwarding Entry</td>
</tr>
<tr>
<td>NHRP</td>
<td>Next Hop Resolution Protocol</td>
</tr>
<tr>
<td>NHS</td>
<td>Next Hop Server</td>
</tr>
<tr>
<td>NBMA</td>
<td>Non-Broadcast Multiple-Access</td>
</tr>
<tr>
<td>NNI</td>
<td>Network-Network Interface</td>
</tr>
<tr>
<td>OA</td>
<td>Ordered Aggregate</td>
</tr>
<tr>
<td>OSPF</td>
<td>Open Shortest Path First</td>
</tr>
<tr>
<td>PBS</td>
<td>Peak Burst Size</td>
</tr>
<tr>
<td>PCR</td>
<td>Peak Cell Rate</td>
</tr>
<tr>
<td>PDR</td>
<td>Peak Data Rate</td>
</tr>
<tr>
<td>PHB</td>
<td>Per-Hop Behavior</td>
</tr>
<tr>
<td>PQ</td>
<td>Priority Queuing</td>
</tr>
<tr>
<td>PSC</td>
<td>PHB Scheduling Class</td>
</tr>
<tr>
<td>PT</td>
<td>Payload Type</td>
</tr>
<tr>
<td>PVC</td>
<td>Permanent Virtual Connection</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>RFC</td>
<td>Request for Comments</td>
</tr>
<tr>
<td>RSVP</td>
<td>Resource Reservation Protocol</td>
</tr>
<tr>
<td>SBM</td>
<td>Subnet Bandwidth Manager</td>
</tr>
<tr>
<td>SCR</td>
<td>Sustainable Cell Rate</td>
</tr>
<tr>
<td>SLA</td>
<td>Service Level Agreement</td>
</tr>
<tr>
<td>SVC</td>
<td>Switched Virtual Connection</td>
</tr>
<tr>
<td>TCA</td>
<td>Traffic Conditioning Agreement</td>
</tr>
<tr>
<td>TCS</td>
<td>Traffic Conditioning Specification</td>
</tr>
<tr>
<td>TOS</td>
<td>Type of Service</td>
</tr>
<tr>
<td>UNI</td>
<td>User-Network Interface</td>
</tr>
<tr>
<td>VBR</td>
<td>Variable Bit Rate</td>
</tr>
<tr>
<td>VC</td>
<td>Virtual Circuits</td>
</tr>
<tr>
<td>VCC</td>
<td>Virtual Channel Connection</td>
</tr>
<tr>
<td>VCI</td>
<td>Virtual Channel Identifier</td>
</tr>
<tr>
<td>VoD</td>
<td>Video On Demand</td>
</tr>
<tr>
<td>VoIP</td>
<td>Voice Over IP</td>
</tr>
<tr>
<td>VP</td>
<td>Virtual Path</td>
</tr>
<tr>
<td>VPI</td>
<td>Virtual Path Identifier</td>
</tr>
<tr>
<td>WFQ</td>
<td>Weighted Fair Queuing</td>
</tr>
</tbody>
</table>
1. Introduction and Literature Review

1.1 Overview

Ensuring QoS for the multimedia applications over the Internet Protocol (IP) datagram based network is a major challenge in the area of computer networks. In this thesis a new flow-based admission control algorithm has been proposed for routing the multimedia traffics with a predefined QoS through a Diffserv-Aware ATM based MPLS network. The routers in the networks work as distributed admission controllers to find a particular path satisfying a particular QoS. A scheme of rerouting some of the flows during the congestion in a particular part of the network is also presented. Integration of different types of traffic in a single network requires an admission control mechanism, which operates according to a resource reservation mechanism. In addition, we consider the ATM shortcut connection in an ATM-based MPLS network has been considered to reduce the end-to-end delay [1-4].

1.2 Motivation

Multimedia refers to composite media, media that contains multiple information streams of various types, such as video, image, sound, text etc. Some multimedia streams such as video and audio require absolute guaranteed QoS. There may be associated time constraints on data rates, delay, jitter and inter stream synchronization. The delivery of multimedia information with absolute guaranteed QoS has presented Internet Service Providers (ISPs) with the following challenges and opportunities [5] that lead us to work for QoS support for multimedia sessions in the Diffserv-Aware ATM based MPLS Network.

- Revenue from commercial multimedia services like VoD, high-quality videoconferencing, interactive video services, Internet phone and WebTV could help rescue the technology sector from the current recession. However, ISPs must be able to provide guaranteed QoS to realize these opportunities.
• Networks must be able to carry multimedia streams with guaranteed QoS. However, the present Internet is based on best effort connectionless datagram service, without any QoS guarantees. This service without any guarantee is not suitable for paid services.

• All the IP datagrams of the session must be routed along the fixed path because best effort datagram service does not guarantee timely delivery of multimedia streams on a particular path.

• Each packet of multimedia stream should contain the information of required class of service that describes latency, jitter, packet loss etc. for reserving bandwidth to meet the user requirements.

• The resources of a network are finite. If these resources are overbooked then QoS must not be maintained. Therefore some form of admission control is necessary to reject some of the applicant session when insufficient resources are available to serve all of them.

• To maximize the profit some multimedia flow can be rerouted to optimize the resources of the network.

1.3 Related Works

The transmission of multimedia streams with a particular and guaranteed level of QoS from the source to destination according to a generalized and flexible form of the traditional SLA [5,6] between users and network owners is a major challenge in multimedia communications over IP networks.

The basic goal of DiffServ architecture is to fulfill the performance requirements of the multimedia sessions that are defined in an SLA [7]. In DiffServ model the way each router in the path treats a packet is referred to as *Per-Hop Behavior* (PHB). The default PHB is best effort type but there are more PHBs already defined or being standardized. Examples of PHB are *Expedited Forwarding* (EF) and *Assured Forwarding* (AF). EF [8] requires the departure rate of the traffic (from the node) to be equal or greater than a configurable rate and it has two independent forwarded classes. AF [9] is a PHB group that provides
packet delivery in four independent forwarded classes with three levels of drop precedence.

The primary objective of MPLS is to integrate label switched forwarding with network layer routing [10]. A fundamental concept in MPLS is that two Label Switching Routers (LSRs) must agree on the meaning of the labels used to forward traffic between them and through them. This understanding is achieved by employing a set of signalling procedures, called a Label Distribution Protocol (LDP) [11], by which one LSR informs another of the label or Forward Equivalent Class (FEC) bindings it has made.

Since MPLS is a core technology, the focus of QoS support in MPLS networks should be “aggregation”, i.e. to provide the means for ensuring individual end-to-end QoS guarantees without maintaining awareness of individual flows on each and every segment of their path. Diffserv is therefore a good candidate to provide QoS within MPLS networks because services are based on a per-hop model and aggregate forwarding resources (buffer space, bandwidth, scheduling policy) would be pre-allocated in each MPLS node for each service type [7].

There considerable work on admission control algorithm has been done already, the references of which may be found in the Internet. Lee [12] has proposed an integrated packet scheduler to accommodate Intserv and Best Effort traffic through the Diffserv-aware MPLS core network. Jamin [13] has worked on measurement-based admission control algorithm for integrated service packet networks. Gibbens [14] demonstrates an algorithm embedded in the end system to block an IP-telephony call when the network load is high. Kelly [15] proposed an Explicit Congestion Notification (ECN) probe based admission control that is fast, scalable and robust. However, no admission control mechanism has ever depicted to address multimedia session to be rerouted to fulfill guaranteed QoS requirements.
1.4 Scope and Focus

The main focus of this thesis is to present a flow based admission control algorithm to transmit multimedia data over the network and a policy of rerouting some of the flows during the congestion in a particular part of the network. The proposed flow based admission control algorithm for the different services model on the ATM-based MPLS network optimizes the network by using admission control according to traffic classification. This algorithm would be suitable for multiple service class environments. The proposed flow-based admission control considers QoS and buffer statistics. This algorithm makes an admission decision in accordance with the conventional measurement-based admission control. Consequently, this algorithm can achieve two objectives: reliable delay bound QoS and high resource utilization. Also, this reduces the blocking probability of high priority service class flow in the operating Differentiated Services. To analyze the performance of the admission control algorithm a discrete event simulation scheme has been developed.

1.5 Outline of the Thesis

A list of acronyms and terms is provided in the glossary. In addition, a list of the figures used throughout the thesis is included at the beginning of the thesis. The main body of the report consists of six chapters including the introduction and the conclusion. Each chapter begins with a brief description of its scope and ends with a brief summary of the outcome. Details of each chapter are provided below.

The thesis describes in detail an admission control algorithm within Differentiated Services aware ATM based MPLS Network. The focus of the contribution is on the rerouting of traffic flows along under utilized Level Switch Path (LSPs) when congestion is detected. This introductory chapter summarizes the contribution by depicting focus and related works.

Chapter 2 relates QoS requirements and issues to transmit multimedia traffic with Differentiated Services, MPLS and ATM. Then in Chapter 3 the proposed network architecture is introduced along with its various components and mechanisms. Admission control algorithm developed by the author is described along with the numerical result in Chapter
4. Simulation model and experimental results are described in Chapter 5. Finally, Chapter 6 concludes with the major contribution of this thesis and presents suggestions for the future research.
2. QoS Requirements and Issues to Transmit Multimedia Traffic

2.1 Introduction

In recent years, the most active area in networking is - data, voice, and video integration. Business users are beginning to combine real-time applications such as voice and video, which have a limited tolerance for network latency, with non-real time data traffic. With Voice Over IP (VoIP) technology - defined as the ability to make telephone calls (real time voice) over IP-based data networks with a suitable QoS and a much superior cost benefit - systems can provide simultaneous voice and Internet access over the same connection, or integrate existing phone connections with the Internet through VoIP Gateways.

QoS refers to the capability of a network to provide better service to selected network traffic over various technologies including Ethernet and 802.1 networks, Wireless networks, IP-routed networks, ATM and Frame Relay (FR) [6]. It can also be interpreted as a method of providing preferential treatment to some arbitrary amount of network traffic, as opposed to all traffic being treated as Best Effort (BE).

2.2 Factors Affecting the Quality of Service (QoS)

2.2.1 Delay

Echo and talker overlap are the problems that result from high end-to-end delay in a voice network. It is a requirement that the round trip delay should be less than 50 ms to avoid echo. Since VoIP has longer delays, such systems must address the need for echo control and implement some means of echo cancellation. Talker overlap (problem of one caller stepping on the other talker's speech) becomes significant if the one-way delay becomes greater than 250 ms. Video conferencing is a real-time application with high interaction that requires average 400 ms delay. Terminal sessions require interactive data transactions with 400 ms delay bound. Web applications should have the same QoS requirements as terminal sessions. Non real time applications (ftp, email) require 1 sec delay bound.
Delay can be attributed to - encoder delay, packetization delay (time to fill an IP packet), input delay (time to join together fragmented IP packets), queuing delay, switching delay (depends on technology such as Layer 3 switches, tag switching system like MPLS and ATM switches) and propagation delay (~5 microsecond/km).

2.2.2 Jitter (Delay Variability)
This is the variation in the inter-packet arrival time (leading to gaps, known as jitter, between packets) as introduced by the variable transmission delay over the network. Removing jitter requires collecting packets in buffers and holding them long enough to allow the slowest packets to arrive in time to be played in correct sequence.
VoIP is a jitter sensitive, real-time high interaction application that supports 50 ms jitter bound. Video Conferencing is also a jitter sensitive real-time application that supports 50 ms jitter bound. Terminal sessions that require interactive data transactions should have no jitter. Web applications should have the same QoS requirements as terminal session. Non-real-time applications (ftp, email) may have larger jitter. Jitter can be attributed to - queuing delay and switching delay [40].

2.2.3 Packet Loss and Out of Order Packets
In traditional network layer routing, as a router receives a packet it makes an independent forwarding decision for that packet. Each router analyses the packet's header and performs a best match routing table lookup to make an independent decision as to what the next hop for the packet should be. As a result, IP networks do not guarantee delivery of packets, much less in order. Packets will be dropped under peak loads and during periods of congestion. Approaches used to compensate for packet loss include interpolation of speech by re-playing the last packet, and sending of redundant information. Out of order packets are treated as lost and replayed by their predecessors. When the late packet finally arrives, it is discarded.
VoIP and Video conferencing require ordered packets with limited packet loss (maximum 1 packet in $10^4$ packets). Terminal session and Web application require ordered packets
with no packet loss. Non real time applications (ftp, email) support out of order packets with no packet loss.

2.3 Efforts to ensure QoS

There are different proposals to enhance the Internet in order to support the different requirements of different types of traffic and they are as follows:

- For service differentiation
  1. Integrated Services (Intserv)
  2. Differentiated Services (Diffserv)
- For multimedia packets switching
  1. MPLS
  2. ATM Switches
- Combination of service differentiation and packets switching
  1. Diffserv aware MPLS
- ATM based MPLS Network
  1. ATM switches as LSR

The following sections present the above-mentioned proposals in greater details.

2.4 Integrated Services (Intserv)

The Integrated Services architecture for the Internet was proposed in RFC 1633 to support real-time traffic as well as BE traffic. Here resources for traffic are explicitly identified and reserved. Network nodes classify incoming packets and use reservations to provide differentiated services. It performs resource reservation using a dynamic signalling protocol and employs admission control, packet classification, and intelligent scheduling to achieve the desired QoS. This model is relatively complex and has difficulties in scaling to large backbones. This architecture is based on the RSVP. RSVP is an IETF Internet Standard (RFC 2205) protocol for allowing an application to dynamically reserve network bandwidth. It enables applications to request a specific QoS for a data flow. Hosts and routers use RSVP to deliver QoS requests to the routers along
the path, and to maintain router and host state to provide the requested service, usually bandwidth and latency. Bandwidth reservation is based on mean data rate, the largest amount of data the router will keep in its queue, and the minimum QoS required. The specific standards and definitions for services developed by the Intserv working group in the IETF fall under Guaranteed QoS. Technologies that can provide guaranteed service for portions of the end-to-end connection include:

- IP-WFQ combined with RSVP signalling or guaranteeing bandwidth on a single link
- Ethernet- subnet Bandwidth Manager (SBM) (when used with a compliant switch)
- ATM-Variable Bit Rate (VBR) and Constant Bit Rate (CBR), and
- Frame Relay- Committed Information Rate (CIR).

