UNIVERSITAT POLITÈCNICA
DE CATALUNYA

Mechanisms to Reduce Routing Information Inaccuracy Effects: Application to MPLS and WDM Networks

Author: Xavier Masip Bruin

Advisor: Dr. Josep Solé-Pareta
Co-Advisor: Prof. Jordi Domingo-Pascual

COMPUTER ARCHITECTURE DEPARTMENT
TECHNICAL UNIVERSITY OF CATALONIA

A THESIS PRESENTED TO THE UNIVERSITAT POLITECNICA DE CATALUNYA
IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor en Enginyeria de Telecomunicació

Barcelona, June 2003

Dr.	
President	

Dr.	
Secretari	

Dr.	
Vocal	

Dr.	
Vocal	

Dr.	
Vocal	

Data de la defensa pública	
Qualificació	

“No s’ha de treballar ni per avui ni per demà, s’ha de treballar per sempre”

Manuel Hugue

“I never think about the future; it comes too fast”

Albert Einstein

Agraïments

Es molt difícil, quasi bé impossible, incloure totes les persones que d'una manera o altra han pres part en l'elaboració d'aquest document. Tot i això, una manera simple, justa i crec jo força acurada és analitzar l'evolució en els anys. Un primer agraïment ha de ser per els meus pares, qui fa ja molts anys em varen permetre començar i continuar els meus estudis.

En aquest punt, el següent pas és donar les gràcies a totes les persones que han contribuït en aquesta Tesi. En aquí he d'incloure els membres del grup de recerca de xarxes del DAC del qual en formo part, els membres del DAC a Vilanova i especialment els estudiants de la EPSVG en Sergi Benach i en Ivan Colomer.

Explícitament vull agrair a En Josep Solé i En Jordi Domingo el suport i la confiança que com a Directors de la Tesi m'han donat. És molt important l'ajut que en moments difícils, quan malauradament les coses no van com seria desitjable, hom pot rebre de qui t'està dirigint.

Com no, un reconeixement per els revisors anònims que han fet possible la "verificació per publicació" dels diferents resultats que configuren aquesta Tesi.

Aquesta Tesi es va començar ja fa uns quants anys a partir d'un tema de treball en el qual hi estàvem treballant dos persones. Finalment, tant en Sergi com jo hem aconseguit acabar el que 14 anys enrera no ens havíem ni imaginat. Tot i que són ja molts anys treballant plegats, Sergi, és un plaer compartir despatx amb tu.

Fins aquí l'evolució cronològica. Només em queda donar les gràcies a una persona. Sense tu, Carmen, això no s'hagués acabat mai. El teu suport ha sigut vital, *imprescindible i necessari per aconseguir aquest objectiu, tal i com ho ha estat per la resta de coses fonamentals de la meva vida.*

Finalment, tan sols complaure'm per l'oportunitat que em dona el fet d'escriure aquests agraïments de poder emprar la llengua del meu país per expressar-me.

Resum

Les xarxes IP tradicionals utilitzen el model de transmissió "best-effort" per transportar tràfic entre clients de la xarxa. Aquest model de transmissió de tràfic no és el més adequat per les aplicacions en temps real com per exemple, vídeo sota demanda, conferències multimedia o realitat virtual que per altra banda tenen cada cop més adeptes entre els clients de la xarxa. A fi de garantir el correcte funcionament d'aquest tipus d'aplicacions, l'estructura de la xarxa ha de ser substancialment modificada amb l'objectiu final de poder optimitzar els seus propis recursos i així poder fer front a aquells tipus de tràfics i de clients que requereixen certes garanties de "Qualitat de Servei" (QoS) per a la seva correcta transmissió.

Aquestes modificacions o millores de la xarxa poden ser perfectament realitzades sota l'entorn d'Enginyeria de Tràfic (Traffic Engineering, TE). Dos són els principals aspectes relacionats amb el funcionament de la xarxa en aquest entorn de TE: els mecanismes de commutació i els mecanismes d'encaminament. Així, per una banda es necessita un mecanisme de commutació molt ràpid en els nodes interns de la xarxa a fi de que els paquets de dades puguin ser processats amb el menor temps possible. En xarxes IP aquest objectiu s'aconsegueix amb el Multiprotocol Label Switching (MPLS). Per altra banda, a fi de garantir certa QoS, les decisions d'encaminament s'han de realitzar tenint en compte quines són les restriccions de QoS sol·licitades per el node client que origina el tràfic. Aquest objectiu s'aconsegueix modificant els esquemes d'encaminament tradicionals, incorporant-hi els paràmetres de QoS en les decisions d'encaminament, generant el que es coneix com algorismes d'encaminament amb QoS (QoS routing).

Centrant-nos en aquest darrer aspecte, la majoria dels algorismes d'encaminament amb QoS existents, realitzen la selecció de la ruta a partir de la informació d'estat de l'enllaç emmagatzemada en les bases de dades d'estat de l'enllaç contingudes en els nodes. Per poder garantir que els successius canvis en

l'estat de la xarxa estiguin perfectament reflectits en aquesta informació d'encaminament, el protocol d'encaminament ha d'incloure un mecanisme d'actualització que faci possible garantir que la selecció de les rutes es fa a partir d'informació acurada de l'estat real de la xarxa. En un entorn IP tradicional, el qual inicialment no inclou paràmetres de QoS, els canvis produïts en la informació d'encaminament són tan sols deguts a modificacions en la topologia i connectivitat de la xarxa. En aquest entorn, donat que la freqüència en la qual s'espera rebre missatges advertint d'aquestes modificacions no és elevada, la majoria dels mecanismes d'actualització es basen en la inclusió d'un cert període de refresc. Així, les bases de dades s'actualitzen periòdicament mitjançant la distribució d'uns missatges que informen a la resta de nodes de l'estat de la xarxa,a fi de que cada node pugui actualitzar la seva base de dades.

No obstant això, hem de tenir en compte que en aquelles xarxes IP/MPLS altament dinàmiques amb requeriments de QoS, aquest mecanisme d'actualització basat en un refresc periòdic no serà útil. Això és degut a la rigidesa que presenta aquest mecanisme, la qual fa que no sigui aplicable a un entorn que presenti contínues variacions del paràmetres dels enllaços cada cop que s'estableixi o s'alliberi una connexió (ara a més de la topologia i connectivitat, s'inclouen paràmetres de QoS, com ampla de banda, retard, variació del retard, etc.). Per tot això, s'haurà de generar un mecanisme d'actualització molt més eficient que sigui capaç de mantenir les bases de dades dels nodes perfectament actualitzades reflectint els continus canvis en l'estat de la xarxa. L'alta granularitat d'aquest mecanisme provocarà una sobrecàrrega de la xarxa, degut a l'enorme quantitat de missatges d'actualització que seran necessaris per poder mantenir informació actualitzada en les bases de dades d'estat de l'enllaç en cada node.

Per reduir aquesta sobrecàrrega de senyalització apareixen les polítiques d'activació (triggering policies) que tenen per objectiu determinar en quin moment un node ha d'enviar un missatge d'actualització a la resta de nodes de la xarxa advertint-los de les variacions produïdes en els seus enllaços. Desafortunadament, l'ús d'aquestes polítiques d'activació produeix un efecte negatiu sobre el funcionament global de la xarxa. En efecte, si l'actualització de la informació de l'estat de l'enllaç en els nodes no es fa cada cop que aquesta informació es veu

modificada, sinó que es fa d'acord a una certa política d'activació, no es podrà garantir que aquesta informació representi de forma acurada l'esta actual de la xarxa en tot moment. Això pot provocar una selecció no òptima de la ruta seleccionada i un increment en la probabilitat de bloqueig de noves connexions a la xarxa.