The Integrated services architecture [16], with RSVP [17,18] as main protocol has severe scalability problems. It has been proven to be unsuitable for large networks since every single micro flow is treated individually. Therefore, this architecture cannot be implemented in the core of the Internet where a large number of flows coexist.

2.5 Differentiated Services (Diffserv)

The basic goal of Differentiated Services architecture [19] is to fulfill the performance requirements of the users. Users request a certain performance level and the network provides it as long as the user traffic has certain characteristics. The performance level provided and the characteristics of the traffic to be injected in the network are defined in an SLA [5].

The part of the SLA dealing with technical details is referred to as Service Level Specification (SLS). Inside the SLS, the Traffic Conditioning Specification (TCS) specifies the expected performance (throughput, drop probability, latency...), the profile of the traffic to be used (peak data rate, burst size) and actions to perform in case of excess traffic.
As a result the differentiated services architecture is based on a simple model where traffic entering a network is classified and possibly conditioned at the boundaries of the network, and assigned to different Behavior Aggregates (BA). Each behavior aggregate is identified by a single Differentiate Services Code Point (DSCP). Within the core of the network, packets are forwarded according to the PHB associated with the DSCP [20].

2.5.1 Per-Hop Behaviors (PHBs)

The way each router in the path treats a packet is referred to as PHB. The default PHB is best effort but there are more PHBs already defined or being standardized. There are also PHBs that are locally defined inside a node and do not correspond to a well-known set of features. All PHBs are local to the node implementing them.

The definition of a PHB does not include the specific algorithm to be employed. It outlines a set of requirements that have to be fulfilled in order to provide that specific PHB. Examples of PHB are Expedited Forwarding (EF) and Assured Forwarding (AF).

EF [8] requires the departure rate (from the node) to be equal or greater than a configurable rate. Traffic to which EF PHB applies is forwarded as soon as possible independently of the state of the node. EF traffic is not delayed in queues as far as it is possible. EF PHB is used for very high priority traffic.

AF [9] is a PHB group that provides packet delivery in four independent forwarded classes with three levels of drop precedence. Packets that belong to a class are not reordered in the node and nodes do not aggregate two classes together.

2.5.2 Differentiated Services Code Points (DSCPs)

Once PHBs have been defined for individual nodes there is a need of providing coherent treatment to the packets along the whole path. Every node traversed by packets that belong to a certain class has to apply the same PHB to the packets. This is achieved by adding a tag to the packets, which describes the PHB required.
This information is placed in the so-called Differentiated Services (DS) field. The DS field supersedes the existing definitions of Type of Service (TOS) in IPv4 and the traffic class octet in IPv6.

The TOS field contains eight bits. Six of them are used to represent the DSCP. Two bits remain unused.

The default best-effort service corresponds to the DSCP equal to zero (DSCP='000000'). The standard DSCP for the EF PHB is '101110'. The table below shows the standard code points for the AF PHB group.

Table 2-1 AF code points

| Low drop Precedence | Class 1 | 001010 | Class 2 | 010010 | Class 3 | 011010 | Class 4 | 100010 |
| Medium drop Precedence | 001100 | 010100 | 011100 | 100100 |
| High drop Precedence | 001110 | 010110 | 011110 | 100110 |

2.5.3 Services provided

Making use of the DSCPs and the respective PHBs triggered in the nodes, it is possible to provide the users with differentiated services. Inside the SLA, the user accepts to send data according to the Traffic Condition Agreement (TCA), and the network compromises to provide a certain level of service.

The edge of the network has to undertake the classification of the flows and check if they are inside the TCA. A set of devices is implemented in the edge routers for this purpose. They are classifiers, meters, markers, droppers and shapers.

Figure 2-1 shows the block diagram of a classifier and traffic conditioner [19]. Note that a traffic conditioner may not necessarily contain all four elements. For example, in the case where no traffic profile is in effect, packets may only pass through a classifier and a marker.
Packet classifiers select packets in a traffic stream based on the content of some portion of the packet header. There are two types of classifiers. The BA classifier classifies packets based on the DSCP only. The Multi-Field (MF) classifier selects packets based on the value of a combination of one or more header fields, such as source address, destination address, DS field, protocol ID, source port and destination port numbers, and other information such as incoming interface [19].

Classifiers are used to "steer" packets matching some specified rule to an element of a traffic conditioner for further processing. Classifiers must be configured by some management procedure in accordance with the appropriate TCA. It is worth while here to mention that in the event of upstream packet fragmentation, MF classifiers, which examine the contents of transport-layer header fields, may incorrectly classify packet fragments subsequent to the first.

Traffic meters measure the temporal properties of the stream of packets selected by a classifier against a traffic profile specified in a TCA. A meter passes state information to other conditioning functions to trigger a particular action for each packet, which is either in- or out-of-profile (to some extent).

Packet markers set the DS field of a packet to a particular code point, adding the marked packet to a particular DS behavior aggregate. According to the state of a meter, the marker may be configured to mark all packets, which are steered to it to a single code point, or may be configured to mark a packet to one of a set of code points used to select a PHB in a PHB group. When the marker changes the code point in a packet it is said to have "re-marked" the packet.
Shapers delay some or all of the packets in a traffic stream in order to bring the stream into compliance with a traffic profile. A shaper usually has a finite-size buffer, and packets may be discarded if there is not sufficient buffer space to hold the delayed packets.

Droppers discard some or all of the packets in a traffic stream in order to bring the stream into compliance with a traffic profile. This process is known as "policing" the stream. It can be mentioned that a dropper can be implemented as a special case of a shaper by setting the shaper buffer size to zero (or a few) packets.

2.5.4 Network Resource Allocation

The implementation, configuration, operation and administration of the supported PHB groups in the nodes of a DS domain should effectively partition the resources of those nodes and the inter-node links between behaviour aggregates, in accordance with the domain's service provisioning policy. Traffic conditioners can further control the usage of these resources through enforcement of TCAs and possibly through operational feedback from the nodes and traffic conditioners in the domain [17]. Although a range of services can be deployed in the absence of complex traffic conditioning functions (e.g., using only static marking policies), functions such as policing, shaping, and dynamic re-marking enable the deployment of services providing quantitative performance metrics.

The configuration of and interaction between traffic conditioners and interior nodes should be managed by the administrative control of the domain and may require operational control through protocols and a control entity. There is a wide range of possible control models.

However, scalability requires that the control of the domain does not require micro-management of the network resources. The most scalable control model would operate nodes in open loop in the operational time frame, and would only require administrative-time scale management as SLAs are varied. This simple model may be unsuitable in some circumstances, and some automated but slowly varying operational control (minutes
rather than seconds) may be desirable to balance the utilization of the network against the recent load profile.

2.6 Multiprotocol Label Switching (MPLS)

A key concept in MPLS is the separation of an IP router's function into two parts: forwarding and control [21-23]. The forwarding part is responsible for how data packets are relayed between IP routers, using label swapping similar to ATM switching's Virtual Path (VP)/Virtual Channel Identifier (VCI). The control part consists of network layer routing protocols to distribute routing information between routers, and label binding procedures for converting this routing information into the forwarding tables needed for label switching. It should be noted that MPLS is not a routing protocol - but is a fast-forwarding mechanism that is designed to work with existing Internet routing protocols such as Open Shortest Path First (OSPF) [24], or the Border Gateway Protocol (BGP) [25]. Figure 2-2 illustrates a typical MPLS domain that consists of Label Switch Router (LSR), Label-Edge Router (LER) and Label Switch Path (LSP) [23].

![Figure 2-2 An MPLS Domain](image)

An LSR is a device that is capable of forwarding packets at layer 3 and forwarding frames that encapsulate the packet at layer 2.

An LER is both a router and a layer 2 switches that is capable of forwarding MPLS frames to and from an MPLS domain. It performs the IP to MPLS Forwarding Equivalence Class (FEC) binding including the aggregation of incoming flows. It also
communicates with interior MPLS LSRs to exchange label bindings. Often referred to as an ingress or egress LSR, because it is situated at the edge of a MPLS domain.

### 2.6.1 Classification of Packets using Labels in MPLS Networks

An FEC is a set of packets that are treated identically by a router, i.e., forwarded out the same interface with the same next hop and label, and assigned the same class of service. When a packet enters the MPLS domain at the ingress node, it is mapped into an FEC. The mapping can be done according to a number of factors, i.e., the address prefix, source/destination address pair, or ingress interface. At the current moment there are three defined FEC elements, an address prefix, router ID and flow (source/destination port and IP addresses). A group of IP packets that are forwarded over the same path and treated in the same manner can be mapped to a single label by an LSR, as shown in the Figure 2-3 [24].

![Figure 2-3 Forwarding Equivalence Class](image)

A label is a short, fixed length, locally significant identifier that is used to identify an FEC. A packet may be assigned to an FEC based on its network layer destination address; however, the label does not directly encode any information from the network layer header. A labelled packet is a packet into which a label has been encoded. The label may reside in an encapsulation header that exists specifically for this purpose, called an MPLS "shim" header, or a stack entry. Alternatively, it may reside in an existing data link as long
as there is a field that is available for that purpose [26,27]. The 32-bit MPLS header contains the following fields:

- Label field (20-bits), carries the actual label value.
- Experimental field (3-bits), not yet defined.
- S (1-bit), Stack supports a hierarchical label stack.
- TTL (8-bits) Time-To-Live an inherent part of IP functionality.

![Figure 2-4 MPLS 'Shim' header](image)

2.6.2 Mechanism of Label Swapping

Label Swapping uses the following procedures to forward labelled and unlabelled packets (a packet into which a label has not been encoded).

In forwarding a labelled packet, an LSR examines the label at the top of the label stack, and examines the *Incoming Label Map* (ILM) to map this label to a *Next Hop Label Forwarding Entry* (NHLFE) [28,29]. With this information it determines where this packet needs to be forwarded to, and performs an operation on the packet's label stack. This operation may mean swapping the incoming label with a new output label, or replacing the label with a new label and pushing a new label on top, or simply "popping" the label and examining network layer header. It then encodes the new label stack into the packet and forwards it.

In forwarding a packet that is received as unlabelled packet, an LSR determines which FEC to assign the packet to by examining the packet's network layer header. Once the packet has been assigned to an FEC, the LSR uses the *FEC-to-NHLFE* (FTN) to map this to an NHLFE. Using the information in the NHLFE, it determines where to forward the packet, and performs an operation on the packet's label stack. It then encodes the new
label stack into the packet and forwards it. Figure 2-5 illustrates forwarding packets in an MPLS domain [30].

![Figure 2-5 Forwarding packets in an MPLS domain](image)

An LSP is an ingress-to-egress switched path built by MPLS nodes to forward the MPLS encapsulated packets of a particular FEC using the Label Swapping forwarding mechanism. It is similar to the concept of virtual channels within an ATM context.

### 2.6.2.1 The Next Hop Label Forwarding Entry (NHLFE)

The NHLFE is used in forwarding a labelled packet. It contains the following information:

- the packet's next hop
- the operation to be performed on the packet's label stack. One of the following operations needs to be performed:
  a) replace the label at the top of the label stack with a specified new label
  b) pop the label stack
  c) replace the label at the top of the label stack with a specified new label, and then push one or more specified new labels onto the label stack.

It may also contain:

- the data link encapsulation to use when transmitting the packet
- the way to encode the label stack when transmitting the packet
- any other information needed in order to properly dispose of the packet.
It should be noted here that at a given LSR, the packet's "next hop" might be that LSR itself. In this case, the LSR would need to pop the top-level label, and then "forward" the resulting packet to it. It would then make another forwarding decision, based on what remains after the label stacked is popped. This may still be a labelled packet, or it may be the native IP packet.

This implies that in some cases the LSR may need to operate on the IP header in order to forward the packet. If the packet's "next hop" is the current LSR, then the label stack operation MUST be to "pop the stack".

2.6.2.2 FEC-to-NHLFE Map (FTN)

The FTN maps each FEC to a set of NHLFEs. It is used when forwarding packets that arrive unlabeled, but which are to be labelled before being forwarded.

If the FTN maps a particular label to a set of NHLFEs that contains more than one element, exactly one element of the set must be chosen before the packet is forwarded.

2.6.2.3 Incoming Label Map (ILM)

The ILM maps each incoming label to a set of NHLFEs. It is used when forwarding packets that arrive as labeled packets.

If the ILM maps a particular label to a set of NHLFEs that contains more than one element, exactly one element of the set must be chosen before the packet is forwarded.

Having the ILM map a label to a set containing more than one NHLFE may be useful if it is desired to do load balancing over multiple equal-cost paths.

2.6.3 MPLS Traffic Engineering Mechanisms

Traffic Engineering within MPLS arises from the network operators' requirement to provide a network infrastructure that is dependable and offers consistent network performance. Traffic Engineering allows network operators the ability to re-route traffic flows away from the "least cost" path calculated by routing protocols and into potentially less congested physical paths through the network. As a result of the unprecedented
growth in demand for network resources and the competitiveness amongst providers, traffic engineering has become the primary application for MPLS. The goal of traffic engineering is to efficiently use the limited network resources such that no individual component i.e., router or a link is over utilized or underutilized [30].

The ability of MPLS to support explicit routes, operate over any media infrastructure and to collect statistics regarding LSPs, suggests it is well suited to provide traffic engineering capabilities.

The IETF have proposed two different protocols for reserving resources within MPLS namely, Constraint-Based Routing Using LDP (CR-LDP) [31], and RSVP with extensions [29,30], to cater for traffic engineering within an MPLS domain.

The route for a given LSP can be established in two ways, control driven (i.e., hop-by-hop LSP), or Explicitly Routed LSP (ER-LSP). When setting up a hop-by-hop LSP, each LSR determines the next interface to route the LSP based on its layer 3 routing topology databases, and sends the label request to the L3 next hop. When setting up an ER-LSP, the route for the LSP is specified in the "setup" message itself, and this route information is carried along the nodes the setup message traverses. All the nodes along the ER-LSP will follow the route specification and send the label request to the next indicated interface. While the hop-by-hop LSP follows the path that normal layer 3 routed packets will take, the ER-LSP can be specified and controlled by the network operators or network management applications to direct the network traffic, independent of the L3 topology. In this way ER-LSP may be used to achieve network traffic engineering.

MPLS also provides flexibility for network operators to manage their traffic with both strict and loose explicit routing. In the case of a strict ER-LSP, a network operator specifies the exact full route (nodes and interfaces) that the ER-LSP will traverse. Loose ER-LSPs allow some flexibility for routing and rerouting options, and minimizes configuration overhead. In addition, a loose segment can be adaptive by moving to a new route according to the changes incurred in the layer 3 routing table. However, this kind of route change is not always desirable due to stability and control requirements of the
operators. In this case, the loose segment provides a "pinning" mechanism, meaning that an alternative route will only be tried when failure happens. The following subsections describe how RSVP and CR-LDP can be used in MPLS networks to reserve resources.

2.6.4 Resource Reservation Protocol (RSVP)

Classical RSVP as specified in RFC 2205 [32,33], allows routers the flexibility to retain their connectionless transport behavior, however RSVP has scalability issues when the number of sessions increases in the network as stated in RFC 2208 [17]. To enable RSVP to be deployed within an MPLS environment, the existing protocol needs to be augmented. RSVP protocol messages are augmented with new objects to support label allocation, distribution and binding, along with explicit routes. Notable changes have been introduced to the existing RSVP protocol infrastructure including modification to the "soft state" mechanism, where messages are sent periodically to maintain the path, and the refresh mechanisms amongst others to enable RSVP to support ER-LSPs, all of which have been documented in [32]. Figure 2-6 illustrates the flow of RSVP messages in establishing a LSP [33].

Figure 2-6 Extensions to RSVP to establish an ERLSP
2.6.5 CR-LDP

CR-LDP has its foundations in the existing LDP protocol, and is extended to incorporate the explicit route information. An explicit route is represented in a label request message as a list of nodes along a constraint-based route. If the requested path can satisfy the resources required, labels are allocated by means of Label Mapping messages. Further details of CR-LDP procedures and features can be found in [31]. Figure 2-7 illustrates the flow of messages when CR-LDP is used to establish a LSP [33].

2.6.6 Comparative Analysis of RSVP and CR-LDP

CR-LDP is a part of LDP and employs the same mechanisms and messages as LDP for discovery, session, establishment, maintenance, label distribution and error handling. Enabling LDP/CR-LDP to provide network providers with a unified distribution and path setup mode for MPLS, thus maximizing operational efficiency.

RSVP with the appropriate extensions is able to operate in the downstream on demand label allocation mode. However, if other MPLS modes are required, i.e., downstream unsolicited, then both LDP and RSVP protocols must be present in the network. This adds complexity and has a negative impact on the planning and operational costs. Another
disadvantage of this situation is the need to manage more than one network, and the objective of MPLS was to move away from that.

CR-LDP uses the reliable transport of TCP, so error notification messages are delivered in an orderly manner. RSVP runs on raw IP transport and cannot guarantee fast failure notification, as a result of this traffic may not be re-routed until the "clean up timeout" interval has expired which is undesirable in communication networks [34,35].

CR-LDP uses "hard state" controlled paths that scale well as the number of ER-LSPs increases in the network. This is because unlike the soft state case, once a path is set up there are no additional messages needed to maintain the path, keeping the number of messages needed to establish, maintain and release ER-LSPs to a minimum, thus allowing CR-LDP to scale well. RSVP on the other hand has a scalability problem, as documented in RFC 2208. As the number of paths through a node increases, the number of soft-state refresh messages to maintain the paths also increases. As documented in RFC 2208 the computational requirement on the routers increase proportionally with the number of sessions.

In summarizing, CR-LDP is an open standard protocol, proposed and accepted by the IETF MPLS working group and also the ITU SG13 [33]. It does not depend on other protocols that are outside the range of the MPLS WG, thus providing a few advantages. It can be enhanced to adopt new network requirements, and it promotes interoperability as documented in [33]. In contrast, RSVP extensions have not yet shown a clear solution for interoperability in the networks. In terms of traffic engineering, technically, both CR-LDP and RSVP provide similar signalling functionality. However, major modifications to make RSVP applicable for traffic engineering has reduced its feasibility in an MPLS network [34,35]. Only Cisco is a strong proponent of the RSVP with extensions approach. As such, this focuses on the former signalling mechanism. Figure 2-8 illustrates the principle difference between RSVP and CR-LDP [36].
2.7 Asynchronous Transfer Mode (ATM)

ATM is a technology designed for the high-speed transfer of voice, video, and data through public and private networks using cell relay technology. ATM is an International Telecommunication Union Telecommunication Standardization Sector (ITU-T) standard. Ongoing work on ATM standards is being done primarily by the ATM Forum, which was jointly founded by Cisco Systems, NET/ADAPTIVE, Northern Telecom, and Sprint in 1991. A cell switching and multiplexing technology, ATM combines the benefits of circuit switching (constant transmission delay, guaranteed capacity) with those of packet switching (flexibility, efficiency for intermittent traffic). To achieve these benefits, ATM uses the following features:

- Fixed-size cells, permitting more efficient switching in hardware than is possible with variable-length packets
- Connection-oriented service, permitting routing of cells through the ATM network over virtual connections, sometimes called virtual circuits, using simple connection identifiers
- Asynchronous multiplexing, permitting efficient use of bandwidth and interleaving of data of varying priority and size

*Figure 2-8 Fundamental difference between RSVP and CR-LDP*
The combination of these features allows ATM to provide different categories of service for different data requirements and to establish a service contract at the time a connection is set up. This means that a virtual connection of a given service category can be guaranteed a certain bandwidth, as well as other traffic parameters, for the life of the connection.

2.7.1 ATM Basics

To understand how ATM can be used, it is important to have a knowledge of how ATM packages and transfers information. The following sections provide brief descriptions of the format of ATM information transfer and the mechanisms on which ATM networking is based.

2.7.1.1 ATM Cell Basic Format

The basic unit of information used by ATM is a fixed-size cell consisting of 53 octets, or bytes. The first 5 bytes contain header information, such as the connection identifier, while the remaining 48 bytes contain the data, or payload (see Figure 2-9). Because the ATM switch does not have to detect the size of a unit of data, switching can be performed efficiently. The small size of the cell also makes it well suited for the transfer of real-time data, such as voice and video. Such traffic is intolerant of delays resulting from having to wait for large data packets to be loaded and forwarded.

![Figure 2-9 Basic format of an ATM Cell](image)

2.7.1.2 ATM Cell Header Formats

The ATM cell includes a 5-byte header. Depending upon the interface, this header can be in either User-Network Interface (UNI) or Network-Network Interface (NNI) format. The UNI cell header, as depicted in Figure 2-10, has the following fields:
• **Generic Flow Control (GFC)**—provides local functions, such as flow control from endpoint equipment to the ATM switch. This field is presently not used.

• **Virtual Path Identifier (VPI) and Virtual Channel Identifier (VCI)**—VPI identifies a virtual path leg on an ATM interface. VPI and VCI together identify a virtual channel leg on an ATM interface. Concatenating such legs through switches forms a virtual connection across a network.

• **Payload Type (PT)**—indicates in the first bit whether the cell contains user data or control data. If the cell contains user data, the second bit indicates whether congestion is experienced or not, and the third bit indicates whether the cell is the last in a series of cells that represent a single AAL5 frame. If the cell contains control data, the second and third bits indicate maintenance or management flow information.

• **Cell Loss Priority (CLP)**—indicates whether the cell should be discarded if it encounters extreme congestion as it moves through the network.

• **Header Error Control (HEC)**—contains a cyclic redundancy check on the cell header.

![Figure 2-10 ATM cell header (UNI Format)](image)

The NNI cell header format, depicted in Figure 2-11, includes the same fields except that the GFC space is displaced by a larger VPI space, occupying 12 bits and making more VPIs available for NNIs.

![Figure 2-11 ATM cell header (NNI Format)](image)
2.7.1.3 ATM Services

There are three general types of ATM services:

- *Permanent Virtual Connection* (PVC) service—connection between points is direct and permanent. In this way, PVC is similar to a leased line.

- *Switched Virtual Connection* (SVC) service—connection is created and released dynamically. Because the connection stays up only as long as it is in use (data is being transferred), an SVC is similar to a telephone call.

- Connectionless service—similar to *Switched Multimegabit Data Service* (SMDS)

Advantages of PVCs are the guaranteed availability of a connection and that no call setup procedures are required between switches. Disadvantages include static connectivity and that they require manual administration to set up.

Advantages of SVCs include connection flexibility and call setup that can be automatically handled by a networking device. Disadvantages include the extra time and overhead required to set up the connection.

2.7.1.4 Virtual Paths and Virtual Channels

ATM networks are fundamentally connection oriented. This means that a virtual connection needs to be established across the ATM network prior to any data transfer. ATM virtual connections are of two general types:

- VPCs, identified by a VPI.

- *Virtual channel connections* (VCCs), identified by the combination of a VPI and a VCI.

A virtual path is a bundle of virtual channels, all of which are switched transparently across the ATM network on the basis of the common VPI. A VPC can be thought of as a bundle of VCCs with the same VPI value (see Figure 2-12).
2.7.1.5 Traffic Contracts and Service Categories

ATM connections are further characterized by a traffic contract, which specifies a service category along with traffic and quality of service parameters. Five service categories are currently defined, each with a purpose and its own interpretation of applicable parameters. The following sections describe the components of the traffic contract, the characteristics of the service categories, and the service-dependent ATM Adaptation Layer (AAL) that supports each of the service categories.

2.7.1.6 The Traffic Contract

At the time a connection is set up, a traffic contract is entered, guaranteeing that the requested service requirements will be met. These requirements are traffic parameters and QoS parameters:

- Traffic parameters—generally pertain to bandwidth requirements and include the following:
  - Peak Cell Rate (PCR)
  - Sustainable Cell Rate (SCR)
  - Burst tolerance, conveyed through the Maximum Burst Size (MBS)
  - Cell Delay Variation Tolerance (CDVT)
  - Minimum Cell Rate (MCR)

- QoS parameters—generally pertain to cell delay and loss requirements and include the following:
  - Maximum Cell Transfer Delay (MCTD)
2.7.1.7 The Service Categories

One of the main benefits of ATM is to provide distinct classes of service for the varying bandwidth, loss, and latency requirements of different applications. Some applications require constant bandwidth, while others can adapt to the available bandwidth, perhaps with some loss of quality. Still others can make use of whatever bandwidth is available and use dramatically different amounts from one instant to the next.

ATM provides five standard service categories that meet these requirements by defining individual performance characteristics, ranging from best effort (Unspecified Bit Rate [UBR]) to highly controlled, full-time bandwidth (Constant Bit Rate [CBR]). Table 2.2 lists each service category defined by the ATM Forum along with its applicable traffic parameters and QoS characteristics.

<table>
<thead>
<tr>
<th>Service Category</th>
<th>Traffic Parameters</th>
<th>QoS Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBR-constant bit rate</td>
<td>PCR</td>
<td>Cell Loss: low, Cell Delay: Low</td>
</tr>
<tr>
<td>VBR-RT-variable bit rate real-time</td>
<td>PCR, SCR, MBS</td>
<td>Cell Loss: low, Cell Delay: Low</td>
</tr>
<tr>
<td>VBR-NRT-variable bit rate non-real Time</td>
<td>PCR, SCR, MBS</td>
<td>Cell Loss: low, Cell Delay: Unspecified</td>
</tr>
<tr>
<td>ABR-available bit rate</td>
<td>PCR, MCR</td>
<td>Cell Loss: Unspecified, Cell Delay: Unspecified</td>
</tr>
<tr>
<td>UBR-unspecified bit rate</td>
<td>No Guarantees</td>
<td>Cell Loss: Unspecified, Cell Delay: Unspecified</td>
</tr>
</tbody>
</table>

The characteristics and uses of each service category are summarized as follows:
• CBR service provides constant bandwidth with a fixed timing relationship, which requires clocking synchronization. Because CBR traffic reserves a fixed amount of bandwidth, some trunk bandwidth might go unused. CBR is typically used for circuit emulation services to carry real-time voice and video.

• VBR-RT service provides only a partial bandwidth guarantee. Like CBR, however, some bandwidth might still go unused. Typical applications include packetized voice and video, and interactive multimedia.

• VBR-NRT service provides a partial bandwidth guarantee, but with a higher cell delay than VBR-RT. This service category is suitable for bursty applications, such as file transfers.

• ABR provides a best effort service, in which feedback flow control within the network is used to increase bandwidth when no congestion is present, maximizing the use of the network.

• UBR service provides no bandwidth guarantee, but attempts to fill bandwidth gaps with bursty data. UBR is well suited for LAN protocols, such as LAN emulation.

2.8 Diffserv Aware MPLS

In an MPLS domain [26], when a stream of data traverses a common path, a Label Switched Path (LSP) can be established using MPLS signalling protocols. At the ingress LSR, each packet is assigned a label and is transmitted downstream. At each LSR along the LSP, the label is used to forward the packet to the next hop.

In a Differentiated Service domain [17] all the IP packets crossing a link and requiring the same Diffserv behavior are said to constitute a Behavior Aggregate (BA). At the ingress node of the Diffserv domain, the packets are classified and marked with a DSCP which corresponds to their Behavior Aggregate. At each transit node, the DSCP is used to select the PHB that determines the scheduling treatment and, in some cases, drop probability for each packet.
RFC 3270 [37] specifies a solution for supporting the Diff-Serv Behavior Aggregates whose corresponding PHBs are currently defined in [8, 9, 17] over an MPLS network. This solution relies on the combined use of two types of LSPs:

LSPs, which can transport multiple Ordered Aggregates, so that the EXP field of the MPLS Shim Header conveys to the LSR the PHB to be applied to the packet (covering both information about the packet's scheduling treatment and its drop precedence).

LSPs which only transport a single Ordered Aggregate (OA), so that the packet's scheduling treatment is inferred by the LSR exclusively from the packet's label value while the packet's drop precedence is conveyed in the EXP field of the MPLS Shim Header or in the encapsulating link layer specific selective drop mechanism (ATM, Frame Relay, 802.1).