Aquesta Tesi es centra en definir i solucionar el problema de la selecció de rutes sota informació inexacta o no acurada de la xarxa, problema conegut com "routing inaccuracy problem". Es consideren dos escenaris de treball, les actuals xarxes IP/MPLS i les futures xarxes òptiques basades en Wavelength Division Multiplexing (WDM). Per ambdós escenaris es proposa un nou mecanisme d'encaminament: BYPASS Based Routing (BBR) per xarxes IP/MPLS i BYPASS Based Optical Routing (BBOR) per xarxes WDM. Els dos mecanismes comparteixen un concepte comú, denominat "bypass dinàmic".

El concepte de "bypass dinàmic", permet que un node intermedi de la xarxa encamini el missatge d'establiment de la ruta que ha rebut del node origen a traves d'una ruta diferent a la que havia estat inicialment calculada per el node font (i explícitament indicada en el missatge d'establiment) quan aquest node intermedi detecti que inesperadament, l'enllaç de sortida no disposa de recursos suficients per fer front a les garanties de QoS requerides per la connexió a establir. Aquestes rutes alternatives, denominades "bypass-paths", són calculades per certs nodes de la ruta principal, simultàniament amb la ruta principal en el node generador del tràfic o d'entrada a la xarxa..

En xarxes IP/MPLS el mecanisme BBR aplica el concepte de "bypass dinàmic" a les peticions de connexions amb restriccions d'ampla de banda, mentre que en xarxes WDM el mecanisme BBOR l'aplica a l'hora d'assignar una longitud d'ona per la qual es transmetrà el tràfic.

Els dos mecanismes i els algorismes d'encaminament que se'n desprenen són avaluats i comparats en diferents escenaris de simulació, per verificar que redueixen de forma més que eficient els efectes negatius produïts sobre el funcionament global de la xarxa, com són la probabilitat de bloqueig i la selecció de rutes incorrectes, degut al fet de realitzar la selecció de rutes sota informació d'encaminament no suficientment actualitzada.

Table of Contents

List of Figures....v
List of Tables....ix
Abbreviations....xi
Abstract.... xiii

PART I. INTRODUCTION....1

1. Objectives and Structure of the Thesis 3
   1.1 Thesis Motivation....5
   1.2 Thesis Objective....7
   1.3 Thesis Structure....8
2. QoS and TE in the new Internet....13
   2.1 Integrated Services model 14
   2.2 Differentiated Services model....16
   2.3 Multiprotocol Label Switching 18
3. QoS Routing 23

PART II. QoS ROUTING IN IP/MPLS NETWORKS....29

4. The Routing Inaccuracy Problem in IP/MPLS Networks....31
5. Review of Existing Solutions 39
   5.1 Safety Based Routing....41
   5.2 Explicit QoS Routing under Inaccurate Network State Information....42
   5.3 QoS Routing with Uncertain Parameters for End-to-End Delay Constrained Traffic .43
   5.4 Ticket-Based Distributed QoS Routing Scheme 45
   5.5 Centralized Server Based QoS Routing 46
   5.6 A localized QoS Routing Approach....48
   5.7 Crankback and Fast Rerouting 49

6. The BYPASS Based Routing Mechanism ... 51
6.1 Description of BYPASS Based Routing ... 53
6.2 Bypass-path Signalling ... 56
6.3 Example for Illustrating BBR Behaviour ... 58
6.4 Performance Evaluation ... 60
7. Applying the BBR Mechanism under Bandwidth Constraints ... 65
7.1 Example Illustrating the BOSP Behaviour ... 67
7.2 Performance Evaluation ... 69
8. BYPASS Discovery Process ... 85
8.1 Example to Illustrate the BDP Performance ... 86
8.2 Performance Evaluation ... 88

PART III. ROUTING IN WDM NETWORKS ... 93

9. Routing and Wavelength Assignment in WDM Networks ... 95
9.1 Introduction ... 95
9.2 Routing in WDM Networks ... 97
10. The Routing Inaccuracy Problem in WDM Networks ... 101
10.1 Problem Definition ... 101
10.2 State of the Art ... 103
11. BBOR: Adaptation of the BBR Mechanism to WDM Networks ... 105
11.1 BBOR Description ... 106
11.1.1 BBOR: A New Triggering Policy ... 106
11.1.2 BBOR: A New Routing Algorithm ... 107
11.2 Example Illustrating How BBOR Works ... 108
11.3 Performance Evaluation ... 111
12. The BBOR Mechanism in a Wavelength Conversion Scenario ... 117
12.1 Wavelength Interchangeable Networks ... 117
12.2 ALG3: Applying the BBOR Mechanism to WI Networks ... 118
12.3 Performance Evaluation ... 119

PART IV. CONCLUSIONS AND FUTURE WORK ... 125

13. Final Conclusions ... 127
14. Future Work ... 131

References 133
APPENDIX A: List of Publications and Projects...137
APPENDIX B: The ns/2 Simulator...141

List of Figures

Figure 1. *Path* and *Resv* messages....15
Figure 2. *PHB* in the *DiffServ* model....17
Figure 3. Routing inaccuracy effects in an *IP/MPLS* scenario....37
Figure 4. Bandwidth acceptance ratio in *SBR*: (a) Threshold; (b) Exponential classes....42
Figure 5. Centralized Server based QoS routing architecture....46
Figure 6. Rule 2 illustration....54
Figure 7. BYPASS Based Routing mechanism....55
Figure 8. IPv4 sub-object of the *ERO*....56
Figure 9. IPv4 sub-object of the *ERO*: (a) non-protected node; (b) protected node....57
Figure 10. Network topology used to illustrate the *BBR* behaviour....58
Figure 11. Network topology used in the simulations....61
Figure 12. Bandwidth Blocking Ratio for the Threshold triggering policy....61
Figure 13. Bandwidth Blocking Ratio for the Exponential class triggering policy....62
Figure 14. Routing Inaccuracy for the Threshold triggering policy....62
Figure 15. Routing Inaccuracy for the Exponential class triggering policy....63
Figure 16. Computed *bypass-paths* for the Threshold triggering policy....63
Figure 17. Computed *bypass-paths* for the Exponential class triggering policy....64
Figure 18. BOSP: The enhanced *BBR* mechanism....66
Figure 19. Routing algorithms inferred from the *BBR* mechanism....67
Figure 20. Network topology used to illustrate the *BOSP* algorithm....68
Figure 21. Bandwidth Blocking Ratio for Scenario 1 (Threshold triggering policy)....69
Figure 22. Bandwidth Blocking Ratio for Scenario 1 (Exponential class triggering policy)....70
Figure 23. Routing Inaccuracy for Scenario 1 (Threshold triggering policy)....71
Figure 24. Routing Inaccuracy for Scenario 1 (Exponential class triggering policy)....71
Figure 25. Computed *bypass-paths* for Scenario 1 (Threshold triggering policy)....72
Figure 26. Computed *bypass-paths* for Scenario 1 (Exponential class triggering policy)....72
Figure 27. The *ISP* topology used in simulations....74

Figure 28. Bandwidth Blocking Ratio for Scenario 2 (Threshold triggering policy)..........74
Figure 29. Bandwidth Blocking ratio and network load for Scenario 2 (*tv* = 20%, *tv* = 40%)..........75
Figure 30. Bandwidth Blocking ratio and network load for Scenario 2 (*tv* = 60%, *tv* = 80%)..........75
Figure 31. *WSP* algorithm behaviour as a function of the network load (Scenario 2)........76
Figure 32. *SSP* algorithm behaviour as a function of the network load (Scenario 2)..........76
Figure 33. *SOSP* algorithm behaviour as a function of the network load (Scenario 2).......77
Figure 34. *BOSP* algorithm behaviour as a function of the network load (Scenario 2)......77
Figure 35. Routing Inaccuracy for Scenario 2 (Threshold triggering policy)..........78
Figure 36. Computed *bypass-paths* for Scenario 2 (Threshold triggering policy)..........78
Figure 37. Bandwidth Blocking Ratio for Scenario 3 (Threshold triggering policy)..........80
Figure 38. Bandwidth Blocking Ratio and network load for Scenario 3 (*tv* = 20%, *tv* = 40%)..........80
Figure 39. Bandwidth Blocking Ratio and network load for Scenario 3 (*tv* = 60%, *tv* = 80%)..........81
Figure 40. *WSP* algorithm behaviour as a function of the network load (Scenario 3)........81
Figure 41. *SSP* algorithm behaviour as a function of the network load (Scenario 3)..........82
Figure 42. *SOSP* algorithm behaviour as a function of the network load (Scenario 3).......82
Figure 43. *BOSP* algorithm behaviour as a function of the network load (Scenario 3)......83
Figure 44. Routing Inaccuracy for Scenario 3 (Threshold triggering policy)..........83
Figure 45. Computed *bypass-paths* for Scenario 3 (Threshold triggering policy)..........84
Figure 46. *BDP* process..........86
Figure 47. *BDP* performance: illustrative example..........87
Figure 48. Bandwidth Blocking Ratio for the Threshold triggering policy..........89
Figure 49. Bandwidth Blocking Ratio for the Exponential class triggering policy..........89
Figure 50. Routing Inaccuracy for the Threshold triggering policy..........90
Figure 51. Routing Inaccuracy for the Exponential class triggering policy..........91
Figure 52. Computed *bypass-paths* for the Threshold triggering policy..........91
Figure 53. Computed *bypass-paths* for the Exponential class triggering policy..........92
Figure 54. Network evolution..........96
Figure 55. *ASON* architecture..........97
Figure 56. Forward reservation: (a) Successful; (b) Unsuccessful..........99
Figure 57. Backward reservation: (a) Successful; (b) Unsuccessful..........100
Figure 58. Routing inaccuracy effects in *WDM* networks..........102

Figure 59. *BBOR* description 109
Figure 60. Network topology used in the *BBOR* illustrative example 109
Figure 61. Topology used in simulations 113
Figure 62. Number of update messages 114
Figure 63. Number of *OSW* as a function of the threshold percentage T_p value 115
Figure 64. Blocking probability for N = 6 and T_p = 50% 115
Figure 65. Blocking probability for N = 10 and T_p = 50% 116
Figure 66. *ALG3* description 120
Figure 67. Number of update messages 120
Figure 68. Blocking probability in *WS* networks 121
Figure 69. Blocking probability in *WI* networks 122

List of Tables

Table 1. IP precedence values mapped to DSCP 16
Table 2. Link QoS attributes...59
Table 3. Feasible routes and selected paths depending on the algorithm in use.................59
Table 4. Link QoS attributes...31
Table 5. Possible *Bypass-paths*...69
Table 6. Cost analysis...79
Table 7. *BBR* Process when including the BDP 87
Table 8. Network State in OXC1..110
Table 9. Routing Table in OXC1...110
Table 10. Illustrative Example..110

Abbreviations

ASON	Automatic Switched Optical Networks
ATM	Asynchronous Transfer Mode
ATMF	ATM Forum
ALG	Algorithm
BA	Behaviour Aggregate
BBR	BYPASS Based Routing
BBOR	BYPASS Based Optical Routing
BDP	BYPASS Discovery Process
BGP	Border Gateway Protocol
BPA	Bypass-Path Address
BOSP	Balance-Obstruct-Sensitive Path Algorithm
BRP	Backward Reservation Protocol
CoS	Class of Service
CR	Constraint Based Routing
D-RSP	Dynamic-Restricted Shortest Path
DiffServ	Differentiated Services
DSCP	Differentiated Services Code Point
ERO	Explicit Routing Object
FRP	Forward reservation Protocol
GMPLS	Generalized Multiprotocol Label Switching
IE	Crankback Information Element
IETF	Internet Engineering Task Force
IntServ	Integrated Services
IP	Internet Protocol
IS-IS	Intermediate System-Intermediate System
ISP	Internet Services Provider
LCP	Least-Congested Path Algorithm
LDP	Label Distribution Protocol
LSP	Label Switched Path
LSR	Label Switching Router
MCP	Multi-constrained Path Problem
MDWCRA	Maximum Delay-Weighted Capacity Routing Algorithm
MIRA	Minimum Interference Routing Algorithm
MP-BCP	Most-Probable Bandwidth Constrained Path
MP-DCP	Most-Probable Delay Constrained Path
MPLS	Multiprotocol Label Switching
MRP	Most Reliable Path
n_bp	Computed *bypass-paths* per route
NTDB	Network Topology Database
OP	Optimal Partition
OP-MP	Optimally Partitioned Most Probable Path
OSL	Obstruct-Sensitive Link
OSPF	Open Shortest Path First
OSSP	Obstruct-Sensitive-Shortest Path Algorithm
OSW	Obstruct-Sensitive Wavelength
OTN	Optical Transport Network
OXC	Optical Cross-Connect

PBR	Profile Based Routing
PHB	Per-Hop-Behaviour
PNNI	Private Network-to-Network Interface
PSR	Proportional Sticky Routing
QoS	Quality of Service
RFC	Request for Comments
RIP	Routing Information Protocol
RSP	Restricted Shortest Path
Rspec	Reservation Specification
RSVP	Resource Reservation Protocol
RTC	Routing Table Cache
RWA	Routing and Wavelength Assignment
SBR	Safety Based Routing
SLA	Service Layer Agreement
SOSP	Shortest-Obstruct-Sensitive Path Algorithm
SP	Shortest Path Algorithm
SSP	Shortest Safest Path Algorithm
SWP	Shortest-Widest Path
TBP	Ticket Based Probing
TE	Traffic Engineering
TED	Traffic Engineering Database
ToS	Type of Service
Tspec	Traffic Specification
VCI	Virtual Channel Identifier
VCR	Virtual Capacity Routing
VPI	Virtual Path Identifier
VPN	Virtual Private Network
WDM	Wavelength Division Multiplexing
WI	Wavelength-Interchangeable Network
WS	Wavelength-Selective Network
WSOSP	Widest-Shortest-Obstruct-Sensitive Path Algorithm
WSP	Widest-Shortest Path

Abstract

Traditional IP networks are based on the best effort model to transport traffic flows between network clients. Since this model cannot properly support the requirements demanded by several emerging real time applications (such as video on demand, multimedia conferences or virtual reality), some modifications in the network structure, mainly oriented to optimise network performance, are required in order to provide Quality of Service (QoS) guarantees.

Traffic Engineering is an excellent framework to achieve these network enhancements. There are two main aspects in this context that strongly interact with network performance: switching mechanisms and routing mechanisms. On one hand, a quick switching mechanism is required to reduce the processing time in the intermediate nodes. In IP networks this behaviour is obtained by introducing Multiprotocol Label Switching (MPLS). On the other hand, a powerful routing mechanism that includes QoS attributes when selecting routes (QoS Routing) is also required.

Focusing on the latter aspect, most QoS routing algorithms select paths based on the information contained in the network state databases stored in the network nodes. Because of this, routing mechanisms must include an updating mechanism to guarantee that the network state information perfectly represents the current network state. Since network state changes (topology) are not produced very often, in conventional IP networks without QoS capabilities, most updating mechanisms are based on a periodic refresh.