As mentioned in [17] "Service providers are not required to use the same node mechanisms or configurations to enable service differentiation within their networks, and are free to configure the node parameters in whatever way that is appropriate for their service offerings and traffic engineering objectives". Thus, the solution defined in RFC 3270 [37] gives Service Providers flexibility in selecting how DiffServ classes of service are Routed or Traffic Engineered within their domain (e.g., separate classes of services supported via separate LSPs and Routed separately, all classes of service supported on the same LSP and Routed together).

Because MPLS is path-oriented it can potentially provide faster and more predictable protection and restoration capabilities in the face of topology changes than conventional hop by hop routed IP systems [37,38].

Such capabilities are known as "MPLS protection". Since the solution presented in RFC 3270 [37] allow Service Providers to choose how Diff-Serv classes of services are mapped onto LSPs, the solution also gives Service Providers flexibility in the level of protection provided to different DiffServ classes of service (e.g., some classes of service
can be supported by LSPs which are protected while some other classes of service are supported by LSPs which are not protected).

Furthermore, the solution specified in RFC 3270 [37] achieves label space conservation and reduces the volume of label set-up/tear-down signalling where possible by only resorting to multiple LSPs for a given FEC [26] when useful or required.

This specification allows support of Differentiated Services for both IPv4 and IPv6 traffic transported over an MPLS network [16].

The solution described in RFC 3270 [37] does not preclude the signalled or configured use of the EXP bits to support ECN simultaneously with Diffserv over MPLS.

2.9 ATM switches as LSR

It will be noted that MPLS forwarding procedures are similar to those of legacy "label swapping" switches such as ATM switches. ATM switches use the input port and the incoming VPI/VCI value as the index into a "cross-connect" table, from which they obtain an output port and an outgoing VPI/VCI value. Therefore if one or more labels can be encoded directly into the fields, which are accessed by these legacy switches, then the legacy switches can, with suitable software upgrades, be used as LSRs. We will refer to such devices are referred to "ATM-LSRs".

There are three obvious ways to encode labels in the ATM cell header:

2.9.1 SVC Encoding

The VPI/VCI field is used to encode the label, which is at the top of the label stack. This technique can be used in any network. With this encoding technique, each LSP is realized as an ATM SVC, and the label distribution protocol becomes the ATM "signalling" protocol. With this encoding technique, the ATM-LSRs cannot perform "push" or "pop" operations on the label stack.
2.9.2 SVP Encoding

The VPI field is used to encode the label, which is at the top of the label stack, and the VCI field to encode the second label on the stack, if one is present. This technique has some advantages over the previous one, in that it permits the use of ATM "VP-switching". That is, the LSPs are realized as ATM SVPs, with the label distribution protocol serving as the ATM signaling protocol.

However, this technique cannot always be used. If the network includes an ATM Virtual Path through a non-MPLS ATM network, then the VPI field is not necessarily available for use by MPLS. When this encoding technique is used, the ATM-LSR at the egress of the VP effectively does a "pop" operation.

2.9.3 SVP Multipoint Encoding

The VPI field is used to encode the label which is at the top of the label stack, part of the VCI field is used to encode the second label on the stack, if one is present, and the remainder of the VCI field is used to identify the LSP ingress. If this technique is used, conventional ATM VP-switching capabilities can be used to provide multipoint-to-point VPs. Cells from different packets will then carry different VCI values.

This technique depends on the existence of a capability for assigning 16-bit VCI values to each ATM switch such that no single VCI value is assigned to two different switches. (If an adequate number of such values could be assigned to each switch, it would be possible to also treat the VCI value as the second label in the stack.)

If there are more labels on the stack than they can be encoded in the ATM header, the ATM encoding must be combined with the generic encapsulation.

2.10 Summary

In this chapter, a review has been performed on the key components of QoS within a differentiated services region, traffic classification and conditioning functions, and how
differentiated services are achieved through the combination of traffic conditioning and PHB-based forwarding.

This chapter has described the mechanisms and procedures that are encompassed by the MPLS protocol and ATM switch to provide QoS. This chapter also describes a flexible solution for support of Diffserv over ATM based MPLS networks. The next chapter describes our proposed network architecture.
3. The Proposed Network Architecture

3.1 Introduction

In this chapter an ATM based MPLS network architecture has been proposed for multiple service class environment of Differentiated Service. A mapping policy of Diffserv flows to ATM based MPLS domain is also presented. Moreover, a packet scheduler has been depicted using the proposed queuing architecture.

3.2 The Diffserv-Aware ATM-based MPLS Network Architecture

Figure 3-1 illustrates the architecture of the ATM-based MPLS network. This architecture is characterized by layering the ATM shortcut network and the MPLS core network with the Diffserv model. The ATM-based MPLS core network is composed of *ATM Label Edge Routers* (ATM-LERs) and *ATM Label Switching Routers* (ATM-LSRs). The ATM-LER is located at the ATM edge switch as an MPLS-aware ingress/egress router. The ATM-LER performs label binding based on the *Label Information Base* (LIB). The ATM-LSR performs a label swapping function and reserves the resources to build an LSP using CR-LDP.

The MPLS Architecture [26] presents the way in which ATM switches may be used as Label Switching Routers. The ATM switches run network layer routing algorithms and their data forwarding is based on the results of these routing algorithms. No ATM-specific routing or addressing is needed because path from source to destination is selected by OSPF or BGP using CR-LDP and then performs label-swapping function of ATM based MPLS domain using VPI and VCI fields.

An ATM-LSR is an LSR with a number of *Label Switching Controlled ATM* (LC-ATM) interfaces, which forwards cells between these interfaces, using labels carried in the VCI or VPI/VCI field, without reassembling the cells into frames before forwarding.
An LC-ATM interface is an ATM interface controlled by the label switching control component. When a packet traversing such an interface is received, it is treated as a labelled packet. The packet's top label is inferred either from the contents of the VCI field or the combined contents of the VPI and VCI fields. Any two LDP peers which are connected via an LC-ATM interface will use LDP negotiations to determine which of these cases is applicable to that interface.

The ATM-LER performs admission control and integrated packet scheduler functions. In the Diffserv model, traffic is classified into different behavior aggregates. Packets, which enter into ATM-LER at the border of the core network, are assigned a single DSCP. They are forwarded as PHBs associated with their code points.

![Figure 3-1 The proposed ATM-based MPLS network architecture.](image)

### 3.3 Constraint Based Routing Over the Proposed Network

CR-LDP is a set of procedures through which ATM-LSRs not only exchange labels and set up LSP but also incorporate constraint based routing. It is an end-to-end setup mechanism of a Constraint-based Routed LSP (CR-LSP) initiated by the ATM-LER.
In CR-LDP, the traffic characteristics of a path are described in terms of a peak rate, committed rate, and service granularity. The peak and committed rates describe the bandwidth constraints of a path while the service granularity can be used to specify a constraint on the delay and jitter that the CR-LDP ATM based MPLS domain may introduce to a path’s traffic.

If an ATM-LSR receives label request messages from ATM-LER and if ATM-LSR can support the CR-LSP Traffic Parameters then the ATM-LSR must reserve the corresponding resources for the CR-LSP. Though ATM-LSRs are reserving resources, due to peak rate delay and jitter may become greater than negotiated QoS. So, when ATM-LSR fail to support the CR-LSP Traffic Parameters then the ATM-LSR must send a Notification Message that contains the "Resource Unavailable" status code. After receiving this Notification Message for a CR-LSP, ATM-LSRs will release any resources that it possibly reserved for the CR-LSP.

3.4 Multimedia IP over the Proposed Network

IP is a connectionless datagram based network layer that performs addressing, routing, and control functions for transmitting and receiving packets. As packets are received by the router, IP addressing information, such as the destination address, is used to determine the best “next hop” the packet should take enroute to its final destination. To introduce QoS capabilities within an IP network there is a need for mechanisms. The following section describes how the IP protocol could implement QoS mechanism in the proposed model.

3.4.1 QoS Routing

QoS routing mechanism of the proposed network provides multilevel service support based on differentiate service classifications. Using the IP Type of Service (TOS) field in the IP header provides a method of marking and distinguishing between different service classes as packets traverse the network. The TOS is composed of 3 precedence bits and 5
TOS bits in IP packet header. The TOS bits are generally regarded as describing the type or class of service to different user.

With differentiated service, IP packets are marked with the desired QoS identifier using DSCP when the packets enter into the ATM based MPLS network.

In the proposed Network, IP packets are forwarded with different DS PHBs using existing Next Hop Resolution Protocol (NHRP) architecture. SVCs will be used bi-directionally by traffic with the same PHBs between the same two points in the network. NHRP shortcuts will be created between the ingress and egress routers and will map DSCP to ATM traffic parameters.

For NHRP operation there has to be one Next Hop Server (NHS) in every Logical IP Subnetwork (LIS) as illustrated in Figure 3-2. All hosts on a LIS, register their Non-Broadcast Multiple-Access (NBMA) and internetwork layer (e.g. IP) address with their NHS when booting.

![Figure 3-2 Next Hop Resolution Protocol](image)

Let us assume that a Source wants to send an IP packet to a Target that lies outside its LIS. To resolve the NBMA address of the target, the source sends a next hop Resolution Request to its NHS. The NHS checks whether the target lies in the same LIS. If the NHS does not serve the target, the NHS forwards the request to the next NHS along the routed path. This forwarding process continues until it reaches the NHS that serves the target. This NHS can resolve the target's NBMA address and sends it back along the routed path. The intermediate NHS can store the address mapping information for the target contained in the Resolution Reply to answer subsequent Resolution Requests.
The main advantage of the NHRP is that it can solve the multiple-hop problem through NBMA networks by offering inter-LIS address resolution, thus enabling the establishment of a single-hop connection through the NBMA network. In the proposed network, a single direct VC can be established across several LISs, bringing QoS in terms of traffic contract guarantees to the IP data flow between the VCs endpoints.

3.4.2 IP Forwarding within MPLS Domain

MPLS allows the mapping of IP packets to FECs to occur only once at the ingress to an MPLS domain. In the case of datagram routing, IP packets would be mapped to a service level that would require packet filtering based on source and destination addresses and incoming interface, etc. Also, some information such as the incoming interface is only available at the ingress node to the network. This implies that the preferred way to offer provisioned QoS is to map the packet at the ingress point to the preferred QoS level, and then label the packet in a way to acknowledge that. MPLS offers an efficient method to label the QoS class associated with any packet.

3.4.3 IP Packets to ATM Cell

IP packets do not work directly with fixed size cells. The AAL segments these packets, transmit the cells individually and reassembles them at the other end.

ATM layer deals with the cells and the transportation of the cells. It defines the layout of a cell and tells what the header fields mean. It also deals with establishment and release of virtual circuits. Congestion control is also located here.

3.4.4 Multiprotocol over ATM (MPOA)

Two possible means of scaling IP to meet the existing and anticipated demand are to either replace the routed infrastructure that exists with newer, faster, cheaper routers that route every single packet at high speed, or to implement a "route once, switch many" strategy. MPOA adopt the latter approach and reduce the latency associated with moving Layer 3 traffic from one subnet to another.
MPOA operates at both Layer 2 and Layer 3, and is capable of providing direct Layer 3 connectivity across ATM when using shortcut flows in order to fully exploit ATM's QoS features. However, it is also capable of using a default flow using LAN Emulation when a suitable shortcut flow does not exist and the Layer 2 traffic is to be transferred within the same subnet or emulated LAN. For destinations outside the subnet, MPOA uses the NHRP to discover the ATM address of the intended destination.

MPOA provides MPOA Clients (MPCs) and MPOA Servers (MPSs) and defines the protocols that are required for MPCs and MPSs to communicate. MPCs issue queries for shortcut ATM addresses and receive replies from the MPS using these protocols. MPOA also ensures interoperability with the existing infrastructure of routers. MPOA Servers make use of routers that run standard internetwork layer routing protocols, such as OSPF, providing a smooth integration with existing networks.

MPOA makes route servers appear as conventional routers to the external IP or other Layer 3 network, and would only be responsible for route calculations. The traditional router function of packet forwarding would be performed by edge devices connected to a route server, to the Layer 3 network, and to the ATM network. At the start of the session, the route server calculates an appropriate end-to-end routes, and computes the best path through the ATM network. The path then emerges from the ATM cloud at an exit point as close as possible to the destination. For short packet flows, the route server would be involved with every packet just like conventional routers. However, for longer flows, such as with an FTP session, or an email transmission, a cut through path can be set up through the ATM network directly via edge devices, bypassing the router server.

At the moment, MPOA cuts paths through the ATM network in the same way, using the NHRP. But the route server has been replaced by the much simpler NHRP server, sometimes called Multiprotocol Server (MPS). The MPS server is merely responsible for computing the path through the ATM network that corresponds best with the Layer 3 hop calculated by a conventional router, using standard routing protocols such as OSPF.
3.5 Packet Scheduler of the Diffserv Model

The Diffserv in the MPLS network needs to carry packets at the desired service category. A separated LSP is created for each FEC and scheduling aggregate pair. Differentiation in treatment of packets from different behavior aggregates has to be implemented by mapping drop precedence. Thus, when the underlying technology is ATM, it can only support two levels of drop precedence. However, by marking the use of the EXP field in the “shim” header for the top label stack entry in the NHLFE, support for all the drop precedence can be provided in MPLS clouds.

A “shim” header cannot be used with ATM because this would involve doing segmentation and re-assembly at each ATM-LSR in order to read the DSCP. Hence, the DSCP in the IP header is not accessible by the ATM hardware responsible for the forwarding. Therefore, two alternative solutions may be considered: either have some part of the ATM cell header mapped to the DSCP, or use an LDP.

In the first approach, the most likely solution is to use the VPI and part of the VCI of the ATM cell header as the label, and to use the remaining eight least significant bits of the VCI to map the DSCP. Then, all that is needed is a functional component in the interior Diffserv-enabled ATM LSRs to perform the appropriate traffic management mechanisms on the cells by interpreting the DSCP correctly with respect to the PHB. In the second

![Figure 3-3 The Diffserv model in the proposed MPLS router system](image-url)
approach, which is more likely for future deployment, the DSCP is mapped to an LSP at the ingress of the MPLS domain. This means that for each DSCP value/PHB a separate LSP will be established for the same egress LSR. The packets belonging to streams with the same DSCP and FEC will be forwarded on the same LSP. In other words, the label is regarded as the behavior aggregate selector [12].

In Figure 3-3 the Diffserv model in the proposed MPLS router system is shown. This is similar to conventional Diffserv architecture but has many additional functions like service mapping.

In Figure 3-3, the Multi-Field (MF) classifier checks Intserv/Diffserv and Best Effort through IP flow classification. A traffic conditioner is a part of a network node that takes the node's ingress packets as its input and places the packets in its output in an order that best satisfies the forwarding requirements set for the packets and uses the network's resources in the best possible way. The Behavior Aggregate (BA) classifier classifies packets based on the DSCP only.

3.6 Queuing Architecture

In the Integrated Packet Scheduler for admission control, there are usually multiple logical queues in ATM-LER (ATM ingress router) and ATM-LSR in the ATM based MPLS domain destined to each outgoing link. Figure 3-4 shows the queuing architecture of the proposed Integrated Scheduler for EF1, EF2, AF1, AF2, AF3, AF4 and Best Effort traffic. Each queue may contain packets from one flow or a class of flows. The basic operation of scheduling is to decide, at a particular moment, which packet in all these queues should be transmitted onto the outgoing link. The decision algorithm directly affects delay performance and the buffering strategy directly affects packet loss performance. The integrated packet scheduler is realized by applying a Priority Queuing (PQ) algorithm and an admission control function.
For the EF and the AF flows, we adopt a First In First Out (FIFO) + queue [36]. The idea behind FIFO+ is to try to limit the accumulation of delay across hops and to decrease the worst-case delay. The flows in every hop are grouped into several classes. Average queuing delay is calculated for each class of each hop. For each packet, the difference between the experienced queuing delays is calculated. This difference is added or subtracted from an offset field in the packet's header. Over the hops, this offset indicates how far ahead or behind the packet is from its class's average. It is required to schedule a packet in its queues as if it had arrived at the packet's real arrival time plus the packet's offset.

It is important to keep in mind that FIFO+, unlike Weighted Fair Queuing (WFQ), is a statistical multiplexing technique. Because it does not provide strict isolation, FIFO+ may occasionally violate its delay requirements. Applications that require strict delay
guarantees will have to be satisfied with the longer delay bounds of queuing schemes like WFQ. For Best Effort flows, we employ a common FIFO shared queue. There is no QoS commitment to each individual Best Effort flow.

We can define seven priorities: EF1, EF2, AF1, AF2, AF3, AF4 and BE. Incoming traffic is classified and stored at seven separate queues. If a higher priority queue has no space for storing incoming traffic, traffic of this class goes to next lower priority queue. But when the space become available to store the traffic that goes to next lower priority queue, the traffic is forced to go back to its appropriate class so that traffic gets the required QoS. The advantage of the priority queue is the absolute preferential treatment that always gives top priority to mission critical traffic in the event of congestion.

3.7 Diffserv Policy in ATM Based MPLS Domain

We have proposed three policy types for Diffserv in ATM based MPLS domain.

1. Traffic Classifier (CL) Policy

A CL Policy specifies the classification of packet flows by assigning labels called CIDs (Classifier Identifiers.). A CL Policy is usually deployed to edge or border interfaces (i.e., interfaces that are connected to points outside the Diffserv domain) and applied to inbound traffic.

2. Traffic Conditioner (TC) Policy

A TC Policy meters, marks, and/or drops packets by using the traffic classifier policy. TC Policies, too, are usually deployed to edge or border interfaces and applied to inbound traffic.

3. Queue Control (QC) Policy

A QC Policy queues, schedules or drops packets according to the CL and TC Policy. A QC Policy is usually deployed to core interfaces (i.e., interfaces that are connected to other interfaces within the Diffserv domain) and applied to outbound traffic. A QC Policy rule can be regarded as (a model of) a queue or scheduler; i.e., traffic control object.
Examples of these policies are given below. The policies are described in a C-like language here for concise description.

1. Example of a CL Policy rule
The following rule (3.1) assigns the CID value "EF CID" to flows coming from IP address 192.168.1.1.

if (Source_IP_address is 192.168.1.1) then
{
    CID = "EF CID";
    MAP CID to EXP field of MPLS domain;
    MAP EXP field to CBR;
}
(3.1)

Marking a CID is the only function of the CL Policy, and it is usually used as a component of a larger policy.
2. Examples of TC Policy rules

A simple (stand-alone) rule

The following rule (3.2) is a simple TC Policy rule and marks DSCP 102 on packets.
if (Source_IP address is 192.168.1.1) then
{   DSCP = 102;
    MAP DSCP to EXP field of MPLS domain;
    MAP EXP to UBR;
}

This rule can be used as a stand-alone rule; i.e., in a given device, no other rule may be applied in cooperation with this rule.

A more complex rule

Rule (3.3), shown below, is applied to flows to which the CID value "EF CID" has been attached to.
if (CID is "EF CID") then
{   if (InformationRate <= 10 Mbps) then
       {   DSCP = "EF";    // marking
         MAP DSCP to EXP Field;
         MAP EXP to CBR;
       } else {
         Discard;    // absolute drop
       }
}

This rule may thus be combined with rule (3.2). Rule (3.3) meters the traffic and marks the DSCP "EF" on the packets in the first 10 Mbps of the traffic, and discards other
packets. This rule can be used for an EF (Expedited Forwarding) service of Diffserv, and it must be combined with a TC Policy rule such as rule (3.1).

3. Examples of QC Policy rules

A simple (stand-alone) rule

The following rule (3.4) applies a bounded priority queuing policy to the queuing and scheduling of packets with a DSCP of "EF"

if (DSCP is "EF") then
{
    SchedulingAlgorithm = "B-PQ"; // bounded priority queuing
    Priority = 6; // means "high"
    ShapingRate = 20 Mbps;
}

(3.4)

The traffic is then shaped to 20 Mbps. Rule (3.4) represents a queue that is connected to a priority scheduler that is not given as a rule.

A more complex rule

Rule (3.5), shown below, represents a scheduling queue, and this rule specifies three discard levels; newly coming packets with a DSCP of "AF1" are discarded only when the queue is filled with 200 packets (100%), while newly coming packets with a DSCP of "AF2" are discarded when the queue contains 140 (70%) or more packets and newly coming packets with a DSCP of "AF3" are discarded when the queue contains 100 (50%) or more packets.

if (DSCP is ["AF1", "AF2", "AF3"]) then
{
    SchedulingAlgorithm = "A-BW"; // aggregated bandwidth fair queuing
    Max_Queue_Size = 200 packets;
    CommittedRate = 64 kbps; // assured minimum rate
    DiscardAlgorithm = "Deterministic Discard";
    if (DSCP is "AF1") then {

DiscardLevel = 100%; //allowed to use whole queue
} elsif (DSCP is "AF2") then {
    DiscardLevel = 70%; //allowed to use 70% of the queue
} elsif (DSCP is "AF3") then {
    DiscardLevel = 50%; // allowed to use 50% of the queue
}; (3.5)

"Deterministic Discard" is specified as the discard algorithm. This specifies a non-random algorithm for dropping packets. This rule can be used for an AF service of Diffserv. A random discard method such as the weighted random early discard can be specified instead of deterministic discard too.

3.8 Justification of the Proposed Architecture

Following are the three main reasons for proposing the new/modified architecture to support QoS of multimedia sessions:

- Diffserv architecture is used for classification of flows to seven classes according to the SLA. If Diffserv architecture is not used, each flow may be treated differently according to the traffic condition and it may become very difficult to manage them. Again if the flows are not classified, all of the flows become BE traffic and this become inefficient for real time multimedia traffic as all the packets compete equally for network resources. If Intserv architecture is used, the network would suffer from scalability. Diffserv policy is introduced by using Queuing architecture and scheduling policy.

- MPLS is used for its fast and efficient forwarding mechanism of packets from source to destination using FEC. Moreover, it is easier to implement and MPLS is faster for forwarding packets than ATM VC/VP mechanism. MPLS ensures the ordered delivery of multimedia packets.

- ATM has been used to introduce the concept of short and fixed sized ATM cell in our proposed network. We have also introduced CBR for the transmission of data.
3.9 Summary

In this chapter the proposed Diffserv-Aware ATM based MPLS network architecture has been described for multiple service class environments to increase the session success rate ensuring QoS for multimedia transmission. Diffserv policy in ATM based MPLS domain and Packet Scheduler of Diffserv Model have also been explained.
4. Admission Control Algorithm and Queuing Analysis

4.1 Introduction

In this chapter a new flow based admission control algorithm through an ATM based MPLS has been introduced for multiple service class environment of Diffserv. A policy of rerouting some of the flows during the congestion in a particular part of the network is also presented to optimize the use of the network resources according to traffic classification, which is suitable for multiple service class environments. The analysis shows that the proposed algorithm achieves reliable delay-bounded QoS performance and reduces the blocking probability of high priority service in the Diffserv model.

4.2 The Proposed Multimedia Session Admission Control Algorithm with One Time Rerouting Option

When a multimedia session starts, it first identifies the optimum routing path from the source to the destination meeting users QoS requirements. The QoS condition parameters are offered on a per flow basis, corresponding to the number of related performance objective. In our proposed technique, we use statistical performance data (delay, jitter) of each network node on the routing path from the source to the destination of a session to determine whether a session fails to meet the user requirements. If it seems that a session may fail to meet the user requirements, a decision for rerouting is taken with a rerouting probability to send the remaining multimedia traffic of the session. If the session gathers bad experience to fulfill QoS it marks the session as failed session. These decisions must be made by the ingress ATM-LER.

Our rerouting algorithm for a multimedia session is considering a practical experience that a large network may face. In a large network, there are so many paths from a source to a destination and when one path is congested, some other paths may remain underutilized i.e., these paths may fulfill the resource requirements for a particular QoS. Hence it would be better if we can reroute this session to a different but currently
optimum physical path. We do not reroute all the congested flows to the current best path, as there is a chance of another congestion if every flow is rerouted to the optimal path. The currently congested path is expected to be congestion free if some of the flows are rerouted to another congestion free path.

Rerouting a flow of multimedia traffic again and again will cause a situation of congestion like best effort datagram service. Eventually that will cause more failure in providing QoS to the sessions. Besides this, rerouting for several times might not be very much appreciated by the customers as it causes interruption of a multimedia session. That is why, only one time rerouting option has been proposed with a predefined rerouting probability for each session that fails to meet QoS.

Figure 4-1 shows the proposed algorithm for admission control of multimedia sessions with rerouting options. Multimedia sessions are divided into service classes such as EF1, EF2, AF1, AF2, AF3, AF4 and BE according to the users' requirements of delay, loss, jitter and bandwidth. Each router checks the required QoS of a Diffserv packet. If the QoS requirement of the packet is not satisfied then the performance statistics of the router is updated accordingly.
Identification of routes from the source to the destination and selecting optimum path considering QoS

Identification of another optimum path
Send rest of the messages using this path

Whether a hop of the selected path from the source to the destination meets the user requirements and resource condition for a packet of different class of service like EF1, EF2, AF1, AF2, AF3, AF4, BE

Overall QoS Condition for session
No Rej the packet
Yes Session is not yet rerouted and within the rerouting probability

Figure 4-1 Proposed algorithm for admission control.

4.3 Congestion Detection in an LSP using the Proposed Algorithm

**Congestion may occur:** Ingress LER establishes LSP for each flow by identifying LSRs and required resources using CR-LDP. However, currently the LER has no knowledge of how those LSRs are being utilised by other LERs. That is why, an LER can set up an LSP using a congested LSR.
Functions of LSR during Congestion: Figure 4-2 depicts the flow of traffic between sources and destinations passing through our proposed queuing architecture at each LSR. It is assumed that over time, as a result of the increased load through LSR, it starts to lose packets from the best effort buffer.

At the input buffer the number of packets belonging to the best effort in the LSP and the number of packets destroyed due to failed insertion are recorded. They are used to calculate the packet loss in the LSP just before the packets are forwarded to the output buffer, whose sole purpose is to act as a transmission buffer through which no loss occurs, as illustrated in Figure 4-3.
Once the packet loss in the best effort buffer on LSP happens more than the predetermined threshold value for an extended time period, the LSR creates an LDP notification message containing the Congestion Indication. The aim of sending the Congestion Indication Notification (CIN) message is to indicate to the ingress LER that there are packets being lost from a particular LSP originating from it.

Rather than all congested LSRs always generating CIN messages, intermediate LSRs upon receipt of a CIN message may append relevant information to it concerning their status if they are also experiencing congestion. If an LSR receives a CIN message shortly after sending one, it checks the Congestion Indication Table (CIT) whose format is shown in Table 4-1, to see if the timer it has set has expired. If it has not expired, it will simply forward the message without appending its own information, otherwise it will include its information before forwarding.

Each LSR experiencing congestion keeps the records of LSPID, current LSP and buffer loss probability in its Congestion Indication Table as shown in Table 4-1.

<table>
<thead>
<tr>
<th>Buffer Stream</th>
<th>LSP ID</th>
<th>Current Loss in LSP</th>
<th>Current Loss in Buffer</th>
<th>Timer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 138.37.32.75</td>
<td>2</td>
<td>3.78e-03</td>
<td>5.63e-04</td>
<td>0</td>
</tr>
</tbody>
</table>

Use of timer: Timers are used to control the responsiveness of the proposed scheme to traffic loading transients. For example, when an LSR is congested it can issue a CIN message. In doing so it sets a retransmission timer. It is not permitted to issue another message until the timer expires, thus avoiding signalling storms whilst improving the robustness of the protocol. Alternatively, if it receives a CIN message on route to the ingress LER from another congested LSR, it can simply append its own congestion information and set the timer accordingly. In doing so, avalanches of congestion notification messages towards the ingress LER are prevented. In addition, stability is
improved by averaging the observed traffic parameters at each LSR and employing threshold triggers.

**Functions of LER during Congestion:** After receiving the CIN messages from the LSR, LER will perform one of the following tasks

- Decide that the packet loss it is currently experiencing remains sufficiently low for it to continue to meet its SLA requirements, allowing/permitting no further action to be taken at this time.
- Renegotiate for new QoS requirements along the existing LSP.
- Negotiate for QoS requirements along an alternative LSP.