In contrast, in highly dynamic large IP/MPLS networks with QoS capabilities a finer updating mechanism is needed. This updating mechanism generates an important and non-desirable signalling overhead if maintaining accurate network state information is pursued. To reduce the signalling overhead, triggering policies are used. The main function of a triggering policy is to determine when a network node must advertise changes in its directly connected links to other network nodes.

As a consequence of reduced signalling, the information in the network state databases might not represent an accurate picture of the actual network state. Hence, path selection may be done according to inaccurate routing information, which could cause both non-optimal path selection and an increase in connection blocking frequency.

This Thesis deals with this routing inaccuracy problem, introducing new mechanisms to reduce the effects on global network performance when selecting explicit paths under inaccurate routing information. Two network scenarios are considered, namely current IP/MPLS networks and future WDM networks, and one routing mechanism per scenario is suggested: BYPASS Based Routing (BBR) for IP/MPLS and BYPASS Based Optical Routing (BBOR) for WDM networks. Both mechanisms are based on a common concept, which is defined as dynamic bypass.

According to the dynamic bypass concept, whenever an intermediate node along the selected path (unexpectedly) does not have enough resources to cope with the incoming MPLS/optical-path demand requirements, it has the capability to reroute the set-up message through alternative pre-computed paths (bypass-paths). Therefore, in IP/MPLS networks the BBR mechanism applies the dynamic bypass concept to the incoming LSP demands under bandwidth constraints, and in WDM networks the BBOR mechanism applies the dynamic bypass concept when selecting light-paths (i.e., selecting the proper wavelength in both wavelength selective and wavelength interchangeable networks).

Finally, the applicability of the proposed BBR and the BBOR mechanisms is validated by simulation and compared with existing methods on different network scenarios. These network scenarios have been selected so that obtained results may be extrapolated to a realistic network.

PART I

INTRODUCTION

In this introductory Part, the motivation, goals and structure of the Thesis are presented. The current Internet scenario including concepts related to Quality of Service, such as routing mechanisms, new network architectures and signalling protocols are also introduced in this Part so that the problem to be addressed in next Part can be easily allocated in the current network scenario.

Chapter 1

Objectives and Structure of the Thesis

The network conception has extraordinarily evolved from that day in September 1969, not so far in the short network history, when Leonard Kleinrock went down so well at the challenge of sending a message from his Host Computer at UCLA. From that day when the first message containing the word "LOGIN" partially succeed on reaching the destination node (indeed destination node only got the word "LOG") to the current Internet, network components, network connectivity, network applications and network utilization have constantly evolved in a way that maybe not even more optimistic might imagine.

Currently, the Internet does not provide *Quality of Service* (QoS) guarantees, and data delivery is based on a simple best-effort transmission model. Emerging real-time applications, such as video on demand, multimedia conferences or virtual reality, cannot be supported under this network definition, due to both the variable delays in the queuing process and the problem of congestion. Before these applications can be used, the network has to be enhanced to support end-to-end QoS. This enhancement lies in optimising network performance to provide QoS guarantees by improving the utilization of network resources (*Traffic Engineering*, *TE*) [1], and providing resilience features for quick recovery from failures. In the *TE* context there are two determinant factors: to have a quick switching mechanism and to have powerful routing mechanisms to select the path by which traffic is sent. The first objective can be reached by implementing *Multiprotocol Label Switching* (*MPLS*) [2] and the second can be achieved by adding new routing mechanisms oriented to improving network performance.

MPLS is an advanced, connection-oriented forwarding scheme, which allows streams from any particular ingress node *(Label Switching Router*, *LSR*, in an *MPLS* domain) to any particular egress node to be individually identified with a simple label. Therefore, *MPLS* provides a straightforward mechanism to forward the traffic associated with each ingress node to egress node pair and substantially improves source routing, since the IP address of the intermediate nodes need not be piggybacked on each packet along the end-to-end path (*Label Switched Path, LSP*, in an *MPLS* domain).

As a basic definition, it is possible to say that the main routing objective is to drive packets to the right destination. In order to perform this goal, the routing process selects the path that can best transport the traffic from the source node to the destination node. This selection is done in accordance with the network state information, namely the current network topology in traditional IP routing, which is mostly obtained from several databases where this information is maintained. In traditional hop-by-hop routing each intermediate node instantly decides where the current packet should be sent in order to reach the destination according to the database information existing in the node. However, if routing is explicitly done in the source nodes (i.e., source routing), only the source node’s routing information is considered. It should be noticed that source routing leads to a reduction in both network control complexity and in the number of databases checked when selecting a path.

Traditional IP routing algorithms are OSPF [3], RIP [4], IS-IS [5] and BGP [6]. Basically, they are based on computing the shortest path by applying either Dijkstra's algorithm or the Bellman-Ford algorithm. The first is a link state algorithm and the second a distance-vector algorithm. The databases mentioned above are filled with link state information or distance information, respectively. In any case, both algorithms ensure only a best-effort performance, which is not recommended for applications demanding specific QoS guarantees. In fact, two network models might be described: on one hand, a network that supports the best effort transmission model for traffic without QoS requirements; and on the other hand, a QoS network model necessary to support traffic with QoS constraints.

Currently, there are several QoS routing algorithms proposed in the literature to cope with the QoS network model that improves the best-effort transmission model currently used in Internet. Unlike traditional IP routing algorithms, for QoS provisioning the routing algorithm must take into account more parameters than exclusively those related with topology and connectivity. QoS routing algorithms include QoS parameters in the path decision process to select the most suitable path in accordance with both traffic requirements and network state, including network topology (quasi-static) and the resources available at each node (dynamic). The management of these parameters is done by the routing protocol, which must include a mechanism to collect, distribute and update all of the parameters needed by the QoS routing algorithm.

1.1 Thesis Motivation

Important factors in the global routing behaviour are where and how routing decisions are taken. Hence, assuming source routing, if resource availability is a key factor in the path selection process, the QoS parameters used by the routing algorithm to decide the routes must perfectly represent the current network state. Two different aspects must be considered to guarantee accurate QoS routing.

First, a mechanism to keep the network state information perfectly updated must be included in the routing protocol. The main function of this mechanism is to decide when a node must send update messages throughout the network to advertise state changes in its directly connected links to all the other nodes. In traditional IP networks, without QoS requirements, network state changes are only due to topology variations, which are not often expected. Hence, maintaining accurate information in the network state databases is easily achieved without many update messages. However, in *IP/MPLS* networks with QoS requirements, keeping the link state databases perfectly updated involves sending update messages whenever a new *LSP* is established or an existing one is released.

Therefore, in a large connection-oriented packet-switched network scenario, where changes are produced very often, this updating process generates a non-desirable signalling overhead. To reduce such a signalling overhead, one of several currently available triggering policies is applied. An unfortunate consequence of applying any of these triggering policies is that the information contained in the network state databases might not represent the current network state at the path set up time. This is only one of the possible causes of inaccurate routing information. When routing is done under inaccurate routing information, the route selected according to this information can be blocked in the path set up process (*routing inaccuracy problem*). This is due to the possibility that the resources required by the incoming *LSP* are not available, contrary to what the database on the source node states.

Second, in addition to the problem above, most QoS routing algorithms utilize the nominal available bandwidth information of the links to select paths. This routing information is obtained from the link state databases stored on each node, which are updated regarding traffic requirements included in new *LSP* demands by applying any triggering policy. However, assuming that most of the clients generating network traffic do not completely use the requested bandwidth, that is, do not strictly fulfil the *Service Layer Agreement* (SLA), a difference exists between the nominal link utilization and the actual link utilization. This gap leads to non-efficient network resource utilization, since the path selection process is performed according to nominal link state information instead of actual link utilization. This problem is addressed in [7], where the authors propose a path selection scheme based on obtaining more accurate link utilization information. Initially, this link utilization information may be obtained from the link state databases. However, this is not possible over time due to scalability concerns. In fact, in this scenario an update message must be sent not only when an *LSP* is established or released, but also when the actual bandwidth used by the traffic flowing by an already established *LSP* changes.

This degree of granularity cannot be supported by any network, so a different method to collect the actual link utilization information is required. The authors propose a method based on estimating the future link utilization. They indeed suggest an algorithm, the *Available Bandwidth Estimation Algorithm,* for computing both an estimation of the real available bandwidth on each link and the duration for which the estimate is valid. This prediction is obtained by sampling the network state periodically, with the period of time changing dynamically depending on the traffic characteristics and the required conservatism of the network domain.

Based on the estimated values obtained by applying this algorithm, a new path selection algorithm is suggested that uses the estimated available bandwidth values as weights of the links, which then are used by a shortest widest path routing algorithm to select the optimal path. Finally, in order to limit network congestion, a threshold parameter is added. Once the path has been computed by using the modified shortest widest path, the routing algorithm computes the available bandwidth on the bottleneck link of the path. Then the threshold parameter is applied to this bottleneck value to compute a benchmark for path selection in such a way that if the bandwidth requested is larger than a certain fraction of the bottleneck link bandwidth, the incoming request is rejected. The authors show that the proposed path selection algorithm performs better than the shortest widest path routing algorithm, because the proposed routing algorithm based on the Available Bandwidth Estimation Algorithm has more accurate information about the actual link load.

1.2 Thesis Objective

Two different aspects have been discussed related to the routing information used to select the routes so far. Both aspects may be summarized as the effects produced in global network performance when using either inaccurate link state information or inaccurate link utilization information. This Thesis focuses on addressing the first scenario, namely the routing inaccuracy problem produced when the path selection process is performed under inaccurate nominal link state information. From this point, the expressions nominal link state information and link state information are used with the same meaning in this Thesis. Hence, explicit QoS routing is considered, where the explicit paths are computed, selected and established by the source node (or ingress node), assuming that the path decision is taken based on inaccurate nominal link state information.

IP/MPLS networks are being considered so far. However, it is well known that the network model is evolving to an *Optical Transport Network* (*OTN*). Optical Transport Networks based on *Wavelength Division Multiplexing* (*WDM*) appear as a potential solution to cope with the increasingly growth of Internet traffic demands. In such systems all-optical *WDM* channels are used to allow the end-to-end users communication. These *WDM* channels are referred as lightpaths, and must be selected in a proper manner in order to optimize the network resources. It is in this point where the routing becomes an important factor in the global network performance. Once more, the accuracy of the network state information is a key aspect to be considered when selecting lightpaths.

The main goal of this Thesis is to propose a new routing mechanism able to reduce the impact on global network performance because of selecting routes under inaccurate routing information. Two mechanisms, *BYPASS Based Routing* (*BBR*) and *BYPASS Based Optical Routing (BBOR)* are proposed in this Thesis to address the routing inaccuracy problem in *IP/MPLS* and *WDM* networks respectively. The *BBR* mechanism contributes to the improvement of global network performance in terms of a substantial connection blocking reduction and a more optimal path selection when the incoming traffic requires bandwidth guarantees. Four routing algorithms inferred from the *BBR* mechanism are evaluated by simulation. The *BBOR* mechanism modifies route selection and wavelength assignment in such a way that connection blocking is reduced as well. Again, three routing algorithms inferred from the *BBOR* mechanism are also evaluated by simulation.

1.3 Thesis Structure

This Thesis is organized in four Parts each one divided into several Chapters. A brief description of each one is now presented.

PART I: INTRODUCTION

Chapter 2: After introducing the network scenario in Chapter 1, this Chapter introduces the main concepts, goals and capabilities of *Traffic Engineering* (*TE*). Main features of proposed QoS network architectures, such as *Integrated Services* and *Differentiated Services* are also briefly introduced. Then, skills and drawbacks of the *Resource Reservation Protocol* (*RSVP*) are also presented. Finally, the use of *Multiprotocol Label Switching (MPLS)* is justified based on the benefits on global network performance due to its application. The main goal of this Chapter is to define the working scenario, focusing on the network modifications in terms of structure and architecture needed to cope with current network challenges.

Chapter 3: A main aspect in the network evolution concerns to QoS Routing. Chapter 3 finely focuses on this topic. Routing evolution, from traditional IP routing to QoS routing, and main routing concepts are described in this Chapter. Most recent routing algorithm proposals existing in the literature are shown to illustrate routing evolution and main routing problems.

PART II: QoS ROUTING IN IP/MPLS NETWORKS

Chapter 4: In Chapter 4, the problem addressed in this Thesis, called *routing inaccuracy problem*, is introduced. Causes and origins that motivate the existence of this problem are clearly presented and justified in this Chapter.

Chapter 5: Some solutions exist in the literature addressing the *routing inaccuracy problem*. A clear and extensive description of these solutions is presented in Chapter 5.

Chapter 6: Then, Chapter 6 presents a routing mechanism called *BYPASS Based Routing (BBR)* as the solution proposed in this Thesis to cope with the *routing inaccuracy problem*. Two new routing algorithms are inferred from the *BBR* mechanism in this *IP/MPLS* scenario. Different simulations are evaluated to verify the benefits of these routing algorithms.

Chapter 7: After that, in Chapter 7, the *BBR* mechanism is applied under bandwidth constraints, generating two new routing algorithms which are also evaluated by simulation and compared with previous *BBR* algorithms.

Chapter 8: The *BYPASS Discovery Process* (*BDP*) is proposed in Chapter 8 to improve the *BBR* performance. The *BDP* extends the *BBR* applicability, so improving the obtained benefits.

PART III: ROUTING IN WDM NETWORKS

Chapter 9: This Chapter shows the expected evolution in network technology, network usage and network requirements that ends in the introduction of optical networks. Main concepts in optical networks, such as those related to routing and the definition of a Control Plane are also described.

Chapter 10: This Chapter serves to introduce the problem. Once the advantages of applying the *BBR* mechanism in an *IP/MPLS* scenario have been analyzed, the problem is extended to optical networks. Finally, work existing in the literature copying with the *routing inaccuracy problem* in optical networks is discussed.

Chapter 11: The *BYPASS Based optical Routing (BBOR)* is proposed as a solution to address the *routing inaccuracy problem* in optical networks. The *BBOR* is based on extending the main concepts of the *BBR* mechanism. Unlike the *BBR* mechanism the *BBOR* also proposes a new triggering policy to reduce the signalling overhead. This Chapter focuses on applying the *BBOR* to wavelength selective networks where wavelength conversion is not permitted.

Chapter 12: After evaluating the *BBOR* behaviour in wavelength selective networks, the *BBOR* is modified to be applied on wavelength convertible networks. Again, the benefits of applying the *BBOR* are verified by simulation.

PART IV: CONCLUSIONS AND FUTURE WORKS

Chapter 13: In this Chapter, the main contributions of the Thesis are pointed out. A short summary is also introduced to conclude the Thesis.

Chapter 14: Finally, the future work is presented. In this Chapter, future lines of work, some of them already planned are briefly described. It is worth to notice that because of the nature of the Thesis which proposes two routing mechanisms to be applied on two different network scenarios, future research is also suggested in both network scenarios and to interoperate between both scenarios.

This Thesis ends with two Appendix. Appendix A lists publications and projects related to this work. Appendix B briefly describes modifications needed on the network simulator to provide it with *BBR* and *BBOR* capabilities

Chapter 2

QoS and TE in the new Internet

Implementing *Traffic Engineering* (*TE*) is fundamental in the current network models, mainly because current routing protocols forward traffic through the shortest paths. *TE* capabilities can be provided by the *Multiprotocol Label Switching* (*MPLS*). In fact, as explained later on the main, direct and most significant benefit provided by the *MPLS* is the provision of *TE* capabilities. The main target of *TE* is to control how traffic flows through one's network so as to optimise network performance and resource utilization [1]. Traditional shortest path selection leads to congestion on some links along the selected path while longer paths are under-utilized. Congestion might be reduced by acting on the routing metric, but this solution may only be suggested on a small network scenario. In a large *Internet Service Provider* (*ISP*) network, new tools are needed for *TE* provisioning. *TE* provides the network with three main aspects: includes a guaranteed QoS, improves the utilization of network resources and provides for quick recovery when a node or link fails. Traditional IP routing mechanisms based on the shortest path selection are extended to balance, distribute and optimise the networks resources turning out QoS routing. Moreover, developing fast rerouting mechanisms is vital to address network components failures. New overall QoS architectures, such as *Integrated Services* (*IntServ*) and *Differentiated Services* (*DiffServ*) are needed to support a QoS network scenario. Even though this Thesis focuses on the first aspect a short description of both architectures is now presented just to introduce the current networking scenario.

2.1 Integrated Services model

The *IntServ* model, developed by the *Internet Engineering Task Force* (IETF), provides an end-to-end QoS solution. *IntServ* defines a set of service classes which specify the potential needs of different clients. *IntServ* follows the signalled QoS model, in which network resources are reserved according to the QoS needs signalled by the end clients. Therefore, end-to-end QoS is obtained by way of end-to-end signalling, state maintenance for each signalled session and admission control at each network element. Two parameters are mainly defined, a traffic specification called *Tspec* and a reservation specification called *Rspec*. The former specifies the kind of application traffic used by the client and so, incoming to the network. The latter specifies the required level of QoS and the reservation of network resources. In accordance with these specifications, *IntServ* requires network elements such as routers, to perform the following:

- policing functions to verify that incoming traffic is conformed to its *Tspec*. Packets that do not meet the *Tspe*c values are dropped.
- admission control functions to check if there are enough resources to support the QoS traffic requirements. If available resources are not enough, the incoming request is rejected.
- packet classification functions, queuing and scheduling mechanisms to separate and properly handle those packets demanding for a specific level of QoS.

There are two service classes defined in the *IntServ*, guaranteed service and controlled load. The guaranteed service class provides for strict bounds on end-to-end delay and assured bandwidth for traffic that meet the requested QoS

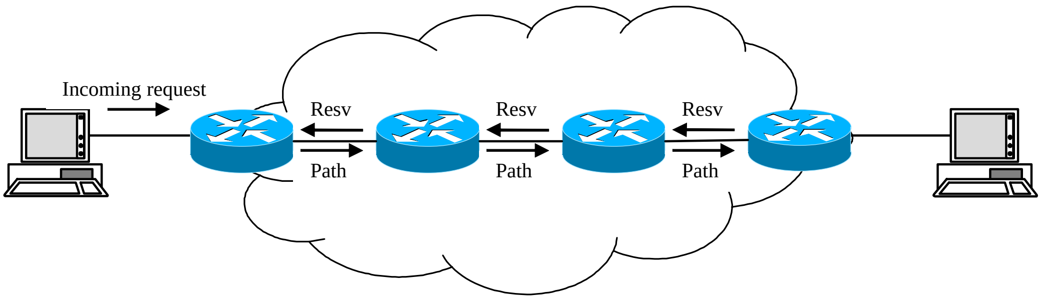

Figure 1. *Path* and *Resv* messages

specifications. The controlled load service class provides a better-than-best effort and low delay service under light to moderate network loads.

A signalling protocol must also be specified in the *IntServ* network model to perform its functions. The *Resource Reservation Protocol (RSVP)* [8] is an *IntServ* signalling protocol used by both the end clients to demand their QoS needs according to the defined *IntServ* service classes and the core network to handle the path establishment. The *RSVP* must be implemented in all the network elements, to allow clients to demand their specific QoS levels. *RSVP* was born to define, establish and maintain reservation of those resources required by a certain *LSP* incoming demand to succeed on flowing traffic. There are two main messages carrying the *RSVP* information: *Path* and *Resv*. Figure 1 illustrates both messages. Once the source node, applying a certain routing algorithm, computes a route reacting to an incoming *LSP* demand, it sends a *Path* message which is forwarded downstream across the selected path to reach the destination. *Path* message includes an *Explicit Routing Object* (*ERO*) including the IP address of the intermediate nodes across the selected path, and *Tspec* and classification information provided by the source node. When the *Path* message reaches the destination node, this node sends a *Resv* message back to the sender along the reverse path, identifying the session for which the reservation is to be made. The reservation process ends with a new path where the traffic associated with the incoming *LSP* demand flows or with a rejection message when there are not enough available resources to cope with the QoS traffic requirements.

However, the *IntServ* and *RSVP* applicability was not really extended due to scalability concerns. In fact, a soft-state must be maintained on all nodes along the selected route to keep the resource reservation alive. This implies a large number of signalling messages flowing throughout the network. Hence, *IntServ* may only be applied to small networks (Intradomain Routing). Even though *RSVP* is really not used to perform resource reservation, it may be used as a signalling protocol.

2.2 Differentiated Services model

As the Internet traffic and the diversity of applications grow, different QoS levels must be applied to different traffic flows demanding for specific differentiated services. The *Diffserv* model [9] for IP QoS provisioning has also been proposed by the *IETF* to allow the network to support different QoS levels according to the QoS required by the end user. This model is very similar to the IP precedence model. The IP precedence model handles traffic by classifying various traffic flows into aggregated classes. The appropriate QoS is required for each aggregate class. Three bits in the *Type of Service* (*ToS*) field stand for eight different aggregated classes. As in the IP precedence model, *DiffServ* model divides traffic into a small number of classes and allocates resources on a per-class basis. This classification is made based on the information contained in the *Differentiated Services Code Point* (*DSCP*) byte. The *DSCP* is carried in the *ToS* or in the *Class of Service* (*CoS*) fields in the IPv4 or IPv6 header respectively. According to [10] only 6 weighted bits are meaningful in the *DSCP* byte while the last 2 bits are currently not used. Hence, even though 64 different classes might be implemented, in practice only a few classes are really implemented. The above defined eight IP precedence levels can be mapped to a fixed *DSCP* classes as shown in Table 1.

Table 1. IP precedence values mapped to *DSCP*

IP Precedence	*Name*	*DSCP*
0	Routine	DSCP 0
1	Priority	DSCP 8
2	Immediate	DSCP 16
3	Flash	DSCP 24
4	Flash override	DSCP 32
5	Critical	DSCP 40
6	Internet control	DSCP 48
7	Network control	DSCP 56

Packets are marked to be properly handled across the path. When the ingress node receives an IP packet it sets the *DSCP* to identify this packet to the desired service class. Then, intermediate nodes across the path check the *DSCP* field value and determine the QoS required by this packet, as shown in Figure 2. This is known as *Per-Hop-Behaviour* (*PHB*). The *PHB* includes all the mechanisms (i.e., packet scheduling, queuing, policing or sapping behaviour) applied in a node to provide the packet for the required QoS. *Behaviour Aggregate* (*BA*) is defined as the set of packets which traversing the same node have the same *DSCP*.

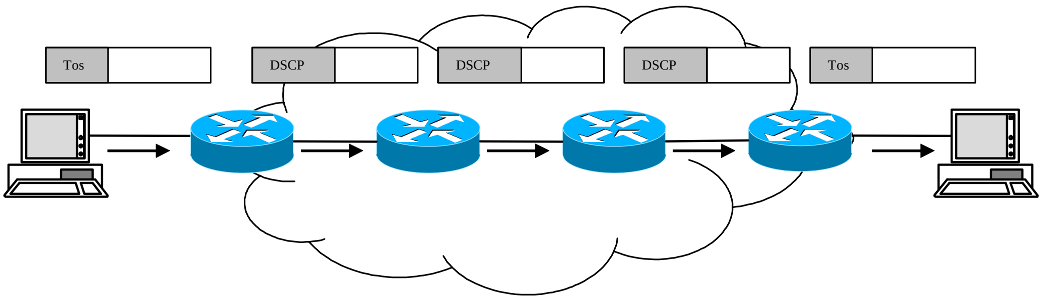

Figure 2. *PHB* in the *DiffServ* model

Four *PHB* implementations are currently available:

- Default *PHB*: In this case only best effort delivery is guaranteed. Packets marked with a *DSCP* value that cannot be mapped to a *PHB* are directly mapped to this *PHB*.
- Class-selector *PHB*: This *PHB* allows IP precedence model to be compatible with the *DiffServ* model.
- Expedited Forwarding (*EF*) *PHB*: Packets marked with the *EF* are prioritized for delivery over others, by providing for low packet loss, low latency, low jitter and guaranteed bandwidth service.
- Assured Forwarding (*AF*) *PHB*: This *PHB* specifies an *AF* class and drop precedence for IP packets, in such a way that different forwarding assurances are given. There are four classes, each one specifying three drop precedence values. In case of congestion packets are dropped based on their relative drop precedence values within the *AF* class.

There are several main characteristics which ease *DiffServ* implementation:

- IP must not be modified as packets are marked in the edge node by using either the *ToS* or *CoS* fields in the IPv4 or IPv6 header respectively.
- Network applications must not be modified to implement *DiffServ* model as an *SLA* has been previously agreed between network provider and network client.
- It is easily scalable due to the aggregation mechanism. In fact traffic matching the same *DSCP* value is handled in the network as an aggregate.
- Intermediate nodes must not maintain a soft-state including information about packet flows as packets are individually handled at intermediate nodes.

2.3 Multiprotocol Label Switching

IntServ and *DiffServ* architectures can be implemented using the *Multiprotocol Label Switching* (*MPLS*). *MPLS* is an advanced label-based switching mechanism which uses the information contained in the labels to compute the routes. In the first case, *IntServ* model, bonds are created between labels and *RSVP* flows, to identify the type of resource reservation associated with the traffic. In the second case, *DiffServ* model, the appropriate *PHB* must be determined from the label. A field is included in the *MPLS* shim header to allow *MPLS* to support eight different *DiffServ* classes.

In an *IP/MPLS* scenario, packets are classified and routed at the ingress *LSR*s. Then, an *MPLS* header is inserted. Depending on the technology, frame-based or cell-based, 32-bit labels are embedded in this header between the Layer 3 header and the Layer 2 header or in the *Virtual Path Identifier* (*VPI*) and *Virtual Channel Identifier* (*VCI*) fields, respectively. When an *LSR* receives a packet it uses the label as the index to look up the forwarding table and the packet is forwarded according to the forwarding table entry. Then the incoming label is replaced by the outgoing label and the packet is forwarded to the next *LSR*. It is worth to notice that the label allocation and distribution is made at the path set-up time.

MPLS uses a label distribution mechanism, the *Label Distribution Protocol* (*LDP*) that handles *LSP*s set-up and any other negotiation between *LSR*s. Currently, a number of different label distribution protocols are being proposed. Existing protocols have been extended so that label distribution can be piggybacked on them (e.g. in MPLS-RSVP [11] the *RSVP* is extended with several additional objects to allow the establishment of explicitly routed *LSP*s using *RSVP* as a signalling protocol). Moreover, new protocols have also been defined for the explicit purpose of distributing labels, such as MPLS-LDP [12] and MPLS-CR-LDP [13].

Although initially the main goal of label-based switching mechanisms were to export the speed of Layer 2 to Layer 3, it is really not the main benefit of such mechanisms because of the low time required by newer Layer 3 switches to perform routing. However, as routing decisions are taken according to labels instead of IP addresses, the main benefits introduced to the IP networks by the *MPLS* are:

- Virtual Private Networks (*VPN*): *MPLS* allows providers to create Layer 3 *VPN*s through their own backbone network for multiple customers, with no encryption required.
- Traffic Engineering: *MPLS* provides traffic engineering capabilities needed for the efficient use of network resources. This feature optimises global bandwidth utilization.
- Quality of Service: *MPLS* allows service providers to provide multiple classes of service with hard QoS guarantees.
- Integration of IP and ATM: *MPLS* allows those carriers networks employing an overlay model in which ATM is used at Layer 2 and IP at Layer 3, to migrate many of the functions of the ATM Control Plane to Layer 3, so reducing scalability issues and simplifying network provisioning.

It is worth to note that *MPLS* includes the main advantages of Layer 3 and layer 2, i.e., performance and scalability respectively. The huge and almost incredible evolution on both, network utilization and network applications lead to a continuous endeavour to keep the network infrastructure up-to-date. *MPLS* allows service providers to differentiate services in a QoS scenario, without requiring continuous modifications in the existing network infrastructure.

MPLS Traffic Engineering uses the *RSVP* to automatic establish and maintain a tunnel, *LSP* tunnel across the backbone. In this scenario the *RSVP* is used only to signal the path set-up. The path of the *LSP* tunnels is selected based on the incoming traffic requirements and the available network resources. The selection process is performed at the source router, by using *Constraint-based Routing* (*CR*). Unlike traditional *IP/MPLS* routing where routes are selected according to network topology, *CR* selects routes considering multiple constraints, such as *LSP* and link attributes. As a consequence, network load is distributed more fairly. From the network client point of view, this traffic traverses the *MPLS* backbone through an end-to-end tunnel which connects the source and the destination nodes. Paths are properly selected whenever network information used by the *CR* accurately represents the real network state in terms of QoS parameters. This constraint-based information is disseminated across the *MPLS* network by extending existing link-state routing protocols such as *OSPF* or *IS-IS* to generate *Traffic Engineering Databases* (*TEDs*). Routing protocols based on link state network information perform better than those based on distance vector network information because unlike the first ones the second ones do not include enough information in their routing tables to compute alternative paths needed by *TE*. Therefore, *OSPF* and *IS-IS* are extended to properly carry constraint-based information.

As a summary, *TE* is essential for service provider and Internet service provider backbones because both backbones must support a high capacity traffic demands and a quick response to network failures. *MPLS* is very useful in the *TE* scenario since it allows service providers to offer traffic engineering networks without substantial network modifications. The main advantages of *MPLS Traffic Engineering* are the following:

- When using *MPLS*, the Layer 3 integrates traffic engineering capabilities, so optimising the path selection process, according to the available bandwidth capacity and the network topology.
- The routing decisions are taken based on the available network resources and the resources required by the incoming *LSP* demand.
- *MPLS Traffic Engineering* uses *Constraint-based Routing*. *CR* selects that path among the shortest ones that meets the QoS requirements.
- *MPLS Traffic Engineering* includes a mechanism to switch traffic flow over when a network element fails.

- *MPLS Traffic Engineering* enables unequal-cost load sharing.
- It computes explicit routes accounting for link bandwidth and for the size of the traffic flow.
- Explicit routing is optimised (in comparison with IP source routing) since a single label instead of the complete list of intermediate IP addresses is sent in the set-up message.
- Explicit paths are dynamically set up by an automated signalling process.

Chapter 3

QoS Routing

New multimedia applications are appearing over the Internet, demanding particular QoS requirements, such as bandwidth, delay, jitter, packet loss and reliability, which must be taken into account when selecting paths. As mentioned in the last Chapter different QoS architectures, such as *DiffServ* and *IntServ*, are proposed to meet these QoS requirements. A key aspect in these QoS architectures is the routing process, i.e., how routes are computed, selected and established.

It is worth to notice that there are two different entities when talking about routing, the routing protocol and the routing algorithm. The routing protocol attends to the matter of collecting the network state dynamics and to flood this information throughout the network. Based on this network state information, the routing algorithm selects the optimal path.

Traditional IP routing algorithms, which are based on the best-effort transmission model, select routes according to the shortest path routing. These algorithms select the path that optimizes the sum of a single value, such as hopcount or delay along the selected path. This routing model is not suitable in a QoS environment. When a certain guarantee is required for sending a particular traffic flow, routing algorithms must add some QoS attributes to the path selection process. Unlike shortest path based routing algorithms, QoS routing algorithms select that route which more precisely meets multiple QoS requirements. Basically, the main goal of QoS routing is to find a route for a particular traffic flow with certain QoS requirements conforming to the QoS cost parameters, which specify the available resources in the network (i.e., conforming to the network resources that can be used to support the incoming traffic request). These QoS requirements may be bottleneck requirements, such as bandwidth, or additive requirements, such as end-to-end delay, in which case the QoS routing process looks for that path that guarantees a minimum available bandwidth or an end-to-end delay bound respectively. In order to perform this route evaluation and selection, the link state databases are extended to include information on available resources, and are often referred to as *Traffic Engineering Databases* (*TED*). In addition to QoS guarantees, there are some other widely sought solutions to common networking challenges that can be perfectly synthesized with the above mentioned goals: optimisation of network utilization, load distribution, the number of paths successfully routed, etc. There is currently a concerted effort in the networking community to achieve all of these objectives.

In the QoS routing context, there are two main issues to be addressed. Firstly, routing decisions are taken based on the network state information. Each node collects this information by implementing a flooding mechanism, which disseminates this information throughout the whole network. Despite its simplicity and reliability, flooding involves unnecessary communications and causes inefficient use of resources, particularly in highly dynamic network where frequent distribution of multiple QoS parameters is expected. Secondly, once nodes have updated network state information a routing algorithm is applied to select the optimal route. It is necessary to point out that a node contains updated network state information when this information perfectly represents the real network state at the moment when path selection process is performed.

Being aware that QoS routing is essential in a network architecture that needs to satisfy traffic and service requirements, it must be assumed that the process required to manage the path selection turns out a difficult problem to be solved.

Computing paths based on multiple QoS constraints is called *multi-constrained path selection problem* (*MCP*). In general *MCP*, is known as an NP-complete problem, therefore it is intractable for large networks. There are several algorithms proposed in the literature to address this problem. Most important MCP algorithms are *Jaffe's Approximate Algorithm* [14], *Iwata's Algorithm* [15], *Self-Adaptive Multiple Constraints Routing Algorithm (SAMCRA)* [16], *Chen's Approximate Algorithm* [17], *Randomized Algorithm* [18], *H_MCOP* [19], *Limited Path Heuristic* [20], and *A*Prune* [21]. A performance evaluation of these algorithms can be found in [22] where fundamental concepts involved in QoS routing are deduced based on the simulation results.

In addition to these algorithms there are other works in the literature aiming at addressing special important sub-problems in QoS routing, such as QoS routing in the context of bandwidth and delay, which is not NP-complete. This covers works such as the *Widest-Shortest Path* (*WSP*) [23], the *Shortest-Widest Path* (*SWP*) [24] and the *Maximum Delay-Weighted Capacity Routing Algorithm* (*MDWCRA*) [25]. In the *WSP* links with residual bandwidth lower than the requested bandwidth are pruned, therefore generating a reduced graph containing only those links supporting the incoming traffic demands. This reduced graph is used to select the shortest path. When there are multiple shortest paths available, the path that maximizes the minimum residual bandwidth on the links in the path is selected. The *SWP* algorithm performs similarly to the *WSP*. It selects the shortest path among the widest ones. In [25] authors concentrate on the specific problem of designing bandwidth-delay constrained algorithms taking into account knowledge of the source-destination node pairs. When QoS routing considers delay and cost is known as *Restricted Shortest Path* problem (*RSP*), which is NP-complete. Focusing on this issue are works presented in [26], [27], [28] and [29]. In [30] a selection of different path selection algorithms based on combining bandwidth, delay and cost (in terms of number of hops) can be found. Reference [31] considers pre-computation of paths with minimum hopcount and bandwidth guarantees. The effects of reserving in advance of the path selection process are addressed in [32].

There are other significant contributions, focusing on other aspects to select routes. The *Maximally Disjoint Shortest and Widest Paths* [33], selects more than one feasible path. In [34] there are proposed bandwidth guaranteed dynamic routing algorithms. The *Minimum Interference Routing Algorithm* (*MIRA*) [35] selects optimal paths not only based on bandwidth guarantees but also considers ingress/egress characteristics in order to avoid negative interferences between routes sharing common links. Inspired by the *MIRA*, the *Profile-Based Routing* (*PBR*) [36] selects routes by using a "traffic profile" of the network as a rough predictor of the future traffic distribution.

QoS routing is generally blamed for increasing the complexity of the path selection process. Several factors can drive one to this conclusion. First, several new parameters used to provide the network with QoS capabilities are added to the routing process. These new parameters generate two implications, namely the database structure must be properly augmented to allocate them and the computational time will also be augmented since more parameters must be considered. Second, the number of update messages needed to maintain perfectly up-to-date link state databases could reach levels that negatively impact correct network performance. The first issue is not of excessive importance, since successive technological advances reduce the cost impact of QoS routing. The second issue has a fundamental influence on global network performance, since the number of update messages cannot be reduced without generating collateral negative effects that can hugely deteriorate network performance. As in IP routing, in QoS routing path selection is performed in accordance with the link state information (assuming Dijkstra's algorithm) contained in the databases of each node in the network. In order to optimise the path selection process the link state information must be correctly updated so that it accurately represents the current network state. In order to perform this, updating mechanisms are incorporated in the routing protocol. These mechanisms allow any node to send update messages that will be flooded throughout the network, advertising the latest changes in its directly connected links. In this way, global network state information is maintained in every node. The important influence that these updating procedures have on global network performance and their impact on the number of paths successfully routed will be later shown.

In fact, most current QoS routing algorithms assume as a condition that the network state databases from which the routing tables are built represent a current picture of the network state. However, it is possible that due to some circumstances this information does not perfectly represent the real network state. In this case a certain degree of routing uncertainty or routing inaccuracy exists in the network state information. Moreover, if update messages are not enough fast flood routing instabilities may occur which produce undesirable routing oscillations. An initial approach to address this problem is based on advertising link weights that are properly quantified instead of instantaneous values [37] and [38]. However, in certain scenarios and under high loads or bursty traffic these mechanisms do not enough reduce routing oscillations. Algorithms for load balancing avoid routing oscillations by providing multiple paths from a source and a destination. Some works related to this can be found in [39], [40], [41], [42], [43], [44] and [45].

As a summary, this Part serves to introduce main aspects related to QoS, such as QoS network architectures and QoS routing algorithms. This part is very useful to describe the network scenario where this Thesis is placed.