### 4.4 Queuing Analysis

To analyze the proposed algorithm with rerouting option for multimedia session, we model the queuing system with a finite buffer, bulk arrival and $M/M/1/K$ queue. We have seven priority queues EFI, EF2, AF1, AF2, AF3, AF4 and BE for investigating the session success rate and average time delay of a multimedia session with different offered load and different rerouting probability.

As the system is modeled as an $M/M/1/K$ queue with bulk arrival, packet flow is exponentially distributed, packet size is modeled by geometric distribution and system can hold at most $K$ customers with finite storage. Let $\lambda$ is the mean arrival rate and $\mu$ is the average service rate.

According to the $M/M/1/K$ model we get the following ideal equations [43]:

$$L = \frac{\rho[1-(K+1)\rho^K+K\rho^{K+1}]}{(1-\rho^{K+1})(1-\rho)}$$  \hspace{1cm}  (4.1) \hspace{1cm} \lambda \neq \mu

$$L = \frac{K}{2}$$  \hspace{1cm}  \lambda = \mu  \hspace{1cm} (4.2)

$$P_\infty = \frac{(1-\rho)\rho^K}{1-\rho^{K+1}}$$  \hspace{1cm}  \lambda \neq \mu  \hspace{1cm} (4.3)
\[ P_K = \frac{1}{K + 1} \]

\[ \lambda = \mu \]  \hspace{1cm} (4.4)

\[ \lambda_e = \lambda (1 - P_K) \]  \hspace{1cm} (4.5)

\[ p = \frac{\lambda_e}{\mu} \]  \hspace{1cm} (4.6)

\[ w = \frac{L}{\lambda_e} \]  \hspace{1cm} (4.7)

Here, \( L \) = Expected number of packets in the system

\( P_K \) = Probability that \( k \) packets are in the system

\( \lambda_e \) = Effective arrival rate

\( p \) = Server utilization

\( w \) = Average time spent in system per packets.

For the above-mentioned model we get the following equations and definitions [12] when no rerouting option is present:

\[ P_{\text{blocking}} = 1 - P_{\text{QoS}} P_{\text{resource}} \]  \hspace{1cm} (4.8)

\[ P_{\text{QoS}} = P_{\text{delay}} P_{\text{loss}} \]  \hspace{1cm} (4.9)

\[ P_{\text{delay}} = P[T_{\text{overall}} \leq D_{\text{bound}}] \]

\[ P_{\text{loss}} = P[R_{\text{loss}} \leq R_{\text{required}}] \]

The explanations of the probability terms in the equations are as follows:

\( P_{\text{blocking}} \) = The probability that a session is rejected during admission or failed when a congestion is found in the network.

\( P_{\text{QoS}} \) = The probability that a session is satisfied with the QoS condition.
\( P_{\text{resource}} \) = The probability that the buffer resource in the system is less than the threshold value

\( P_{\text{delay}} \) = The probability that a delay on a particular path is less than the negotiated threshold value

\( P_{\text{loss}} \) = The probability that the packet loss is less than the threshold value \( R_{\text{required}} \)

\( T_{\text{overall}} \) = The estimated overall delay

\( R_{\text{loss}} \) = The estimated loss ratio

\( D_{\text{bound}} \) = The negotiated delay bound

We can calculate \( P_{\text{resource}} \) approximately as

\[
P_{\text{resource}} = \rho \left[ \lambda^* \leq \mu c \right]
\]

where \( c \) is the channel capacity and \( \lambda^* \) is instantaneous estimated arrival rate (i.e. packets arrive at a particular moment). So throughput \( \rho \) can be obtained by

\[
\rho = \frac{\lambda (1 - P_{\text{blocking}})}{\mu c}
\]

The expression of \( P_{\text{blocking}} \) is given as follows

\[
P_{\text{blocking}} = 1 - \prod_{i \in \gamma} (1 - L_i) = 1 - \prod_{i=1}^{h} (1 - L_i)
\]

Where \( L_i \) is the probability that the \( i \)-th link is blocked on the root \( \gamma \) and \( h \) is the hop distance on the root \( \gamma \).

\[
\bar{T} = \frac{1}{\lambda \gamma (1 - P_{\text{blocking}})} \sum_{i=1}^{h} \lambda_i (1 - L_i) \bar{T}_i
\]

Where \( \bar{T} \) is the average flow transfer delay and \( \bar{T}_i = 1/(\mu_i c_i - \lambda_i (1 - L_i)) \).
The overall average end-to-end delay can be calculated by summation of the average system delay and average flow transfer delay. In a non-preemptive priority queuing system, the average waiting time of messages of Priority Class $p$ is given by

$$E[W_p] = E[	au_0] + \sum_{p=1}^{P} E[T_p] + \sum_{p=1}^{P-1} E[T_p].$$

Where $\tau_0$ is a residual lifetime to complete the current service, $T_p$ is the message service time of priority class $p$ and $\bar{T}_p$ is the message service time of higher priority during the waiting interval of messages with Priority Class $p$. Here $E[\tau_0]$ is the weighted sum overall priority classes.

Consider the total number of packets of a multimedia session is $M_S$. So according to Equation (4.8), the total number of packets that will not meet the QoS requirements due to congestion can be expressed by

$$M_L = M_S \cdot P_{\text{blocking}} \quad (4.13)$$

And the total number of packets reaching the destination after fulfilling QoS requirement of a multimedia session will be

$$M_D = M_S \cdot P_{\text{QoS}} \cdot P_{\text{resource}} \quad (4.14)$$

The probability that a session will be successful can be expressed as follows:

$$P_{SS} = P\left(\frac{M_D}{M_L} \geq M_T\right) \quad (4.15)$$

Where, $M_T$ is the negotiated threshold ratio.

From the equation (4.8), (4.15) we get

$$P_{SS} = P\left(\frac{P_{\text{QoS}} \cdot P_{\text{resource}}}{(M_T + 1)} \geq M_T\right) \quad (4.16)$$

$P_{SF} =$ The probability that a session will fail to meet the user requirements = $1 - P_{SS}$. 
Let $P_{SR}$ be the predefined rerouting probability for a nearly failed session. Hence, the probability that a nearly failed session will be rerouted is $(1 - P_{SS}) P_{SR}$.

Now let us consider a scenario where a session has been rerouted from Path $\gamma$ to Path $\beta$ after facing the difficulties to meet the QoS. The probability that a rerouted multimedia session will be successful can be expressed as follows:

$$P_{SSR} = P_{QoS\gamma} \cdot P_{resource\gamma} \cdot P_{QoS\beta} \cdot P_{resource\beta} \cdot (1 - P_{SS}) P_{SR}$$

(4.17)

From the equations (4.16, 4.17), the probability that a session will be successful with or without rerouting will be,

$$P_{TSS} = P_{SS} + P_{SSR}$$

(4.18)

Now from Equation (4.17) it has been found that $P_{SSR}$ increases with the increase in the rerouting probability $P_{SR}$. But this will increase the total arrival rate $\lambda$ in the new path $\beta$. Eventually $P_{resource}$ will decrease when $\lambda$ will exceed the total capacity [Equation (4.12)]. Thus $P_{SS}$ will decrease as a result of flow transfer delay. The rerouting policy for different offered loads has been demonstrated in the next chapter using simulation.

4.5 Numerical Results

Table 4-2 and Table 4-3 show the different performance results such as delay, packet loss and blocking probability calculated from the above mentioned equations in queuing analysis for different utilization ratio, service rate, buffer size and number of hops. These data have been plotted in the graphs shown in Figures 4-8, 4-9 and 4-10.

**Table 4-2 Average time spent in system per packets**

<table>
<thead>
<tr>
<th>Utilization, $\rho$</th>
<th>Service Rate (Packets/sec), $\mu$</th>
<th>Buffer (Packets), N</th>
<th>Expected Number of Packets, L</th>
<th>Effective Arrival Rate, $\lambda_e$</th>
<th>Service Time (\mu s) /Packet s</th>
<th>Packet Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$10^6$</td>
<td>100</td>
<td>50</td>
<td>$9.901 \times 10^5$</td>
<td>50.5</td>
<td>0.009901</td>
</tr>
<tr>
<td>Utilization, $\rho$</td>
<td>Service Rate (Packets/sec), $\mu$</td>
<td>Buffer (Packets), $N$</td>
<td>Expected Number of Packets, $L$</td>
<td>Effective Arrival Rate, $\lambda_e$</td>
<td>Service Time ((\mu)s) /Packet</td>
<td>Packet Loss</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>---------------------------------</td>
<td>-----------------</td>
<td>--------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>1</td>
<td>$10^6$</td>
<td>200</td>
<td>100</td>
<td>$9.9502 \times 10^5$</td>
<td>100.5</td>
<td>0.0049751</td>
</tr>
<tr>
<td>1</td>
<td>$10^6$</td>
<td>300</td>
<td>150</td>
<td>$9.9668 \times 10^5$</td>
<td>150.5</td>
<td>0.0033223</td>
</tr>
<tr>
<td>0.98</td>
<td>$10^6$</td>
<td>100</td>
<td>33.912</td>
<td>$9.7701 \times 10^5$</td>
<td>34.71</td>
<td>0.0030486</td>
</tr>
<tr>
<td>0.98</td>
<td>$10^6$</td>
<td>200</td>
<td>45.475</td>
<td>$9.7965 \times 10^5$</td>
<td>46.42</td>
<td>0.00035793</td>
</tr>
<tr>
<td>0.98</td>
<td>$10^6$</td>
<td>300</td>
<td>48.31</td>
<td>$9.7999 \times 10^5$</td>
<td>49.3</td>
<td>4.6757 \times 10^{-5}</td>
</tr>
<tr>
<td>0.95</td>
<td>$10^6$</td>
<td>100</td>
<td>18.429</td>
<td>$9.4972 \times 10^5$</td>
<td>19.40</td>
<td>0.0002977</td>
</tr>
<tr>
<td>0.95</td>
<td>$10^6$</td>
<td>200</td>
<td>18.993</td>
<td>$9.5 \times 10^5$</td>
<td>19.99</td>
<td>1.7527 \times 10^{-5}</td>
</tr>
<tr>
<td>0.95</td>
<td>$10^6$</td>
<td>300</td>
<td>19</td>
<td>$9.5 \times 10^5$</td>
<td>20</td>
<td>1.0377 \times 10^{-8}</td>
</tr>
<tr>
<td>0.90</td>
<td>$10^6$</td>
<td>100</td>
<td>9</td>
<td>$9 \times 10^5$</td>
<td>9.99</td>
<td>2.6562 \times 10^{-6}</td>
</tr>
<tr>
<td>0.90</td>
<td>$10^6$</td>
<td>200</td>
<td>9</td>
<td>$9 \times 10^5$</td>
<td>10</td>
<td>7.0551 \times 10^{-11}</td>
</tr>
<tr>
<td>0.90</td>
<td>$10^6$</td>
<td>300</td>
<td>9</td>
<td>$9 \times 10^5$</td>
<td>10</td>
<td>1.8739 \times 10^{-15}</td>
</tr>
<tr>
<td>0.85</td>
<td>$10^6$</td>
<td>100</td>
<td>5.6667</td>
<td>$8.5 \times 10^5$</td>
<td>6.66</td>
<td>1.3122 \times 10^{-8}</td>
</tr>
<tr>
<td>0.85</td>
<td>$10^6$</td>
<td>200</td>
<td>5.6667</td>
<td>$8.5 \times 10^5$</td>
<td>6.67</td>
<td>1.1478 \times 10^{-15}</td>
</tr>
<tr>
<td>0.85</td>
<td>$10^6$</td>
<td>300</td>
<td>5.6667</td>
<td>$8.5 \times 10^5$</td>
<td>6.67</td>
<td>1.0041 \times 10^{-22}</td>
</tr>
<tr>
<td>0.80</td>
<td>$10^6$</td>
<td>100</td>
<td>4</td>
<td>$8 \times 10^5$</td>
<td>5</td>
<td>4.0741 \times 10^{-11}</td>
</tr>
<tr>
<td>0.80</td>
<td>$10^6$</td>
<td>200</td>
<td>4</td>
<td>$8 \times 10^5$</td>
<td>5</td>
<td>8.299 \times 10^{-21}</td>
</tr>
<tr>
<td>0.80</td>
<td>$10^6$</td>
<td>300</td>
<td>4</td>
<td>$8 \times 10^5$</td>
<td>5</td>
<td>1.6905 \times 10^{-30}</td>
</tr>
</tbody>
</table>
Figure 4-4 Average time spent by a packet in the System

Figure 4-5 Packet loss for a given utilization & buffer size

Table 4-3 Packet loss for different utilizations & hop numbers using buffer size of 100 packets.
Utilization & P\text{blersing} & Hop & Total P\text{blersing} \\
\hline
1 & 0.009901 & 2 & 0.019704 \\
1 & 0.009901 & 4 & 0.03902 \\
1 & 0.009901 & 6 & 0.057955 \\
.98 & 0.0030486 & 2 & 0.0060879 \\
.98 & 0.0030486 & 4 & 0.012139 \\
.98 & 0.0030486 & 6 & 0.018153 \\
.95 & 0.0002977 & 2 & 0.00059531 \\
.95 & 0.0002977 & 4 & 0.0011903 \\
.95 & 0.0002977 & 6 & 0.0017849 \\

\textbf{Conclusion} \\
In this chapter the proposed Admission Control Algorithm has been described for Diffserv-Aware ATM based MPLS network ensuring QoS for multimedia transmission. The congestion detection and path renegotiation policy during congestion has been described in ATM based MPLS domain. The queuing analysis has been done and
numerical results for the proposed model has been calculated. Results shows that delay and packet loss become exponential when offered utilization exceeds 98%.
5. Simulation Model for the Proposed System and Results

5.1 Introduction

In the previous chapter a new admission control algorithm for multimedia transmission in ATM based MPLS network was proposed. Having described the functional behaviour of the algorithm, it is necessary to investigate effect of different physical implementations, traffic loading and session success. This is achieved by simulation.

Having proposed the signalling changes between the LSRs, the volume of signalling traffic through the core network can be calculated theoretically for any particular traffic situation. However the nature of the networks such that the estimation of reactionary performance becomes too complex for mathematical analysis. That is why the results by mathematical analysis are not guaranteed to be a true representation of the system. Due to statistical variation of the multimedia sessions, it is not possible to determine the delay, packet loss and session success rate analytically. Therefore simulation models are required to obtain performance data for various scenarios.

5.2 Network Modelling

To analyze the performance of admission control algorithm we ran a simulation for a network depicted in Figure 5-1. This is a simple nine-node network, which is loosely based on UUNet's backbone [41]. This backbone is generally used for video telephony and video conferencing. Each node may act as LER and/or LSR. The LSR and LERs are connected via 155 Mb/s links. Each node is associated with a tagged source module.
5.3 Tagged Source Module

The tagged source module initiates the establishment of a connection between itself and the destination. In this simulation study, the aggregation of a number of individual flows are modelled as a single connection request. Only the positive outcomes for the initial connection requests are modelled. The tagged source module generates a constant bit rate, 33.92 Mb/s flow of packets, setting the traffic parameters i.e., bandwidth and loss probability that were requested for within each packet.

Each packet is a fixed size of 424 bits. In this research, MPLS was assumed to operate over ATM. The cell size of 53 bytes equates to 424 bits. This usually involves composing a signalling message and transmitting the message to the network with a packetisation delay and a transmission delay. The processing delay is set to 1 ms and the transmission delay for a 155.52 Mb/s link is 2.726 μs and renegotiation process take 12.5 ms.
A CBR source was chosen to represent the connection between the source and destination. Due to its deterministic behaviour any variation in its performance will be as a result of the network. It is this predictable behaviour that has made a CBR source the ideal choice as the tagged source.

The tagged source module is the parent process. On transmitting an initial request message, it invokes the child process to generate the tagged source. The tagged source module creates an instantaneous tagged source child process for each successful connection. Figure 5-2 and Figure 5-3 illustrates the finite state machines for the two respectively.

Once initialisation of the module is complete at the beginning of the simulation, the module goes to the “idle” state. It remains in the “idle” state until a self-interrupt is received to send an initial packet requesting a certain amount of bandwidth and a ceiling loss probability for a connection set-up.

The transition to the next state is determined by the type of interrupt, e.g., a stream interrupt indicating a Hello message informing it of who its neighbour is or a self-interrupt to generate the tagged source.

Once the finite state machine transfers to a particular state, the processing for that state is carried out. In the case of receiving a Hello message, the ID of the sender is recorded in the associated neighbouring database. If the interrupt was a self-interrupt it generates the child process that starts the tagged source. Once the processing is complete the transition is back to the “idle” state.

Figure 5-2 Parent process source module processing states
Figure 5-3 Child process tagged source processing states

Figure 5-4 shows the relationship between the parent and child process, a child process exists for each successful connection request. However, in the simulation one child process is spawned for simplicity.

Figure 5-4 Source module process relationship with Tagged sources' child process.

5.4 Background Source Module

The role of the background source module is to cause statistical variation in the traffic traversing the buffers i.e., to cause the loss probability experienced by the individual buffer streams to vary randomly. This is engineered to cause losses to exceed a predetermined threshold as required. Figure 5-5 shows an example of how the buffer stream is configured.
Figure 5-5 Buffer stream

On off sources were chosen to represent the background traffic in the simulation model. A single on off source is fed into the signalling buffer stream i.e., the buffer stream through which the signalling packets traversed, and into the same buffer stream as the tagged source to provide some delay jitter along the signalling path.

The background source module comprises of a parent node process that invokes a child process for each instance of an on off source. The finite state machine that represents them is shown in Figures 5-6 and 5-7 respectively.

Figure 5-6 Parent process background source state machine
The on off source chosen for the simulation has the following traffic parameters;

On time = 0.00025 seconds;

Off time = 0.00225 seconds;

Peak Cell Rate = 93500 packets / second;

Distribution = negative exponential for the on and off times.
5.5 Background Source Sink

The background sources feeding the individual buffer streams are destroyed at the output of the Scheduler Module. This prevents an avalanche of packets occurring at the egress LER.

5.6 LER and LSR Node Models

This simulation model is a signalling protocol based model, so it is necessary to generate the appropriate signalling messages in accordance with the CR-LDP mechanisms. To achieve this, every request for a connection must be processed on an individual basis and the progress of each connection request must be tracked. The simulation model of the LSR and LER as it appears at the node is shown below in Figure 5-9.

![Simulation model of LER/LSR](image)

*Figure 5-9 Simulation model of LER/LSR*

The node (LSR or LER) is the parent process. This is a static process that will last throughout the simulation. The primary functions of the node parent process are described below:

- The behaviour of the LSR and LER are both contained within the node process. The LER has a few additional mechanisms that are not required in the LSRs. In a
real network an LSR may act as both a LER/LSR at the same time for different LSPs.

- Initialisation of the simulation environment is done by sending Hello messages to all surrounding nodes and gathering the necessary information from received Hello messages to formulate the topology of the network. This information is used to formulate the routing table that is necessary for the duration of the simulation. This operates as a simplified routing process similar to OSPF.

- When a particular connection request comes from a source the LER will start the process for setting up LSP from source to destination meeting the QoS. It will be easily calculated from the network topology available to every node.

- During Congestion the LSR is responsible for generating and receiving CIN, Renegotiation Success and Renegotiation Failure notification messages. Each connection request has a unique LSP ID assigned by the ingress LER. All signalling messages generated by a request will contain this ID; the reply to the signalling messages will also contain this ID. The LER maintains a list of connection requests and their corresponding IDs.

- The ingress LER is also responsible for transmitting Status Request messages into the network, either periodically or on receipt of a CIN message.

The following subsections describe the individual components of the LSR/LER.

5.6.1 Root Module

The responsibility of LER Root module is to respond to the request for a connection. It determines an explicit route from itself to the destination. It then generates a Label Request message that is sent to the next hop in the explicit route. On receipt of Label Request each LSR determines whether it can meet the request and if so it continues to forward the Label Request message to the next LSR reserving the necessary resources. On receiving a Label Request the egress LER allocates a label and sends the value in a Label Mapping message to the source of the Label Request message. On receipt of a Label
Mapping message the LSR uses the flow ID to identify the connection this label refers to. The LSR assigns a new label that is forwarded to the next hop along the route and forms a binding with the received label.

During congestion, a congested LSR generates a CIN message containing the loss probability of the relevant buffer and flows and forwards the message to the next hop towards the ingress for the flow. On receipt of a CIN message other LSRs along the path each determine if they are suffering loss.

On receipt of a CIN message the ingress LER takes one of the following decisions:

- Negotiate along an alternative path;
- Accept the current losses and do nothing.

In our simulation model we have not considered the renegotiation process for new QoS requirements along the existing LSP because during congestion all class of traffics may become EF traffic and may violet the basic principal of DiffServ.

The following section explains how the Root Module was implemented.

The finite state machine for the Root Module is shown in Figure 5-10. Once initialisation of the module is complete at the beginning of the simulation, the module goes to the idle state. During initialisation each LSR/LER formulates Hello messages with its address and sends them to all the surrounding LSRs. On receiving a Hello message the ID and the stream connecting this LSR to its neighbour is recorded in a database. Transition to a new state occurs on receipt of an 'initiator' message in the case of the LER and on receiving a Label Request in the case of an LSR. The Root Module handles all the major signalling processing. It receives the initialisation message to establish a LSP and the associated Label Request and Label Mapping messages. On receiving CIN messages it formulates renegotiation procedures and it is responsible for generating and maintaining Status Request messages.
5.6.2 QoS Buffer Module

The QoS Buffer Module is dimensioned to have a number of buffer streams catering different minimum loss probability thresholds. Throughout the simulation the number of packet loss by the buffer is calculated and the calculated value passed to the Root Module where it is updated in the Buffer Table maintained by each LSR/LER. Initially each buffer stream is dimensioned for a given minimum loss probability threshold as shown in Figure 3-4.

There are eight buffer streams used in the simulation, one for signalling messages dimensioned for a very low loss threshold probability e.g., $10^{-9}$. The threshold probability
for EF1, EF2, AF1, AF2, AF3, AF4 and BE are $10^{-8}$, $10^{-7}$, $10^{-6}$, $10^{-5}$, $10^{-4}$, $10^{-3}$ and $10^{-2}$ respectively.

The tagged source along with the background sources are passed into one of the buffer streams. The QoS Buffer module calculates the loss probability for every packet loss of every class of service (e.g., 50 packets lost at a particular buffer stream). Additional background traffic is fed into the buffer stream containing the tagged source to maintain the loss probability in the buffer less than a predefined threshold. At this point a CIN message is generated and transmitted towards the ingress LER. For every arrival events (such as 500 packets arriving at a buffer stream) the long-term loss probability is calculated for that particular buffer stream. This value is sent to the Root module to update the Buffer Table with the current loss probability available within that buffer. This information is included in the periodic Status Request messages sent back to the ingress LER. The following describes how the QoS Buffer Module is implemented.

The QoS Buffer is developed as a finite state machine. Once initialisation of the module is complete at the beginning of the simulation, the module either goes to the “idle” state or receives a packet. On receipt of a packet it determines which buffer stream it belongs to and attempts to insert it. The state diagram for the QoS Buffer is shown below in Figure 5-11.
5.6.3 Scheduler Module

The Scheduler performs both scheduling and behaves as a transmission buffer. The Scheduler resides in the LSR/LER and receives packets transmitted from the QoS Buffer Module and buffers them before forwarding them towards the next hop. The scheduler has only one buffer stream and is dimensioned so no loss occurs within it. It serves at a rate of 155.52 Mbps as shown below in Figure 5-12. Every 2.726 μs it "visits" the buffer stream, if there is a packet present it will serve it immediately i.e., forwards it towards the next hop LSR.
The scheduler is developed as a finite state machine. Once initialisation has occurred at the beginning of the simulation, the module goes to the “idle” state or receives a packet. Every 2.726 μs it serves one if there is any packet in the buffer. The state diagram for the scheduler is shown in Figure 5-13.

![Finite state machine for the scheduler](image)

**Figure 5-13 Finite state machine for the scheduler**

5.6.4 Sources

A high value service rate was chosen for the buffers to allow variability in the sources feeding them. It was found that low values of service rate resulted in a very rapid change in loss probability.

Consider the service rate of a buffer serving the tagged source was set at 100000 packets/second. The Peek Cell Rate of the tagged source fed into the buffer was set to produce 80000 packets/second. This value was chosen to be less than the service rate of the buffer to prevent loss. The number of on off sources required to load the buffer to a certain utilisation are estimated as follows:

The individual buffers are loaded by the on off sources as shown in Figure 5-14. The objective is to maintain a minimum loss probability threshold for each of the buffer
streams. To ensure that the sources were loading the buffer streams correctly the following procedure was carried out: A number of on off sources were fed into a single buffer and the number of packets served by the buffer over a given time period was recorded and compared against the theoretical value (Figure 5-29).

The simulation parameters of an individual on off source are shown in Table 5-1.

### Table 5-1 On Off source simulation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak Cell Rate during the on period</td>
<td>33.94Mb/s (93500 packets/second)</td>
</tr>
<tr>
<td>$T_o$ (On Time Duration)</td>
<td>0.00025 seconds</td>
</tr>
<tr>
<td>$T_{off}$ (Off Time Duration)</td>
<td>0.00225 seconds</td>
</tr>
<tr>
<td>On Time Distribution</td>
<td>Negative exponential</td>
</tr>
<tr>
<td>Off Time Distribution</td>
<td>Negative exponential</td>
</tr>
</tbody>
</table>

5.7 Packets Forwarding using MPLS

Traditional IP routers analyze the network layer header of each packet and parse their routing table to match with the IP destination address for choosing the next hop. MPLS is a technology that integrates label-swapping paradigm with network-layer routing within LSRs. The essence of MPLS is the generation of a short fixed-length "label" that acts as a shorthand representation of an IP packet's header that is subsequently used for packet forwarding within the MPLS network resulting in high-speed switching. The ingress LSR analyses the contents of the incoming packet IP header and selects the appropriate MPLS header with which to "encapsulate" the packet. The MPLS 'shim' header contains a 20-bit label, a 3-bit $exp$ field (initially defined as EXPerimental and currently used as Class of
Service (COS field), a 1-bit label stack indicator, and an 8-bit Time-To-Live (TTL) field. When the LSRs receive a labelled packet, they use the label as an index in an ILM table containing entries of the form \( \langle \text{in label}, \text{out label} \rangle \) called NHLFE. The incoming label is replaced with the new outgoing label and the packet is forwarded to the next hop. Finally, the egress LSR "decapsulate" the MPLS header as the MPLS packets leave the network. The sequence of labels from an ingress LSR to an egress LSR is called LSP, which is similar to a unidirectional ATM Virtual Circuit, whereas the labels have a local significance. LSP setup can be control-driven (i.e., triggered by control traffic such as management, routing updates or resource reservation) or data-driven (i.e., triggered by the presence of a specific flow). Also, LSP setup can follow a downstream approach whereas the downstream LSR at the end of the link initiates the LSP, or a downstream-on-demand approach where the downstream LSR generates the labels in response to requests made by an upstream LSR. A mapping between IP packets and an LSP must take place at the ingress LSR by specifying a FEC to a label. A FEC is defined as a group of packets that can be treated in an equivalent manner for the purposes of forwarding. The ingress LSR uses a FTN which is used when forwarding packets that arrive unlabeled and are to be labelled before forwarding. An example of an FEC is the set of unicast packets whose destination addresses match a particular IP address prefix. FECs can also be defined at different levels of granularity (source, destination, port level).

In the simulation program, we have set up LSP and generate stack of NHLFE for each session using Bellman and Ford shortest path algorithm. In this algorithm, \( \text{dist}^L \) (Egress LER) be the length of the optimum path from the Ingress LER to Egress LER under the constraint that the optimum path contains at most \( L \) link. Then \( \text{dist}^L \) (Egress LER) = \( \text{cost} \) (Ingress LER, Egress LER). Our goal then is to compute \( \text{dist}^{n-1} \) (Egress LER) for all Egress LER. Here \( n \) = total number of LSRs in the network. This can be done using the dynamic programming methodology. First we make the following observations:

- If the optimum path from Ingress LER to Egress LER with at most \( k \), \( k>1 \), links has no more than \( k-1 \) links, then \( \text{dist}^k \) (Egress LER) = \( \text{dist}^{k-1} \) (Egress LER).
• If the optimum path from Ingress LER to Egress LER with at most \( k \), \( k > 1 \), links has exactly \( k \) links, then it is made up of the optimum path from Ingress LER to some LSR \( j \) followed by the link \(<j, \text{Egress LER}>\). The path from Ingress LER to \( j \) has \( k-1 \) links and its cost is \( \text{dist}^{k-1}(j) \). All LSR \( i \) such that the link \(<i, \text{Egress LER}>\) is in the network are candidates for \( j \). Since we are interested in the optimum path, the \( i \) that minimizes \( \text{dist}^{k-1}(i) + \text{cost}(i, \text{Egress LER}) \)

These observations result in the following recurrence for \( \text{dist} \):

\[
\text{dist}^k(\text{Egress LER}) = \min \{ \text{dist}^{k-1}(\text{Egress LER}), \min \{ \text{dist}^{k-1}(i) + \text{cost}(i, \text{Egress LER}) \} \}
\]

This recurrence is used to formulate the following algorithm:

Algorithm BellmanFord(\(v, \text{cost}, \text{dist}, n\))

\{
  for \(i := 1\) to \(n\) do
    \(\text{dist}[i] := \text{cost}[v, i]\);
  for \(k := 2\) to \(n-1\) do
    for each \(u\) such that \(u \neq v\) and \(u\) has at least one incoming link do
      for each \(<i, u>\) in the network do
        if \(\text{dist}[u] > \text{dist}[i] + \text{cost}[i, u]\) then
          \(\text{dist}[u] := \text{dist}[i] + \text{cost}[i, u]\);
\}

5.8 MPLS Service Mapping

Table 5-2 proposes the possible mappings for all service combinations across the MPLS network to obtain the equivalent service. Taking into account that the 7 BAs within the MPLS network offer less service granularity than the Diffserv classes (one EF and four AF with three possible drop precedence levels in each class), traffic flows requiring similar service are grouped together into a single class, while the system's admission control and class selection rules ensure that the service requirements for flows in each of the classes are met.

Table 5-2 Proposed QoS mapping

<table>
<thead>
<tr>
<th>PHB</th>
<th>MPLS EXP Field</th>
</tr>
</thead>
</table>

5.9 Random Number Generation

The purpose of random number generation is to produce a sequence of numbers, drawn from the uniform distribution over the range 0 to 1, which appears to be independent. A good random number generator (RNG) should appear to be uniformly distributed on [0,1] and should not exhibit any correlation between generated numbers. It must be fast and avoid the need for much storage. A random number sequence must be reproducible; this aids debugging, and can be used to increase the precision of results.

Random function in Java program generally generates a sequence of uniform random numbers.

5.10 Mean Calculation

Mean delay is calculated from the following recurrence relations:

\[
\begin{align*}
\overline{X}_0 &= 0 \\
\overline{X}_n &= \frac{(n-1)\overline{X}_{n-1} + x_n}{n} \\
\end{align*}
\]

Here, \(\overline{X}_0\) = Average delay where number of packets=0

\(\overline{X}_n\) = Average delay where number of packets=n

\(\overline{X}_{n-1}\) = Average delay where number of packets=n-1

\(x_n\) = Delay in the nth packet
5.11 Variance Calculation

Variances is calculated from the following recurrence relations:

\[ T_0 = 0 \]
\[ T_n = T_{n-1} + x_n^2 \]
\[ V_n = \frac{1}{n} - \bar{X}_n \]  

(5.2)

Here, \( V_n \) = Delay variance for n packets.

5.12 Experimental Results and Discussion

To analyze the performance of admission control algorithm we ran a simulation for a network depicted in Figure 5-1. Each node is identical and is given a token capacity of 700K packets, mean service time is 2.73 μs and available bandwidth is 155Mbps. The basic unit of the packet size is 53 octets for the bulk arrival model. In this simulation we estimate the session generation rates from the nodes for different offered loads (from 10% to 100%) in the links of the network. Each node generates different multimedia sessions with different QoSs. The source of a session must be the generating node and the destination may be any other node in the network. The simulation has been repeated for different rerouting probabilities. The average size of a session is 760MB. The applications are assigned different levels of priorities described as follows [40].

VoIP (Voice over IP) is a jitter sensitive and real-time high interaction application with low drop precedence. This is classified as EF traffic with 100 ms delay and 50 ms jitter bound. Video Conferencing is a jitter sensitive and real-time application with high interaction. This is classified as AF1 traffic with 400 ms delay and 50 ms jitter bound. Terminal sessions that require interactive data transactions are classified as AF2 traffic with no jitter and 400 ms delay bound. Web applications are classified as AF3 with the same QoS requirements. Non real time applications (ftp, email) are classified as AF4 and they require 1 sec delay bound. Bandwidth requirement for AF class is defined by a
committed and a peak rate, whereas it is only a committed rate for EF class. The maximum allowed packet loss ratio for each class is $10^{-3}$.

Failing to meet the user requirements or QoS is detected by the queue size, packet loss and delay. Each node/hop stores the information when it successfully transmits a packet to the destination or fail to transmit a packet for each session.

Figure 5-15 to 5-21 compare theoretical average queuing delays and experimental (from simulation) queuing delays for EF1, EF2, AF1, AF2, AF3, AF4 and BE with different utilizations. Theoretical average queuing delays have been plotted for 100%, 98%, 95% and 90% using the equations in section 4.4. Queuing delays from the simulation has been plotted for the same utilizations. Results show that queuing delays of EF1, EF2 are less than theoretical average queuing delay and queuing delays of AF3, AF4 and BE are higher than average theoretical queuing delay. Queuing delays of AF1 and AF2 are almost equal with the average theoretical queuing delay. The reason behind the result is that EF1 enjoys the highest priority and BE enjoys the lowest priority.

---

![Figure 5-15 Queuing delay (simulation) for EF1 and average delay (theory) with different utilization](image-url)
Figure 5-16 Queuing delay (simulation) for EF2 and average delay (theory) with different utilization

Figure 5-17 Queuing delay (simulation) for AF1 and average delay (theory) with different utilization
Figure 5-18 Queuing delay (simulation) for AF2 and average delay (theory) with different utilization

Figure 5-19 Queuing delay (simulation) for AF3 and average delay (theory) with different utilization
Figure 5-20 Queuing delay (simulation) for AF4 and average delay (theory) with different utilization

Figure 5-21 Queuing delay (simulation) for BE and average delay (theory) with different utilization

Figure 5-22 to 5-28 compare theoretical average packet loss and experimental (from simulation) packet loss for EF1, EF2, AF1, AF2, AF3, AF4 and BE with different
utilizations. Theoretical average packet losses have been plotted for 100% and 98% using the equations in section 4.4. Packet loss from the simulation has been plotted for the same utilizations. Results show that packet losses of EF1, EF2, AF1, AF2, AF3 are less than theoretical average packet loss and packet losses of AF4 and BE are higher than average theoretical packet loss. The reason behind the result is that EF1 enjoys the highest priority and BE enjoys the lowest priority.

![Diagram](Figure 5-22 Packet loss (simulation) for EF1 and average packet loss (theory) with different utilization)
Figure 5-23 Packet loss (simulation) for EF2 and average packet loss (theory) with different utilization

Figure 5-24 Packet loss (simulation) for AF1 and average packet loss (theory) with different utilization
Figure 5-25 Packet loss (simulation) for AF2 and average packet loss (theory) with different utilization

Figure 5-26 Packet loss (simulation) for AF3 and average packet loss (theory) with different utilization
Figure 5-27 Packet loss (simulation) for AF4 and average packet loss (theory) with different utilization

Figure 5-28 Packet loss (simulation) for BE and average packet loss (theory) with different utilization
Figure 5-29 depicts Utilization Vs Data served based on theory and simulation. In our proposed network, there are 13 links and bandwidth for each link is 155 Mbs. So theoretical value of data served in the network for 60%, 80% and 90% are 1209 Mbs, 1612 Mbs, 1824 Mbs respectively. However in simulation, it has been found that total data served in the network for 60%, 80% and 90% are 1325 Mbs, 1576 Mbs, 1743 Mbs respectively.

![Graph showing Data served vs. Utilization](image)

**Figure 5-29 Data served vs. Utilization**

Figure 5-30 shows the end-to-end packet transfer delay for different rerouting probability. We have found that average delay increases if we increase the rerouting probability.
Figure 5-30 Average end to end delay for various utilizations vs. rerouting policy.

Figure 5-31 shows the Number of packet losses in 100 second from the system for different utilizations. If we increase the rerouting probability for higher offered load packet loss increases exponentially. The reason behind it is that when we reroute the flow to a new path, that path may become also congested.

Figure 5-31 Packet loss for various utilizations
Figure 5-32 shows the behavior of session success rate with the increase in offered utilization/throughput and the rerouting probability. Here we get different optimal session success rates with different rerouting probabilities for different values of offered utilizations. For a particular offered utilization (98%) session success rate increases with the increase in rerouting probability (below 35%) but if the rerouting probability is further increased the newly rerouted path is congested and finally session success rate decreases.

![Figure 5-32 Session success rate vs. % Reroute](image)

In Figure 5-33 we observe the behavior of average time delay and we find that it is sensitive to rerouting probability. Average time delay increases with the increase in rerouting probability. The profile becomes almost exponential for higher offered load. This is due to high congestion in the network.
Figure 5-33 Average waiting time with different loads and routing probability

Figure 5-34 shows the session success rates of different CoS for different offered loads with 30% rerouting probability. As EF gets the highest priority, it enjoys the highest session success rate. Session success rate decreases with the increase in offered load as it creates more congestion in the network. Less priority classes are more affected by this congestion than the high priority classes. We get about 72% session success rate on the average for BE class when the offered load is about 100%, whereas it is 99% for EF1 class.
5.13 Summary

In this chapter, a new flow-based admission control algorithm has been proposed with rerouting option through a DiffServ-Aware ATM based MPLS network for multiple service class environments to increase the session success rate ensuring QoS for multimedia transmission. The simulation demonstrates the performance of our algorithm in routing multimedia traffic. The numerical result shows that we get higher session success rate if we increase the rerouting probability for smaller offered utilization. But if the offered load is more than 0.6 we need to do admission control with lower rerouting probability to get higher session success rate.
6. Conclusion

This chapter summarizes the major contributions of this thesis and presents suggestions for future research.

6.1 Major Contributions

The central challenge being faced by the network operators is to maintain customer guarantees whilst sustaining profitability. A fundamental approach to this problem adopted by many providers has been to map traffic flows onto a physical topology using least cost metrics calculated by Interior Gateway Protocols. The limitations of this mapping were often resolved by over provisioning bandwidth along links that were anticipated being heavily loaded or to employ limited forms of load-sharing. However, as networks grow larger and the demands of the customer becomes greater (in terms of required bandwidth and QoS), the mapping of traffic flows onto physical topologies needs to be approached in a fundamentally different way so that the offered load can be supported in a controlled and efficient manner.

This thesis proposes a new flow-based admission control algorithm that solves the problem of dynamically managing traffic flows through the network by re-balancing streams during periods of congestion. It proposes mechanisms and procedures that will allow label switched routers (LSRs) in the network to utilise mechanisms within MPLS to indicate when flows may be about to experience possible frame/packet loss and to react to it. Based upon knowledge of the customers' SLAs, together with instantaneous flow information, the label edge routers (LERs) can then instigate changes to the LSP route to circumvent congestion that would hitherto violate the customer contracts.

A new flow-based admission control algorithm has been presented with rerouting option through the proposed Diffserv-Aware ATM based MPLS network for multiple service class environments to increase the session success rate ensuring QoS for multimedia transmission. Simulation demonstrates the performance of our algorithm in rerouting
multimedia traffic. This algorithm supports fully distributed controlled architecture in the enterprise network. Moreover, though rerouting IP packets by the router is an old concept but rerouting the flow is the new one.

We have proposed ATM based MPLS network architecture for Differentiated Service with a mapping policy of Diffserv flows to ATM based MPLS domain. The proposed network supports the virtual shortcut connection or MPLS label swapping using an underlying ATM capability. This has a number of advantages on throughput, end-to-end delay, and flow utilization. The proposed approaches guarantee a hard QoS using an ATM shortcut connection for different service class and increase resource utilization according to the per class QoS condition.

In policy-based networking, policies sometimes have to be combined and applied in cooperation to represent such programmable and customizable network functions such as Diffserv. We have designed and implemented three types of policies and three types of virtual flow labels (VFLs) to connect the policy rules to enable the representation of complex Diffserv policies. Policy combination also allows sub-classing of DSCP-based service classes, and the separation of service and subscriber policies. The careful design of Diffserv policies has enabled simple Diffserv policies to be represented in a simple form. We have also depicted packet scheduler for multiple class environments using our queuing architecture.

6.2 Future Scope of Work

We suggest the following research plans for Diffserv aware ATM based MPLS network as future research work.

- Further work may investigate the performance evaluation of heterogeneous source, as WANs will not typically be composed purely of sources with similar characteristics. Here heterogeneous source implies different link bandwidth, different transmitting rate, different buffer size etc. In our simulation program we have considered fixed session size. But in real world it varies significantly.
Further simulations could determine optimum number of rerouting a flow that will support QoS and optimize the profit for Internet Service Provider according to SLA. Relating optimum number of rerouting with offered and actual utilization mathematically will give new dimension to the researchers.

Finding mathematical equations of this network architecture for different queuing models might be a good topic for the researchers. Mathematical equations can be derived to find delay, jitter, packet loss for heterogeneous system and can be used those equation to calculate session success rate.

Another area of further work involving intelligent agents or artificial intelligence could be applied together with traffic engineering matrices to predict the congestion prior to it actually arising and to take proactive action accordingly. Intelligent agents can be used to find optimum admission of sessions, dynamic resource utilization and higher session success.
References


[41] L. Hobbs. UUNet's backbone, Nortel Networks.

Curriculum Vita

Surname: Ali

Given Name: Mohammad

Father Name: Md. Fazlul Rahman

Place of Birth: Rangpur, Bangladesh.

Date of Birth: January 5, 1974

Degree Awarded:

BSc. in Computer Science & Engineering 1999

Publication: