UNIVERSITAT POLITÈCNICA DE CATALUNYA
Departament d'Arquitectura de Computadors

Offset Time-Emulated Architecture for Optical Burst Switching - Modelling and Performance Evaluation

Ph.D. thesis by
Mirosław Klinkowski

Advisor:
Dr. Davide Careglio
Co-advisors:
Prof. Josep Solé i Pareta, and
Prof. Marian Marciniak

PhD Program: Arquitectura i Tecnologia de Computadors

November 2007

ACTA DE QUALIFICACIÓ DE LA TESI DOCTORAL

Reunit el tribunal integrat pels sota signants per jutjar la tesi doctoral:

Títol de la tesi: **Offset Time-Emulated Architecture for Optical Burst Switching - Modelling and Performance Evaluation**
Autor de la tesi: **Mirosław Klinkowski**

Acorda atorgar la qualificació de:

- ☐ No apte
- ☐ Aprovat
- ☐ Notable
- ☐ Excel·lent
- ☐ Excel·lent Cum Laude

Barcelona, …………… de/d'…..............…………….. de ....….

El President	El Secretari	
.... (nom i cognoms)	.... (nom i cognoms)	
El vocal	El vocal	El vocal
.... (nom i cognoms)	.... (nom i cognoms)	.... (nom i cognoms)

To Ola and my parents

Acknowledgments to

Marian, Josep, Davide, and Michał
for their measureless support, inspiration
and for guiding me through the research of this thesis

Wala, Barbara, Andrzej and Michał
for encouraging me at the very beginning of this thesis

And to my family
for their love, support, and for being always with me

Contents

List of Figures v

List of Tables ix

Summary xi

Resum xv

Structure of the thesis xix

1 Introduction 1
- 1.1 Optical Transport Networks 2
- 1.2 Overview of the Thesis 4

PART I: Background 7

2 Optical switching architectures 9
- 2.1 Principle of operation 10
  - 2.1.1 Optical circuit switching (OCS) 10
  - 2.1.2 Optical packet switching (OPS) 10
  - 2.1.3 Optical burst switching (OBS) 11
  - 2.1.4 OPS vs. OBS 11
- 2.2 Main characteristics and comparison between OCS, OBS, and OPS . 12
- 2.3 Summary 18

3 Optical burst switching 21
- 3.1 Overview of general OBS concepts 22
  - 3.1.1 Signalling 22
  - 3.1.2 Architectures and functions of OBS nodes 24
  - 3.1.3 Offset time provisioning 27
  - 3.1.4 Resources reservation 30
  - 3.1.5 Contention resolution 31
  - 3.1.6 Burst scheduling 33
  - 3.1.7 Quality of service provisioning 34
  - 3.1.8 Network routing 34

PART II: Offset Time-Emulated OBS Architecture 37

4 E-OBS architecture 39
4.1 Principles of E-OBS 39
4.1.1 Node architecture 39
4.1.2 Control operation 41
4.2 Characteristics of E-OBS and C-OBS 42
4.3 Summary 48

5 Modelling of E-OBS control plane 49
5.1 Introduction 49
5.2 Modelling of control plane 50
5.2.1 Control plane impacting factors 50
5.2.2 A queuing model of OBS switch controller 52
5.3 E-OBS controller with a single processor 52
5.3.1 Queuing models 52
5.3.2 Results 54
5.4 Summary 57

PART III: Quality of service provisioning 59

6 QoS provisioning in OBS networks 61
6.1 Basic concepts of QoS in OBS networks 62
6.1.1 QoS metrics 62
6.1.2 Absolute vs. relative QoS guarantees 62
6.1.3 QoS in connection-oriented and connection-less OBS 63
6.2 Categories of QoS mechanisms in OBS networks with one-way signalling 63
6.2.1 Control plane-related mechanisms 63
6.2.2 Edge-based mechanisms 64
6.2.3 Core-based mechanisms 65

7 Performance of QoS mechanisms in E-OBS 69
7.1 Overview 69
7.1.1 QoS scenario details 69
7.1.2 Simulation scenario 71
7.2 Threshold selection in BD-W mechanism 71
7.3 Performance results 73
7.3.1 Burst loss probability and throughput 73
7.3.2 Burst preemption vs. offset time differentiation 74
7.4 Summary 76

8 Effective burst preemption in E-OBS 77
8.1 Preemption rate in a buffer-less OBS node 77
8.2 Preemption Window (PW) mechanism 79
8.2.1 Principles 80

8.2.2 The length of preemptive window 81
8.3 A single-wavelength model of PW mechanism 82
8.3.1 Some remarks 85
8.4 Computer simulation of PW mechanism 86
8.4.1 Simulation scenario 86
8.4.2 Numerical results 87
8.4.3 PW and FDL buffering 89
8.5 Summary 90

PART IV: Routing 91

9 Routing in OBS networks 93
9.1 Introduction 93
9.1.1 Routing terminology 93
9.1.2 Reactive and proactive burst loss reduction techniques 95
9.1.3 Hop-by-hop vs. explicit routing 96
9.2 State of the art 97
9.2.1 Alternative routing 97
9.2.2 Multi-path routing 98
9.2.3 Single-path routing 99
9.3 Summary 100

10 Isolated alternative routing strategies for labelled E-OBS networks 101
10.1 Scenario under study 101
10.2 Algorithms 104
10.3 Results 106
10.4 Summary 106

11 Optimization of multi-path routing 111
11.1 Routing scenario 111
11.2 Formulation 112
11.2.1 Loss models of OBS network 112
11.2.2 Optimization problem 115
11.3 Partial derivatives 116
11.3.1 NR-LL model 116
11.3.2 R-LL model 118
11.3.3 Remarks 118
11.4 Implementation issues 119
11.5 Performance 120
11.5.1 Comparison of routing schemes 120
11.6 Summary 123

PART V: Conclusions and future works **125**

12 Conclusions and future works **127**

Acronyms **131**

Bibliography **133**

A Related publications **149**
A.1 Papers 149
A.2 Papers under submission 152
A.3 Contribution to European projects 152
A.4 Other publications 152

List of Figures

1.1 Generic architecture of optical DWDM network. 2
1.2 The trend of migration in optical networking. (courtesy of [Pro05]) . 3

2.1 Optical circuit switching network. 10
2.2 Optical packet switching network. 11
2.3 Optical burst switching network. 12
2.4 Overview over key parameters determining circuit/burst/packet granularity and required switching technology. 13

3.1 Signalling protocols in OBS networks. 23
3.2 OBS ingress edge node. 25
3.3 Burst control packet format. 26
3.4 OBS core switching node. 27
3.5 Offset time provisioning architectures. 29
3.6 Resources reservation schemes. 31
3.7 Contention resolution mechanisms. 32
3.8 Burst scheduling algorithms. 34

4.1 General E-OBS core node architecture. 40
4.2 Fiber delay coils; a) a single component, b) a part of an OBS test-bed. (courtesy of Newport Corp., and OITDA) 40
4.3 Time dependencies in E-OBS. 41
4.4 Unfairness in conventional OBS. 42
4.5 Burst loss probability vs. remaining hops number. 43
4.6 Burst loss probability vs. offered traffic load. 44
4.7 JIT resources reservation in a) C-OBS, and b) E-OBS. 45

5.1 Data and control networks of OBS. 50
5.2 Exemplary controller architectures. 51
5.3 General OBS control-plane queuing model. 52
5.4 Queuing models: a) M/M/1 with reneging, b) M/D/1/K. 53
5.5 Intensity of control packet arrival. 54
5.6 Loss probability of control packets. 55
5.7 Delay budget vs. normalized mean burst duration. 56
5.8 Delay budget vs. average burst length. 56

6.1 Categories of QoS mechanisms in OBS networks. 64

6.2 Selected QoS mechanisms in OBS networks. 65

7.1 Evaluated QoS network scenario. 70
7.2 Performance of BD-W mechanism ($c = 8$), a) HP class BLP, b) LP class BLP, c) throughput, d) threshold value guaranteeing $BLP_{HP} \leq 10^{-4}$. 72
7.3 Performance of QoS mechanism vs. link dimensioning ($\rho = 0.8, \alpha_{HP} =$ 30%), a) HP class BLP, b) LP class BLP, c) overall BLP, d) effective data throughput. 74
7.4 Burst loss probabilities vs. HP class relative load in OTD and BP mechanisms ($\rho = 0.8, c = 8$), a) HP class, b) LP class. 75
7.5 Effective throughout vs. HP class relative load in OTD and BP mechanisms, with overall traffic load: a) 0.5, b) 0.8. 75

8.1 Percentage of additional signalling necessary to release preempted burst at each node, with HP class load: a) 30%, b) 50%. 79
8.2 Principles of the preemption window mechanism. 80
8.3 The length of preemptive window in PW mechanism. 81
8.4 Markov chain representing a single-wavelength model of PW mechanism. 82
8.5 (a) Successful preemption and (b) preemption fail cases; the processing times are neglected for simplicity, T is the duration of the Preemption Window, l_{LP} and l_{HP} are the durations of the LP and HP bursts respectively, t is the arrival time of the HP control packet. 83
8.6 Simulation vs. modeling results ($\rho = 0.8, \alpha = 0.3, \mu = 2$). 86
8.7 Burst blocking probability as a function of T comparing Gaussian and Exponential traffic models ($\alpha = 30\%$, $\rho = 0.8$, $W = 16$). 87
8.8 Burst blocking probability as a function of T and of W ($\alpha = 30\%$, $\rho = 0.8$, Gaussian traffic model). 88
8.9 Burst blocking probability as a function of ρ comparing Gaussian and Exponential traffic models and different α ($T = 10\mu s$ and $W = 32$). . 88
8.10 Burst blocking probability as a function of T (normalized to $1/\mu$) for different W and FDL buffer size ($\alpha = 25\%$). 89

9.1 Routing algorithms. 94

10.1 Network topologies; a) SIMPLE, b) NSFNET, and c) EON. 103
10.2 Isolated alternative routing algorithms: a) PER, and b) BPR. 105
10.3 Burst loss probability in PER, a) SIMPLE (32λ), b) NSFNET (32λ), and c) EON (64λ). 107
10.4 Burst loss probability in BPR, a) SIMPLE (32λ), b) NSFNET (32λ), and c) EON (64λ). 108
10.5 Amount of bursts experiencing given number of hops in BPR, a) SIMPLE (32λ, $\rho = 0.8$), b) NSFNET (32λ, $\rho = 0.8$), and c) NSFNET (32λ, $\rho = 0.5$). 109

11.1 Example of OBS network with multi-path source-based routing; x_1 and x_2 are the splitting factors and $x_1 + x_2 = 1$. 112

11.2 Link load models: a) non-reduced OBS, b) reduced OBS, and c) reduced CS. 114
11.3 Accuracy of NR-LL model, in SIMPLE, NSFNET, and EON topologies, with 8, 32, and 64 wavelengths per link, respectively. 115
11.4 Burst loss probability in OR, a) SIMPLE (32λ), b) NSFNET (32λ), and c) EON (64λ). 121
11.5 Comparison of optimized multipath source routing with isolated alternative routing strategies, a) SIMPLE (32λ), b) NSFNET (32λ), and c) EON (64λ). 122

List of Tables

2.1 Advantages, drawbacks and foreseen for future implementations . . . 19

4.1 Advantages and drawbacks of offset-time provisioning architectures . 48

5.1 Performance of queuing models. 53

6.1 Characteristics of QoS mechanisms in OBS networks with one-way signalling 67

9.1 Classification of literature on alternative routing in OBS networks. . . 97
9.2 Classification of literature on adaptive multi-path routing in OBS networks. 98
9.3 Classification of literature on adaptive single-path routing in OBS networks. 99

10.1 Network topologies 102

Summary

The evolution of the transport networking is driven by continuously increasing traffic demand due to the introduction of broadband Internet access and new end-user business applications as well as the continuing paradigm shift from voice to data services. These trends have emerged at the same time as the advance in optical technologies which has enabled the development of high-capacity transmission systems. The role of optics in communication networks is often limited to the realization of transmission functions, however, the next-generation networks will perform either some or all the switching and control functions in the optical domain. As a result the optical transport networks will provide a global transport infrastructure for legacy and new IP services (IP over DWDM).Optical burst switching (OBS) technology is a promising solution for reducing the gap between transmission and switching speeds in future networks. It offers both flexibility and efficiency through the exploitation of statistical multiplexing in optical domain.

This thesis presents the analysis, modelling, and evaluation of the optical burst switching network with an emulated offset-time provisioning (E-OBS). E-OBS defines an OBS network architecture to transport and switch optical data bursts in a core network. On the contrary to a conventional offset-time provisioning OBS (C-OBS) architecture, where a transmission offset time is introduced in the edge node, in an E-OBS network the offset time is provided in the core node by means of an additional fibre delay element. The architecture is motivated by several drawbacks inherent to C-OBS architectures.

Due to the limitations in optical processing and queuing, OBS networks need a special treatment so that they could solve problems typical of data-centric networks. Contention resolution in optical domain together with quality of service (QoS) provisioning for quality demanding services are, among other things, the main designing issues when developing OBS networks. Another important aspect is routing problem, which concerns effective balancing of traffic load so that to reduce burst congestion at overloaded links. Accounting for these requirements, the design objectives for the E-OBS architecture are (i) feasibility of offset-time provisioning, (ii) an overall high quality of service, and (iii) reduction of network congestion. These objectives are achieved by combining selected concepts and strategies, together with appropriate system design as well as network traffic engineering.

Part I of this thesis provides the background information for the design of E-OBS. First, in order to motivate the application of optical burst switching concept, Chapter 2 reviews the general characteristics, requirements, and trends of switching architectures considered for next generation optical networks. In particular, it introduces *optical circuit switching*, *optical packet switching*, and *optical burst switching* architectures. Then Chapter 3 discusses the features of optical burst switching as well as the-state-of-research solutions that are considered for OBS networks.

Part II provides the discussion on functionality and feasibility of an E-OBS architecture. Chapter 4 introduces principles of E-OBS operation as well as it demonstrates that C-OBS possesses many drawbacks that can be easily avoided in E-OBS. Some of the discussed issues are the problem of unfairness in resources reservation, difficulty with alternative routing, complexity of resources reservation algorithms, efficiency of burst scheduling, and complexity in QoS provisioning. The feasibility of offset time provisioning with the assistance of a fibre delay element is investigated in Chapter 5. First, several factors that have impact on the control plane operation are discussed. In a poorly-engineered network the congestion in control plane may delay excessively the processing of control packets in an electronic core-node controller and as a result lead to the loss of data bursts. In order to approach this outcome effectively two queueing models, which represent the operation of an exemplary OBS node controller, are introduced. The analyzed models allow to expose some relations which exist between key OBS system parameters. Using obtained results the feasibility of E-OBS operation with commercially available fibre delay elements is confirmed.

Part III addressed the problem of QoS provisioning. Chapter 6 discusses some basic concepts of QoS as well as it presents the state of the art mechanisms dealing with QoS in OBS. The discussion is supported by a qualitative comparison of the mechanisms. Chapter 7 complements the study with a quantitative comparison of the performance of selected, most addressed in the literature, QoS mechanisms in an E-OBS scenario. As an outcome a burst preemption mechanism, which is characterized by the highest overall performance, is qualified for operating in E-OBS. Since the preemptive mechanism may produce the overbooking of resources in any OBS network, Chapter 8 addresses this issue. Particularly, it proposes a *preemption window* mechanism to solve the problem. Then it provides an analytical model which legitimates correctness of the solution.

Part IV concerns the routing problem in OBS networks. Chapter 9 introduces general routing terminology, and based on these terms it classifies different routing strategies that have been proposed in the OBS literature. As a continuation, Chapter 10 studies several adaptive routing algorithms, which are based on isolated alternative routing, designed for *labelled* E-OBS networks. The routing objective is to help the contention resolution algorithms in the reduction of burst losses by balancing the link loads and avoiding excessive congestion. Since the proposed algorithms might have some difficulty with the congestion reduction, Chapter 11 provides another solution, which is based on a non-linear optimization framework. In the scope of this proposal, two optimization models for multi-path source routing are formulated and solved. Then some related implementation issues are discussed.

Concluding, E-OBS is shown to be a feasible OBS network architecture of profitable functionality, to support efficiently the QoS provisioning, and to be able to operate with different routing strategies and effectively reduce the network congestion.

It should be emphasized that the work is a part of the research activities performed by the Advanced Broadband Communication Center (CCABA) at The Department of Computer Architecture of the Technical University of Catalonia. In particular, the work is carried on within five relevant research projects: the IST-2002-506760 NOBEL (Next Generation Optical Network for Broadband in Europe) project, the IST-2002-001933 E-PHOTON/ONE (Optical Networks: Towards Bandwidth Manageability and Cost Efficiency) project, the COST 291 (Towards Digital Optical Networks) action and the COST 293 (Graphs and Algorithms in Communication Networks) action, all off them founded by the European Commission, and the CATARO project founded by of the Spanish Ministry of Education and Science.

Resum

L'evolució de les xarxes publiques de transport de dades destaca per el continu augment de la demanda de tràfic a la que estan sotmeses. La causa és la imparable popularització d'Internet i del seu ús per a tot tipus d'aplicacions, i en concret de les de banda ampla i interactives (des de veu sobre IP fins a nous serveis de dades com Storage Area Networks i Grid Computing), tant per part dels usuaris residencials com des de negocis.

Aquesta tendència ha esdevingut paral·lela a l'avenç de les tecnologies òptiques que han permès el desenvolupament dels sistemes de transmissió d'alta capacitat que alhora la fan possible. De totes maneres, el paper de la tecnologia òptica en les xarxes de telecomunicacions es limita encara molt sovint només a la transmissió/recepció, tanmateix, en la pròxima generació de xarxes aquesta tecnologia prendrà un paper cada vegada més rellevant també en la commutació i les funcions de control i gestió de la xarxa. Com a resultat d'això les xarxes de transport òptiques proporcionaran una infrastructura de transport global per a tota mena de serveis basats en IP (IP sobre DWDM). En aquest sentit, les xarxes de commutació de ràfegues òptiques (OBS: Optical Bursts Switching) són una solució extraordinàriament prometedora tant per la flexibilitat que ofereixen com per el seu alt rendiment fruit de l'explotació de la multiplexació estadística en el domini òptic.

Aquesta tesi presenta l'anàlisi, modelització i avaluació de les xarxes de commutació de ràfegues òptiques basades en l'emulació del temps de compensació (emulated offset time: E-OBS). El concepte d'E-OBS defineix una arquitectura de xarxa OBS per al transportar i commutar ràfegues òptiques en una xarxa troncal en la que, al contrari de l'arquitectura convencional (C-OBS) en la que el temps de compensació s'introdueix des dels nodes d'accés, el temps de compensació s'introdueix en cadascun dels nodes de la xarxa per mitjà d'un retardador de fibra addicional. L'arquitectura E-OBS permet superar algunes de les desavantatges inherents a arquitectures C-OBS, però la seva gran virtut és la compatibilitat amb les xarxes de commutació de circuits òptics (OCS: Optical Circuit Switching) actuals i les futures xarxes de commutació de paquets òptics (OPS: Optical Packet Switching), de manera que les xarxes OBS basades en una arquitectura E-OBS) poden facilitat enormement la transició de unes a les altres.

A ala vista dels principals requeriments de disseny de les xarxes OBS, que són la resolució de contencions en el domini òptic, la provisió de qualitat de servei (QoS) i l'òptim encaminament de les ràfegues per tal de minimitzar la congestió de la xarxa, . en aquesta tesi es proposa un disseny de l'arquitectura E-OBS basada en (i) un mètode viable per a la provisió del temps de compensació, (ii) una qualitat alta global de servei, i (iii) un mecanisme d'encaminament que minimitzi congestió de xarxa.

La primera part d'aquesta tesi proporciona la informació documental necessària per al disseny d'E-OBS. Primer es revisen les característiques generals, requisits, i tendències de les principals arquitectures de commutació òptica. En particular, ens centrem en la commutació de circuits òptica, la commutació de paquets òptica i la commutació de ràfegues òptiques; per a, mes endavant centrar-nos en l'estat de l'art de les xarxes de commutació de ràfegues.

La segona part se centra en l'estudi de la funcionalitat i viabilitat de l'arquitectura E-OBS. S'introdueixen els principis d'operació d'E-OBS i s'identifiquen els principals esculls que presenten les arquitectures C-OBS i que deixen de ser-ho en una arquitectura E-OBS. Alguns d'aquests esculls són la dificultat d'utilitzar un algorisme d'encaminament amb rutes alternatives, la complexitat dels algoritmes de reserva de recursos i la seva falta d'equitat, la complexitat en la provisió de la QoS, etc. En aquesta segona part es constata que l'arquitectura E-OBS redueix la complexitat dels de reserva de recursos i es verifica la viabilitat d'operació i de funcionament de la provisió del tremps de compensació en aquesta arquitectura a partir de figures de comportament obtingudes amb retardadors de fibra comercialment disponibles.

La tercera part encara el problema de la provisió de la QoS. Primer s'hi revisen els conceptes bàsics de QoS així com els mecanismes de tractament de la QoS per a xarxes OBS fent-ne una comparació qualitativa i de rendiment de tots ells. Com a resultat s'obté que el mecanisme que presenta un millor comportament és el d'avortament de la transmissió de les ràfegues de més baixa prioritat quan aquestes col·lisionen amb una de prioritat més alta (es l'anomenat Burst Preemption mechanism), el qual en alguns casos presenta un problema de senyalització innecessària. Aquesta tercera part es conclou amb la proposta d'un mecanisme de finestra a afegir al esquema de Burst Preemption que només funciona sobre una arquitectura E-OBS i que soluciona aquest problema.

En la quarta part s'afronta el problema de l'encaminament en xarxes OBS. Es comença per revisar la terminologia general d'encaminament, i en base a aquesta terminologia es classifiquen les principals estratègies d'encaminament que s'han proposat en la literatura per a ser utilitzades en xarxes OBS. A continuació, s'estudia el comportament dels algoritmes d'encaminament adaptatius, els aïllats amb rutes alternatives i els multicamí distribuïts, sobre xarxes E-OBS. A la vista dels resultats no massa satisfactoris que s'obtenen, es planteja una solució alternativa que es basa en model d'optimització no lineal. Es formulen i resolen dos models d'optimització per als algoritmes encaminament de font multicamí que redueixen notablement la congestió en les xarxes OBS.

Finalment, aquesta tesi conclou que l'arquitectura E-OBS és factible, que és més eficient que la C-OBS, que proveeix eficaçment QoS, i que és capaç d'operar amb diverses estratègies d'encaminament i de reduir eficaçment la congestió de xarxa.

Structure of the thesis

Environment

Optical Burst Switching (OBS): is a photonic network technology which overcomes the wavelength switching inefficiency by a proper exploitation of the statistical multiplexing in the optical layer. On the contrary to optical packet switching, OBS handles large data bursts aggregated from the client packets in order to reduce the processing and switching requirements. Moreover, a burst control packet is transmitted in a dedicated control channel and delivered with some offset time prior to the data burst.

In this dissertation

The emulation of offset-times by means of a fiber span applied in the core node is a possible way of provisioning the offset time in OBS networks. On the contrary to a conventional OBS, where the offset time is introduced in the edge node by delaying the transmission of a burst with respect to its control packet, the offset time-Emulated OBS (E-OBS) has not been studied extensively. In this dissertation we show the advantages of E-OBS architecture and address several issues related to network modelling, QoS provisioning and routing over this architecture.

Offset-time provisioning

Related work:

- C-OBS (Conventional OBS) and E-OBS architectures

Contributions:

- Functional analysis of offset time provisioning methods

Modelling of E-OBS

Related work:

- State of the art on C-OBS node architecture/network dimensioning/desing issues

Contributions:

- Modelling of the OBS control plane operation
- Estimation of feasible E-OBS system parameters

QoS provisioning

Related work:

- Existing QoS mechanisms for OBS networks
- Burst preemption mechanisms with signalling overhead, in detail

Contributions:

- Qualitative and quantitative comparison of the most referenced QoS mechanism
- Estimation of the signalling overhead in a burst preemption mechanism
- Proposal of a new preemption-based mechanism without signalling overhead

Routing

Related work:

- Alternative, multi-path, single-path routing mechanisms
- LP formulations for routing problem

Contributions:

- Performance evaluation of isolated alternative routing
- Non-linear optimization of multi-path source routing

Chapter 1

Introduction

The telecommunication networks are experiencing a continuous increase in demand for transmission capacity. This trend is strictly related with the exponential growth of the Internet. The evolution of Internet technologies is accompanied by development of miscellaneous user and network applications; peer-to-peer (P2P) data/multimedia file exchange, video broadcasting, grid services are among the most bandwidth-demanding applications. Simultaneously, we can observe a big progress in the deployment of broadband access network technologies (e.g., ADSL, WLAN, or FTTH) which place immense traffic on the metro and core transport networks [KMM$^+$05]. As a consequence, the next generation networks should offer high transmission and switching capacities in order to cope with such increasing traffic demands.

Another consequence of the expansion of the Internet is the continuing paradigm shift from voice to data services. This trend is followed by the migration of telecommunication industry from the voice-optimized to the IP-centric networks. In traditional voice-centric networks (such as SONET/SDH [G.700]) the connection-oriented circuit-switched operation is both powerful and efficient. In particular, the traffic characteristic of a superposition of several different constant-bit-rate connections is the total sum and is still constant bit rate. In sharp contrast to, the traffic characteristics in data-centric networks are typically characterized as being very bursty. Also, the packets, which are the basic data transport units, are often variable in the length. Additionally, while a typical voice call is quite predictable in its duration (i.e. the length of a conversation) the duration of the data transmission session in the Internet can vary by orders of magnitude and is often characterized by heavy-tailed distributions. All this results in so-called self-similar nature of Internet traffic [LTWW94] which has a direct impact on network dimensioning; in particular, buffer sizing is crucial.

Electronic IP router architectures are not scalable enough and will suffer from technological limitations when trying to reach the multi-terabit throughput range [PPP06]. Indeed when the bit rates increase the density of integration of electronic circuits increases as well. This may produce unwanted (parasitic) capacitances and impedances at small dimensions. Another problem is power consumption and heat dissipation at high integration level. Finally, electronic technology has its speed limitations.

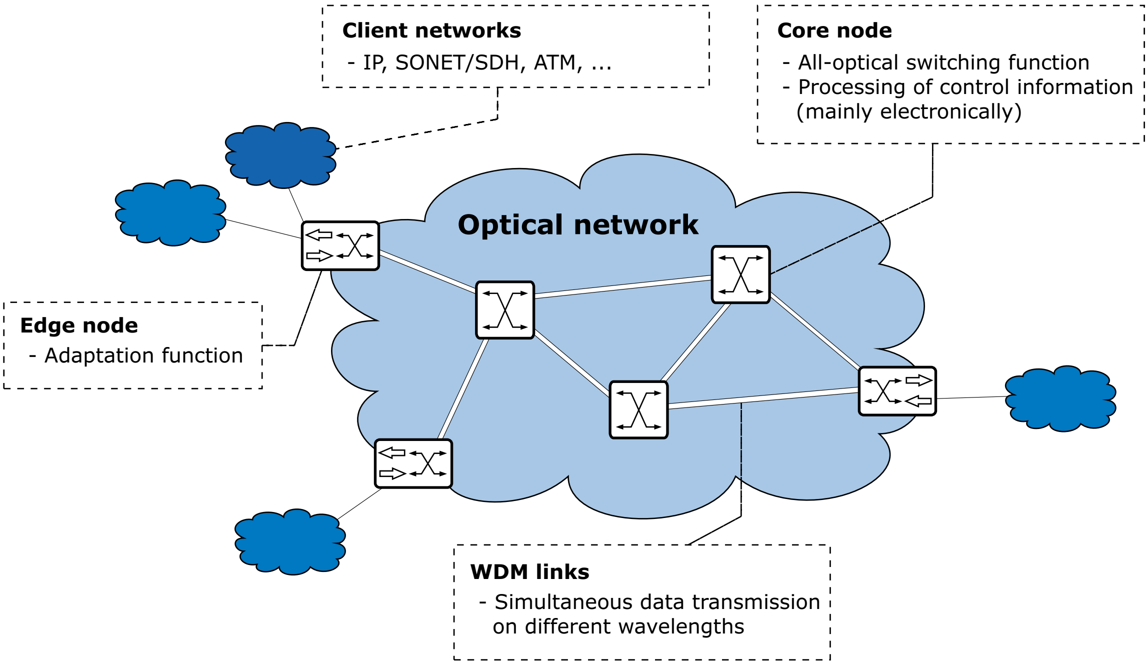


Figure 1.1: Generic architecture of optical DWDM network.

The observed trends have emerged at the same time with advances in optical transmission technologies which led to wavelength division multiplexing (WDM) systems (e.g., see [Cav00][Kar02]). The dense WDM (DWDM), which is an extension of WDM, is able to accommodate up to hundreds of wavelengths; hence providing huge transmission capacities. Accelerated development of optical networking has been possible due to feasibility of integrated optics for both passive and active optical components [Chi01].

1.1 Optical Transport Networks

A generic architecture of optical transport network (see Figure 1.1) consists of source and destination edge nodes and intermediate core nodes that are connected by WDM links. Client networks (IP, ATM, SONET/SDH, etc.) are connected to the edge nodes where there is some adaptation function placed, which is responsible for conversion of data signal from its input form to an optical format used in the optical network. This function can perform for instance a simple conversion of wavelength, if the client network is an optical network, or more sophisticated data aggregation/assembly operation. The data in its optical form is transmitted through WDM links towards core nodes; the WDM technology allows to transmit several data signals on different wavelengths at the same time. The core nodes are responsible for processing of control information and all-optical switching of data signals. In most cases the processing of control information is performed electronically. When the data reaches the destination edge node it is converted back to its client signal format.

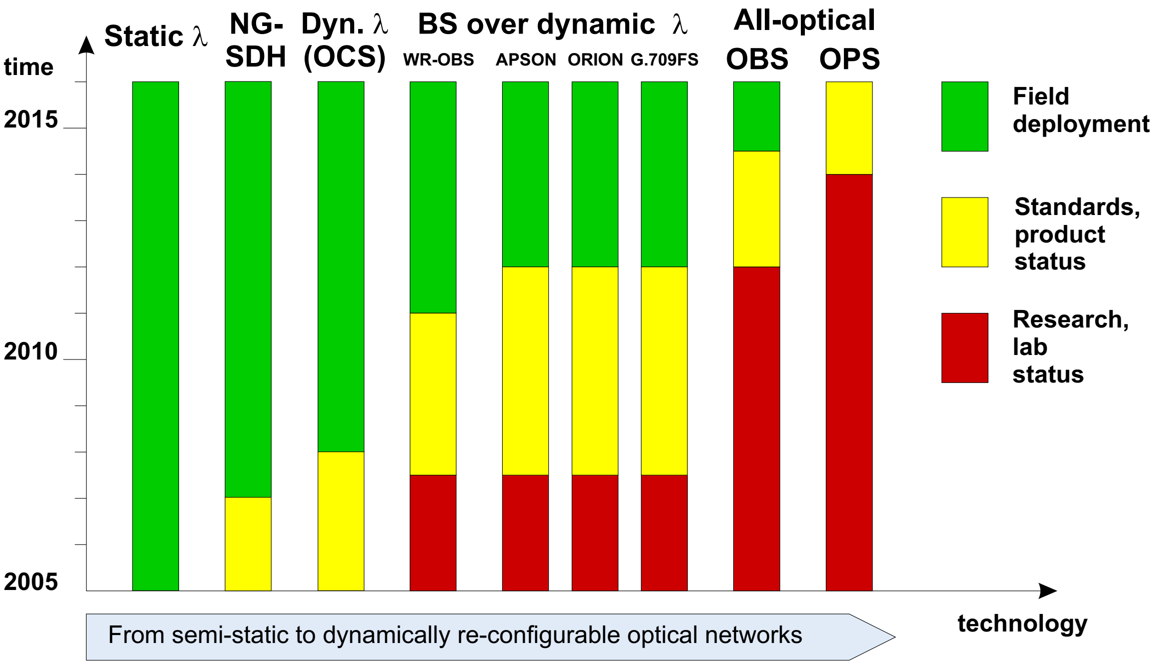


Figure 1.2: The trend of migration in optical networking. (courtesy of [Pro05])

Looking for a transport architecture of future optical Internet a few network architectures have been proposed (e.g., see [LES00][PPP06]). These optical network architectures differ with respect to the degree of optical transparency and the flexibility of optical interconnection [GDW03]. Taking into account the current status of optical technologies the short-term solutions will take advantage of a less flexible circuit-switching model in optical circuit-switched (OCS) networks [CGK92][VKM$^+$01]. Nonetheless, the fact that the Internet is a connection-less packet-based network is the main driver to develop, in a longer perspective, data-centric optical transport networks. In this context two other switching architectures are considered by the research community, namely, optical packet switching (OPS) and optical burst switching (OBS) [XPR01]. In the perspective of network optimization the implementation of packet/burst switching techniques directly in the transport network will bring more statistical sharing of physical resources and will reduce the connection costs.

The expected migration of switching functions from electronic to optics will be gradual and will take place in several phases (see Figure 1.2) [KM01]. Nowadays, transport networks (SDH/ATM/IP) are based mainly on static point-to-point and ring connections. While the role of optics in these networks is limited mainly to the realization of transmission functions the next-generation networks are expected to perform dynamic optical switching as well. The scenario of such a migration anticipates different mid-term optical network solutions like for instance next-generation SDH (NG-SDH), wavelength-routed optical burst switching (WR-OBS), automatic switched optical networks (ASON), the Ontario research and innovation optical network concept (ORION), or G.709 approach [Pro05]. Although, the main functions related with node and network control will remain in the electronic domain, still, it is very likely that some simple control operation will be also performed by means of optical processing [DHL$^+$03].

Considering these trends the optical burst switching technology can be seen as a promising solution. In particular, it reduces the gap between transmission and switching speeds in future networks as well as it offers both flexibility and efficiency through the exploitation of statistical multiplexing in optical domain. It should be underlined that great research activity has been undertaken in the field of OBS over the past years. A number of contributions have been focused on the functionality and performance of OBS node and network architectures. On the other hand, the limitations of optical technologies bring some operational difficulties in conventional OBS architectures. The burst contention problem, unfairness in access to transmission resources, complexity of control, the inherent lack of quality of service (QoS) guarantees, poor performance of shortest-path routing, are some of the examples.

Taking into account these trends and requirements, this thesis analyzes, models, and evaluates an offset time-emulated OBS architecture (E-OBS). E-OBS facilitates both node and network operation with the assistance of local offset time provisioning in core switching nodes. In addition, E-OBS supports efficient QoS provisioning as well as it eliminates some constraints in OBS routing management.

The particular contributions of this thesis include:

- functional analysis of offset time provisioning methods,
- modelling of the OBS control plane operation,
- estimation of feasible E-OBS system parameters,
- qualitative and quantitative comparison of the most referenced QoS mechanisms,
- estimation of the signalling overhead in a burst preemption mechanism,
- proposal of a new preemption-based mechanism without signalling overhead,
- performance evaluation of isolated alternative routing,
- non-linear optimization of multi-path source routing.

1.2 Overview of the Thesis

This thesis is structured into the background part on optical switching architectures (Chapter 2) and optical burst switching (Chapter 3) as well as the parts on the analysis and modelling (Chapters 4-5), QoS provisioning (Chapters 6-8), and routing (Chapters 9-11) in E-OBS.

- Chapter 2 introduces OCS, OPS, and OBS architectures. It also reviews their general characteristics, requirements, and trends.
- Chapter 3 overviews the main features of optical burst switching technology.
- Chapter 4 carries out a discussion on C-OBS and E-OBS architectures. It demonstrates that C-OBS possesses many drawbacks that can be easily avoided in E-OBS. Some of the issues discussed in this chapter are: the problem of unfairness in resources reservation, difficulty with alternative routing, complexity of resources reservation algorithms, efficiency of burst scheduling, and complexity in QoS provisioning.
- Chapter 5 studies feasible system parameters for E-OBS operation. The factors that have impact on the control plane operation are discussed. In order to approach the problem of excessive processing delays in an OBS switch controller, two queueing models are studied. The obtained results demonstrates some relations which exist between key OBS system parameters. Also the feasibility of commercially available fibre delay elements for E-OBS operation is verified.
- Chapter 6 discusses some basic concepts of QoS as well as it reviews the state of the art mechanisms dealing with QoS in OBS. The discussion is supported by a qualitative comparison of the mechanisms.
- Chapter 7 complements the study with a quantitative comparison of the burst loss probability performance of selected QoS mechanisms. As an outcome a burst preemption mechanism is qualified to be the most efficient QoS method for E-OBS networks.
- Chapter 8 introduces a preemption window mechanism which prevents from the overbooking of resources in preemption-based OBS networks. An analytical model of the mechanism is derived.
- Chapter 9 introduces general routing terminology. Then it classifies different routing strategies considered for OBS networks.
- Chapter 10 studies two isolated alternative routing algorithms designed for *labelled* E-OBS networks. The objective is to help the node in the contention resolution problem, and thus, to reduce burst losses.
- Chapter 11 provides another routing solution which is based on non-linear routing optimization. In the scope of this proposal, two optimization models for multi-path source routing are formulated and solved. Also some implementation issues are discussed.
- Chapter 12 summarizes this thesis and presents an outlook to future work.

PART I

Background

Chapter 2

Optical switching architectures

Optical networks adopt several switching models that have been deployed successfully in electronic networks.

In traditional voice-communication networks, which apply a circuit-switching model of operation, the communication between end users is achieved with the assistance of dedicated, and established in the connection-setup phase, circuit (or channel) connections. Such dedicated circuits cannot be used by other users during the connection is active even if no communication is taking place at the moment. Thanks to relatively simple maintenance of circuit-connections the adaptation of circuit-switching model to optical circuit switching (OCS) networks was considered from the very beginning [CGK92].

On the contrary to the voice-oriented networks, the data-centric networks apply a packet-switching paradigm. In such networks the end-to-end communication is achieved thanks to the transmission of small packets that carry portioned information. The packets are routed between network nodes over data links that are shared with other traffic. In order to preserve from the collision of packets a packet buffering is applied. As a consequence of the success of electronic data packet networks, an immense effort has took place, beginning from the mid of 1990s, in the research on optical packet switching (OPS) network technology [Chi95][CVG$^+$98].

A burst switching, applied initially in the block transfer of asynchronous transfer mode (ATM) networks [I.300], is another switching paradigm adopted to optical networks. A burst switched network is a packet-like network where each switching node extracts control information from incoming packet header in advance in order to establish and maintain the appropriate switch connection for the duration of incoming burst of data packets. Optical burst switching (OBS) architectures were proposed in the late 1990s [QY99][Tur99] with the objective to overcome both the low flexibility of OCS architectures and technological limitations of OPS architectures.

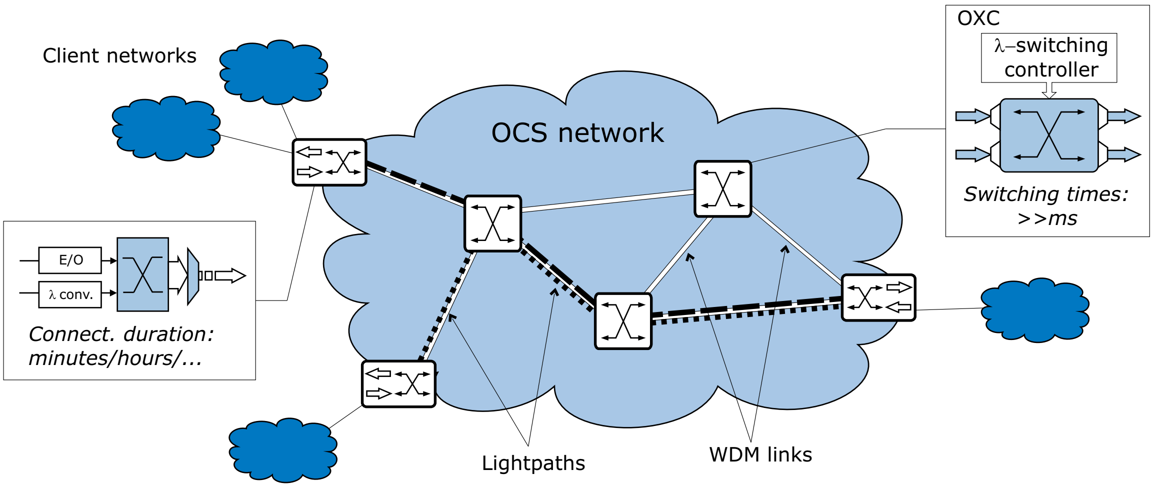


Figure 2.1: Optical circuit switching network.

2.1 Principle of operation

2.1.1 Optical circuit switching (OCS)

The operation of optical circuit switching networks is connection-oriented. In particular, the transmission of data from a source node to a destination node is realized on pre-established paths called the light-paths (see Figure 2.1). OCS switching nodes are referred to as optical cross-connects (OXC). An OXC is responsible for all-optical switching of data carried on an input wavelength (usually denoted as λ) in the input port to an output wavelength in the output port. In OCS networks the smallest switching entity, called later the *granularity*, is a wavelength.

Typical connection durations are expected to be even as low as some seconds and the connection setup and release can be performed during some ms.

2.1.2 Optical packet switching (OPS)

In optical packet switching networks data packets are statistically multiplexed in optical domain and link wavelength resources are shared between packets belonging to different connections (see Figure 2.2). Control information is carried in packet headers and it is extracted in each OPS node (router). An entirely optical OPS router is supposed to process this control information in an optical way (e.g., see [DHL$^+$03]). Nevertheless, due to still immature all-optical processing the header is usually converted to its electrical form and processed in an electronic node controller. The controller configures a switching matrix so that the packet payload is switched and buffered in an all-optical way.

The transmission times of typical IP packet at $10Gbps$ range from tens of ns to approx. $1\mu s$ and further decreases at higher bit-rates.

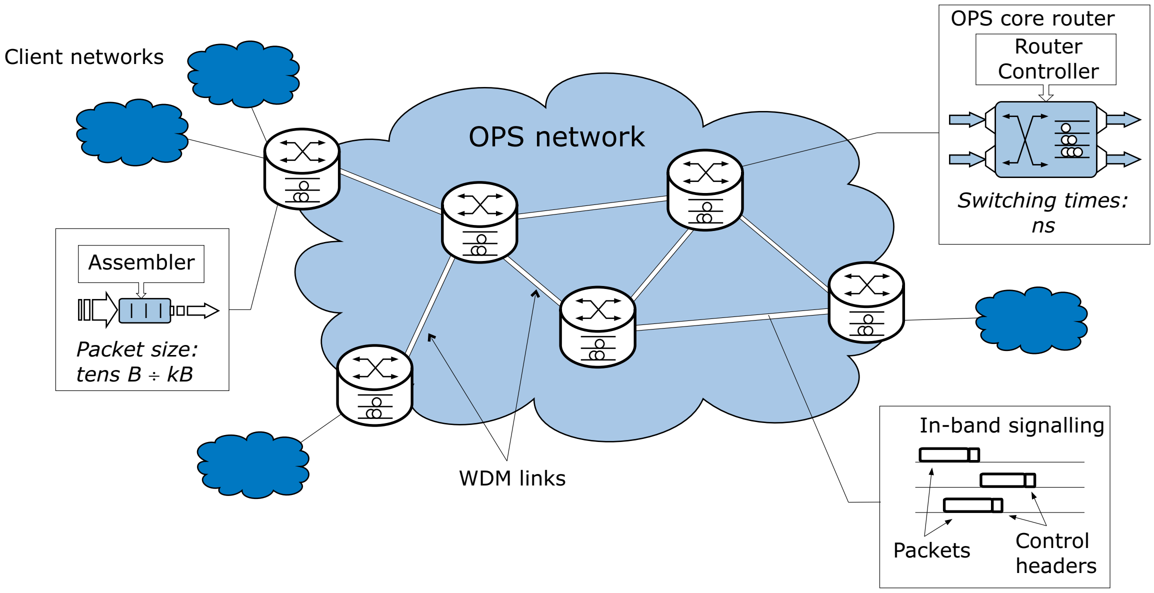


Figure 2.2: Optical packet switching network.

2.1.3 Optical burst switching (OBS)

In the optical burst switching the wavelength resources are shared between different connections, similar to OPS. At the edge of an OBS network, the packets coming from legacy networks (e.g., IP, ATM networks) are aggregated into large optical data bursts which are further transmitted and switched in the network (see Figure 2.3). Each burst has assigned a control packet. The burst control packet and its data payload are transmitted separately on dedicated wavelengths. The control packet is delivered to a core switching node with some offset time prior to the burst payload. In such a way an electronic controller of the core node has time both to process the control information and to setup the switching matrix for the incoming burst. The burst crosses the configured nodes remaining all the way in optical domain.

The duration of typical burst, which aggregates a group of packets, can last from some μs to several hundreds of ms.

2.1.4 OPS vs. OBS

During the past years the definition of OBS and OPS has become less clear because of the large number of proposals claiming either name. Both burst switching and packet switching models provide sub-channel granularity by employing asynchronous time division multiplexing. In case switching is performed all-optically and data stays in the optical domain until the destination edge node the concepts can be referred to as optical burst switching and optical packet switching.

Following characteristics either individually or in combination can be regarded defining for OBS in contrast to OPS:

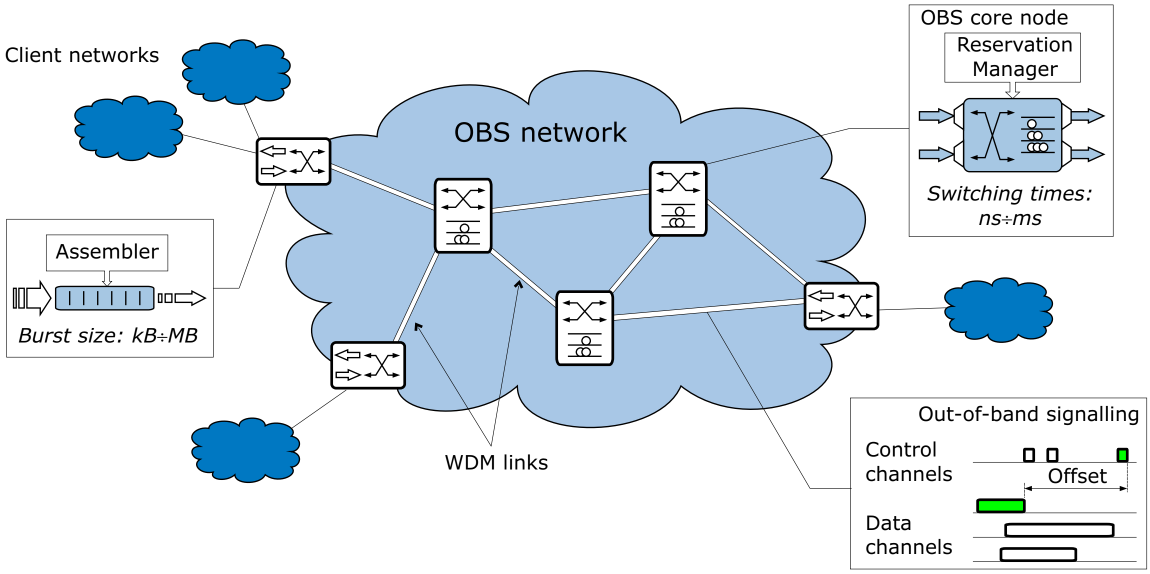


Figure 2.3: Optical burst switching network.

1. client layer data is aggregated and assembled into larger variable length optical data units in edge nodes,
2. control information is signalled out-of-band, processed electronically in all core nodes and used to set up the switch matrix before the data bursts arrive.

2.2 Main characteristics and comparison between OCS, OBS, and OPS

In order to recognize the main characteristics of optical switching architectures we lead a detailed discussion on several control and data plane-related issues.

Hardware requirements

There is a significant difference in the switching time requirements of each optical switching architecture (see Figure 2.4). This together with the various switching granularity (circuits/bursts/packets, i.e., switching speeds of ms for burst switching with end-to-end setup, μs for burst switching with one-pass reservation, and ns for packet switching) is reflected in different requirements for both applied optical-switching components [PPP03] and electronic node controllers [VGPMGH$^+$07].

The switching function in OCS nodes can be achieved with commercially available optical switch technologies such as for instance *micro-electro-mechanical systems* (MEMS) [CLP02]. In fact, relatively long MEMS switching times are sufficient for low-dynamic OCS operation. On the other hand, a dynamic character of optical burst and packet switching requires a fast-switching operation. Ns-scale switching times

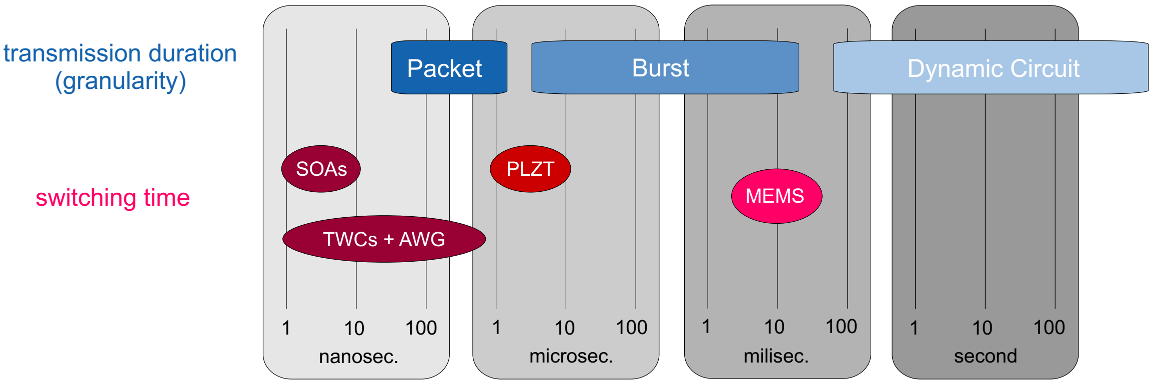


Figure 2.4: Overview over key parameters determining circuit/burst/packet granularity and required switching technology.

can be achieved with commercially available *arrayed waveguide grating* (AWG) and *tuneable wavelength converter* (TWC) technologies or the *semiconductor optical amplifier* (SOA) technology (e.g., see [CDB$^+$03][TZ99]). The *lead lanthanum zirconate titanate* (PLZT) technology [NTL$^+$05] with sub-μs switching times can be a good choice for optical burst switching.

The huge amount of quasi-simultaneously arriving OPS packets may intensify the congestion in control plane. As a result the OPS node controller, which cannot sustain the control traffic load, starts dropping optical packets. One method to alleviate this effect is to introduce high-speed network processors [VGPMGH$^+$07]. Although OBS architectures with the aggregated burst transmission reduce the congestion of control processing, still, they have to be designed carefully in order to prevent from data losses [BD07].

The complexity of hardware is also related with the node architecture, functionality of optical components and equipment dimensions. The highest demands of the above-mentioned issues become apparent for the OPS [BIPe99] which is still looking forward to more advanced compact optical elements. OBS architectures with processing offsets and long data burst transmissions have relatively moderate technological requirements comparing to OPS.

Further we can distinguish the following technological requirements for optical packet and burst switching:

- Suitable switch fabric technology with low loss, low crosstalk, low polarization dependence and low power consumption.
- Burst mode receivers that allow for fast synchronization of clock and adjustment of decision threshold, synchronization/adaptation speed requirements depending on and adapted to switching speed.
- Optical regenerators including wavelength conversion which are tuneable (arbitrary or tuneable/selectable input wavelength and/or tuneable output wavelength) with tuning speed equal to switching speed.

- Tuneable transmitters with tuning speed equal to the switching speed at high output power and high side mode suppression ratio.
- Efficient fibre delay line structures for optical buffering and synchronization especially for OPS.

Processing complexity

The processing complexity of control algorithms at the node level is much higher at OBS/OPS than in OCS. In the formers, specific algorithms for contention resolution, data assembling (at the edge node) and QoS provisioning have to be implemented to be able to fully exploit the potential benefits of the inherent statistical multiplexing principle.

Essentially, the complexity at the network level corresponds to the complexity of routing algorithms or a path calculation process. In the case of OCS, a calculation of the light-path has to occur prior to its establishment in the network, although once established, little maintenance is required at intermediate core nodes. In a connectionless OPS/OBS environment a path has to be calculated at each node and for each packet or burst, respectively. Since each packet/burst potentially has TE and QoS related attributes it needs to be taken into account during the forwarding process. This complexity can be reduced if an MPLS-like connection-oriented environment is implemented. In such case packets or bursts are routed over pre-established virtual network paths.

An important requirement for OBS switch controller is to keep strict the timing of a burst payload arrival relative to its control packet arrival. Thus the management of offset times in OBS networks involves additional complexity.

Performance

Several performance parameters can be envisioned. Below a non-conclusive list is briefly discussed. It contains network utilization, throughput, burst/packet loss probability, and transmission delay.

A disadvantage of OCS is its inefficiency in transportation of traffic which has not been groomed or statistically multiplexed (e.g., see [CEJ05c][AEBS05][LQYG06]). Indeed the OCS switching model, which is useful in carrying highly aggregated long-lived traffic streams, does not fit well within the Internet paradigm of packet switching. Therefore the network utilization or more particularly link utilization in OBS/ OPS can be higher than in OCS due to the statistical multiplexing that allows for a better exploitation of network resources and allows to fit to the actual traffic demands. As a result, the amount of network resources (like e.g., number of interfaces to legacy networks, number of consumed wavelengths) necessary to transport specific amount of IP traffic might be larger in OCS networks than in OBS/OPS networks.

When considering the throughput of OBS/OPS networks, one has to deal with the burst contention problem which is further complicated due to the lack of optical random access memories (RAM). Introduction of fibre delay line (FDL) buffering and wavelength conversion in OPS networks reduces the packet loss probability. Although similar techniques can be applied in OBS networks, still, considerable burst durations put limitations on effective employment of FDL buffers in these networks. As a result we experience high bursts losses in buffer-less OBS networks.

The connection blocking probability is irrelevant for OBS/OPS networks, while it is an important parameter for OCS networks, operating on circuits. It can be shown that introducing sparse wavelength converter pools decreases the connection blocking probability significantly [KA98].

Delays in optical networks may arise due to several reasons. The main transmission delay factor is the propagation delay of optical signal through the optical medium, i.e., optical fibre. Another delay factor is related to the processing and switching operations in intermediate nodes. It is envisioned that the delays produced in intermediate nodes in OPS will be significantly small, to be neglected. In OBS networks, in addition, a burst assembly process may add considerable delays. The delays caused by the setup of offset times might be either small or long depending on the switching and processing technologies used in intermediate OBS nodes. The delays in OCS networks concern mainly delays produced during the connection establishment process. When the connection is established the propagation delay is dominant delay factor.

Finally, an important aspect is the performance at the TCP layer. There is no problem in the case of OCS switching model because the transmitted information does not experience delay variation or losses in optical domain. On the contrary, both OBS and OPS architectures need for dedicated TCP implementations to overtake the degradation of performance due to the higher burst and packet drop rates as well as a possible problem of burst and packet reordering (e.g., see [SPG05][Gun07]). Indeed, the 'edge' problem of OPS/OBS networks is the out-of-order packet/burst arrival which is produced either by the contention resolution and QoS algorithms at intermediate nodes or due to multi-path routing. This issue raises the additional requirement of supplementary hardware (i.e., large memories) at the destination (i.e., sink) nodes due to the reordering operations, which adds further packet transmission delays.

Flexibility

Flexibility in this context expresses the capability to adapt client signals with different bit-rates and data formats in order to transport them through the same network infrastructure.

Appropriate functions are required in edge nodes to adequately adapt different client signals into a common underlying optical transport network. In the OCS architecture such adaptation is achieved either by using an E/O conversion (if the interface of the legacy network element is electrical) or with wavelength conversion (using a transponder/adapter). The adaptation issue in OBS architectures is more complex and involves the aggregation/assembly function. This function is responsible for burst formation at the ingress edge node and for the burst disassembly at the egress node for a given burst flow. The design issue is the proper choice of a burst aggregation scheme with its associated parameters, namely the size of the burst and acceptable burst formation time.

The granularity in the OCS model is very coarse. It corresponds to the bit rate of the wavelength and it is determined at the connection time and fixed for the duration of the connection. OBS/OPS granularity can be very flexible compared to the OCS. These architectures allow for a packet-level switching, moreover, theoretically each packet can have a different bit rate.

A scalability factor that expresses the facility of a new connection/path establishment (MPLS paradigm) gives the OPS/OBS architectures an advantage over the OCS switching model.

QoS

An advantageous feature of connection-oriented circuit switching architectures is that they have no concept of QoS. In the network there is no need to performing QoS-based queue management, as the necessary and sufficient resources (from ingress to egress) are assigned prior to the transmission of the actual data. The issue lies in the fact that there might be contention in the access to transmission resources for the connection requests of different QoS classes.

In OBS/OPS architectures each individual burst/packet has particular QoS attributes, and thus each individual unit requires to be processed. The QoS attributes can be encoded in the reservation mechanism or imbedded in each burst/packet as a code-point that triggers consequent scheduling actions at each intermediate node. Therefore, additional mechanisms based on properly designed algorithms with the hardware on node level as well as network QoS mechanism (e.g., QoS differentiation mechanism, QoS routing algorithm) have to be implemented. These mechanisms should consider burst/packets prioritization (scheduling), resource reservation, and admission control capabilities. Therefore, the complexity (and costs) to integrate QoS in OBS/OPS networks is usually high.

Control plane

GMPLS/ASON has been proposed as control plane (connection management, protection/ restoration, etc.) architecture for OCS networks. Equally, GMPLS/ASON architecture is considered as the control plane architecture for OBS/OPS networks, nevertheless some modifications to the proposed architecture are required and necessary to be taking into account to allow the protocols to operate over the OBS or OPS networks (e.g., see [PSPCK07]).

Signalling overhead (i.e., the volume of signalization/control data) in optical networks is related to the amount of managed data demands. Each circuit switched connection oriented demand and the OPS packet or OBS burst essentially hold the same information. In the circuit demand, the source and destination address is signalled, a circuit is established and all data is transmitted over the established circuit. In the OBS/OPS, each individual burst can be considered as a, although very short, connection circuit. Thus the highest overhead is observed in OPS networks, where each optical packet carries header control information (corresponding to the routing tables in each node). Due to the inherent nature of the OBS architectural model the signalling overhead is significantly reduced. Another aspect is the network's ability to adopt to varying traffic demands. In OCS networks, there is significantly more time necessary to configure the necessary resources when the establishment of a new light-path is performed, which was let's say triggered by an increase in traffic demand, compared to the OBS/OPS approach. In the OPS/OBS network architecture, an increase in traffic demands will require significantly less time for resource configuration, due to the inherent statistical multiplexing principle.

Under high traffic load network scenarios, however, advanced routing and admission control algorithms are expected to be implemented: firstly, in order to avoid contention, by distribution of the traffic over the network in a more engineered manner and thus using the available resources optimally, secondly, to keep the burst blocking rates in some operational bounds.

For connection-less burst and packet switching, the UNI (user-network), I-NNI, and E-NNI (network-network) interfaces need to be adapted to pass burst and packet switching specific information. Besides building the related routing tables, also the address translation may need to be implemented in case of specific addressing schemes.

For connection-oriented burst and packet switching, the routing of virtual or physical circuits with QoS constraints needs to be supported.

For hybrid solutions (wavelength switching combined with OBS), signalling and resource reservation concepts need to be elaborated.

Further tasks for the design of the Control Plane are e.g., the distribution of topological information, the used addressing scheme in the OBS/OPS networked layer, enhanced protections schemes (packet 1+1) for packet based transport technology. Besides, the design of a suitable DCN (control channel) for the UNI, NNI has to be defined.

Routing management

In general, the objectives of routing management are the traffic load distribution over the network (i.e., traffic engineering) as well as preserving the high priority traffic from the best effort traffic (i.e., QoS routing).

The dynamics of OCS networks are incomparable lower than of OBS/OPS networks. The OCS networks operating in a circuit-switched mode create optical connections for the long-term end-to-end data transmission without wavelength sharing capabilities. In such networks, we can consider quasi-static network state (i.e., wavelengths/bandwidth occupancy on particular links is known) when we deal with the so-called routing and wavelength assignment (RWA) problem [ZJM00]. This allows for solving the routing problem on the basis of actual network state information.

The highly dynamic characteristics of OBS/OPS networks, due to very fast burst and packet transmission, produces inaccuracy effects in the network state information. This involves network performance degradation (i.e., increased burst/packet loss probability). Moreover, there is a need to deal with a big number of relatively small transmission units (bursts/packets) in OBS/OPS networks. It makes the problem more close to the routing in IP networks with the additional issue of the lack of optical memories causing that the switch has to use complex contention resolution algorithms in order to provide an acceptable burst/packet loss rate level when bursts/packets contention occurs (a related problem is providing of QoS guarantees). Another issue is the high throughput of optical switching technology that involves huge requirements for processing capacities of a switch control unit (e.g., looking-up the routing tables). All these factors increase the network complexity an involve additional functionality requirements. The application of connection-oriented MPLS environment with its constrained labelled-switched logical paths (for TE and QoS purposes) and fast labels look-up can support the routing management.

The common objective of routing management in OCS networks is minimizing the connection blocking probability by means of explicit routing algorithms. Regarding the RWA problem, the main goal is to establish optical circuits in the network with optimization of the wavelength resource utilization. Routing algorithms in OBS/OPS networks have to balance the traffic load (and so burst/packet loss probability) by means of adaptive routing algorithms in order to help the nodes in the contention resolution problem.

In the OCS model, some strategies are defined to protect or restore the light-paths in case of failure while the problem is still not enough addressed in OBS and OPS.

2.3 Summary

According to the specific features of optical circuit, packet and burst architectures, we can identify their advantageous and drawbacks, which are summarized in Table 2.1. We can also foresee the application of these switching models in the future network deployments.

With respect to the flexibility as well as the network utilization and wavelength consumption, optical circuit-switched networks lag behind compared to optical burst and packet-switched networks. Nevertheless, technology availability makes OCS networks an upcoming solution for the next generation transport networks. The only factor that can delay their introduction by network operators seems to be the need for regaining the costs of investments of already exploited transmission systems (SDH/SONET).

The operation of optical packets switching is very effective, particularly, due to the packet-level transmission and switching granularity. Nevertheless, realization of such bandwidth-efficient, flexible and data centric all-optical networks faces significant challenges. The complexity of the OPS and the high technological requirements can significantly shift the development of OPS networks into the future. This prognosis can undergo modifications as some technological breakthrough in the photonics occurs.

Optical burst switching combines the best characteristics of both OPS and OCS architectures. The bandwidth granularity of OBS networks lies between the bandwidth granularity of OCS and OPS, and relatively relaxed technological requirements (especially for switching and processing components) make it an interesting solution for next generation optical networks. However, the high blocking probability is con-

	OCS	OPS	OBS
Advantages	Natural QoS; reliability Components & subsystems commercially available	Very high flexibility (traffic dynamics) Very efficient network utilization Reduced node size	High flexibility (traffic dynamics) Efficient network utilization Possible using of lower speed switching elements
Drawbacks	Low flexibility and network utilization Very high wavelength consumption, large node sizes	Only preliminary components & subsystems available High control complexity (processing effort. QoS, routing etc.) Resilience more complex Requires more effort for packet reordering	Components & subsystems partly available Control complexity (QoS, routing etc.) Effort for traffic aggregation Resilience more complex
Foreseen for the future	A short term deployment	Waiting for technological breakthrough, especially for compact, low-cost optical components A longer term deployment	OBS is a viable solution for efficient optical networks A mid term deployment

Table 2.1: Advantages, drawbacks and foreseen for future implementations

sidered a serious challenging issue of OBS. Therefore, there is a strong requirement for contention resolution mechanisms (wavelength conversion, FDL buffers, and deflection routing, opto-electronic solutions), also with QoS support, applied in hardware or as an accurately operating control algorithm.

Moreover, routing algorithms should be enhanced with TE functions in order to alleviate the contention resolution problem by appropriate traffic load distribution over the network.

Chapter 3

Optical burst switching

The idea of optical burst switching (OBS) has arisen as an alternative to a low-flexible optical circuit-switching network operation and technological immaturity of optical packet switching solutions (we have discussed these and some other issues in details in Chapter 2).

The principal design objective for an OBS network is that aggregated user data is carried transparently as an optical signal, i.e., without an O/E/O conversion. This optical signal goes through the switches that have either none or very limited buffering capabilities. Besides, the control information is carried on a dedicated channel and separately from the user data. In such a network the wavelength resources are allocated temporally and shared between different connections. It increases network flexibility and its adaptability to the bursty characteristics of IP traffic. Moreover, the aggregation of user data helps to reduce the scale of control information processed in the network as well as it relaxes the switching requirements. Since the control information and the user data are separated they can be encoded with different modulation formats as well as transmitted at different rates. Such division improves network management and provides additional flexibility.

Other justification for OBS concept comes from the network user side. Yet not long ago the predictions on expected services talked mainly about a meaningful participation of real-time multimedia applications with streaming video and broadcasted TV services in packet networks. Instead, the dominance of multimedia and data file transfers (e.g., MP3/divx) using various P2P services together with still limited streaming traffic modifies previous goals [Odl04]. With such P2P services, the typical methods being planned for controlling networks do not fit to user expectations well. The matter to users now is getting a quite big amount of bits quickly, with low transaction latency. OBS concept with fast optical transmission of huge amounts of data seems to match to these expectations well.

Similar objectives of high capacity and usually long-distance data transfers are in grid networks. A grid network is a distributed collection of heterogeneous computational, storage and network resources (see Figure **??**). Most of current operational grids are dedicated to a limited set of computationally and/or data intensive scientific problems, like e.g., energy physics, astronomy weather forecast or high performance computing/visualization. The requirements of grid applications comprise among other things high bandwidth transmission, low connection set-up times and varied transmission granularity for both short and long grid jobs. Network flexibility and huge optical capacity of OBS technology are the appropriate characteristics for actual and future grid applications.

3.1 Overview of general OBS concepts

An OBS network consists of a set of electronic *edge* nodes and optical *core*, or *intermediate*, nodes connected by DWDM links (see Figure 2.3). Ingress, or *source*, edge nodes aggregate data coming from client networks and assemble them into optical *bursts*. Each burst is composed of a *data payload* and a *control packet*. The burst control packet is generated when the assembly process of the burst data payload is finished. The burst control packet carries all the information necessary to discriminate the burst inside the network, like for instance, the burst class or its length. In OBS networks there is a strong separation between data and control planes. In particular, the burst data *payload*, which is the carrier of user data, is transmitted on one of data wavelengths, whilst the burst control packet with its *signalling* message is transmitted on a dedicated *control channel* (wavelength). The control channels can be either *out-of-fibre* or *in-fibre*. In the former a dedicated fibre is provided only for the transmission of control information, whilst in the latter the control channels use the same fibre as the data channels. Inside an OBS network the control information is processed electronically, whilst the data burst payload is transmitted all-optically, without optical to electrical conversion.

In an OBS network the burst control packet is delivered to the core node with some *offset time* prior to its data payload. The offset time is introduced in order to give time for both processing of burst control information and reconfiguration of the switching matrix. The control packet is processed in an electronic *controller* of the switching node. The controller performs several functions, among others the burst *forwarding* and *resources reservation*. The forwarding function, which is related to the network *routing*, is responsible for determination of an output link (port) the burst is destinated to. The resources reservation function makes a booking of a wavelength in the output link for the incoming burst. In case the wavelength is occupied by another burst a *contention resolution* mechanism, if exists, is applied. The contention resolution mechanism may require a *scheduling* policy if alternative resources can be provided for the burst transmission. Also, a *quality of service (QoS) provisioning* function, if implemented, may involve particular treatment of higher priority bursts. In case no resources are available for the incoming burst it is lost.

After the burst transmission is finished in a node the resources can be released for other connections.

3.1.1 Signalling

OBS signalling adapts the ATM block transfer (ABT) standard proposed for burst-switching ATM networks [I.300]. There are two versions of ABT protocol, namely:

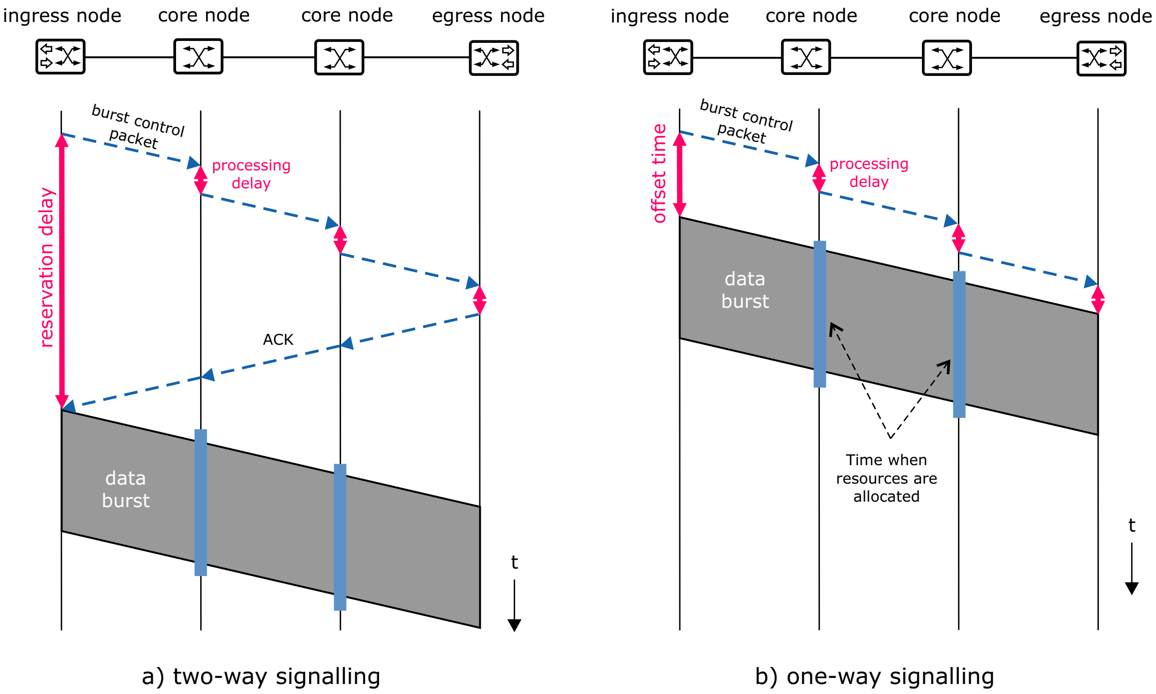


Figure 3.1: Signalling protocols in OBS networks.

- with delayed transmission, which is known as a *tell-and-wait* (TAW) signalling in OBS [Wid95], and
- with immediate transmission, which is called a *tell-and-go* (TAG) signalling [Wid95][VS97].

The TAW protocol, which is recognized sometimes as the *two-way signalling* protocol, performs an end-to-end resources reservation with acknowledgment (see Figure 3.1a) [DKKB00]. In particular when an ingress edge node has a burst ready to be sent it dispatches a request burst control packet towards the network. If all the core nodes on the routing path can accommodate the burst the request is accepted and the ingress node is allowed to go ahead with transmission of the burst payload. Otherwise, the request is refused and the ingress node has either to send another request later or to drop the burst.

The TAG protocol operates with a *one-way signalling* and it allocates transmission resources on-the-fly, a while before the burst payload arrives to a switching node (see Figure 3.1b) [QY99][Tur99]. In TAG signalling the ingress edge node sends a request burst control packet and after that, immediately, without receiving any confirmation, it transmits its data payload. If any core node along the routing path cannot carry the burst because of its congestion the burst is drooped.

A disadvantage of two-way signalling protocols concerns the latency produced during the connection establishment process [KB02][WZSZ03]. For this reason the TAW signalling is oriented more towards metro networks. In such networks, short transmission distances allow keeping low the connection setup times.

The one-way reservation signalling model allows operating in large-distance networks. In such architectures the problem of synchronization between the burst control packet and its data payload arises in the network. Therefore each switching node has to keep updated the information about relative time-scale position of the control packet and the payload. Another issue is the problem of burst contention in the network. Indeed, a burst is released towards the network even it is not guaranteed there are transmission resources available to deliver it to the destination node. For this reason several contention resolution mechanisms have been proposed to alleviate this problem, as we discuss later.

A great feature of OBS concept is the possibility of operating with two-way and one-way signalling protocols simultaneously. In particular, in a two-way resources reservation mode one can setup aside some wavelengths to be used as in an OCS scenario, whilst one-way reservation messages make a statistical use of the rest of available resources. In this way the same optical infrastructure can support both static (by wavelength switching) and dynamic (by burst switching) traffic.

In order to make a distinction, our further discussion concerning OBS assumes a one-way signalling protocol, whilst a two-way signalling model is reflected well in an OCS network.

3.1.2 Architectures and functions of OBS nodes

An OBS edge node

An inter-working function between client networks and an optical OBS network is provided in OBS edge nodes (see Figure 3.2). The client network can be any legacy network like e.g., IP, ATM, SONET/SDH, or other network. An ingress edge node is responsible for adaptation of the client data signals to the format used in the OBS network; accordingly, an egress edge node performs the opposite operation.

A few functions can be distinguished to be performed by an OBS ingress edge node:

- aggregation of data from client networks,
- assembly of a burst payload,
- generation of a burst control packet,
- (optionally) set-up of an offset time,
- burst transmission,
- other functions (e.g., burst segmentation).

Data from client networks is aggregated according to a *forwarding equivalence class* (FEC). Each FEC describes client data of similar or identical characteristics, like for instance their destination, QoS class, or transmission time window. A burst payload is assembled from the data of the same FEC and according to a given burst

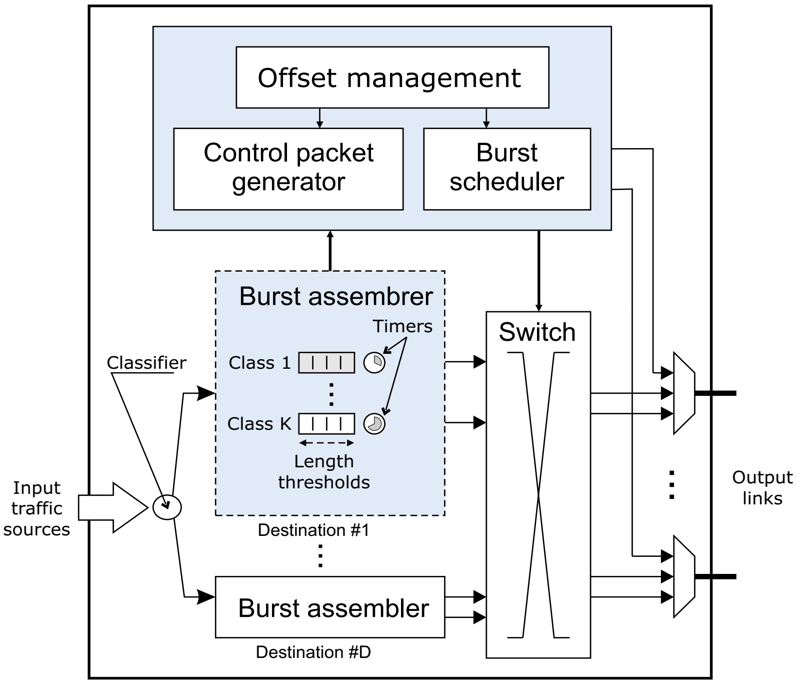


Figure 3.2: OBS ingress edge node.

assembly algorithm. The algorithm takes a decision about when the burst aggregation process should be finished. Several burst assembly algorithms have been proposed for OBS networks (e.g., see [RG04][YLC^{+}04]):

- *timer-based* - define the maximum amount of time for the burst assembly process,
- *burst length threshold-based* - specify the maximum, permissible length of the burst,
- *hybrid* timer/length-based algorithms,
- other algorithms (e.g., with exponentially distributed burst inter-arrivals guaranteed).

A burst assembly algorithm influences the overall network performance. In fact it allows a network designer to control the burst traffic characteristics, e.g., such as burst arrival process to core nodes or burst length distribution. Timer-based algorithms have the burst inter-arrival times determined, whilst length threshold-based algorithms have the burst lengths determined.

An edge node should equip the burst control packet in the burst relevant information, sufficient to handle the burst payload in core nodes. An exemplary burst control packet, shown in Figure 3.3, comprises information about the burst duration, the payload arrival time (relative to the control packet arrival), the class of burst and

Message type	QoS	Input identifiers (port, wavelength, label)	
Burst arrival time	Burst duration	*other functions*	FEC/CRC

Figure 3.3: Burst control packet format.

some routing/forwarding information (input wavelength, an identifier of the routing path).

In a common OBS scenario the ingress edge node introduces an offset time between the burst control packet and its payload. In the simplest case such offset is fixed and equal to the time necessary for processing and switching operations in all the nodes laying on the longest routing path. The problem of offset time provisioning is addressed later in more details.

A burst segmentation is another (optional) function that can be found in the edge node. This function performs a partition of the burst payload onto several data segments. In case the burst collides with another one in a core node its data contending segments can be dropped.

An OBS core node

A transparent switching/routing of optical bursts from one fibre link to another is performed in an OBS core node.

The following functional blocks of an OBS core node can be distinguished (see Figure 3.4) [[Nor03]]:

- an input interface,
- an electronic switch controller,
- an optical switching core, and
- an output interface.

Main function of the input interface is the extraction of control and data channels. Each control channel is connected to a burst mode receiver. The burst mode receiver retrieves the control information from control packets, converts it to electrical form and passes it down to the switch controller. Simultaneously, the data bursts carried on different wavelengths are de-multiplexed and delivered to the optical switching core. Some OBS architectures possess a fibre coil element introduced into the data path. The fibre coil provides some offset time for the control packet processing operation. The input interface also monitors incoming signals and conditions them as required, e.g., through power equalization and regeneration.

The switch controller processes control packets. In particular, it makes a forwarding table lookup, and reserves transmission resources for the incoming data payload.

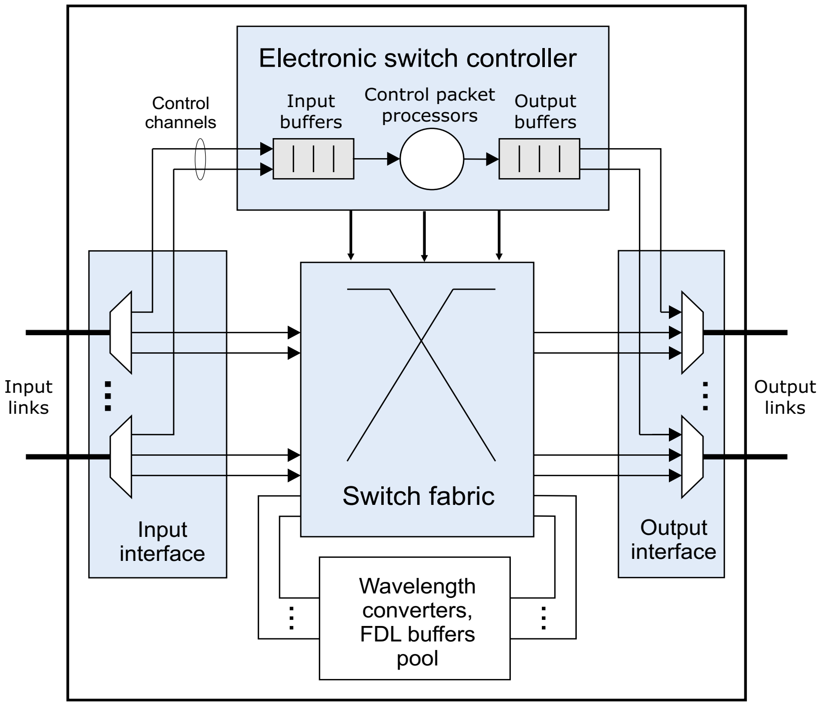


Figure 3.4: OBS core switching node.

The resources reservation is preceded by identifying a suitable switching matrix path and resolving the contention problem, possibly with some QoS policy introduced. The controller usually updates information encoded in the control packet. It is responsible also for sending, in proper instants of time, control signals to the switching core and other switch components in order to handle optical data bursts.

The optical switching core is built with a switching matrix and other dedicated components. The switching matrix can be characterized by the mode of its operation (asynchronous/synchronous), dimension, switching time, internal blocking properties (e.g., non-blocking) and signal degradation (e.g., see [PPP06]). The dimension of switching matrix should be $(NxW)(NxW)$ if N is the number of output/input ports and W is the number of wavelengths per port (link). Other components that can be found in the optical switching core are e.g., wavelength converters and fibre delay lines (FDL); they are used as burst contention resolution mechanisms.

The output interface implements an update of control information, DWDM multiplexing of data and control channels and conditions for optical output signal.

3.1.3 Offset time provisioning

An important feature of OBS architectures is provisioning of offset times, which separate the burst control packets and their payloads. The offset time gives some *delay budget* for processing and switching operations in core nodes, without the need for buffering of optical data burst payload. The burst is lost if effective processing time, the control packet undergoes in the controller, is lower than the delay budget. Therefore appropriate setup of offset times is crucial in OBS networks.

The offset time can be introduced, either

- in an electronic ingress edge node, by delaying the transmission of data burst payload ([QY99]), or
- in an optical core switching node, by means of an additional fixed-length fibre delay element introduced into the data path (as e.g., in [AST$^+$06]).

Three different offset-time provisioning architectures can be distinguished, with regard to the place where the offset time is introduced (see Figure 3.5):

- a *conventional* OBS (C-OBS), with processing offsets introduced in edge nodes,
- an *offset time-emulated* OBS (E-OBS), with processing offsets introduced in core nodes,
- a *hybrid* OBS (H-OBS), with processing offsets introduced both in edge and core nodes.

Later we can distinguish four models of offset-time provisioning in OBS networks with respect to the changes of delay budget a burst experiences during its trip through the network:

- the delay budget decreases - proper to a C-OBS architecture,
- the delay budget is fixed - proper to an E-OBS architecture in which the burst control packet is released together with its data payload in consecutive core nodes (OPS-like operation),
- the delay budget increases - proper to an E-OBS architecture in which the burst control packet is released immediately after its processing in each core node,
- the delay budget fluctuates - proper to a hybrid architecture.

In C-OBS architectures the offset time is setup in a soft-way, by delaying the transmission of data burst payload with respect to its control packet. The offset should compensate all switching and processing times of all the nodes lying on a routing path; hance it can be seen as a global offset, which is setup only once. An important property of C-OBS architectures is that the offset varies inside the network. Indeed it decreases after each hop by the time the control packet spends in the node controller.

In E-OBS architectures the offset time is introduced in a hard-way, by means of a fibre delay coil element which postpones the arrival of the data burst payload to the switching matrix. The fibre delay coil is a passive piece of fibre of fixed length. Such element is responsible only for compensating the switching and processing times produced in the corresponding node; the offset is local and it has to be introduced in each switching node. On the contrary to C-OBS, in E-OBS it is possible, in principle, to keep the offset times fixed in consecutive nodes.

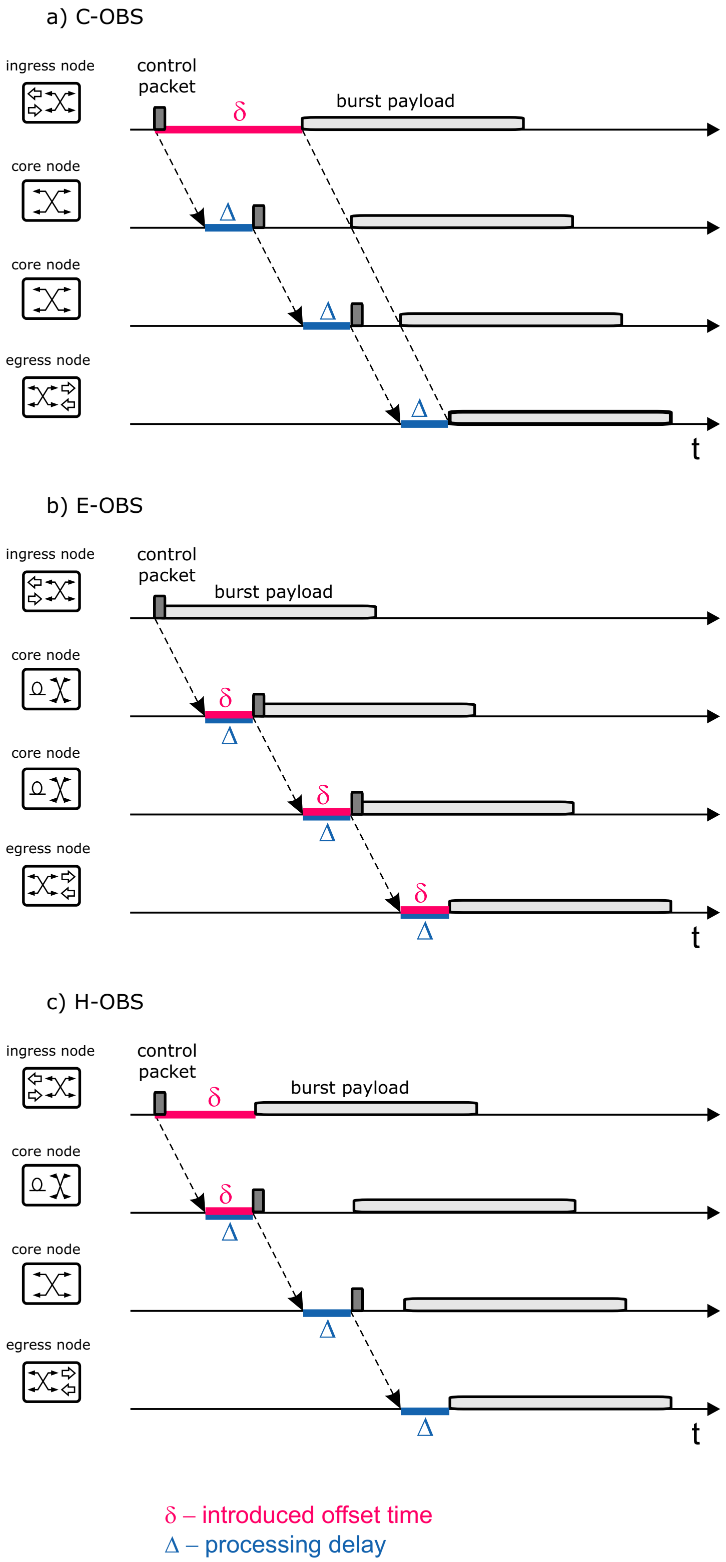


Figure 3.5: Offset time provisioning architectures.

3.1.4 Resources reservation

A resources reservation process in the core node concerns the reservation of resources necessary for undisturbed switching and transmission of data bursts from input to output ports. This process includes reservation of switching resources (i.e., a switching path in the switching matrix), a wavelength in the output link, and other shared resources, e.g., wavelength converters or FDL buffers, depending on capabilities the node is enhanced with.

Separation of data and control channels together with the offset-time provisioning allows using different resources reservation schemes in OBS networks. Each reservation starts from the setup and finishes after the resource release entity. Both resources setup and release can be either *explicit* or *estimated* [BP03]:

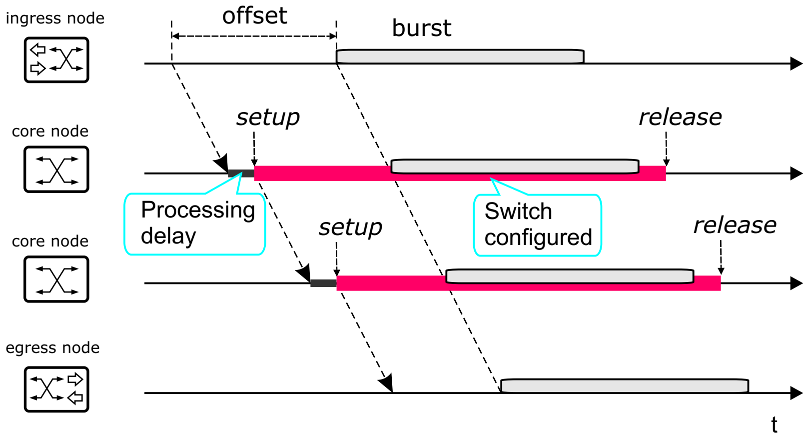
- *explicit setup* - the resources are configured immediately upon processing of the control packet,
- *estimated setup* - the reservation of resources and configuration of the switching matrix is delayed until the actual burst arrival,
- *explicit release* - the source sends an explicit trailing control packet to signify the end of a burst transmission,
- *estimated release* - the end of a burst transmission is known from the burst length, and therefore the moment of release can be calculated.

Different resources reservation algorithms have been proposed by adopting the above presented rules:

- *Just-In-Time* (JIT) [WM00] - performs an immediate resource reservation (see Figure 3.6a). It checks for the wavelength availability just at the moment of processing of control packet. It adopts either explicit or estimated resources release. The advantage of this algorithm is its simplicity, however, at the cost of worsen efficiency due to the overprovisioning of resources.
- *Horizon* [Tur99] - performs estimated setup and resources release. It is based on the knowledge of the latest time at which the wavelengths are currently scheduled to be in use.
- *Just-Enough-Time* (JET) [YQ97] - performs estimated setup and resources release (see Figure 3.6b). It reserves resources just only for the time of burst transmission. It is one of the most efficient mechanisms, with improved burst blocking probability when comparing to other algorithms. A disadvantage of JET algorithm is its high complexity as long as it allows for filling the voids that occur between already performed reservations.

In case of an estimated resource reservation the control packet should carry exact information about burst payload arrival and its length.

a) Just-In-Time



b) Just-Enough-Time

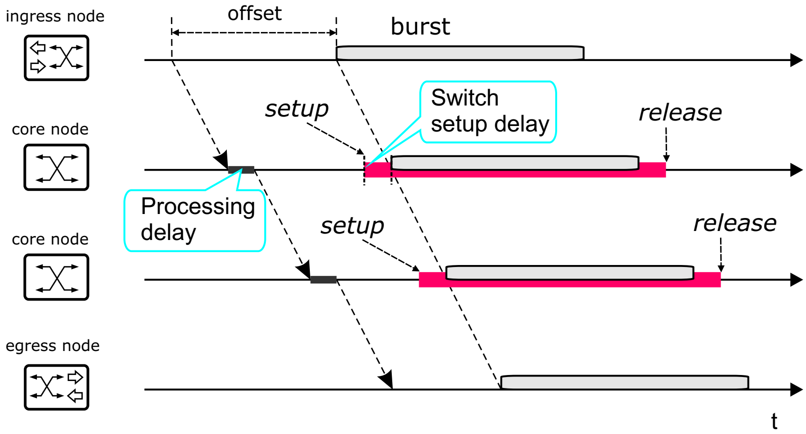


Figure 3.6: Resources reservation schemes.

3.1.5 Contention resolution

A burst contention occurs when more then one burst solicit for the same resources at the same moment. Resolution of the burst contention is a crucial problem in OBS networks. Two factors that complicate the contention resolution are unpredictable and low-regular burst statistics [LES00] and the lack of optical memories. Loosing a burst that aggregates a number of packets may have worse effect than loosing a single packet. The case might be really serious if the burst carries packets belonging to TCP connections [CR06].

Similarly like in OPS networks, the burst contention can be resolved with the assistance of following mechanisms (see Figure 3.7):

- *wavelength conversion* (WC) [ELP03] - converts the frequency of a contending burst all-optically to other, available wavelength;

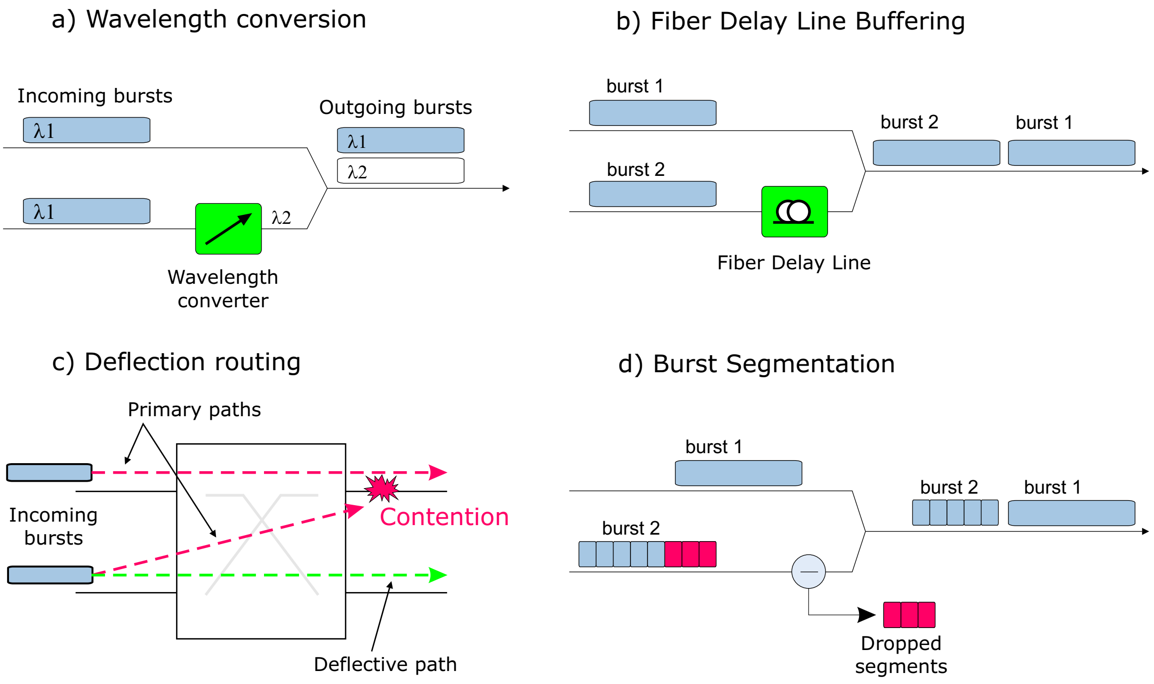


Figure 3.7: Contention resolution mechanisms.

- *deflection routing* (DR) [CZZ04] - forwards a burst spatially, in the switching matrix, to another output port (fibre).
- *fibre delay line* (FDL) *buffering* [HCA98] - operates in time domain and resolves the contention by delaying the departure of one of bursts by a specific period of time.

In case none of mechanisms can resolve the contention the burst is dropped.

The wavelength conversion is a natural way to resolve contentions in OBS networks. A drawback of this mechanism, however, is high cost of WC devices, especially, in case of a full-wavelength conversion, which is performed in the wide frequency range.

Application of the deflection routing in OBS networks is almost cost-less since no additional devices are necessary in order to run this mechanisms. On the other hand, the operational complexity may be high since the mechanism should assure that a deflected burst reaches its destination, even when forwarded to another node output link. Efficiency of this mechanism depends heavily on the network topology and routing strategy as long as the contention is resolved by re-routing the traffic to adjacent nodes. It was shown that the deflection routing can improve network performance under low and moderate traffic loads whilst it may intensify burst losses under high loads [ZVR$^+$04]. Another difficulty that has to be managed properly is the out-of-order burst arrival problem, which occurs when bursts traveling over variable-length paths arrive to the destination in a disordered sequence.

Even if one of the principal design objectives for OBS was to build a buffer-less network the application of FDL buffering is considered as well. Both feed-forward and feed-back FDL buffer architectures can be used [Gau02]. In [Gau03] it was shown that combined application of FDL buffering with WC can significantly reduce burst losses in OBS networks. As a buffering tool the FDLs are bulky and not scalable. Comparing to electronic buffers and their role in current packet networks, the FDL offers only a limited buffering capability. For a typical fibre span of $80km$ length the corresponding maximum delay that can be introduced by FDL buffer, without need for optical signal amplification, is $266\mu s$ [Gau03]. In order to work effectively an FDL buffer has to provide several delays. Therefore a basic delay unit of such buffer, which for efficiency reasons should correspond to the average burst duration [CC01], can not exceed some tens of μs. Considerable burst durations significantly limit the application of FDL buffers in OBS networks, when comparing to OPS networks.

Another technique that aims at the reduction of data losses is a burst segmentation technique [VJ02a]. In this mechanism each burst is divided into a few segments. If a burst contention occurs, instead of loosing the entire burst either the head or tail segments of one of the contending bursts are dropped. The burst segmentation increases the operational complexity due to additional information about the burst segments that have to be transported and processed. Moreover, when the node loses some packets of the burst head and the following packets in the same burst belong to the same flow, it breaks the correct packet sequence what causes a degradation problem at the end-to-end transport protocol (such as TCP).

3.1.6 Burst scheduling

A scheduling algorithm undertakes a decision which wavelength or FDL delay has to be assigned to a given burst in case there are more resources available. The simplest scheduling schemes can be based on either a random or a round-robin resources selection.

More advanced scheduling policies, which are based on Horizon and JET resources reservation mechanisms (see Figure 3.8), are:

- the *latest available unused channel* (LAUC) [XVC99], which is a Horizon-type algorithm; it keeps a track of the latest unscheduled resources and searches for a wavelength with the earliest available allocation;
- the *void-filling* (VF) [XVC00], which is a JET-based algorithm; it keeps a track of the latest unused resources and allows putting short bursts into time gaps before the arrivals of future scheduled bursts. Several variations of the VF algorithm can be found in literature (e.g., see [MRZ04]). A VF algorithm can achieve better performance than a Horizon-based one, however, at the cost of high processing complexity. In an OBS with FDL buffering, this complexity can be decreased when using a FDL-batch algorithm [XQLX03].

There is a group of scheduling techniques which are enhanced with so called *look-ahead processing window* capability (e.g., see [FJ03a][JECA03]). The look-ahead

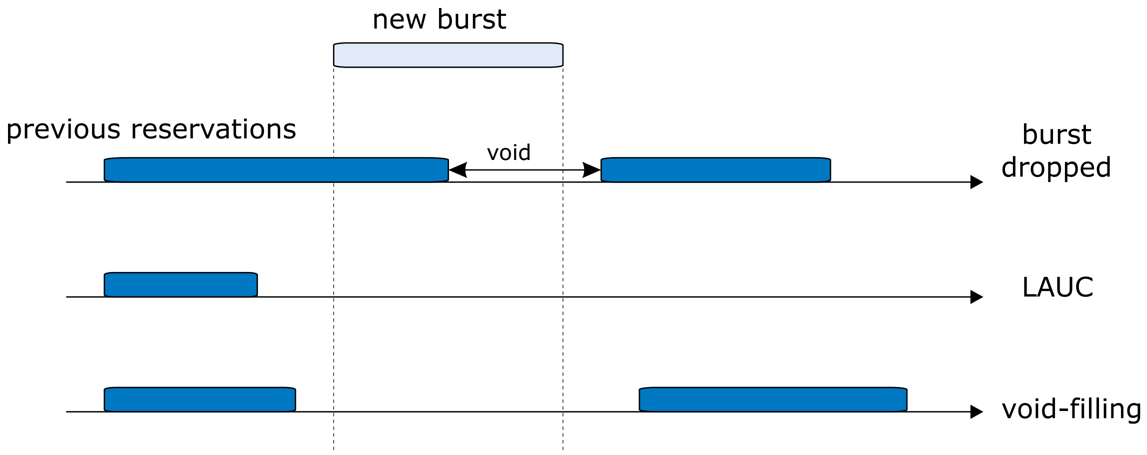


Figure 3.8: Burst scheduling algorithms.

window allows collecting more information about incoming burst reservations, and thus their optimized processing can be performed.

3.1.7 Quality of service provisioning

OBS architectures need for dedicated QoS mechanisms in order to preserve the quality-demanding applications from the best-effort data traffic. Since optical networks do not have an equivalent to electronic random access memories the problem of burst-loss quality guarantees is very challenging. On the other hand, almost buffer-less and fast transmission in OBS networks may result in lower latency than in traditional data networks.

The detail of QoS provisioning in OBS networks as well as some particular issues are addressed in Part 3 of the thesis.

3.1.8 Network routing

OBS architectures with no buffering capabilities are sensitive to burst congestion. A proper routing strategy, enhanced with some *traffic engineering* (TE) capabilities, may help in the congestion reduction. A highly dynamic character of burst traffic, however, may result in the inaccuracy of network state information. Moreover, there is a need to deal with a big number of relatively small data bursts. Other issue is the high throughput of optical switching technology which involves additional requirements for processing capacities of switch controllers (e.g., fast looking-up of routing tables). All these factors increase network complexity and involves additional funcional requirements in OBS.

Application of a connection-oriented *multi-protocol label switching* (MPLS) architecture [Ros01], with its explicit routing and fast labels look-up, can help in the discussed issues. As a consequence several routing methods (e.g., see [ZLW$^+$04] [ZWZ$^+$04][LY06][HHM05]) apply the concept of *labelled* OBS (LOBS) for TE, as proposed in [Qia00].

An important issue related to the routing problem is end-to-end QoS provisioning; in this context, several solutions have been proposed in the literature (e.g., see [VJ02b][LKSG03][KG03a][LYH$^+$06][ACP04]). Some routing strategies support network resilience by the computation of backup paths ([CMC06][Bou03][HHM05][GZ06] [JQX00]). A study on multicast routing in OBS networks can be found in [JXC$^+$00] and [JQX00].

Part 4 of the thesis is devoted to the routing problem in OBS networks.

PART II

Offset Time-Emulated OBS Architecture

Chapter 4

E-OBS architecture

From the very beginning there have been considered two distinct concepts of offset time provisioning in OBS networks [QY99]. In a conventional OBS (C-OBS) the offset time is introduced in the edge node by delaying the transmission of the burst payload with respect to its control packet. On the contrary, in an offset time-emulated OBS (E-OBS) the offset time is provided in each core node by means of additional fibre delay element. Although C-OBS has attracted lots of attention it possesses many disadvantages that can be avoided in E-OBS.

The intention of this Chapter is to point out the strengths and weaknesses of C-OBS and E-OBS architectures. At the beginning we introduce operational principles of an E-OBS architecture. Then we lead comparative discussion on several issues related to both functional and performance characteristics of E-OBS and C-OBS.

4.1 Principles of E-OBS

4.1.1 Node architecture

E-OBS core node is a typical OBS node (e.g., [XVC00]) with additional pool of fibre delay coils (FDC) introduced into the data path of the input interface (see Figure 4.1). The control channels are provided either out-of-fibre (i.e., in a dedicated fibre) or in-fibre (i.e., in the same fibre as data channels). In the case the in-fibre control channels are used, they should be filtered before the pool of FDCs. To perform this function a passive optical device like e.g., a band splitter module, or an optical drop multiplexer (ODM) can be used.

The input control channels pass through the optical to electrical (o/e) converters and are directed to the switch controller. The controller is equipped with input buffers to store the incoming control packets before their processing in the processor unit(s) (CPU). After that and some output buffering, control packets are converted back to the optical signal form and transmitted through the output control channels to the output interface. If in-fibre control channels are used, the output interface combine both data and control channels in a multiplexer into an output fibre.

The input data fibres, after separation of control channels, pass through the pool

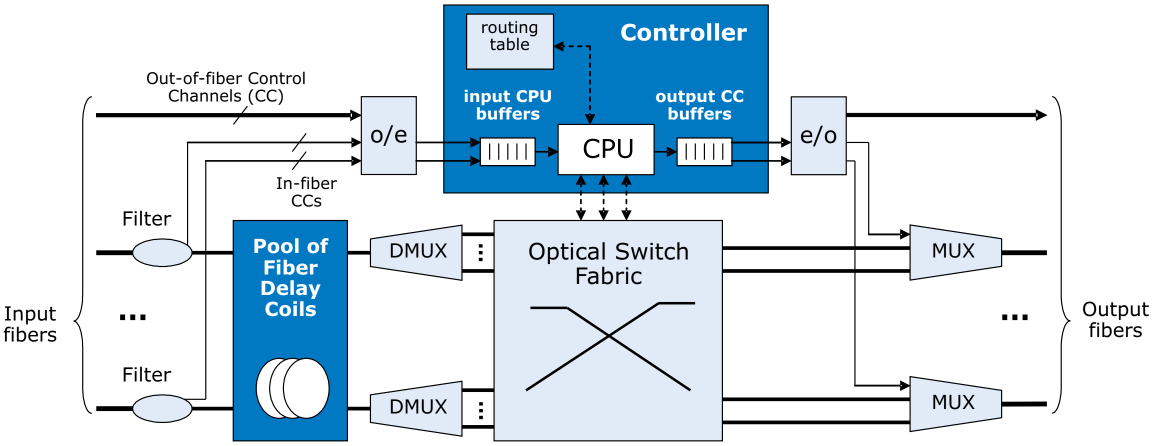
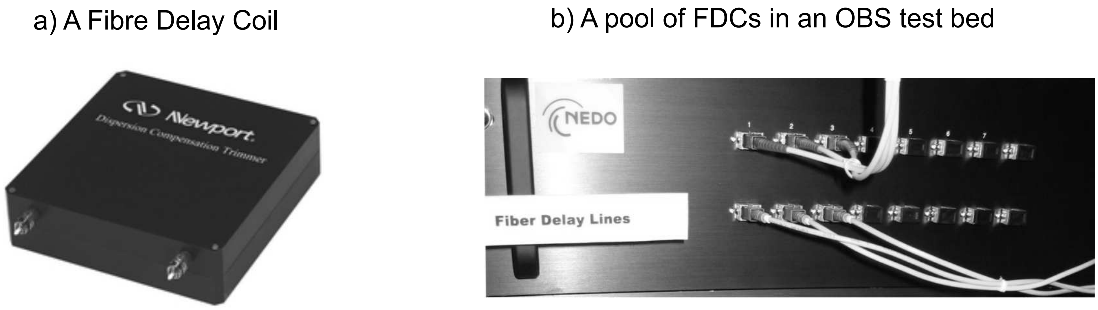


Figure 4.1: General E-OBS core node architecture.


a) A Fibre Delay Coil

b) A pool of FDCs in an OBS test bed

Figure 4.2: Fiber delay coils; a) a single component, b) a part of an OBS test-bed. (courtesy of Newport Corp., and OITDA)

of FDCs - each data fibre passes through one FDC. Then the data channels (wavelengths) are de-multiplexed and the data bursts, from each data channel, undergo all-optical switching operation in the switch fabric to appropriate output ports.

The complexity of FDC is much lower than of any FDL buffer. In fact the FDC is a piece of fibre of quite limited, fixed length and it does not require any switching capability. Such components are commercially available (e.g., see [Fib07a][Fib07b]); exemplary parameters of a FDC presented in Figure 4.2a are: the insertion loss $< 0.3db/km$, fibre length up to $4km$ what gives $20\mu s$ of delay, operating wavelengths $1260 \sim 1650nm$, dimension $6.00'' \times 6.00'' \times 1.59''$ with enclosure.

It worths to mention that there is a need for only one FDC per each input port. The maximum nodal degree in the most referenced mesh network topologies (see Section 10.1 for more details, and also [RFL05]) does not exceeds 5, thus the introduction of a pool of FDCs into an OBS node should not cause much troubles. Indeed some OBS test-beds operating with FDCs (see Figure 4.2b) are already available [AST$^+$06].

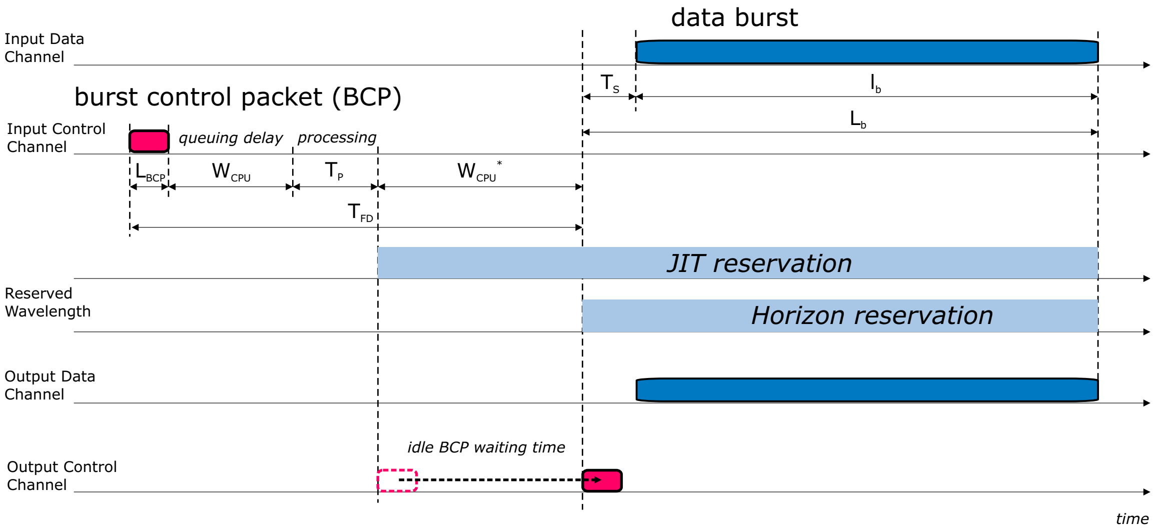


Figure 4.3: Time dependencies in E-OBS.

4.1.2 Control operation

In principle, E-OBS is considered to operate with a one-way signalling. In a proposed E-OBS scenario we assume the burst control packet is dispatched from the edge node prior to its data payload with a small offset introduced just to compensate the switch re-configuration delay.

When the burst reaches a core node, the control packet goes directly to the switch controller, whilst the payload is delayed in FDC by some fixed time. During this period the control packet is queued and processed in the controller so that to reserve the switching and transmission resources for the arriving data payload. This operation is repeated in each core node. When the burst reaches its egress node it is disassembled and data are delivered to the higher layer protocol.

The control packet after its processing can be forwarded to the next node either immediately or it remains in the controller memory until the processing offset expires (an example is presented in Figure 4.3). Both solutions have their advantages and disadvantages.

- In an **immediate control packet forwarding**, the offset time between the control packet and the payload increases hop-by-hop what may result in two effects. On the one hand the increasing offset gives more chances to reserve the resources and it may help the burst to accomplish its trip (e.g., see [YQD01] [KCM04]). On the other hand there might be additional waste of resources in the case of JIT resources reservation since the reservation periods increase in consecutive nodes.

- In a **delayed control packet forwarding**, the time distance between the control packet and the payload, in principle, is kept fixed from link to link

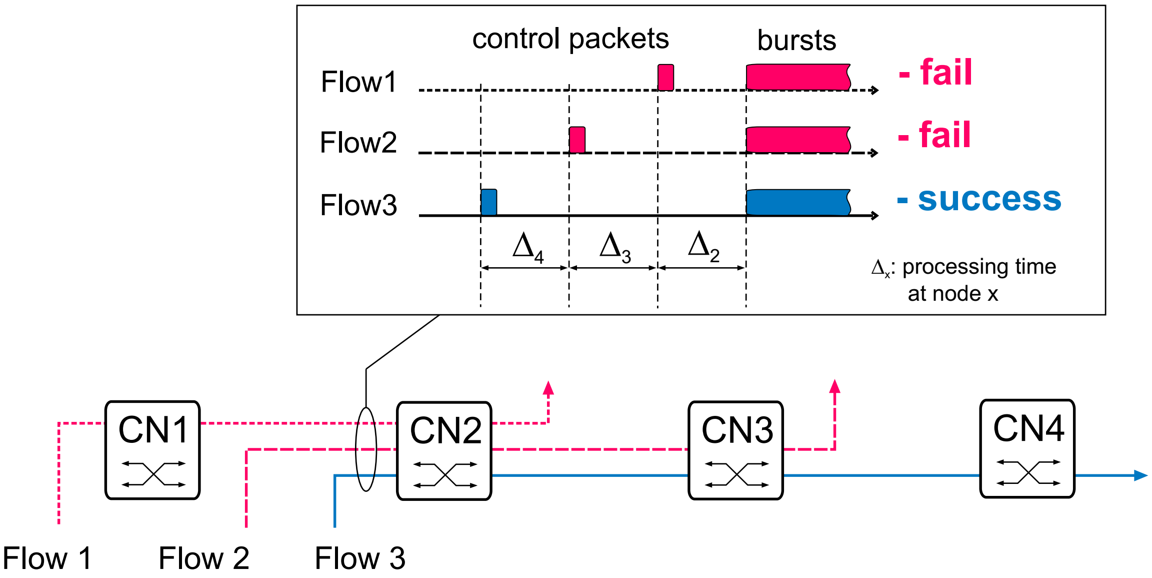


Figure 4.4: Unfairness in conventional OBS.

inside the network. Thanks to this feature there is no variation of offset times in E-OBS. The only inconvenience is a possible contention of control packets in the output control channel. Therefore some output buffers has to be used and the emerging buffering delays should be compensated by the FDC.

In this thesis we concern on the delayed control packet forwarding.

An important requirement for the controller is to keep strict the timing of a burst arrival relative to its control packet arrival. We assume that on entry to the switch each burst control packet is time-stamped. Then after its processing and scheduling to an output queue the relative burst arrival time is re-calculated and updated in the control packet.

4.2 Characteristics of E-OBS and C-OBS

Fairness

In conventional OBS, whilst the control packet is forwarded through the network its global offset time decreases successively at each hop by processing time, which is the time the control packet spends in the node controller. The emerging variation of offset times can produce *unfairness* in access to transmission resources (see Figure 4.4). Indeed a burst of higher number of hops remaining to reach the destination, and thus of larger offset time, has more chances to reserve an output wavelength than a burst of smaller offset time. The described effect starts to play role if the offsets assigned to the bursts are larger than the burst durations (e.g., see [DG01]). It worths to mention that this feature has been used in an offset-time differentiation mechanism designed for QoS provisioning [YQD01].

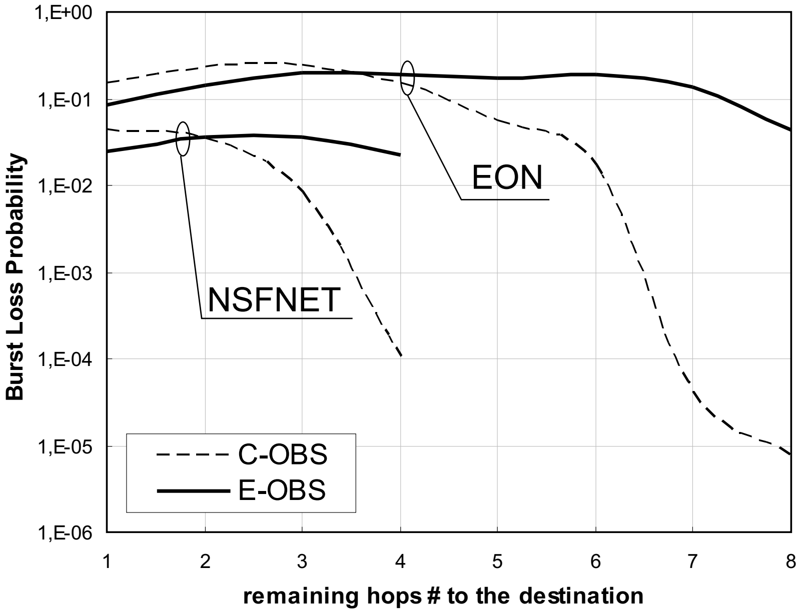


Figure 4.5: Burst loss probability vs. remaining hops number.

Another negative aspect related to the unfairness is the *path length priority effect* (see [KCM04]). This effect corresponds to the increased loss probability of bursts that approach their destination and at the and of their trip have small offsets. In particular such bursts can be easily overtaken by the bursts of higher offsets, e.g., which have just been expedited from the ingress node. As a consequence, we could have unnecessary waste of transmission resources that were already utilized in all the nodes traversed by the lost bursts.

In order to illustrate the unfairness effect, in Figure 4.5 we present some exemplary simulation results. We consider two network topologies called NSFNET (an American backbone network) and EON (a pan-European network) of 15 and 28 nodes, and 23, and 39 links respectively (see Section 10.1 for more details about the network scenarios). Each link has 32 data wavelengths and the transmission rate is $10Gbps$. Each node is an edge node generating $25.6Erlangs$ (0.8 load, when normalized to the link capacity). Bursts have exponentially distributed inter-arrival times and lengths (mean duration of $32\mu s$). $1\mu s$ and $10\mu s$ are the times considered for the switching and processing operation respectively. The JET resources reservation with the LAUC-VF scheduling is used. Shortest path routing is applied.

We can see that the bursts that begin their trip, i.e., of high number of remaining hops to the destination, undergo lower losses than the bursts which have just the ultimate hops to reach the destination in C-OBS. On the other hand, in E-OBS each burst has the same chances to reserve the transmission resources as long as the offset times that are determined by the length of a FDC are the same. The results presented in 4.5 confirm this observation. In particular, we can see that the burst loss probability

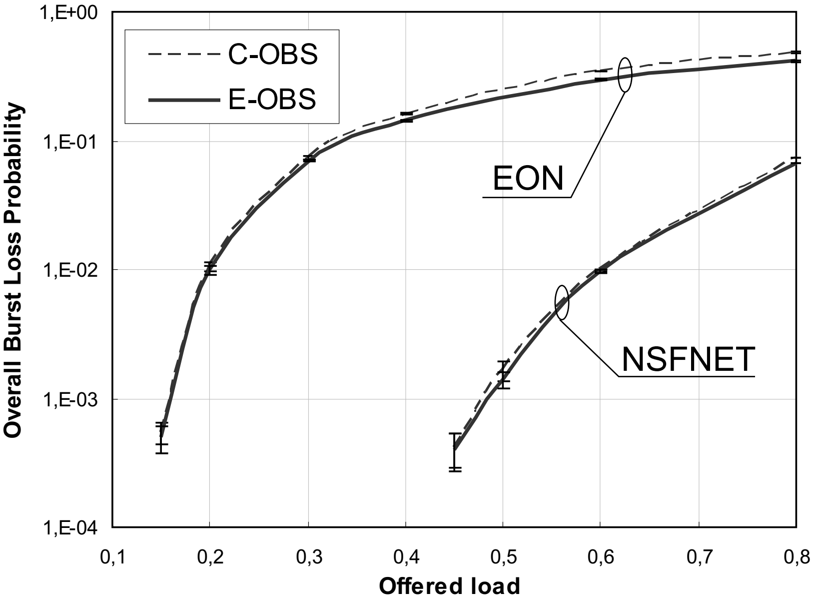


Figure 4.6: Burst loss probability vs. offered traffic load.

is much more stabilized without regard to the number of hops remaining to reach the destination; its slight variation is due to an unbalanced traffic-load distribution. As a consequence, the unfairness in access to transmission resources experienced by the bursts belonging to different connections disappear in E-OBS.

Burst loss and delay performance

Figure 4.6 compares the overall burst loss probability obtained as a function of offered traffic load (normalized to the link capacity) under different network scenarios. As we can see both C-OBS and E-OBS architectures offer similar performance.

Transmission delay produced in OBS networks is due to the link propagation delay d_l (approx. $1ms$ in $200km$ link) and the offset provided for the processing δ_p (up to several μs) and switching δ_s (below μs if a fast switching matrix is used, as e.g., in [GWL+05][AST+06]) purposes. We have already assumed that in E-OBS the switching offset between the burst control packet and the data payload is introduced in the edge node. Hence, the end-to-end transmission delay D the burst undergos is the same in both C-OBS and E-OBS architectures, and it can be expressed as:

$$D = (n+1)d_l + n\delta_p + \delta_s \tag{4.1}$$

where n is the number of intermediate switching nodes

Note that the propagation time is still a dominant delay factor.

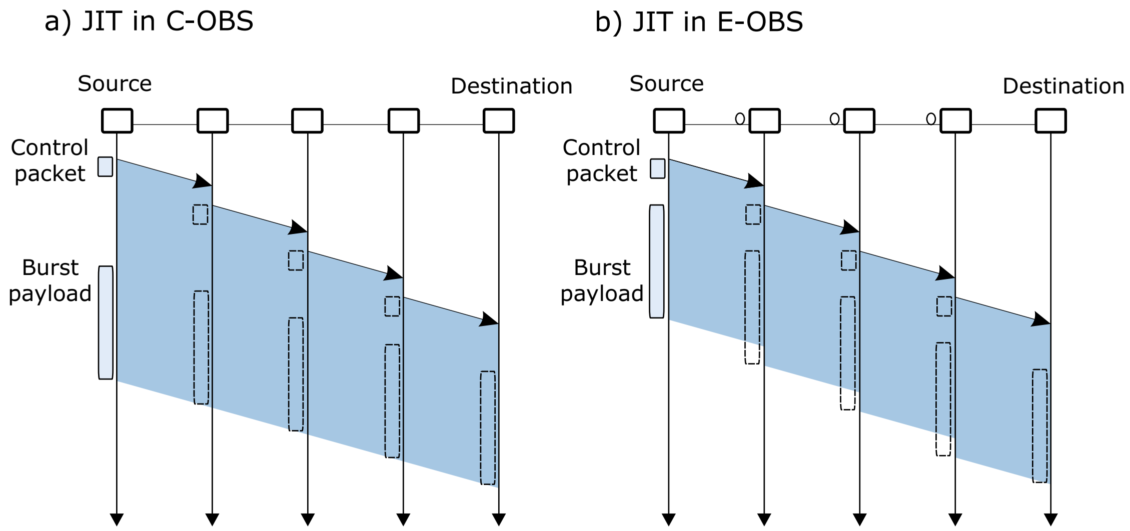


Figure 4.7: JIT resources reservation in a) C-OBS, and b) E-OBS.

Resources reservation

The Just-in-Time resources reservation algorithm performs an immediate resource reservation as it checks the wavelength availability just at the moment of the control packet processing. On the contrary both the Horizon and the JET perform a delayed resources reservation for the period of time starting from the burst arrival time. The difference between these algorithms is that the Horizon searches for a wavelength that does not have any later reservations while the JET allows for filling the voids that occur between reservations.

E-OBS can operate with any resources reservation algorithm. Nevertheless, the JIT and the Horizon algorithms seem to be the most appropriate ones, whilst the JET algorithm is frequently considered for C-OBS networks.

The advantage of JIT is its low complexity since the only information that has to be kept record of in network nodes is whether a wavelength is currently available or not. Over-provisioning of resources due to early reservations is the main drawback of JIT (see Figure 4.3). As a result, burst losses can occur even when there are no transmission conflicts on the same wavelength. Nevertheless, the reservation periods are shorter in E-OBS than in C-OBS due to smaller offsets times (see Figure 4.7). Hence we can expect that the over-provisioning effect of JIT algorithm will have much lower impact on the performance in E-OBS than in C-OBS.

E-OBS, in principle, does not experience the offset variation inside the network. If we consider the switching nodes are not enhanced with FDL buffering, a resources reservation algorithm does not need to be void-filling aware. For this reason the Horizon algorithm can be used instead of the more complex JET algorithm without any degradation of performance.

Burst scheduling

One of the key problems of OBS is to schedule the bursts efficiently so that the throughput is maximized and the burst losses are minimized. In the OBS networks without FDL buffering the performance of an online best-effort scheduling algorithm depends among other things on the offset time and the burst length distributions. In particular, the best performance is achieved when all bursts have the same offsets and the same lengths [LQXX04]. Whilst C-OBS is characterized by variable offsets, E-OBS can provide fixed offset times.

Another benefit from the core node-introduced offsets in E-OBS is some facility in the application of look-ahead processing window techniques. These techniques need for soma extra offset in order to constitute a look-ahead processing window. The processing window allows for more efficient burst scheduling in both contention resolution [FJ03a][JECA03] and QoS provisioning [FJ03b][KCMSP06]. Since the processing window can be provided easily in E-OBS, by means of additional FDC delay, its introduction in C-OBS may seriously aggravate the unfairness.

Quality of Service provisioning

Several strategies have been considered in the literature to provide QoS capabilities in OBS networks. Among them a burst preemption technique (e.g., see [KA03]) and an offset-time differentiation technique [YQD01] can offer the utmost performance with regard to the class differentiation (see Chapter 7 for a detailed analysis). The former allows overwriting the resources reserved for LP bursts by HP reservations in case of burst conflicts. The latter assigns an extra offset time to HP bursts, what favors them during the resources reservation process.

The general drawback of burst preemptive-based mechanisms in OBS is the overbooking of resources in the downstream nodes in case of a successful preemption. Therefore there is a need for additional signalling procedure to be used in order to release them, or the resources are wasted. This problem is addressed in Chapter 8, and we show that the overbooking of resources can be effectively avoided in E-OBS nodes enhanced with the processing window capability.

Performance of the offset-time differentiation mechanism may be affected by the *multiplication of effective classes* due to the offset variation [DG01]. In order to diminish this effect the offset times should be low enough in C-OBS. E-OBS does not have such limitations thanks to its fixed offset-time provisioning.

Routing and network survivability

C-OBS architectures have some difficulties with alternative/deflection routing. In particular, edge nodes should know the routing path prior to the control packet transmission in order to calculate and setup the offset times accurately. When allowing for alternative routing inside the network, an *insufficient offset* problem may emerge. Indeed in case an alternate route is longer than a primary route the burst is dropped if the control packet does not have enough time to reserve resources ahead of the data burst. For this reason, the offset time should be either calculated for the worst case, i.e. for the longest possible alternative path, what may result in superfluous burst delay, or additional hardware (an output FDL like in [HLH02b]), or control [CEJ05a] mechanisms have to be involved in order to diminish this effect.

Some OBS restoration schemes presented in the literature consider deflection routing to coop with link failures (e.g., [XTGE$^+$04], [HSE04]). Again, an important factor that has to be taken into consideration here is the insufficient offset effect. Therefore, the choice of the offset time is very critical due to its influence on the burst losses in OBS networks.

In E-OBS the routing path can be created freely inside the network with any alternative routing algorithm as long as the offset time is introduced in each core node by means of the inlet FDC.

Hardware complexity

Fibre delay coil There is some additional hardware complexity in E-OBS due to the need for FDCs to be introduced at the input ports of core nodes (we have already discussed this in Section 4.1). Typical FDC delays necessary for E-OBS operation range from some μs to tens of μs, depending on switching and control processing technologies (e.g., see [BBE$^+$05]) used as well as particular choices for control algorithms (resources reservation, scheduling, etc.). Therefore considered lengths of FDC can be between $1 \div 5km$, as e.g., in [AST$^+$06].

The attenuation of optical signal (below $0.3dB/km$) should be taken into account when analyzing the power budget and designing the amplification stages. It is important to say that there is a need for only one FDC per node input port which compensates offset-times for all the data channels simultaneously. The control channel should be extracted before the FDC module and brought to the switch controller. The application of FDC might be advantageous in the context of signal regeneration since this fibre could act as a dispersion compensation unit for the optical signal entering the node.

Memory requirements The requirements related to the amount of electronic memory installed in C-OBS nodes are higher than in E-OBS nodes. C-OBS edge nodes require for output electronic buffers to store the assembled bursts for the offset period. The capacities of such buffers greatly depend on the burst assembly parameters as well as on the offset times itself. In some OBS scenarios the burst payloads are considered to carry some $Mbytes$ of data. Moreover the offsets, which comprise the processing times for all core nodes laying on the routing path, might be very large. As a result the memory requirements in C-OBS might be really high.

In E-OBS the burst after its assembly has to wait in the edge node only for a short switching offset period. Then it is sent towards the network as soon as there are free transmission resources in the output link. Some additional buffers are necessary in E-OBS switching nodes, because the burst control packets might need to be stored after their processing. Nevertheless, the memory requirements in this case are moderate as long as the lengths of control packets are very small.

	C-OBS	E-OBS
Fairness	No	Yes
Performance	BLP slightly better in E-OBS, end-to-end delay the same	
Resources reservation, scheduling complexity	High	Low/Medium
QoS	Some difficulties	Some facilities
Alternative routing	Limited	Not limited
Hardware complexity	Memory (at the edge)	Fibre delay element (in the core)

Table 4.1: Advantages and drawbacks of offset-time provisioning architectures

4.3 Summary

Table 4.1 summarizes both the qualities and drawbacks of the discussed offset-time provisioning architectures. The E-OBS surpasses the C-OBS in many aspects as we discussed it in this Chapter. Therefore there is a motivation for recognizing the E-OBS as an efficient and functional solution for OBS networks.

Although not mentioned before, there is one more great benefit of E-OBS. Thanks to the application of fiber delay elements, which makes the operation of E-OBS and OPS are very similar, an E-OBS architecture can be seen as an immediate migration step towards OPS architectures. Still, the differences between both technologies lay in the length of transmission units (burst vs. packet) and signalling mode (out-of-band vs. in-band), and thus, higher hardware and processing requirements of the OPS. Nevertheless, as the progress in optical technologies will continue these dissimilarities should disappear in the future. As a result, the application of E-OBS may facilitate the migration from OBS networks to OPS networks.

Since E-OBS architectures need for additional fibre delay elements, the study on their feasibility is provided in the next Chapter.

Chapter 5

Modelling of E-OBS control plane

5.1 Introduction

Due to the separated transmission of burst control packets and data payloads both opto-electronic control and all-optical data planes can be seen as two parallel networks, namely a data and a control network (see Figure 5.1). The burst is lost if either its control packet or its payload is lost; it occurs, in principle, due to resources occupancy in congestion states. Both burst contention resolution mechanisms and scheduling algorithms deal with the problem of congestion in data plane (e.g., see [CCXV99][XVC00]). The congestion in control plane can be solved by packet queuing in electronic buffers of the controller (e.g., see [KA04]).

Burst losses can be also due to early burst payload arrivals. This effect arises if an effective processing delay the control packet undergoes in the controller is larger than a delay budget given by the offset time; in such case the burst is lost. The effective processing delay is determined by the queuing delay and processing time of control packet as well as the switch setup time. While control packets experience variable queuing delays, depending on the congestion situation, the effective processing delays vary as well. As a result the determination of appropriate delay budget and setup of offset times that would prevent burst losses is not a trivial task. Notice that excessive over-provisioning of offset times is undesired in OBS networks since it results in extended burst delays and puts constraints on the application of fibre delay elements in E-OBS.

Although, there are some studies that consider the impact of congestion in control plane on OBS node/network performance (see e.g, [WZV02], [BD07]), still, few of them address the problem of sufficient offset time provisioning. In [BD06] an initial discussion on some factors which constitute the processing delay budget is provided. In [KCK04] a control packets scheduling algorithm reducing the effect of insufficient offsets is proposed. Finally, in [CCXV99] an M/M/1 queuing model is used to compute an approximation for the complementary distribution of the control packet processing delay. Since the results presented in these works are very preliminary the study has to be continued.

In order to address thoroughly the problem of sufficient offset time provisioning

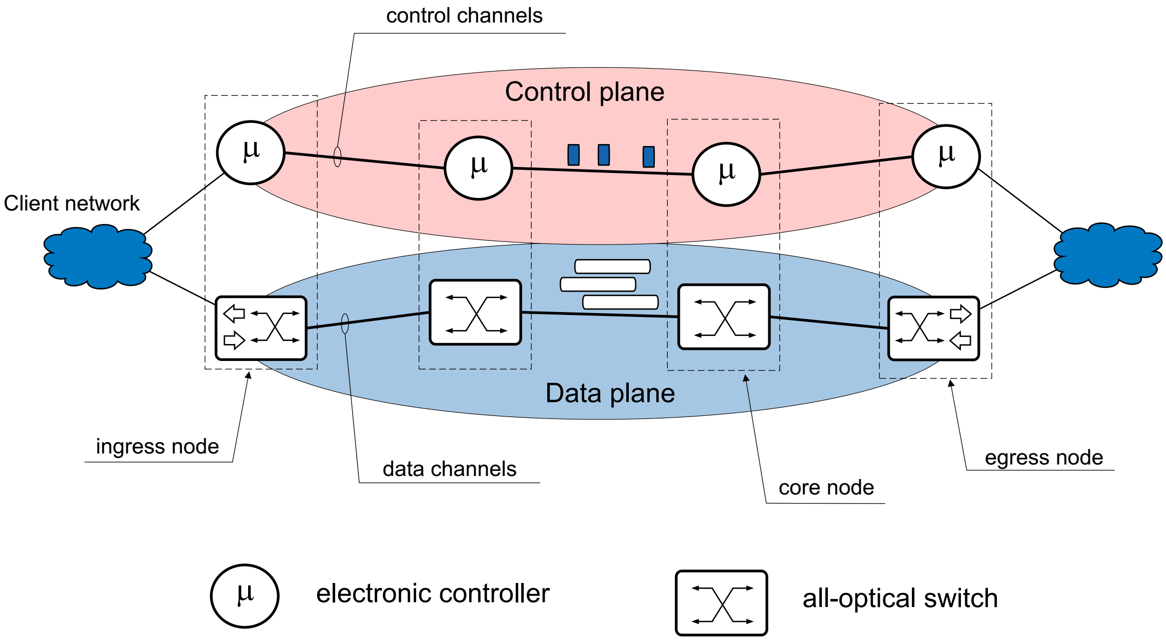


Figure 5.1: Data and control networks of OBS.

the operation in control network has to be analyzed. In particular one has to build a queuing model of control plane taking into account actual system parameters.

In this Chapter we provide a discussion on several factors that have impact on the control plane operation. Moreover, we build two exemplary models of E-OBS control plane which allow as to estimate the delay budget that have to be provided to the bursts in order to achieve certain target burst loss probability.

5.2 Modelling of control plane

Before elaborating a model of OBS control plane one has to identify both the modelling objectives and the model impacting factors. In particular, the fidelity of model depends on the phenomenon one wants to study. Some control-plane stability constraints in OBS (see [WZV02][BD07]) can be obtained with a simple algebra based on basic system parameters. On the other hand a more complex queuing analysis has to be applied when elaborating a model which involves time dependencies.

5.2.1 Control plane impacting factors

There are many factors that influence the OBS control plane operation and performance; below we list the main of them.

- *Network architecture* - depending on the use of either a conventional OBS architecture or an offset time-emulated OBS architecture, or some hybrid solution

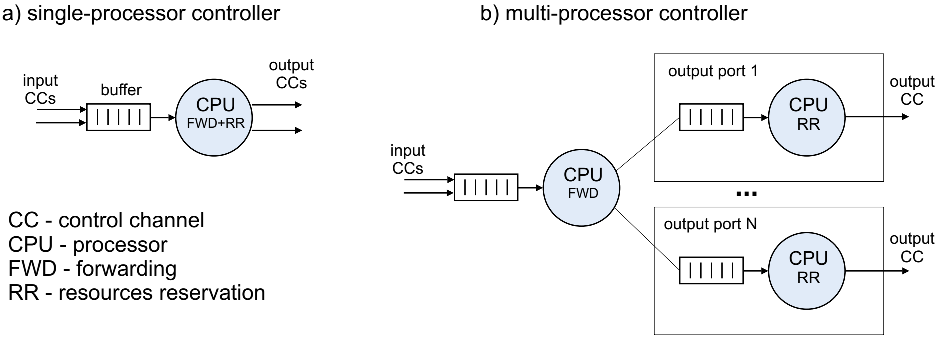


Figure 5.2: Exemplary controller architectures.

the offset time may either vary or do not inside the network. As a result the delay budget of bursts entering the node, in principle, is either variable or fixed.

- *Node controller architecture* - a simple controller can consist of a single processor unit with a buffer handling all the burst control packets in a centralized way. More advanced controllers can use distributed, pipelined, and parallelized operation onto multiple processors (e.g., see Figure 5.2). Such architectures speed-up the processing of control packets.

- *Functions and algorithms* - the main functions performed by the controller processors are: forwarding of burst control packets, resources reservation (with contention resolution and QoS functions) for incoming burst payloads, and configuration of the switching matrix. These functions may be realized with algorithms of different complexity and performance. The algorithm implementation can be either memory-based, where the processing time depends on the seeking time in the memory map, or combinatorial, where the processing time is constant. Both selection and implementation of algorithms influence the service time distribution of the controller.

- *Processing technologies* - several alternatives exist for the processor implementation, starting from relatively slow processors of general purpose, through the *field programmable gate arrays* (FPGA) and *network processors* (NP), to the fastest but also the less flexible *application-specific integrated circuits* (ASIC) (e.g., see [BBE$^+$05]). The first three technologies allow for both memory-based and combinatorial algorithm implementations, while the ASIC may be limited only to combinatorial solutions.

- *Queuing discipline* - either simple first-in, first-out (FIFO) or more advanced disciplines, for instance with ordering the burst control packets according to their offsets, can be used in the buffers.

- *Data plane-related parameters* - the number of both node input/output ports and data wavelengths has an impact on the amount of burst control traffic delivered to the controller.

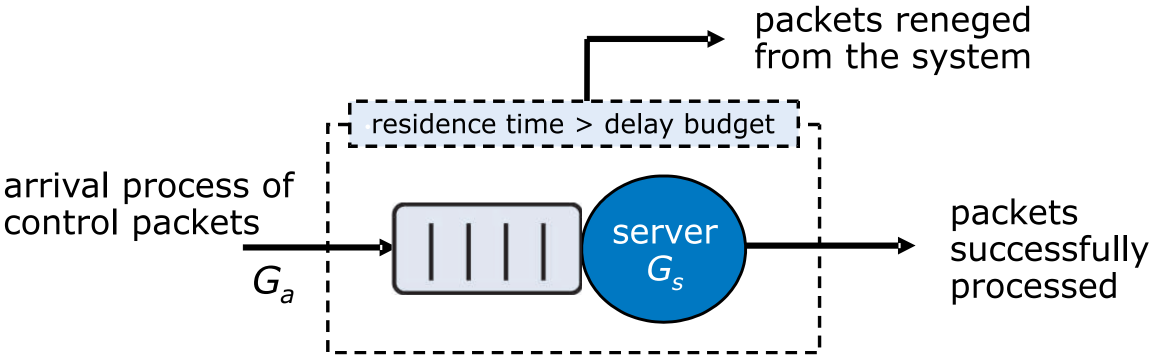


Figure 5.3: General OBS control-plane queuing model.

- *Characteristics of burst control traffic* - the arrival process of burst control packets depends on the burst traffic load, the burst assembly algorithm, in particular on the distribution of both the payload and the control packet lengths, the number of control channels, and the transmission rates in both control and data channels.

5.2.2 A queuing model of OBS switch controller

In general, OBS control network is a network of node controllers connected by control channels. Each controller can be seen as a queuing system. There is some burst control traffic offered to the controller. The arrival process of control packets is closely related to the arrival process of data bursts; therefore according to [IA01] it can be modelled as a Poisson process.

Construction of an accurate queuing model of node controller may be a difficult, if not impossible, task. The controller service time distribution largely depends on its features (as discussed before). In particular, the controller architecture could be represented as a queuing network of buffer-processor systems; some approximation techniques like for instance a two-moment analysis [Whi83] could be applied here.

The operation of OBS controller can be seen as a queuing with reneging [Boc05]. In particular, a burst control packet, when accepted to the queue, leaves the system non-served if its delay budget τ is lower than the effective processing delay R, or (in other words) the residence time (see Figure 5.3). The delay budget is equal to actual offset-time of the burst. In a well-designed system this offset should be long enough in order to reduce the probability of burst losses due to their reneging, $P = p\{R > \tau\}$.

5.3 E-OBS controller with a single processor

5.3.1 Queuing models

We concern on a simple controller with one processor unit and one FIFO buffer, which handle all the burst control packets arriving to the node. The processing times of the

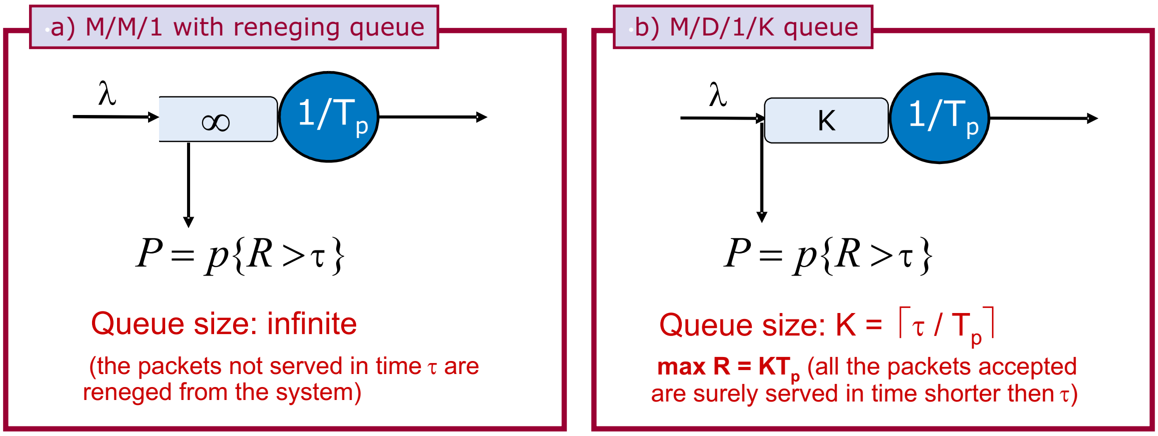


Figure 5.4: Queuing models: a) M/M/1 with reneging, b) M/D/1/K.

	Packet loss probability	**Delay budget**
M/M/1 with reneging	$P = \frac{(1-\rho)e^{\tau(\rho-1)/T_p}}{1-\rho e^{\tau(\rho-1)/T_p}}$	$\tau = \left(\frac{\ln\left(\frac{P}{1-\rho+P\rho}\right)}{\rho-1}\right) T_p$
M/D/1/K	$P \approx \frac{\rho^{\left(\frac{2\tau/T_p-\sqrt{\rho}}{2-\sqrt{\rho}}\right)}(\rho-1)}{\rho^{\left(2\frac{\tau/T_p+1-\sqrt{\rho}}{2-\sqrt{\rho}}\right)}-1}$	$\tau \approx \left(\frac{\left(\ln\left(\frac{P}{1-\rho+P\rho}\right)-\ln(\rho)\right)\left(2-\sqrt{\rho}\right)}{2\ln(\rho)}+1\right) T_p$

Table 5.1: Performance of queuing models.

processor are either *exponentially* distributed (EXP) or *deterministic* (DET), with the mean denoted as T_p. We focus on the E-OBS control operation, thus the delay budget τ of all the bursts entering the node is constant.

Having such a scenario, for each processing time distribution we can consider a different queuing model, respectively:

- for EXP - **M/M/1 queue with reneging** (see Figure 5.4a), where all control packets are accepted to the queue; a packet is lost if period τ expires before the packet is served.

- for DET - **M/D/1/K queue without reneging** (see Figure 5.4b), where control packets are accepted to the queue only if there is free space; when accepted all those packets are served. The system (queue and server) capacity $K = \left\lceil\frac{\tau}{T_p}\right\rceil$ guarantees that all the packets entering the queue are served before period τ. Notice that K gives an upper bound on the packet loss probability of a M/D/1 queue with reneging.

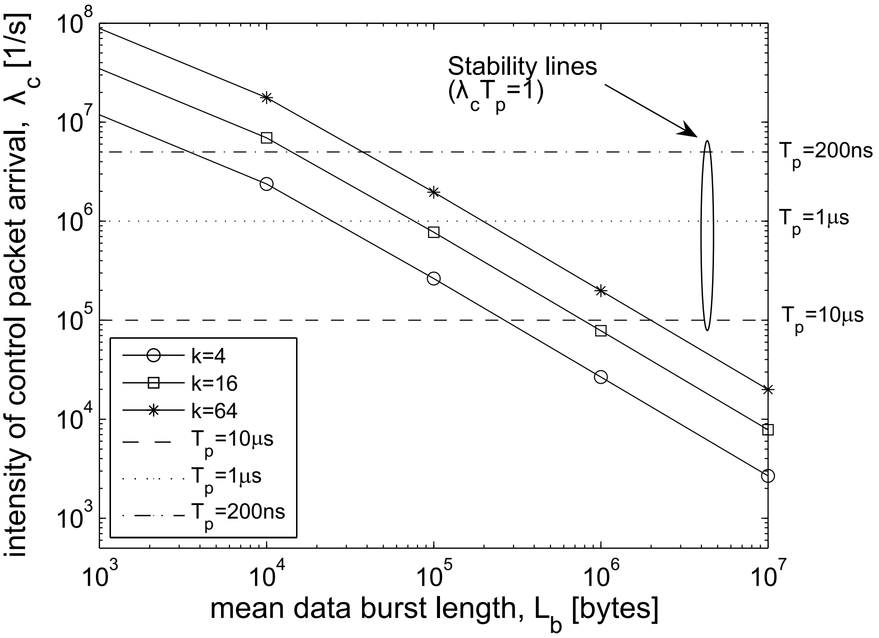


Figure 5.5: Intensity of control packet arrival.

A burst loss probability function $P(\tau)$ and its inverse form $\tau(P)$ are presented in Table 5.1. We use a fine approximation of M/D/1/K queue which was proposed by [SC05], whilst we have exact results for M/M/1 queue with reneging [Bar57]. In the notation, ρ is the processor load ($\rho = \lambda T_p$, where λ is the intensity of control packet arrival). In the case of M/D/1/K queue we will consider τ to be a multiple of T_p ($\tau = KT_p$).

5.3.2 Results

The E-OBS node under study has $N = 4$ input/output ports. The transmission bit rate of data channel is $r_b = 10Gbps$. We consider fast switch operation with the switching time $T_S = 1\mu s$. The analyzed mean processing times are $T_p = \{10\mu s, 1\mu s, 200ns\}$ (as in [BBE+05]). We assume the number of control channels is high enough to carry entire control traffic and to have the packet contention effect in a control channel negligible.

Control-plane stability

In Figure 5.5 we present the intensity of control packet arrival λ_c in the function of average data burst length L_b for the systems with different number of data wavelengths k per port. The burst traffic load ρ_b is such that the target burst loss probability in data plane $P_{Td} = 10^{-4}$; with the Erlang B-loss formula (see 7.1) we find it equal to $\rho_b = \{0.33, 0.49, 0.62\}$ per wavelength, respectively for the system with

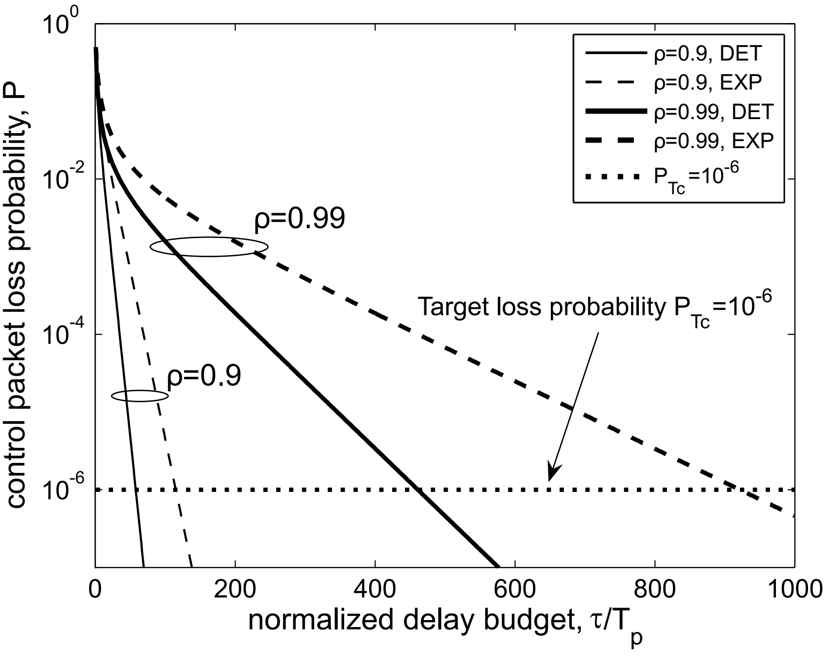


Figure 5.6: Loss probability of control packets.

$k = \{16, 32, 64\}$ wavelengths. As we can observe the intensity of packet arrival increases with the number of wavelengths and is inversely proportional to the burst length.

Moreover for different processing times T_p we plot the boundary $\lambda_c = 1/T_p$ of the control-plane stability constraint $\rho = \lambda_c T_p < 1$ (see [BD07]). Taking this into account, for each pair of k and T_p we can find the minimum average burst length which assure the stability of controller operation. Note that with shorter T_p (what means faster processor operation) this limit can be lowered.

Control-plane loss

We study the impact of delay budget τ (normalized to the processing time T_p) on the loss probability P of control packets, for the system with different processor (controller) load $\rho = \{0.9, 0.99\}$ and processing time distribution (EXP or DET). As we can observe in Figure 5.6, P decreases if either τ increases or ρ decreases. With deterministic processing times we need smaller τ to achieve a certain level of packet loss probability than in case of exponentially distributed processing times; however, this difference is reduced with lower ρ. The dotted line delimit a minimum τ which guarantees a target loss probability in the control plane $P_{Tc} = 10^{-6}$; for instance for EXP and $\rho = 0.9$ such τ is equal to about 100 times of T_p.

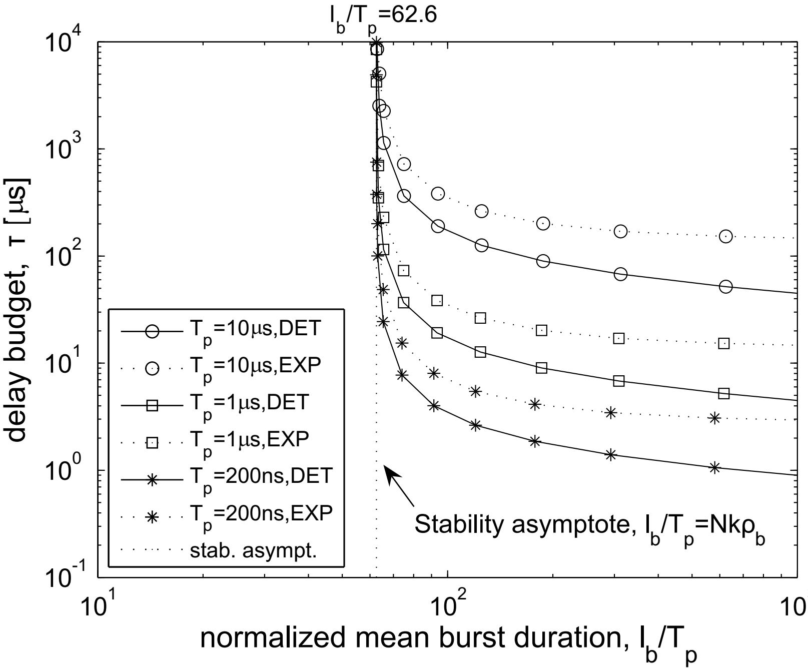


Figure 5.7: Delay budget vs. normalized mean burst duration.

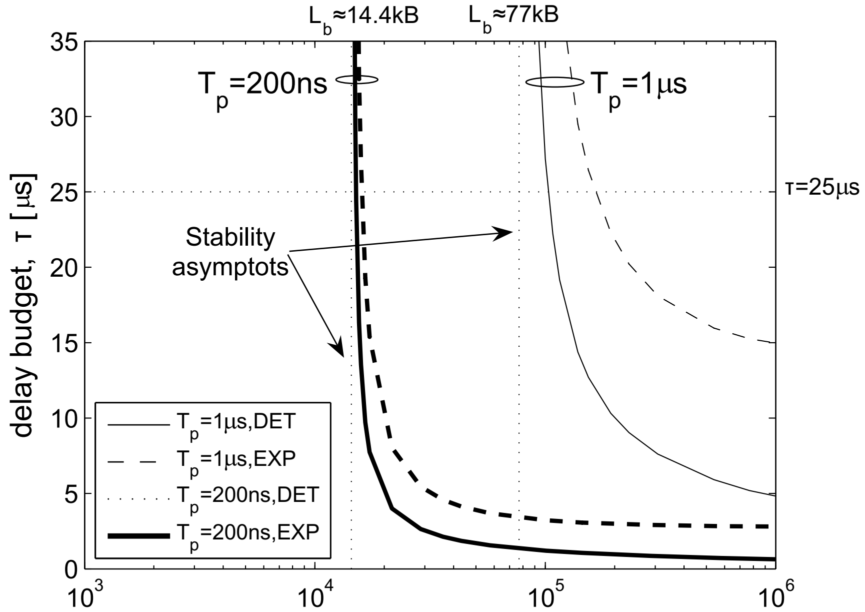


Figure 5.8: Delay budget vs. average burst length.

Delay budget vs. burst size

In Figure **??** we investigate the impact of normalized mean burst duration l_b/T_p on delay budget τ in the system with $k = 32$ (and the total number of data wavelengths $Nk = 128$), different T_p, and target loss probabilities $P_{Td} = 10^{-4}$ ($\rho_b = 0.49$) and $P_{Tc} = 10^{-6}$. We can see that if l_b/T_p approaches the stability asymptote ($Nk\rho_b = 62.6$) we have $\tau \to \infty$ for all curves.

Finally, in Figure 5.8 we plot a reference (dotted) line $\tau = 25\mu s$ corresponding to the offset provided by a feasible fiber delay coil (see [AST$^+$06]). With such target τ we can find a lower bound on average burst length L_b which preserves the system performance. In particular it is about $100kbytes$ under moderate, deterministic processing times ($T_p = 1\mu s$), while in the case of fast processing ($T_p = 200ns$) the limiting value of average burst length is very close to the one determined by the stability constraint.

5.4 Summary

In this Chapter we address the problem of congestion in the control plane of OBS network. In order to approach this issue a queuing model of control plane operation is studied. Since several factors have an impact on the OBS control-plane operation the elaboration of such model may be a difficult task.

We give some preliminary results for an exemplary E-OBS system with a single processor performing in the node controller. Depending on the distribution of processing times we model such system either as M/M/1 queue with reneging or as M/D/1/K queue without reneging. The obtained results show that by appropriate setup of the minimum mean burst length the congestion in control plane can be effectively limited. Moreover for the analyzed system with moderate processing times we show that a feasible fibre delay element can both effectively provide the offset times and concurrently preserve the system performance.

PART III

Quality of service provisioning

Chapter 6

QoS provisioning in OBS networks

The problem of data loss is not uncommon in packet-switching networks. As the network, or even some of its links and nodes, becomes congested, router buffers fill and start to drop packets. Another cause can be the changes of routes as a result of inoperative network links. For non-real-time applications, such as file transfer and e-mail, packet loss is not critical as long as packet protocols provide retransmission to recover dropped packets. However, in the case of real-time information, for instance in voice, video, telemedicine applications, packets must arrive within a relatively narrow time window to be useful to reconstruct the multimedia signal. Retransmission in this case would add extensive delay to the reconstruction and would cause clipping or unintelligible speech as well as discontinuous picture. Packets lost means lost of some information for these services. Discussed questions led to the introduction of different *quality of service* (QoS) classes.

This Chapter addresses the problem of quality of service (QoS) provisioning in OBS networks. The lack of optical memories results in quite complicated operation of OBS networks, especially, in case when one wants to guarantee a certain level of service quality. Indeed, quality demanding applications like for instance real-time voice or video transmissions need for additional QoS differentiation mechanisms in order to preserve them from low priority data traffic. In this context the burst blocking probability metric is perhaps of the highest importance in OBS networks.

QoS differentiation can be provided either with respect to forwarding performance (e.g., the burst loss rate), or with respect to service availability. In the former case, certain quality guarantees are expected during a normal, fault-less operation. The latter case concerns QoS-enhanced protection mechanisms in the resilience problem. In this thesis we focus on QoS differentiation strategies with respect to the forwarding performance.

6.1 Basic concepts of QoS in OBS networks

6.1.1 QoS metrics

Effective QoS provisioning in OBS engages both the definition of specific QoS classes to be given for higher level applications and some dedicated mechanisms in order to provide such classes. In general, each class can be characterized by a specific statistical traffic profile and has to satisfy distinct QoS requirements. In particular, the requirements concern to ensure a certain upper bounds on end-to-end *delay*, *delay variation* (also called the jitter) and *burst loss probability*.

The delays arise mostly due to the propagation delay in fibre links, the introduced offset time, edge node processing (i.e., burst assembly) and optical FDL buffering. The first two factors can be easily limited by properly setting up the maximum hop distance allowed for the routing algorithm. Also the delay produced in the edge node can be imposed by a proper timer-based burst assembly strategy. Finally the optical buffering, which in fact has limited application in OBS, introduces relatively small delays. Since there are many factors that influence the end-to-end data delay in OBS networks the problem of jitter is more complicated and needs a special treatment. This topic, however, is out of the scope of this thesis.

In a well-designed OBS network the data losses should arise only due to resources (wavelength) unavailability in a fibre link. The probability of burst blocking in the link strongly depends on several factors, among other things on the implemented contention resolution mechanisms, burst traffic characteristics, network routing, traffic offered to the network and relative class load. Since a joint relation between these factors is usually very complex in formulation the control of burst losses may be quite awkward in buffer-less OBS networks.

6.1.2 Absolute vs. relative QoS guarantees

There can be distinguished two basic models of QoS provisioning in OBS networks, namely *relative* QoS and an *absolute* QoS. In the former case, the performance of a class is defined with respect to other classes, for instance it is guaranteed that the loss probability of bursts belonging to HP class is lower than the loss probability of bursts belonging to LP class. In the latter case, an absolute performance metric of quality as for example the maximal acceptable level of burst losses is defined for a class. The performance of given class in relative QoS model usually depends on traffic characteristics of the other classes, whilst the absolute QoS model aims at irrelative quality provisioning. The absolute QoS model requires more complex implementations in order to achieve desired levels of quality in a wide range of traffic conditions while at the same time to preserve high output link utilization.

Absolute QoS guarantees are expected by upper level applications. The lack of optical memories, however, complicates the implementation of absolute QoS model in OBS networks, comparing for instance to electrical data networks. For this reason the most of QoS mechanisms considered for OBS networks, basically, offer relative QoS guarantees.

6.1.3 QoS in connection-oriented and connection-less OBS

The problem of QoS guarantees in connection-oriented OBS networks (i.e., with a two-way signalling) is similar to the one existing in dynamic wavelength-switched networks. In particular it concerns providing low establishment delays and low connection blocking probabilities, especially for HP connection requests. The establishment delay specifically is critical in such networks. The reason is that the bursts have to wait in the electrical buffers of the edge nodes until the termination of the connection establishment process. This may produce the buffer overflow and, as a consequence, data losses. When the connection is established there is no data loss inside the network and the transmission delay is only due to the optical signal propagation delay. In this context the connection-oriented OBS operation can provide absolute quality guarantees.

On the contrary, the one-way signalling model needs for additional support in QoS provisioning in order to preserve HP traffic from LP traffic during both the resource reservation process and the burst transmission.

In this thesis we focus on QoS guarantees in one-way signalling OBS networks.

6.2 Categories of QoS mechanisms in OBS networks with one-way signalling

In this section we provide a general classification of QoS mechanisms considered for OBS networks. In most cases, the contention resolution-based QoS mechanisms have a similar application in both OBS and OPS networks. Moreover, OBS possesses its inherent characteristics like for instance the use of pre-retransmission offsets and the ability to operate with different signalling modes. Such capabilities enable implementation of other QoS schemes that are particular only for OBS networks.

In general, several components can contribute to QoS provisioning in OBS networks with one-way signalling (see Figure 6.1). They are related to the control plane operation, through signalling and routing functions, and to the data plane operation both in edge nodes and in core nodes.

6.2.1 Control plane-related mechanisms

Two mechanisms involving control plane operation can provide service differentiation. On one hand, a hybrid signalling protocol (e.g., see [MGK$^+$04]) that consists of a co-operation of two-way and one-way signalling modes can support absolute QoS. In this scenario the establishment of end-to-end transmission paths, by the two-way signalling, can provide the guarantees such as no losses and negligible delays inside the network, while the unreserved resources can be used to transmit the best-effort burst traffic, with the one-way signalling.

On the other hand, similarly like it was proposed in OPS networks (e.g., see [ZCC$^+$04][YMY01]), a routing function can support QoS provisioning. In particular, a properly designed routing protocol may minimize the path lengths for delay-sensitive

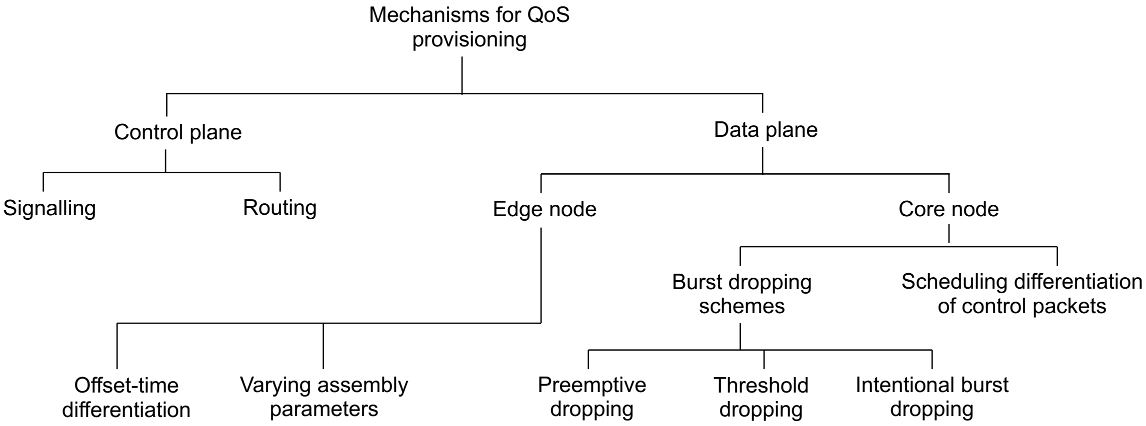


Figure 6.1: Categories of QoS mechanisms in OBS networks.

applications, and even preserve the selection of overloaded parts of the network for loss-sensitive ones, for instance thanks to a deflection routing operation (as e.g., in [LKSG03][KG03a][LYH$^+$06]).

6.2.2 Edge-based mechanisms

Edge nodes are responsible for the burst assembly process so that the incoming client packets are aggregated into bursts in the electronic buffers, according to their class and destination. Solutions like [VJ03], where bursts are unaware class assembled, involve additional complexity and they are used only in particular cases, e.g., together with a burst segmentation mechanism.

Then QoS can be achieved in the following ways:

- *Offset Time Differentiation* [YQ98], which is probably the most addressed QoS technique in OBS networks. The idea here is to assign an *extra offset-time* to high priority bursts, which results in an earlier reservation, in order to favor them while the resources reservation is performed (see Figure 6.2a). The offset time differentiation mechanism allows to achieve absolute isolation between HP and LP classes, i.e., no HP class burst is blocked by a LP class burst. To achieve such feature, however, the extra offset has to be as large, at least, as a few average LP burst durations. The main advantage of this technique is its simplicity; it reduces the loss probability of HP bursts by their postponed transmission from the edge node and no differentiation mechanism is needed in the core nodes. The disadvantages are both the sensitivity of HP class to burst length characteristics [DG01] and extended pre-transmission delay that may not be tolerated by some time-constrained applications. Another problem of the offset-time differentiation mechanism in C-OBS networks is the multiplication of effective classes due to the offset variation [DG01] which may impair the class isolation. In order to reduce this effect a processing offset, which gives a

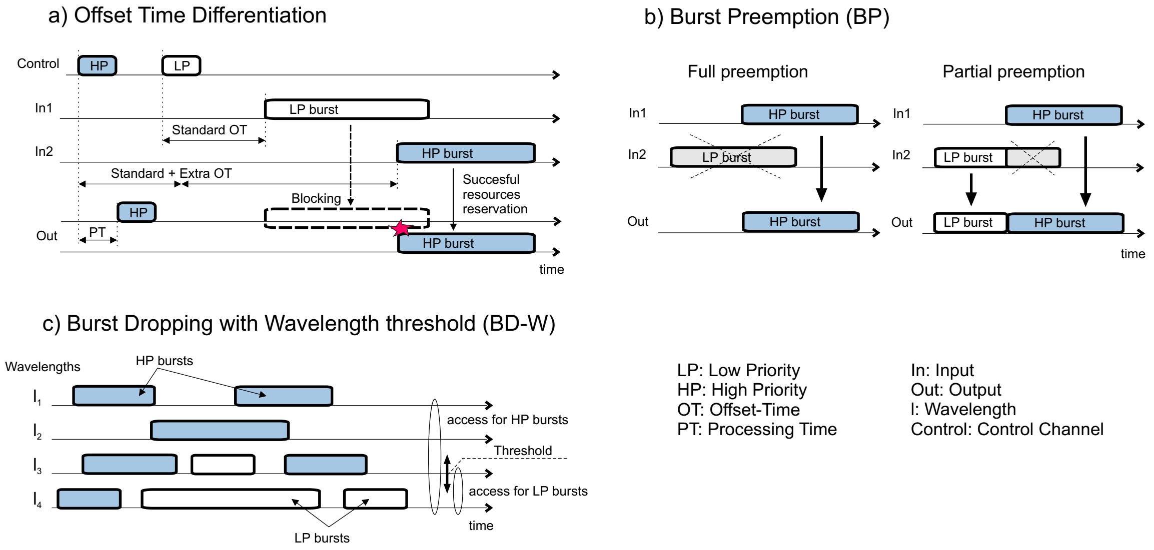


Figure 6.2: Selected QoS mechanisms in OBS networks.

margin to the processing and switching operation in core nodes, should be small enough.

- *Varying burst assembly parameters* like, e.g., preset timers or burst lengths. In particular, the packets belonging to HP class can be aggregated with shorter burst assembly timers than LP packets [BS04]. In this way the latency experienced by HP traffic can be minimized. Also, in the networks with FDL-buffering and void-filling capabilities shorter bursts can improve the BLP performance of a HP class [KCSSP05]. Designing of a burst assembly algorithm is a delicate task as long as the resulting traffic characteristics may influence overall network performance.

Another QoS function at the edge node is the classification of traffic with assignation of specific attributes to the bursts like, e.g., labels, or priorities. These attributes are carried by burst control packets with the purpose of their further discrimination and processing in core nodes.

6.2.3 Core-based mechanisms

First of all, QoS provisioning in core nodes takes place when resolving the contention problem and is achieved with the assistance of a burst drooping technique. The contention resolution usually is supported by some mechanism(s) like wavelength conversion, FDL buffering or deflection routing. The following burst dropping techniques have been proposed for QoS differentiation in OBS:

- *Preemptive dropping*, which in case of the burst conflict overwrites the resources reserved for a lower priority burst by a higher priority one; the preempted, LP burst is discarded (see Figure 6.2b). Several variations of the preemption mechanism can be found in the literature and both relative and absolute QoS models are supported by this technique (e.g., see [OS06][YJJ03]). In general the preemption can be either *full* or *partial*. The full preemption concerns the entire LP burst reservation [KA03] while the partial preemption overwrites only the overlapping part of LP reservation [VJ03]. The partial preemption allows for more efficient resources utilization comparing to the full preemptive scheme. Its drawback, however, is additional complexity in the burst assembly process since this technique requires additional information about the burst data segments to be carried and processed in core nodes.

- *Threshold-based dropping*, which provides more resources, like wavelengths or buffers, to HP bursts than to LP ones according to a certain *threshold* parameter (see Figure 6.2c). If the resources occupation is above the threshold, the LP bursts are discarded and the HP bursts are accepted until there are some resources available. Likewise the OPS networks, in which some threshold-based algorithms have been proposed to be used for the wavelength and FDL buffer assignment problem [CCRZ04], similar solutions can be applied easily in OBS networks [ZVJC04].
- *Intentional bursts dropping*, which maintains the performance objectives of the higher priority bursts on certain levels by intentional dropping the lower priority bursts. As a discarding scheme, a *random early detection* (RED) technique can be used [ZVJC04]. The intentional burst dropping may be classified as an absolute QoS technique.

Another group of mechanisms, which support QoS provisioning in core nodes, is based on queuing and scheduling management of burst control packets that arrive to the node controller. These mechanims make use of the observation that by a proper ordering of burst control packets some reservation requests can be processed earlier; as a result, there is more chances to encounter free transmission resources. Some of proposed burst control packet scheduling mechanisms are taken directly from well-studied electrical packet networks. For instance, in [WR04] the burst control packets are processed according to their priorities, while in [KA04] a *fair packet queuing* algorithm, which regulates access to the reservation manager for different classes of services, is applied. A disadvantage of priority scheduling techniques in OBS networks is extended burst delay. Indeed an additional offset time has to be introduced in order to give enough time for gathering the burst control packets, for the purpose of their prioritized scheduling in the controller.

QoS mechanism	Implemented QoS model	Supported QoS parameter	Advantages	Disadvantages
Hybrid signalling	absolute	delay / burst losses	- absolute end-to-end loss and delay guarantees for HP	- lower statistical multiplexing gain, inefficient usage of bandwidth (less resources available for LP traffic)
QoS routing	absolute (delays) relative (burst losses)	delay / burst losses	- introduces QoS guarantees on network level	- controlling burst losses may be challenging (need the knowledge about network state)
Offset-time differentiation	relative	burst losses	- simple, soft operation - no need for any differentiation mechanism in core nodes	- sensitivity of HP class to burst length characteristics - extended pre-transmission delay
Varying burst assembly parameters	absolute (delays) relative (burst losses)	delay / burst losses	- assembly parameters can be easily setup	- the resulting traffic characteristics may influence network performance
Preemptive dropping	relative / absolute	burst losses	- fine class isolation - improved link utilization in scheme with a partial preemption - absolute QoS can be achieved with a probabilistic preemptive scheme	- overbooking of resources in consecutive nodes (in case of successful preemption) - additional complexity involved in the burst assembly process in case of partial preemption
Threshold-based dropping	relative	burst losses	- can be easily implemented	- the efficiency of bandwidth usage strongly depends on threshold adaptability to traffic changes
Intentional burst dropping	absolute	burst losses	- can provide absolute QoS	- the link utilization may suffer - complex implementation
Scheduling differentiation of control packets	relative	burst losses	- priority queuing in electrical buffers is a feasible and well studied technique	- extended delay (need for longer queuing windows and so larger offset times to perform effectively)

Table 6.1: Characteristics of QoS mechanisms in OBS networks with one-way signalling

In Table 6.1 we summarize the main features of discussed QoS mechanisms.

Chapter 7

Performance of QoS mechanisms in E-OBS

When examining the literature one can find several proposals of QoS mechanisms in OBS networks (see Chapter 6). Usually it is difficult to compare their performance as each mechanism is evaluated in a specific node/network scenario. Nevertheless, a few works provide comparative performance results of selected QoS mechanisms. For instance, in [ZLW$^+$04] Zhang analyzes different QoS provisioning scenarios, which apply either a wavelength threshold-based or an intentional burst dropping principle, with the purpose of absolute quality guarantees. In [VJ03] Vokkarane compares the performance of different QoS schemes with a burst segmentation approach applied. Also the evaluation of different optical packet-dropping techniques in an OPS network scenario is provided in [OS06].

In this Chapter we make an extension to these studies. In particular, we confront the performance of a frequently referenced offset time differentiation (OTD) mechanism with two burst-dropping techniques, namely, with a burst-preemptive dropping (BP) and a wavelength threshold-based dropping (BD-W). In general, all these mechanisms aim at differentiation of burst loss rates in OBS networks that operate with a one-way signalling. We confront the performance of mechanisms in the E-OBS architecture and under a connection-less UDP traffic scenario. Two classes of traffic are considered, namely, a high priority (HP) class and a low priority (LP) class.

7.1 Overview

7.1.1 QoS scenario details

All QoS mechanisms are evaluated in a unified single core-node scenario (see Figure 7.1). There is a number of edge nodes which generate some HP class and LP class burst traffic pattern. The burst traffic is handled in the core node according to a given resources reservation and burst drooping policy. The performance of QoS mechanisms is evaluated using the output burst traffic characteristics at the node output link.

Our core node implements the E-OBS architecture, i.e., the offset times are intro-

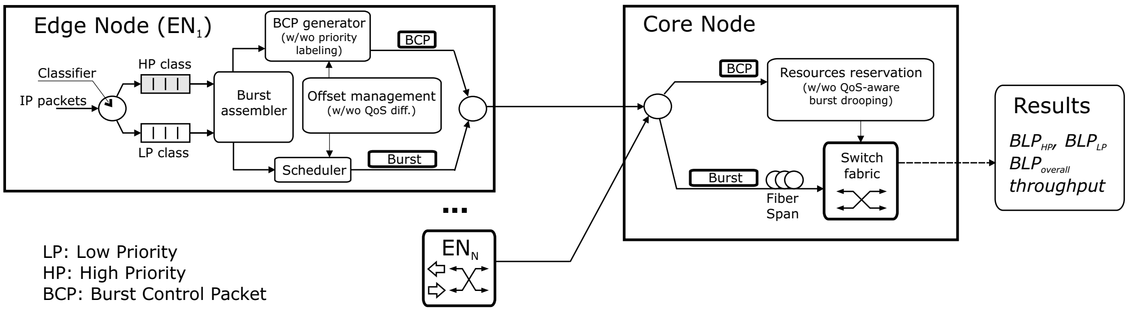


Figure 7.1: Evaluated QoS network scenario.

duced by means of additional fibre delay coils. On the contrary to C-OBS, there is no additional offset, except an optional extra QoS offset, introduced in the edge node between a burst control packet and its data payload. Thus we avoid the impact of variable-offsets on the scheduling operation [LQXX04] and we can get deeper insight into the mechanisms behavior. Since the scheduling operation affects all the mechanisms equally we can expect that their relative performance will be also preserved in C-OBS.

We focus on a (nowadays) technologically feasible OBS core node (e.g., [GWL$^+$05] [AST$^+$06]), which operates with relatively low number of input ports and wavelengths but with fast, sub-microsecond switching matrix and short burst durations.

We consider that in the core node the burst scheduler uses a void filling-based algorithm. Our implementation of the algorithm searches for a wavelength that minimizes the time gap which is produced between currently and previously scheduled bursts. We assume that the searching procedure is performed according to a round-robin rule, i.e. each time it starts from the less-indexed wavelength.

The implementation of QoS mechanisms is as following:

- The size of extra offset time assigned to HP bursts in the offset time differentiation mechanism is equal to 4 times of the average LP burst duration. According to [YQD01] it assures quasi-absolute class isolation.

- We consider a simple full-preemptive scheme when implementing the burst preemption mechanism. Particularly, each HP burst is allowed to preempt at most one LP burst if no free wavelength is available. The preemption concerns a LP burst the dropping of which minimizes the gap produced between the preempting HP burst and the rest of burst reservations.

- The wavelength threshold-based burst dropping mechanism performs according to a *restricted* approach (e.g., see [OS06]). In particular, the threshold value specifies the maximum number of wavelengths that can be simultaneously occupied by LP bursts. On the contrary, the HP bursts are allowed to access the whole pool of wavelengths. The threshold selection problem is discussed in the next subsection.

In both scenarios, the preemptive burst dropping and the wavelength threshold-based dropping, there is a traffic classification function implemented in the edge node, which assignes priorities to the bursts.

The metrics we evaluate are:

- *burst loss probability* (BLP), which corresponds to the amount of data bursts lost in the core node among all the data bursts generated in the edge nodes, and
- *effective data throughput,* later called the *throughput*, which represents the percentage of data volume served with respect to overall data volume offered to the core node.

The burst loss probability is obtained with respect to both HP and LP class traffic as well as to the overall traffic.

In this study we are interested in quantitative comparison of QoS mechanisms more than in the system design or dimensioning. Therefore some of the simulation parameters are setup so that to obtain the evaluation results, especially in case of the HP class performance, in reasonable simulation times.

7.1.2 Simulation scenario

We set up an event-driven simulation environment to evaluate the performance of QoS mechanisms. The simulator imitates an E-OBS core node with no FDL buffering capability, full connectivity, and full wavelength conversion. It has 4×4 input/output ports and 8 data wavelengths per port (if not specified otherwise), each one operating at $10Gbps$. The switching times are neglected in the analysis.

The traffic is uniformly distributed between all input and output ports. In most simulations the offered traffic load per input wavelength is $\rho = 0.8Erlang$ (i.e., each wavelength is occupied in 80%) and the percentage of HP bursts over the overall burst traffic, also called HP class relative load α_{HP} , is equal to 30%.

The burst length is normally distributed with the mean burst duration $L = 32\mu s$ and the standard deviation $\sigma = 2 \cdot 10^{-6}$. In further discussion we express the burst lengths in *bytes* and we neglect the guard bands. Thus the mean burst duration L corresponds to $40kbytes$ of data (at $10Gbps$ rate). The burst arrival times are normally distributed with the mean that depends on the offered traffic load and the standard deviation $\sigma = 5 \cdot 10^{-6}$.

All the simulation results have 99% level of confidence.

7.2 Threshold selection in BD-W mechanism

One of designing problems in any threshold-based mechanism is how to specify the threshold value. The selection of threshold can be supported by an appropriate analysis of the mechanism performance. An analytical queuing model of the wavelength threshold-based dropping (BD-W) mechanism was presented in [OS06]. Based on

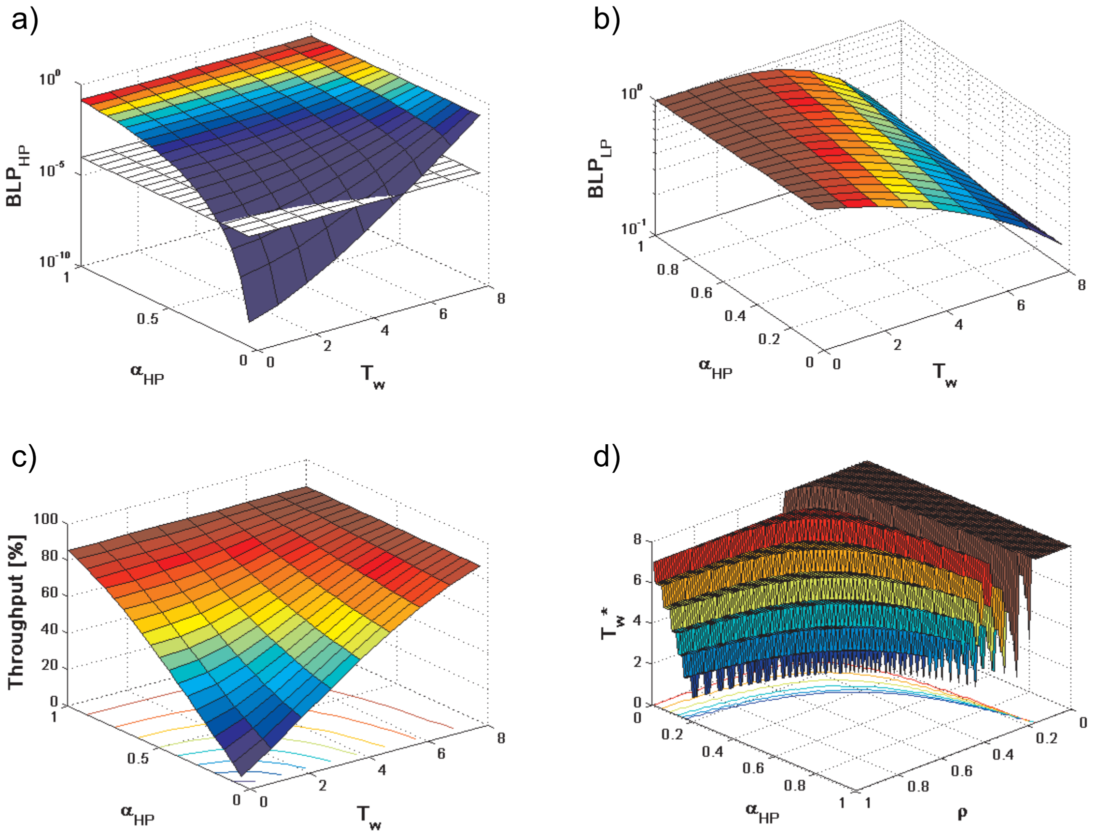


Figure 7.2: Performance of BD-W mechanism ($c = 8$), a) HP class BLP, b) LP class BLP, c) throughput, d) threshold value guaranteeing $BLP_{HP} \leq 10^{-4}$.

this model, we present some performance characteristics of BD-W mechanism as well as we present a method of threshold selection. In our analysis we consider a system with 16 wavelengths, 0.8 traffic load and exponentially distributed burst arrivals and burst lengths. We use T_w to denote the threshold value.

Let $Erl(\cdot)$ be the Erlang's loss formula:

$$Erl(A, c) = \frac{A^c}{c!}\left[\sum_{i=0}^{c} \frac{A^i}{i!}\right]^{-1} \tag{7.1}$$

where A is the offered load and c is the number of wavelengths.

In Figure 7.2 we present exemplary results of HP and LP class burst loss probabilities and the throughput. We can see that the performance of BD-W mechanism depends on both HP relative traffic load (α_{HP}) and threshold value (T_w). For given α_{HP} one can adjust BLP_{HP} by a proper selection of the threshold, however at the cost of effective throughput. On one hand, the minimum BLP_{HP} is delimited by $b_1 = Erl(\alpha_{HP}\rho, c)$, and achieved with $T_w = 0$. In this case the LP class traffic is not served. On the other hand, the maximum BLP_{HP} is equal to $b_2 = Erl(\rho, c)$, and obtained for $T_w = c$. Such a scenario means that no class differentiation is provided by the mechanism.

Assuming that some target burst loss probability BLP^*_{HP} is higher than b_1, we can find some threshold T^*_w that complies $BLP_{HP}(T^*_w) \leq BLP^*_{HP}$ and maximize the throughput. Figure 7.2d presents such evaluation as a function of ρ and α_{HP} for $BLP^*_{HP} = 10^{-4}$ and $c = 8$ wavelengths.

7.3 Performance results

7.3.1 Burst loss probability and throughput

In our study both the offset time differentiation (OTD) and the burst preemption (BP) can be characterized by absolute class isolation. In particular the extra offset time of OTD mechanism assures that the contention of HP bursts is only due to the other HP burst reservations. Therefore, if we assume the exponentially distributed burst arrivals and lengths, the burst loss probability of HP class can be modelled with formula (7.1), and it corresponds to $Erl(\alpha_{HP}\rho, c)$. Similarly, the BP mechanism allows to preempt any LP reservation by a HP one, and the loss of a HP burst happens only if all the wavelengths are occupied by HP reservations. As a result, the loss probability of HP bursts can be expressed as $Erl(\alpha_{HP}\rho, c)$ again. Note that both schemes can successfully transmit LP bursts if either there are free wavelengths, not occupied by any earlier HP reservations in the case of OTD mechanism, or they are not preempted by HP bursts in the case of BP mechanism.

As it was already discussed, the BD-W mechanism achieves the best HP class performance if there is no threshold established ($T_w = 0$); i.e. only HP bursts are transmitted at the output port. In such case, the HP class burst loss probability of BD-W mechanism is the same as of both OTD and BP mechanisms. However, the throughput of BD-W mechanism is very low since none LP burst can be served. As Figure 7.2 shows, by increasing the threshold we can improve the throughput, but at the cost of worsen HP class performance.

In Figure 7.3 we provide comparative simulation results for the scenario described in the previous section, namely for $\rho = 0.8$ and $\alpha_{HP} = 30\%$. The evaluation is performed for different number of data wavelengths (c) in a link. We establish T_w, the wavelength threshold of BD-W mechanism, equal to 50% of c so that LP class bursts can access at most the half of all the available wavelengths simultaneously.

As we can see in Figure 7.3a, by increasing the number of wavelengths in a link we improve the effectiveness of QoS differentiation. The improvement of BLP_{HP} in both OTD and BP mechanism can be so high as, for instance, of three orders of magnitude when having 16 instead of 8 wavelengths. Also, we can see that BD-W offers the worst HP class performance among the evaluated mechanisms.

Comparing BLP_{LP}, the overall BLP and the effective throughput (Figure 7.3b-d), although the shape of OTD and BP performance characteristics is similar, still, the results are in favor of the BP mechanism (see the next subsection for more details). Regarding the BD-W mechanism, we can see that is exhibits very poor performance, which hardly depends on the number of wavelengths. The reason is that BD-W

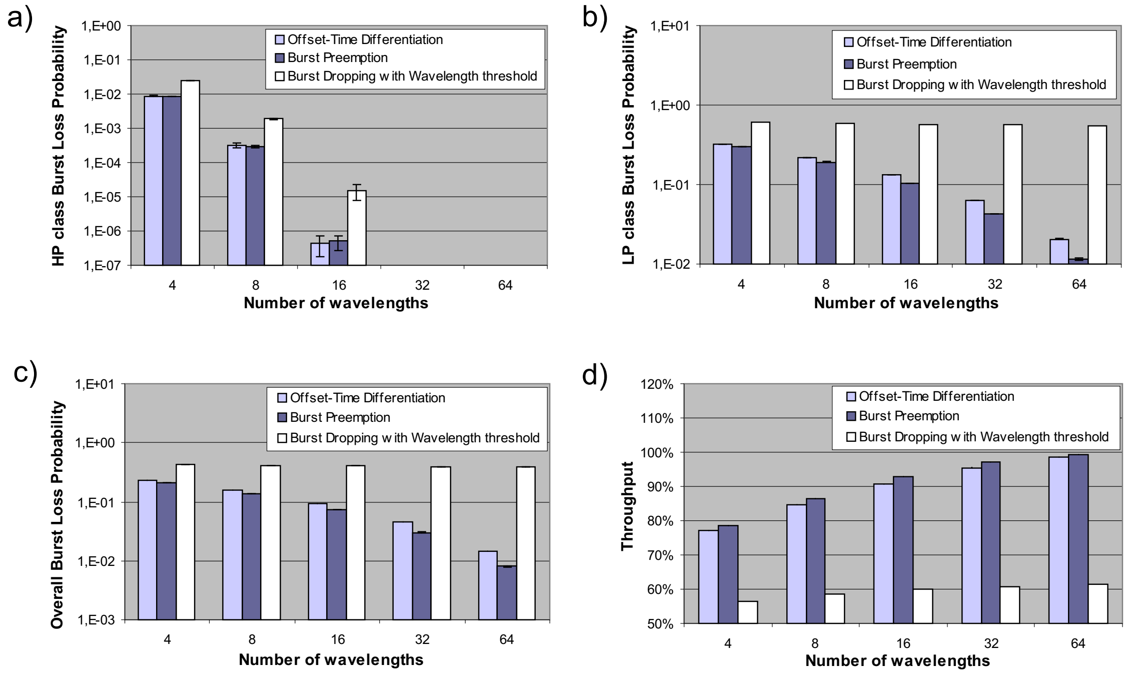


Figure 7.3: Performance of QoS mechanism vs. link dimensioning ($\rho = 0.8, \alpha_{HP} = 30\%$), a) HP class BLP, b) LP class BLP, c) overall BLP, d) effective data throughput.

mechanism has effectively fewer wavelengths available for the burst transmissions in the output link than the other mechanisms. Indeed, BD-W provides only 50% of wavelengths for LP class bursts, while at the same time it attempts to serve the same amount of offered burst traffic. As a result, both the LP class burst loss performance and the throughput are seriously deteriorated.

Although the FDL buffering, in principle, is hardly considered in OBS networks, the utilization of short data bursts may enable its application in the contention resolution problem. The application of FDL buffering should improve the utilization of link resources, and thus the node throughput, as well as it should decrease the loss probability of bursts belonging to each priority class (e.g., see [BNH$^+$03]).

7.3.2 Burst preemption vs. offset time differentiation

The results of HP class burst loss performance presented in Figure 7.3a and Figure 7.4a confirm the correctness of argumentation provided in the previous section. In particular, we can see that BLP_{HP} of both OTD and BP is comparable under any traffic load conditions (Figure 7.4a) as well as with different link dimensioning (Figure 7.3a).

On the other hand, in Figure 7.4b we can see that LP traffic is served more efficiently in BP mechanism than in OTD mechanism. The explanation to this fact can be found in [LQXX04], where it is shown that the scheduling operation may be impaired by the variation of offset-times, the feature which is inherent to OTD

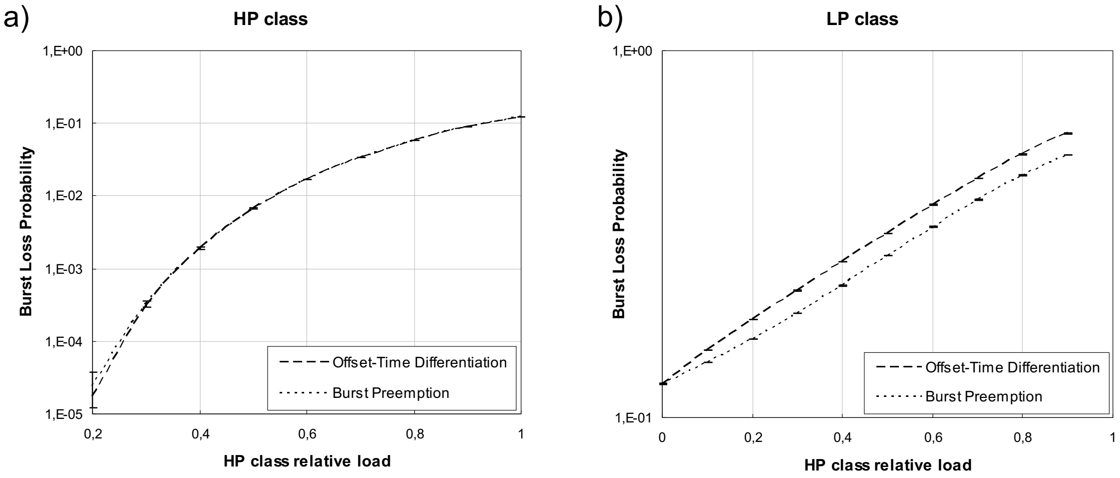


Figure 7.4: Burst loss probabilities vs. HP class relative load in OTD and BP mechanisms ($\rho = 0.8, c = 8$), a) HP class, b) LP class.

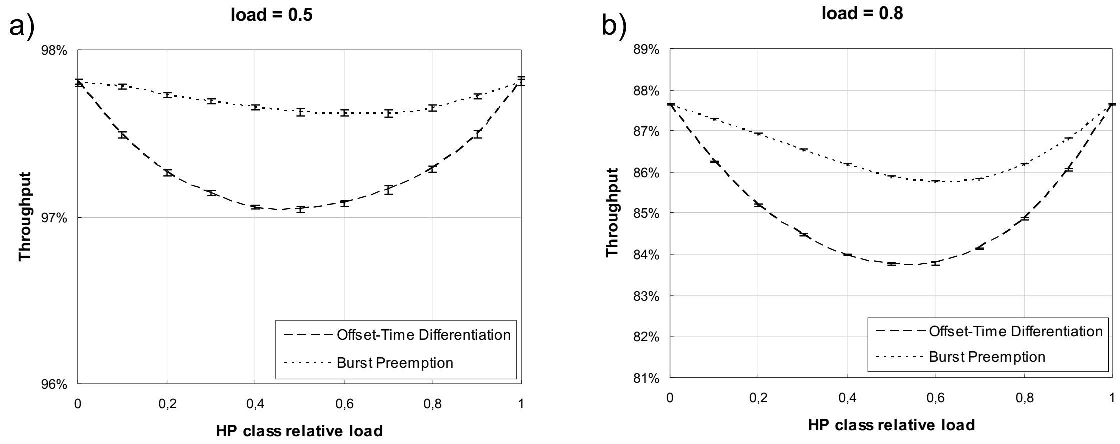


Figure 7.5: Effective throughout vs. HP class relative load in OTD and BP mechanisms, with overall traffic load: a) 0.5, b) 0.8.

mechanism. As Figure 7.5 shows, the use of variable offsets impairs the effective data throughput of OTD mechanism, especially, if the classes are equally loaded. Comparing Figure 7.5a and 7.5b we can see that the aggravation of throughput is more serious in highly loaded nodes.

There is some deterioration of the effective data throughput in the BP mechanism. It is a result of the preemptive operation which allows dropping of a LP burst even if it is being transmitted. Since the front (already transmitted) part of this burst has used some link bandwidth, it also increased the effective traffic load offered to the link; hence, the burst blocking probability increases.

7.4 Summary

In this Chapter we study performance of the most addressed mechanisms providing relative QoS differentiation in OBS networks with one-way signalling. In particular, we show that a burst preemptive mechanism concurrently achieves efficient resources utilization and offers highly-effective QoS differentiation. An offset time differentiation mechanism, which is frequently invoked in the literature, also provides high HP class performance, however, its scheduling efficiency, and so the throughput, is aggravated by the variation of offset times. Finally, a wavelength threshold-based mechanism can be characterized by the poorest overall performance that significantly depends on the threshold value. The application of this mechanism may be reasonable only in the networks with a large number of wavelengths in the link, where the wavelength threshold parameter would be relatively high (in order to serve the LP traffic efficiently) and it could adapt according to the traffic changes. Although the performance of QoS mechanisms has been evaluated in a single node scenario, still, we can expect their similar behavior in a network scenario.

Chapter 8

Effective burst preemption in E-OBS

Burst preemption is one of the most frequently applied techniques to provide burst differentiation in OBS networks. This technique can be used both in QoS provisioning (see Chapter 6) as well as in OBS routing [CZZ04][LY06]. In the previous Chapter we show that in a single node scenario a preemption-based burst dropping mechanism is characterized by high overall performance. Analyzing the mechanism behaviour in a network scenario, however, we encounter the problem of so called *phantom* bursts. The phantom burst is a burst, which was dropped in a node, but its control packet is still travelling through the network reserving the transmission resources. Such situation takes place for instance in the conventional OBS networks with a burst preemption mechanism applied. This effect may lead to the wastage of transmission resources and a superfluous control processing effort. In fact either the overbooked resources in downstream nodes are wasted or an additional signalling procedure should be carried out in order to release them. Moreover the control packets belonging to the phantom bursts burden the controllers of switching nodes unnecessarily.

In this Chapter we estimate the amount of additional signalling necessary to release the overbooked resources in a single buffer-less OBS node. Then we present a novel control mechanism which efficiently applies the burst preemption technique in an E-OBS node without the resources overbooking. Analytical and simulation results prove the effectiveness of our proposal.

8.1 Preemption rate in a buffer-less OBS node

In order to estimate the amount of additional signalling, in this section we calculate a *preemption rate* metric (R) that expresses the amount of preempted bursts over all successfully transmitted bursts at the node output port. Since each preemption would involve a signalling message to release the resources on the ongoing path such metric corresponds well to the signalling overhead produced in a node. In this analysis we consider a full-preemptive scheme, which means that the preemption concerns the entire burst reservation. Moreover we assume the exponentially distributed burst arrivals.

Let $n_{preempt}$ be the number of successful preemptions, $n^{(np)}_{lost\_HP}$ and $n^{(p)}_{lost\_HP}$ be the number of HP bursts lost in a non-preemptive scenario (without burst preemption) and a preemptive scenario (with full burst preemption) respectively, n_{in_HP} be the number of incoming HP bursts, n_{in} be the total number of incoming bursts and n_{out} be the total number of bursts transmitted at the output in a given period of time.

Since each preemption means the acceptance of a HP burst instead of a LP burst, $n_{preempt}$ can be also interpreted as a difference between all the HP bursts lost in the non-preemptive scenario and the HP bursts lost in the preemptive scenario:

$$n_{preempt} = n^{(np)}_{lost\_HP} - n^{(p)}_{lost\_HP}. \tag{8.1}$$

Obviously:

$$n^{(np)}_{lost\_HP} = n_{in\_HP} \cdot B^{(np)}_{HP}, \tag{8.2}$$

$$n^{(p)}_{lost\_HP} = n_{in\_HP} \cdot B^{(p)}_{HP}, \tag{8.3}$$

where $B^{(np)}_{HP}$ and $B^{(p)}_{HP}$ are the HP burst loss probabilities in the non-preemptive scenario and the preemptive scenario.

From the previous equations we obtain:

$$n_{preempt} = n_{in\_HP} \cdot \left(B^{(np)}_{HP} - B^{(p)}_{HP}\right) = \alpha_{HP} \cdot n_{in} \cdot \left(B^{(np)}_{HP} - B^{(p)}_{HP}\right), \tag{8.4}$$

where $\alpha_{HP}$is the HP class load ratio.

Than the preemption rate is equal to:

$$R = \frac{n_{preempt}}{n_{out}} = \frac{\alpha_{HP} \cdot n_{in} \cdot \left(B^{(np)}_{HP} - B^{(p)}_{HP}\right)}{n_{in} \cdot (1 - B^{(p)})}. \tag{8.5}$$

Notice that the overall burst loss probability in the preemptive scenario ($B^{(p)}$) and the HP class burst loss probability in the non-preemptive scenario ($B^{(np)}_{HP}$) are the same. Moreover, $B^{(p)}_{HP}$ depends only on the HP class load due to the absolute class isolation principle. Finally, assuming exponentially distributed burst arrivals, we can use (7.1) to calculate burst loss probabilities. Therefore, by proper substitution we obtain the following estimation of the preemption rate in a node:

$$R = \frac{\alpha_{HP}\left[Erl\left(\rho, c\right) - Erl\left(\alpha_{HP}\rho, c\right)\right]}{1 - Erl\left(\rho, c\right)} \tag{8.6}$$

where ρ, α_{HP}, c, respectively, are the overall traffic load, HP class load ratio and the number of wavelengths in a link, and $Erl(\cdot)$ is given by (7.1).

The numerator of the formula 8.6 indicates the reduction of burst losses of the HP class due to the preemption with respect to the non-preemptive scenario, while the denominator conditions the preemption only to those bursts that are successfully allocated.

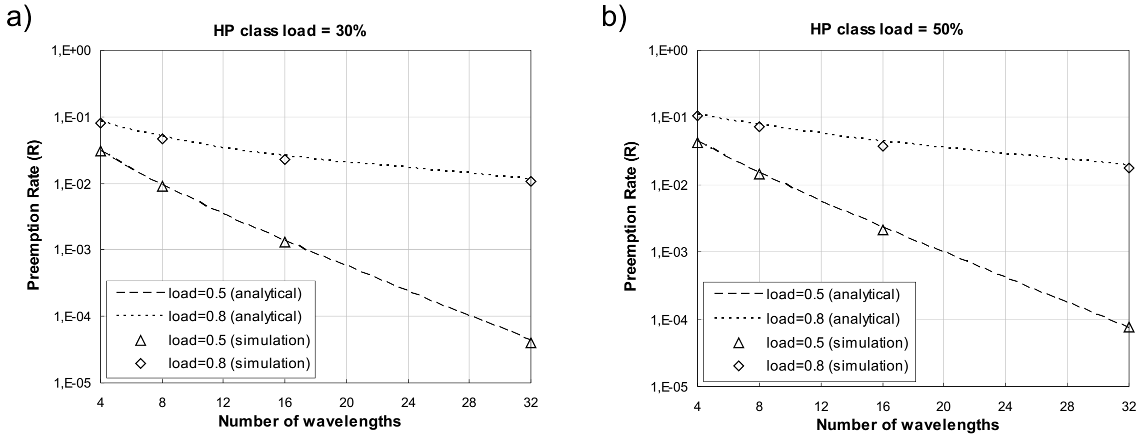


Figure 8.1: Percentage of additional signalling necessary to release preempted burst at each node, with HP class load: a) 30%, b) 50%.

Figure 8.1 presents both analytical and simulation results of the preemption rate in a single node scenario. As we can see, R increases significantly in the systems with lower number of wavelengths as well as at higher traffic loads. A small disparity between analytical and simulation results comes from the fact that the simulated bursts are stream-like arranged in a data channel (bursts do not overlap each other) and their arrivals are not more exponentially distributed.

As we already said, R corresponds to the percentage of additional signalling required (at each node) to release the preempted bursts. If such a signalling procedure is not provided, there is waste of transmission resources due to the phantom reservations in all the nodes on the ongoing routing paths. In large networks, of high number of nodes, the problem might be intensified since all nodes undergo a similar effect.

A particular attention should be paid to preemptive-based routing mechanisms. For instance, a deflection routing mechanism proposed in [CZZ04] as well as a burst cloning mechanism proposed in [LY06] assume that the bursts carried over alternative (duplicate) paths can be preempted by the bursts carried over primary paths. In such a scenario the amount of preempted bursts would be really high as long as both ρ and α_{HP} are high. As a consequence the phantom burst reservations, which occupy the transmission resources unnecessarily, decrease the effectiveness of routing mechanisms.

Evaluation of the phantom burst effect in a network scenario is out of the scope of this thesis.

8.2 Preemption Window (PW) mechanism

In this section we propose a control mechanism that overcomes the problem of resource overbooking due to the burst-preemptive operation. For the purpose of our mechanism we define a time window in which the preemption of LP burst is allowed.

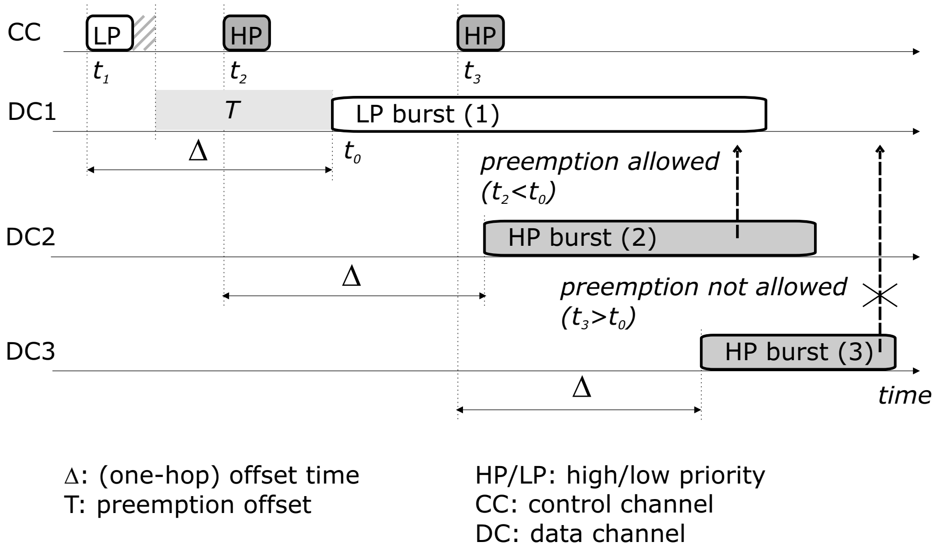


Figure 8.2: Principles of the preemption window mechanism.

The preemption window (PW) mechanism expands look-ahead processing window techniques to the burst preemption context.

8.2.1 Principles

In the PW mechanism a control packet is delivered to the switch controller with some extra offset, besides the processing offset time. This additional offset constitutes a preemptive window T during which the controller can preempt the lower priority reservation by the higher priority reservation.

An important rule of the PW mechanism is that the control packet, after its processing, is waiting in the memory of the controller until T expires and only then it can be sent to the next node (if the burst has not been preempted) or dropped (in case of successful preemption). After the control packet is sent, the preemption of its burst is not allowed in the node. Thanks to these rules each control packet has its corresponding data burst existing in the network (no phantom bursts are present), and there is no need for any signaling procedure to be carried out in order to release the resources on the outgoing path in case of successful burst preemption.

Figure 8.2 shows an illustrative example of the PW mechanism. In this example, a preemption of the LP burst (1) can be performed only by the HP burst (2) since the control packet of the latter burst arrives in preemptive window T. On the other hand, the HP burst (3) is not allowed to preempt the LP burst (1) because its control packet arrives out of window T.

The main advantage of the proposed mechanism is the lack of signalling overhead in case the preemption occurs. Indeed when the control packet reserves resources it already knows that its burst will arrive to the node. There is no resources overbooking in downstream nodes, and so, there is no need to release them. It should be pointed

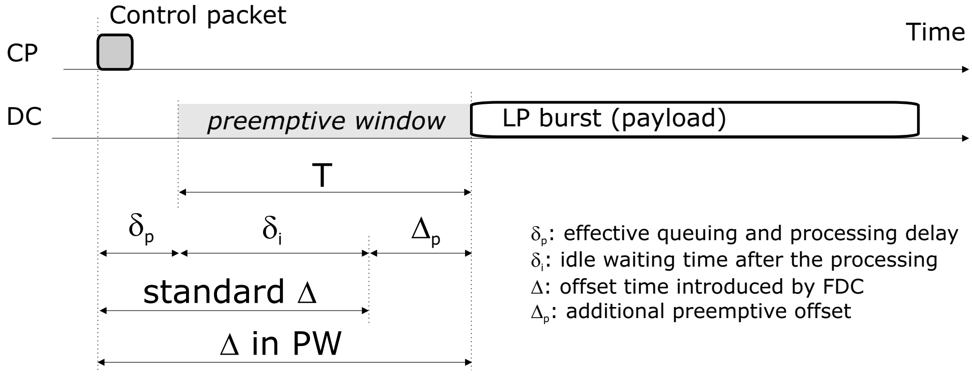


Figure 8.3: The length of preemptive window in PW mechanism.

out that the PW mechanism can work with any, either full or partial, burst preemption principle.

The preemption offset can be provided in both C-OBS and E-OBS architectures. In the former the edge node adds an additional offset, which accounts the preemption windows in all the nodes of the routing path. A disadvantage of this solution is the increase of variation of offset times, which may further intensify the unfairness in access to transmission resources (see Chapter 4). For this reason we consider the PW mechanism more appropriate to be used with E-OBS architectures.

8.2.2 The length of preemptive window

Preemptive window T begins after the end of processing of the burst control packet and lasts until the arrival of its payload (see Figure 8.3). In further discussion, for simplicity, we assume that the payload incorporates a guard band for the switching operation.

Period T can be calculated as:

$$T = \Delta - \delta_p \tag{8.7}$$

where Δ is the offset introduced by inlet FDC in E-OBS node, and δ_p is the effective queuing and processing delay of control packet.

Since δ_p is variable (see Chapter 5 for more details) period T is variable as well. In the simplest case, T corresponds to the idle waiting time period δ_l after the processing of control packet. In order to increase this period, the FDC can add some additional preemptive offset Δ_p. In this case T could be also expressed as:

$$T = \delta_i + \Delta_p \tag{8.8}$$

In the context of burst differentiation, the value of T becomes an important trade-off between high burst delay (too large preemptive window) and ineffective burst preemption (too short preemptive window). Scope of the following sections is to determine the minimum value of T that provides optimal blocking probability.

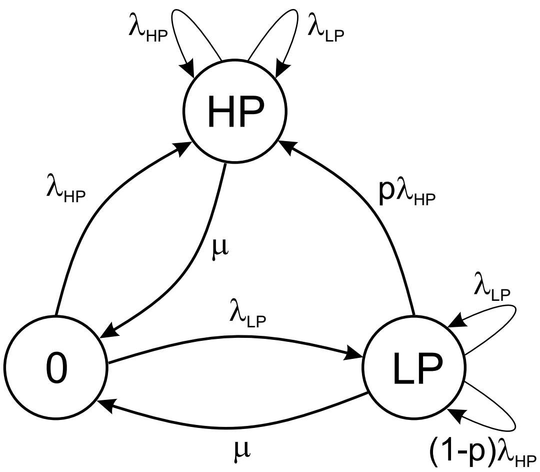


Figure 8.4: Markov chain representing a single-wavelength model of PW mechanism.

8.3 A single-wavelength model of PW mechanism

In this section, we analyze the blocking probabilities of the two classes of bursts, namely a high priority (HP) and a low priority (LP) class, in a single channel system when a full-burst preemption mechanism with PW principle is applied.

According to [IA01] we assume a Poisson process for the HP and LP burst arrivals with rates λ_{HP} and λ_{LP} respectively; the overall arrival rate will be $\lambda = \lambda_{HP} + \lambda_{LP}$. Let t be an i.i.d. exponential random variable which denotes the burst inter-arrival time. Also, let l denotes the burst duration, which follows an exponential distribution with mean value $1/\mu$. Although, we assume the same distribution for both classes, still, in further analysis we use l_{LP} in order to emphasize that we mean the duration of an LP burst.

Let denote the possible wavelength states as: Q_0 for free wavelength, Q_{LP} if occupied by an LP burst, and Q_{HP} if occupied by an HP burst. A corresponding Markov chain, which determines the states' transition, is presented in Figure 8.4.

The equlibrium equations are the following:

$$p\lambda_{HP}P_{LP} + \lambda_{HP}P_0 = \mu P_{HP}, \tag{8.9}$$

$$\lambda_{LP}P_0 = (\mu + p\lambda_{HP})P_{LP}, \tag{8.10}$$

and

$$P_0 + P_{HP} + P_{LP} = 1, \tag{8.11}$$

where P_0, P_{LP}, and P_{HP} denote the probabilities of being in states Q_0, Q_{LP}, and Q_{HP} respectively.

By solving this system of equations we can easy determine the state probabilities

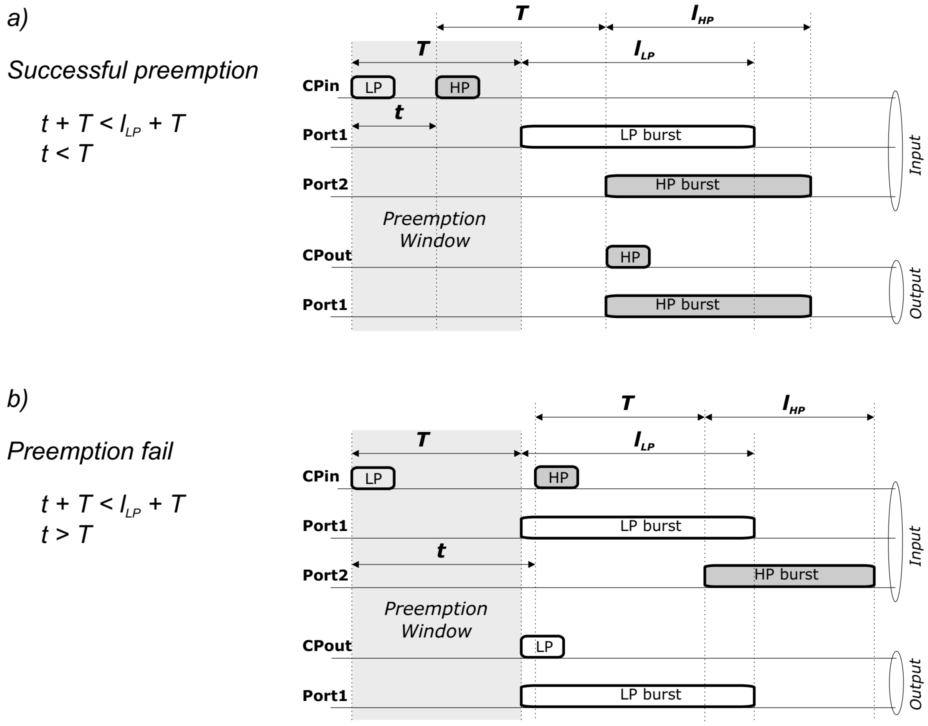


Figure 8.5: (a) Successful preemption and (b) preemption fail cases; the processing times are neglected for simplicity, T is the duration of the Preemption Window, l_{LP} and l_{HP} are the durations of the LP and HP bursts respectively, t is the arrival time of the HP control packet.

$$P_0 = \frac{\mu}{\lambda + \mu}, \tag{8.12}$$

$$P_{LP} = \frac{\lambda_{LP}\ \mu}{(\mu + p\ \lambda_{HP})\ (\lambda + \mu)}, \tag{8.13}$$

$$P_{HP} = \frac{\lambda_{HP}\ (\mu + p\ \lambda)}{(\mu + p\ \lambda_{HP})\ (\lambda + \mu)}. \tag{8.14}$$

Consequently, the burst blocking probability of LP bursts (B_{LP}) and HP bursts (B_{HP}) can be calculated as

$$B_{LP} = P_{LP} + P_{HP} + p\ \frac{\lambda_{HP}}{\lambda_{LP}} P_{LP,} \tag{8.15}$$

$$B_{HP} = P_{HP} + (1 - p)P_{LP}, \tag{8.16}$$

where p is the probability of successful preemption (referred to as $P(Y)$) with respect to all attempts of preemption (referred to as $P(A)$).

Figure 8.5 helps to discriminate successful and failed preemptions. In particular, in Figure 8.5a the preemption of an LP burst can take place since the control packet of the HP burst arrives in the preemption window. The HP burst in Figure 8.5b is not allowed to preempt the LP burst because its control packet arrives out of the preemption window.

Considering the i.i.d. exponential random variable, the probability of successful preemption $P(Y)$ can be calculated as

$$\begin{aligned} P(Y) &= P((t_{HP} < l_{LP}) \cap (t_{HP} < T)) \\ &= \int_0^T \int_y^\infty \mu\, e^{-\mu x}\ \lambda_{HP}\, e^{-\lambda_{HP} y}\, dx\, dy \\ &= \frac{\lambda_{HP}}{\lambda_{HP} + \mu}\ (1 - e^{-(\mu + \lambda_{HP})T}). \end{aligned} \tag{8.17}$$

The probability $P(A)$ is equal to

$$P(A) = P(t_{HP} < l_{LP}) + P((t_{HP} < l_{LP}) \cap (t_{HP} > T)) \sum_{i=1}^{\infty} P(t_{HP} < l_{LP})^i, \tag{8.18}$$

and it represents a first HP burst arrival when the wavelength is occupied by an LP burst ($t_{HP} < l_{LP}$) plus all further HP arrivals (the summation) in case of preemption fail ($t_{HP} < l_{LP}$ and $t_{HP} > T$).

By solving (8.18) we obtain

$$\begin{aligned} P(A) &= \frac{\lambda_{HP}}{\mu + \lambda_{HP}} + \frac{\lambda_{HP}}{\mu + \lambda_{HP}}\ e^{-(\mu + \lambda_{HP})T} \sum_{i=1}^{\infty} (\frac{\lambda_{HP}}{\mu + \lambda_{HP}})^i \\ &= \frac{\lambda_{HP}}{\mu + \lambda_{HP}}\ (1 + \frac{\lambda_{HP}}{\mu}\ e^{-(\mu + \lambda_{HP})T}). \end{aligned} \tag{8.19}$$

The probability of successful preemption p is therefore

$$p = \frac{P(Y)}{P(A)} = \frac{1 - e^{-(\mu + \lambda_{HP})T}}{1 + \frac{\lambda_{HP}}{\mu}\ e^{-(\mu + \lambda_{HP})T}}. \tag{8.20}$$

Taking into account (8.15), (8.16) and (8.20)

$$B_{LP} = \frac{\lambda}{\lambda + \mu} + \frac{\lambda_{HP}\ \mu\ (1 - e^{-(\lambda_{HP} + \mu)T})}{(\lambda_{HP} + \mu)(\lambda + \mu)}, \tag{8.21}$$

$$B_{HP} = \frac{\lambda}{\lambda + \mu} - \frac{\lambda_{LP}\ \mu\ (1 - e^{-(\lambda_{HP} + \mu)T})}{(\lambda_{HP} + \mu)(\lambda + \mu)}. \tag{8.22}$$

8.3.1 Some remarks

Having the burst blocking probabilities for LP class (8.21) and HP class (8.22), we can derive overall burst blocking probability $B_{overall}$ that is given by

$$B_{overall} = \frac{\lambda_{HP}}{\lambda_{HP}+\lambda_{LP}} B_{HP} + \frac{\lambda_{LP}}{\lambda_{HP}+\lambda_{LP}} B_{LP} = \frac{\lambda_{HP}+\lambda_{LP}}{\lambda_{HP}+\lambda_{LP}+\mu} = \frac{\lambda}{\lambda+\mu}. \tag{8.23}$$

As we could expect, the obtained result corresponds to the Erlang loss formula (7.1). Indeed, the PW mechanism does not impair the total blocking probability and even in the case of preemption, when a LP burst is replaced by a HP one, the number of lost bursts is preserved. Also, notice that the formula does not involve T parameter, contrary to both B_{HP} and B_{LP}.

Now, let's check the burst blocking probabilities under the boundary conditions. For $T = 0$ we obtain

$$B_{LP} = B_{HP} = \frac{\lambda_{HP}+\lambda_{LP}}{\lambda_{HP}+\lambda_{LP}+\mu}, \tag{8.24}$$

which is also equal to $B_{overall}$. It is clear, because since $T = 0$ there is no preemption (NP) and the mechanism performs as a poor scheduling mechanism without QoS differentiation.

Now, for $T \to \infty$ we have

$$\lim_{T\to\infty} B_{LP} = \frac{\lambda_{HP}+\lambda_{LP}}{\lambda_{HP}+\lambda_{LP}+\mu} + \frac{\lambda_{HP}\mu}{(\lambda_{HP}+\mu)(\lambda_{HP}+\lambda_{LP}+\mu)}, \tag{8.25}$$

and

$$\lim_{T\to\infty} B_{HP} = \frac{\lambda_{HP}+\lambda_{LP}}{\lambda_{HP}+\lambda_{LP}+\mu} - \frac{\lambda_{LP}\mu}{(\lambda_{HP}+\mu)(\lambda_{HP}+\lambda_{LP}+\mu)} = \frac{\lambda_{HP}}{\lambda_{HP}+\mu}. \tag{8.26}$$

We can see that for $T \to \infty$ both B_{LP} and B_{HP} exponentially approach their asymptotes defined by constant functions of λ_{HP}, λ_{LP} and μ parameters. In particular the second asymptote for B_{HP} could be also derived from the Erlang loss formula with only the HP traffic taken into account. The explanation is that since $T \to \infty$ the length of LP bursts would be almost surely less than T and therefore an HP burst can be blocked only by another HP burst. In this case the mechanism behaves like a classical preemption algorithm (CP), where an HP burst can always preempt an LP burst.

Figure 8.6 presents the characteristics of the discussed model (PW model), validated by simulation results (PW sim). Notice, that the x-axis on the graph is normalized by the mean burst duration $(1/\mu)$ and α is the HP traffic ratio.

The PW model gives a glance on the mechanism's behavior in a single-wavelength system. To complete the study in the next section we provide simulation results of PW mechanism in a multi-wavelength scenario.

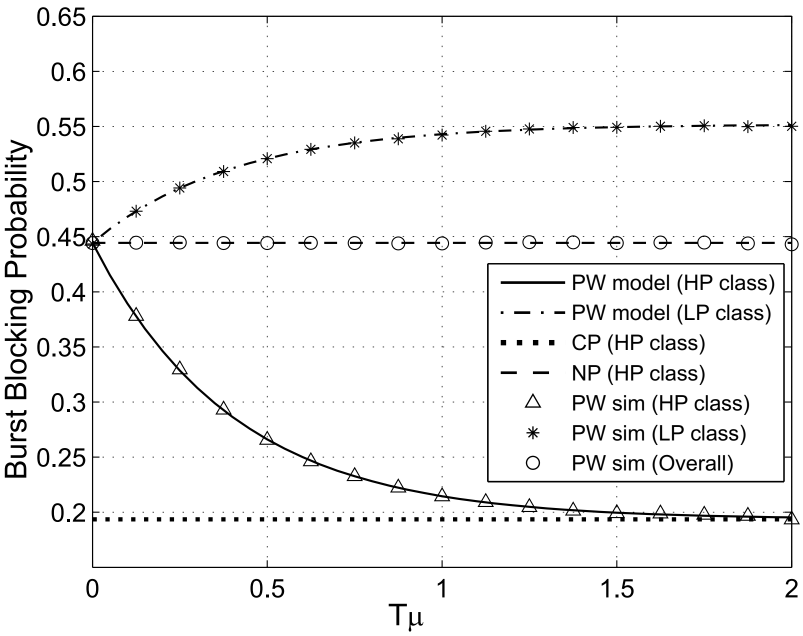


Figure 8.6: Simulation vs. modeling results ($\rho = 0.8, \alpha = 0.3, \mu = 2$).

8.4 Computer simulation of PW mechanism

We use an event-driven simulator to evaluate the performance of a full-burst preemptive mechanism with PW applied. We look for an effective offset, introduced by means of the inlet FDC, which is a trade-off between offering high performance and minimizing the delay. We assume there are two classes of traffic, namely a HP class and a LP class.

8.4.1 Simulation scenario

We evaluate the PW mechanism in a single buffer-less OBS node with full wavelength conversion, 4×4 input/output ports and operating at $10Gbps$. The LAUC scheduling with full-burst preemption is applied. For the purpose of simplicity the processing and switching times are set to 0. The traffic is uniformly distributed between all ports. We consider two traffic models: a general Exponential and a specific Gaussian burst length and inter-arrival time distributions; the latter represents the traffic generated by a hybrid time-length burstifier. Both models use $40kbytes$ ($32\mu s$) as mean bursts length; for Gaussian model we set up the standard deviation to $2\mu s$, and minimum and maximum burst lengths to $4kbytes$ and $4Mbytes$, respectively. The mean burst inter-arrival times depends on the offered load ρ. The HP burst traffic ratio over overall one is denoted as α. All the simulation results have 99% level of confidence.

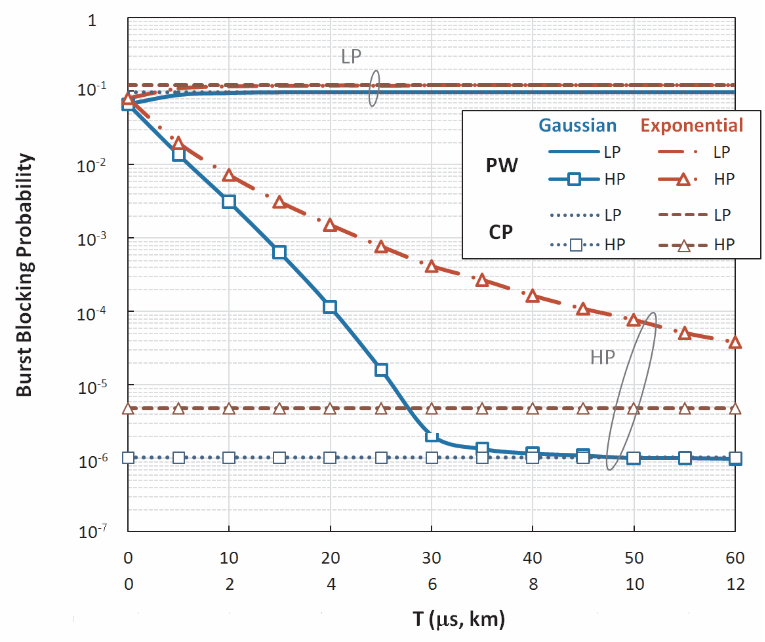


Figure 8.7: Burst blocking probability as a function of T comparing Gaussian and Exponential traffic models ($\alpha = 30\%$, $\rho = 0.8$, $W = 16$).

8.4.2 Numerical results

In Figure 8.7, we firstly compare the Classical Preemption (CP) with our Preemption Window (PW) solution as a function of the delay T. When $T = 0$, there is no possibility of preemption and PW performs as a simple scheduling without burst differentiation. When T increases, HP (LP) burst blocking probability decreases (increases) and approximates to an asymptote, which corresponds to the results obtained with CP. In case of Gaussian traffic, PW quickly reaches the CP performance (T larger than $30\mu s$), while worse results are obtained with Exponential one (T larger than $60\mu s$). This is because the former generates a concentration of burst durations more closed to the length of the fibre delay coil than latter; it has to be underlined that this Gaussian traffic model can be easily obtained well tuning the time/length thresholds of the burstifier [YLC+04].

As Fig. 8.8 shows, burst blocking probability would be further reduced in the systems with more wavelengths. We can discern that for $T \geq 30\mu s$ ($6km$ of FDC) and $W \geq 16$ wavelengths, HP burst blocking probability is less than 10^{-6}.

In Figure 8.9, we analyze the blocking probability as a function of the offered load ρ and of the percentage of HP burst traffic load α. The T window is fixed to $10\mu s$ ($2km$) and 32 wavelengths are considered. We can observe that PW achieves very low HP burst blocking probabilities, e.g. 10^{-5} at $\rho = 0.65$ and $\alpha = 40\%$. Again, PW behaves better when the burst generation follows the Gaussian model.

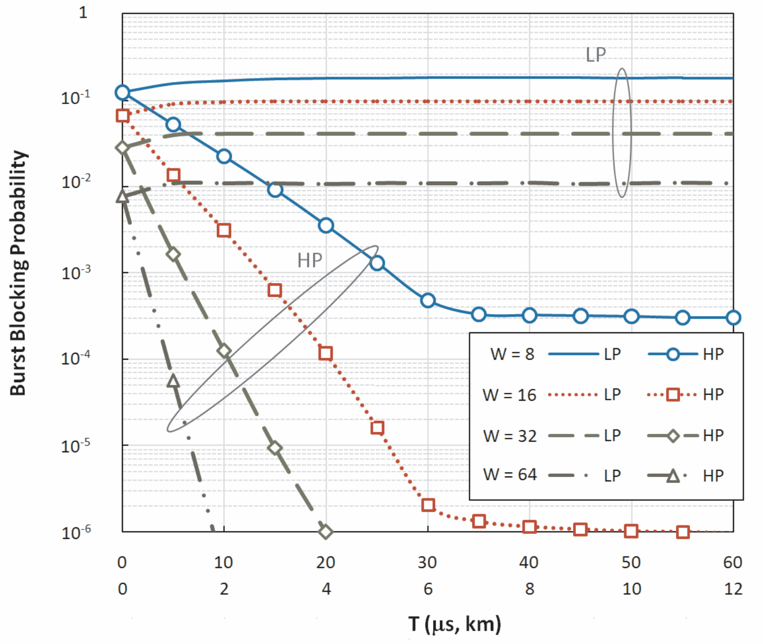


Figure 8.8: Burst blocking probability as a function of T and of W ($\alpha = 30\%$, $\rho = 0.8$, Gaussian traffic model).

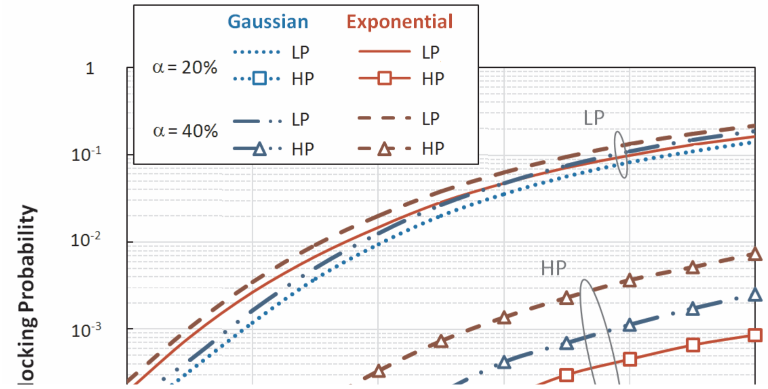


Figure 8.9: Burst blocking probability as a function of ρ comparing Gaussian and Exponential traffic models and different α ($T = 10\mu s$ and $W = 32$).

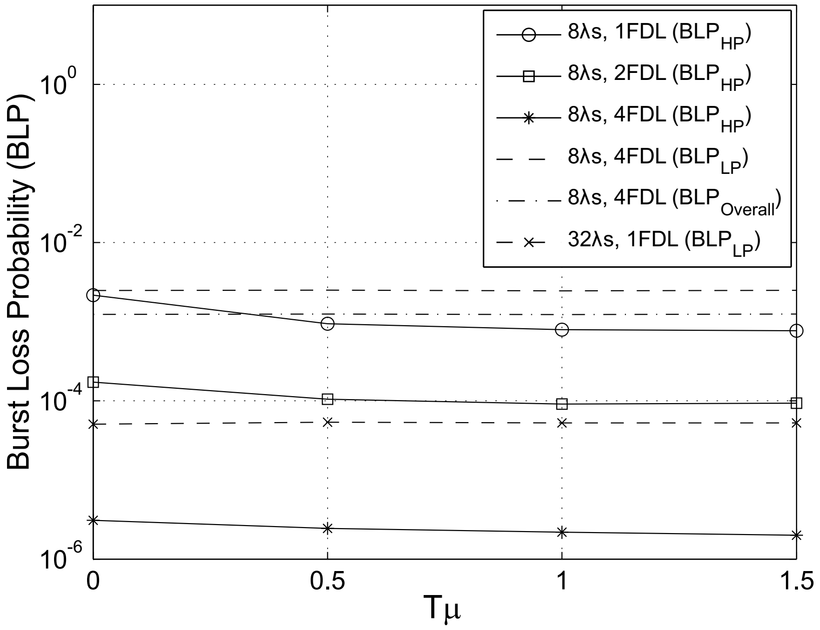


Figure 8.10: Burst blocking probability as a function of T (normalized to $1/\mu$) for different W and FDL buffer size ($\alpha = 25\%$).

8.4.3 PW and FDL buffering

Now we consider that the switching node is enhanced with a feed-back FDL buffer (e.g., as in [HCA98]). The feed-back node architecture allows us to preempt any LP burst even, at the moment, it is transmitted through the FDL buffer. In fact, when a preemption occurs we know that the LP burst has not reached the output port yet, thanks to the PW rule. Hence, we can easily block this burst in the switching matrix (after its looping through the FDL), and thus make impossible its further propagation towards an output link. It have to be noticed that the preemption of a burst which is transmitted through a feed-forward FDL buffer might result in the propagation of a part of optical signal, which has not been blocked by the switching matrix. Since this useless part of the burst would reach the next node it could cause false optical signal detections and therefore additional information such as jam sequence might be required.

In our analysis we assume that the feed-back buffer emulate N output feed-forward buffers, each one operating with 8 optical channels, where N is equal to the number of output ports. The number of delay lines is between 1 and 4 depending on the simulation. The provided delays are linearly increasing with a basic delay unit equal to $32\mu s$, which corresponds to the mean burst duration.

In Figure 8.10 we show the results of BLP for different buffer size and number of wavelengths as a function of T (normalized to mean burst duration $1/\mu$). We see that even with one FDL used there is no significant gain in the performance when increasing T. It is due to the fact that the buffer itself introduces some variable preemption window and therefore no extra preemptive offset in the inlet FDC is necessary. This also explains why, even with T equal to 0, the results of BLP_{HP} are much lower than BLP_{LP} and $BLP_{Overall}$. Therefore the length of input FDC and its consequent delay produced in the node can be reduced. Notice that the controller still has to postpone the transmission of control packets in order to send them together with the corresponding bursts.

Finally, we can observe that the application of FDLs decreases blocking probability of LP bursts, In particular, in the system with 32 wavelengths (λs) and just with only 1 FDL the BLP_{LP} can be below 10^{-4} in a node.

8.5 Summary

The high overall performance of a preemption-based burst dropping mechanism designates it to be a suitable burst differentiation mechanism in OBS networks. However, in this thesis we are concerned on relative quality guarantees, the preemption technique can be extended to absolute QoS provisioning as well. Such a study can be found e.g., in [OS06]. There the superiority of a packet preemptive technique over other packet differentiation techniques, including an intentional packet dropping technique, is demonstrated again.

In OBS the main drawback of a burst preemption mechanism is the overbooking of resources in case of a successful burst preemption.

In this Chapter we propose a dedicated control mechanism, called the preemption window mechanism, which preserves from resources overbooking. PW allows for preemption of a low priority burst only in a specific preemptive window period. Although, our mechanism can be applied in C-OBS, still, it benefits more from the E-OBS control architecture.

Both modeling and simulation results show that the PW mechanism achieves the same performance of the conventional preemptive scheme. The obtained values of the preemption window show the feasibility of its application; e.g., a fibre of about $6km$ is enough when a Gaussian distributed burst traffic model is applied.

Furthermore, the PW mechanism can be used with any other preemptive technique like burst segmentation.

Finally, in the scenarios with FDL buffering there is no need for extra preemption offset in order to obtain QoS differentiation since it is provided by the FDL buffer itself.

PART IV

Routing

Chapter 9

Routing in OBS networks

OBS architectures with no buffering capabilities are sensitive to burst congestion. The existence of a few highly congested links may seriously aggravate the network throughput (e.g., see [KHCSP05]). A *burst loss probability* (BLP) which adequately represents the congestion state of entire network is the primary metric of interest in OBS networks.

The congestion can be reduced either by appropriate network dimensioning or by proper network routing. The dimensioning approach fits the node and link capacities according to the matrix of actual traffic load demands and after such optimization it needs only either a simple shortest path algorithm or a similar mechanism (e.g., see [KG03b][GKZM05]). Some parts of such network, however, may encounter the congestion problem if the traffic demands change. On the contrary, the routing approach introduces some operational complexity since it often needs advanced mechanisms with signalling protocols involved. Nevertheless, the advantage is that it facilely adapts to the changes in traffic demands. Since both presented solutions complete rather than substitute each other, an OBS network should be designed with both a proper link dimensioning and an adequate routing strategy operating inside the network.

In this part of the thesis we address the problem of network routing in the context of burst loss performance and congestion reduction.

9.1 Introduction

9.1.1 Routing terminology

Routing algorithms can be grouped into two major classes: **non-adaptive** and **adaptive** (see Figure 9.1) [Tan88]. Non-adaptive, also called **static**, ones do not base their routing decisions on measurements or estimates of the current traffic and topology, whereas adaptive, or **dynamic**, ones do.

In static routing the choice of the route to use to get from node A to node B is computed in advance, off-line, and downloaded to the nodes when the network is booted. Thus, routing variables do not change during the time. The simplest

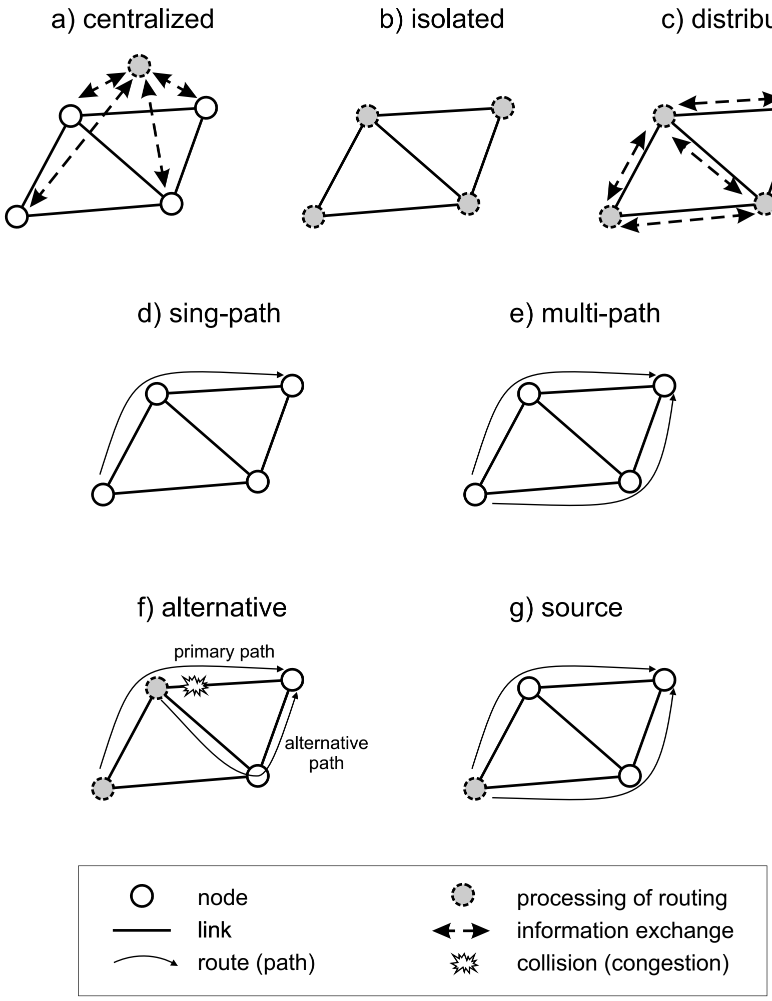


Figure 9.1: Routing algorithms.

technique for static routing is based on a *shortest path routing* algorithm, where the routing objective is to find a routing path of minimum length. The path length in the shortest path routing can be calculated in several ways: the number of hops and the geographic distance are the easiest metrics; for the former we will alternatively use the term *shortest hop routing*.

On the other hand, adaptive algorithms, attempt to change their routing decisions to reflect changes in topology and the current traffic. Adaptive algorithms can be further divided into three families, which differ in the information they use, namely:

- **centralized** (or **global**) - a single entity uses information collected from the entire network in an attempt to make optimal decisions,
- **isolated** (or **local**) - a local algorithm runs separately on each node, which only uses information available there, such as e.g., output link congestion,
- **distributed** - uses a mixture of global and local information.

So far we have tacitly assumed that there is a single path between any pair of nodes and that all traffic between them should use it. Such routing approach is usually called **single-path** routing. In many networks, there are several paths between pairs of nodes that are almost equally good. Better performance can frequently be obtained by splitting the traffic over several paths, to reduce the load on each of the communication links. The technique of using multiple routes is called **multi-path** routing. An advantage of multi-path routing over single path routing is the possibility of sending different classes of traffic over different paths. It can also be used to improve the reliability of the network, in particular, if the routing tables contain disjoint routes.

Alternative routing, often referred to as **deflection** routing, is a special case of multi-path routing. Later we distinguish alternative routing as a technique where all the traffic is sent over a primary routing path. In case the primary path is unavailable for soma period of time a secondary, alternative path is selected.

Another distinction in routing algorithms can be with respect to the place where the routing decision is taken. Whilst most of routing algorithms can perform in each node, in **source** routing only the source makes most or all of the routing decisions. Thus, with source routing the entire path to the destination is known to the sender and is included when sending data. Source routing allows a source to directly manage network performance by forcing data to travel over one path to prevent congestion on another.

9.1.2 Reactive and proactive burst loss reduction techniques

To reduce burst losses *reactive* and *proactive* techniques are applied in the network [TVJ03].

- **Reactive techniques**, e.g., wavelength conversion, FDL buffering, deflection routing, attempt to resolve burst contentions rather than avoid the contentions. Usually, they are based on a local information at the node.

- **Proactive techniques**, reduce the number of burst contentions, by policing the traffic at the source (buffering or dropping data), or by routing traffic in a way that the congestion in the network is minimized. A proactive policing at the source may be controlled by feedback information that indicates congestion in the network.

Most routing-based proactive techniques involve two stages; **route calculation** and **route selection**.

The route calculation can be divided into two categories, namely *static* and *dynamic*. In a static-route calculation, one or more routes are calculated ahead of time, based on some static metric, such are physical distance or number of hops. For instance, paths can be computed using Dijkstra's shortest-path algorithm. In general these static techniques are suitable when the traffic if fairly steady; however they may suffer if traffic is fluctuating over time.

In dynamic route calculation techniques the routes are computed periodically based on certain transient (dynamic) traffic information such as link congestion or number of contentions. Route computation can be performed either centrally in a predestinated node or distributively in individual network nodes.

The information necessary to make the route computation can be obtained in two ways, namely *probe*-based or *broadcast*-based. In the probe-based approach, the source node sends a probe message into the network. The core nodes respond to the probe and return necessary information to the source. A particular case of probe messaging could be a feed-back notification about successful (ACK) or failed (NACK) burst transmission. In the broadcast approach, the core node is responsible for transmitting relevant congestion information periodically to all the edge nodes. The probe can either be sent once for every connection request or periodically based on some interval. The second option is preferable in OBS networks since the duration of data bursts is usually short. In order to reduce the control traffic in the broadband approach, the feedback information can be sent only if there is a change in the congestion status of a link from the previous value.

Once the routes are computed, one of the routes is selected for the data transmission. In single-path routing, the route-selection stage is omitted. In a static route-selection, the traffic is splitted so that its fixed fraction is sent on each of the routing paths. Dynamic route-selection policies are based on feedback information, like in dynamic route-calculation techniques. For each route a given cost function is performed so that the routes are ranked according to their congestion state. Both the traffic splitting vector and the route ranking techniques should react to link congestion states and adopt accordingly in order to shift some part of traffic to less-loaded links.

Stabilization is a significant issue in dynamic route calculation and selection. In particular, multiple sources when reacting to congestion simultaneously, may result in oscillation between congested and un-congested states on particular links. Hence, such effect should be avoided in the network.

9.1.3 Hop-by-hop vs. explicit routing

Routing of data through the network can be performed either *hop-by-hop*, like e.g., in connectionless IP networks, or *explicitly* from source-to-destination, like e.g., in connection-oriented *multi-protocol lambda switching* (MPLS) networks.

- In **hop-by-hop** routing, or *datagram*-based routing, a control packet contains the destination address of the burst, based on which layer 3 forwarding (or routing) is done at every intermediate node.

- In **explicit** routing, or *virtual connection*-based routing, a logical connection, also called *label switched path* (LSP), is set-up first over an explicit physical route. Each control packet carries a label (an LSP identifier), based on which layer 2 forwarding (or switching) is done at every intermediate node. As a result, all bursts sent on an explicit route will follow the path through to the destination. The collection of LSPs between various pairs of nodes essentially forms a *virtual network* on top of the physical fibre network topology.

	Type	**Inform.**	**Routes**	**Deflection**	**Other**
[WMA02]	S	I	S	F	-
[KKK02]	S	I	S	F	-
[HLH02a]	S	I	S	F	-
[CWXQ03]	S	I	S	P	-
[ZVR$^+$04]	S	I	S	F	-
[VJ02b]	S	I	S	F	Q
[CZZ04]	S	I	S	F	-
[LKSG03]	A	C	O	F	Q
[KG03a]	A	C+D	O	T	Q
[LYH$^+$06]	A	C	O	P	Q
[CEJ05b]	A	D	S	R	-
[HAM$^+$05]	A	D	S	T	-

Type: non-adaptive, static (S), adaptive (A)
Information: isolated (I), centralized (C), distributed (D)
Routes: static (S), dynamic, optimized (O)
Deflection: fixed (F), threshold-based (T), probabilistic (P), rank-based (R)
Other: QoS-aware (Q)

Table 9.1: Classification of literature on alternative routing in OBS networks.

Normally, layer 2 forwarding is based on finding an exact match between the label carried by a packet and a label created during the LSP set-up process and accordingly, it is faster than layer 3 forwarding. As well, the simplicity of layer 2 forwarding can facilitate traffic engineering (i.e., intentional distribution of traffic over the network) and end-to-end QoS. These capabilities fit well to both high-speed processing requirements of node controllers and the need for constrained routing, in order to preserve from link overloads, of buffer-less OBS architectures. As a result, the use of *labelled optical burst switching* (LOBS) has been proposed in [Qia00] as a natural control and provisioning solution under the MPLS framework.

9.2 State of the art

9.2.1 Alternative routing

A great part of research on routing problem in OBS networks concerns alternative (or deflection) routing. In alternative routing, when the burst contention occurs, a deflective mechanism reacts to it and re-routes a blocked burst from the primary to an alternative route. Deflection routing can be combined with other burst contention resolution mechanisms (e.g., see [VJ02b][GKS04]).

Routing strategies considered for alternative routing in OBS networks can be either non-adaptive or adaptive.

In non-adaptive alternative routing both primary and alternative routing paths

	Inform.	**Routes**	**Selection**	**Selection Method**	**Other**
[TR05]	C	O	P	O	-
[PMP07]	C	S	P	O	-
[OA05]	D	S	P	H	-
[LMC05]	D	S	P	H	-
[AdDA07]	D	S	P	H	-
[LLGC06]	D	S	P	O	-
[TVJ03]	D	S	R	H	-
[ACP04]	D	S	-	H	Q
[GBIQ04]	D	S	R	H	-
[IYS05]	D	S	R	H	-
[HTM06]	D	S	-	-	-
[YR06a]	D	S	R	H	-

Information: centralized (C), distributed (D)
Routes: static (S), optimized (O)
Selection: static, probabilistic (P), dynamic, rank-based (R)
Selection Method: optimized (O), heuristic (R)
Other: QoS-aware (Q)

Table 9.2: Classification of literature on adaptive multi-path routing in OBS networks.

are fixed (static), and in most cases calculated with the Dijkstra algorithm. A number of alternative paths can be given from a node to the destination. Routing decision is taken in isolation, based only on a local node congestion state information.

Adaptive alternative routing strategies apply a proactive calculation of alternative paths as well as their dynamic selection. The calculation of alternative paths is performed in an optimized way with the assistance of linear programming formulations. These methods need for the information about network topology and traffic demands. In the case of dynamic alternative route selection some heuristics methods are used. In particular, either threshold-based or path-rank (priority) or probabilistic route selection techniques are applied. Dynamic route selection methods need for distribution of some link/node state information between respective nodes. Some of alternative routing strategies support QoS provisioning by routing differentiation with respect to the quality class.

Table 9.1 summarize the key literature on the alternative routing in OBS networks.

9.2.2 Multi-path routing

Multi-path routing strategies in OBS networks aim in adaptive distribution of traffic over a number of routing paths in order to reduce network congestion. Although, some proactive optimization techniques can be found [TR05], still Dijkstra's shortest-path algorithm is the most explored method for pre-calculation of routing paths. In most cases a small number of disjoint SPs with respect to the number of hops is

	Inform.	**Method**	**Inform. type**	**Other**
[HN04]	C	O	T	-
[ZLW$^+$04]	C	O	T	-
[ZWZ$^+$04]	C	O	T	F
[LY06]	C	O	-	-
[TR05]	C	O,H	T	-
[CMC06]	C	O,H	T	F
[OTYC05]	C	H	-	-
[DPZQ06]	C	H	T	-
[Bou03]	D	H	B	F
[HHM05]	D	H	B	F
[GZ06]	D	H	B	F

Information: centralized (C), distributed (D)
Method: optimization (O), heuristic (H)
Information type: topology with traffic demands given (T), broadcasted (B)
Other: failure-recovery implemented (F)

Table 9.3: Classification of literature on adaptive single-path routing in OBS networks.

calculated between each source-destination pair of nodes. In OBS multi-path routing, the selection of routing path is performed in the source. The path selection can be either according to given probability, like in the multi-path routing with traffic splitting, or according to the path congestion rank. Some authors propose centralized, optimization methods for the calculation of traffic splitting vector whilst the others apply distributed heuristic methods. A ranking of the less-congested paths usually is obtained with some distributed heuristics algorithms. In both cases the distributed methods need for updates about the network state information from intermediate/destination nodes to the source nodes. Such signalling messages can be either broadcasted or based on some events, like for instance, the burst dropping event.

Table 9.2 summarize the key literature on the multi-path routing in OBS networks.

9.2.3 Single-path routing

Both non-adaptive (static) or adaptive (dynamic) strategies are considered for single-path routing in OBS networks. Static routing is usually based on Dijkstra's shortest path calculation with respect to the number of hops (e.g., see [YR06b]).

Adaptive single-path routing aims in burst congestion avoidance thanks to a proactive path calculation. The path calculation can be performed either in a centralized or in a distributed way. Centralized (or pre-planed) routing in OBS, in most cases, makes use of the optimization theory with (mixed) integer linear programming formulations. In each case it is supposed that a route computation unit has a knowledge about network topology and (long-term) traffic demands. On the contrary, distributed routing uses some heuristics. Node state statistics are broadcasted, usually in a periodical manner, and used to calculate link weights (costs) in respective nodes. Then a Dijkstra-like calculation is applied in order to find the lowest cost route. Some of adaptive single-path routing strategies support network resilience by the computation of backup paths.

Table 9.3 summarize the key literature on the adaptive single-path routing in OBS networks. The table is structured by horizontal lines according the the main criteria.

9.3 Summary

Dijkstra's shortest path algorithm is the primary routing strategy, frequently explored in OBS networks. Shortest path routing reduces overall network utilization when calculated with respect to the number of hops. On the other hand, some links may be overloaded, while others may be spare, leading to excessive burst losses. Therefore several both reactive and proactive routing strategies have been proposed with the objective of the reduction of burst congestion.

First research studies concerned alternative routing with a static route calculation and selection. Although deflection routing improves network performance under low traffic conditions [WMA02], still, it may intensify the burst losses under moderate and high loads [ZVR$^+$04]. Indeed the problem of alternative routing in buffer-less OBS networks is over-utilization of link resources, if an alternative route has more number of hops than a primary path. Therefore in the next step the objective was to optimizate the set of alternative routes as well as to introduce some adaptive path selection techniques (see Table 9.1). Assigning of lower priorities to deflected bursts with their possible preemption is another important technique, which preserves from excessive burst losses on primary routes [CZZ04].

Multi-path routing represents another group of routing strategies, which aim in traffic load balancing in OBS networks (see Table 9.2). Most of the proposals are based on a static calculation of the set of equally-important routes with Dijkstra's algorithm. Then the path selection proceeds adaptively according to some heuristic or optimized cost function. Both traffic splitting and path ranking techniques are used in the path selection process.

The issue related to any multi-path routing is the problem of out-of-order burst arrivals. The burst reordering is common for both multi-path and alternative routing scenarios, in which the routing paths differ with respect to a physical distance. To cope with this problem either some dedicated mechanisms (e.g., see [GBIQ04][LMC05] [PMP07]) or single-path routing have to be used.

Network congestion avoidance in single-path routing is achieved thanks to a proactive route calculation. Since most of the strategies proposed for OBS networks consider a centralized single route calculation some authors study distributed routing algorithms (see Table 9.3). Both optimization and heuristic methods are used. Moreover, several works address the problem of network resilience and failure recovery.

Chapter 10

Isolated alternative routing strategies for labelled E-OBS networks

In this Chapter we propose two isolated alternative routing algorithms. Our objective is to find the algorithm that at the same time can be easily implemented in a connection-oriented, labelled E-OBS network and improves the overall burst loss performance. As a reference we use a simple shortest hop routing (SPR) algorithm. The evaluation is performed in an event-driven simulator environment. All simulation results have 99% level of confidence.

10.1 Scenario under study

Network architecture

We consider an OBS network with one-way signalling, Horizon resources reservation, LAUC burst scheduling and E-OBS architecture. The application of E-OBS facilitates routing management. In particular, there is no constraint on the length of an alternative path as well as the offsets do not have to be computed in source nodes in advance but they are introduced accordingly in immediate core nodes. Notice that since the offset time is fixed in corresponding E-OBS nodes there is no need for a void filling-based burst scheduling algorithm.

Each network node is both an edge node and a core switching node capable of generating bursts destined to any other nodes. In the analysis we assume that the source nodes do not buffer the bursts after completing their aggregation. Also, the nodes are not enhanced with FDL buffers.

Number of data wavelengths c is the same for each link and equal to $c = \{32, 64\}$, depending on the scenario. The transmission bitrate of data wavelength is $10Gbps$.

	SIMPLE	**NSFNET**	**EON**
number of nodes N	6	15	28
number of links K	8	23	41
minimum node degree	2	2	2
maximum node degree	4	4	5
average node degree	2.67	3.07	2.928
minimum link length $[km]$	500	247	218
maximum link length $[km]$	500	2831	1500
average link length $[km]$	500	1022	625
network diameter (hops)	3	4	8

Table 10.1: Network topologies

Network topologies

Our routing strategies are evaluated with three logical network topologies (see Figure 10.1):

- the SIMPLE mesh network topology,
- the NSFNET network topology, which represents an American backbone network [Nsf], and
- the EON network topology, which is a pan-European network defined in European COST 266 action [RI03].

The SIMPLE network has 6 nodes and is the smallest network. On the other hand the EON network (28 nodes) is the largest network. The number of nodes (N) and links (K) in the NSFNET network (15 nodes) can be placed in-between. Average node degree ($2K/N$) is approximately the same for both the NSFNET and the EON networks. Maximum and average link length is significantly larger in the NSFNET network compared to the other two networks. The NSFNET network contains both rather short and very long links.

Network diameter, which is the maximum distance between node pairs based on the number of hops, is a good indicator for the amount of through traffic in network nodes.

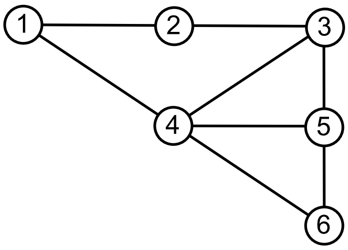
The details on the topologies can be found in Table 10.1.

Traffic model

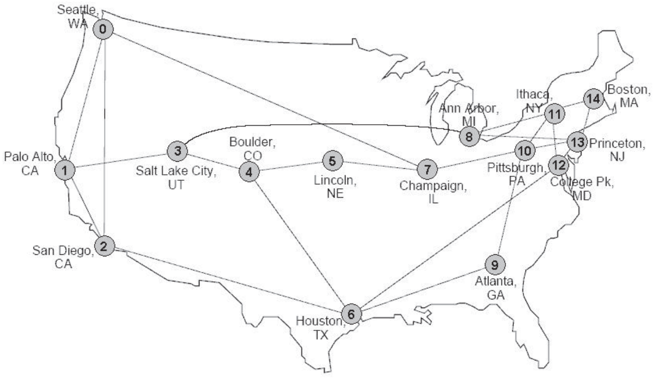
The traffic is uniformly distributed, i.e., the following matrix of demands $\mathbf{T}$ is defined:

$$\mathbf{T} = \frac{c\rho}{N-1}(\mathbf{E} - \mathbf{I}). \tag{10.1}$$

a) SIMPLE network



b) NSFNET network



c) EON network

Figure 10.1: Network topologies; a) SIMPLE, b) NSFNET, and c) EON.

where N is the number of network nodes, c is the number of wavelengths in the network link, ρ is the traffic load offered to edge node normalized to the link capacity, $\mathbf{E}$ is the unit matrix, $\mathbf{I}$ is the identity matrix, and all the matrixes has dimension $N \times N$. In other words, between each pair of source-destination nodes there is a traffic offered. The volume of traffic is equal to the amount of traffic load entering the edge node divided by the number of corresponding destination nodes.

We consider a Poisson arrival process for generating bursts with exponentially distributed lengths. As several authors already observed it (e.g., see [CEJ05b]), the length distribution does not have a significant effect on the results in buffer-less OBS networks.

Route calculation and selection

We assume the routing paths are calculated according to Dijkstra's shortest hop algorithm. In all studied routing algorithms we consider that there are k pre-established LSPs between all source-destination pairs of nodes available. The routes are not necessarily disjoint.

An LSP selection is performed according to a given routing algorithm. Particularly, isolated alternative routing allows to select an LSP, from the set of all available LSPs, in any network node. We consider per-burst routing decision.

10.2 Algorithms

We propose two isolated alternative routing algorithms, namely, a **path excluding routing** (PER) algorithm and a **by-pass routing** (BPR) algorithm. Each algorithm performs a deflection of transmitted data burst from a primary to an alternative routing path if there are no transmission resources available on the primary path. The routing decision is taken only using local (isolated) output link state information.

Path excluding routing algorithm

In PER algorithm the edge node selects the first available path from the set of paths to the destination. This selection determines the next hop and excludes from the set of available paths all those paths that not include this hop in their route. Hence, from the k original paths, each node removes some paths as long as there remains only one path.

Figure 10.2a-b show an example. A burst is generated in node A and destinated to node E. $k = 3$ paths are setup: the shortest ones are 1. $A - D - E$, then 2. $A - B - C - E$ and 3. $A - D - F - E$. If the first (shortest) path from the list is congested on its output port, A selects the $A-B-C-E$ path definitely excluding the other possibilities. This means that the rest of the nodes in the selected path cannot take other routing decisions. If the output port of A toward D is not congested, both $A - D - E$ and $A - D - F - E$ are selected while the other is removed. The next node D will take the path decision in the same way. If the output port of D toward

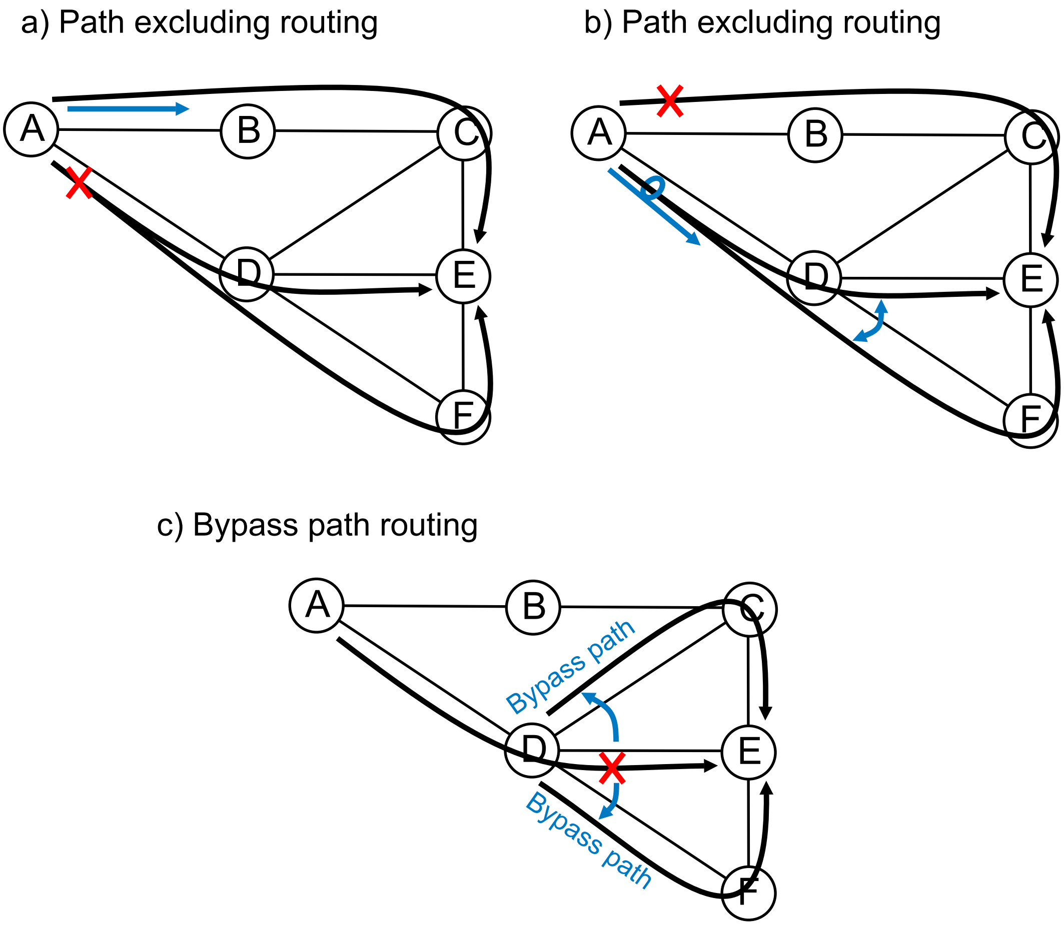


Figure 10.2: Isolated alternative routing algorithms: a) PER, and b) BPR.

node E is not congested, it chooses the path $D-E$; otherwise $D-F-E$ is selected. It is evident that when all output ports are congested, the burst is lost.

Bypass path routing algorithm

In BPR algorithm, for each burst, the source node selects a single path as a function of the state of its output queues. The route can be modified only when travelling burst finds a congested link. In this case, the node tries to *by-pass* it using the shortest available path to the destination.

Figure 10.2c shows an example of this algorithm behavior. Node A transmits a packet/burst to node D with destination node E (the path is $A-D-E$). When burst arrives to node D, no resources are available to reach node E. Therefore, node D finds two by-pass paths in its forwarding table: $D-C-E$, and $D-F-E$. It selects the first available one.

10.3 Results

Our isolated alternative routing algorithms are evaluated in the network scenarios described in Section 10.1; in particular, each link has $c = 32$ data wavelengths in SIMPLE and NSFNET network, and $c = 64$ wavelengths in EON network. We consider the scenarios with $k = \{2, 4, 6\}$ LSPs between each pair of nodes available in SIMPLE and NSFNET networks, while there can be $k = \{2, 4, 6, 8, 10\}$ LSPs in EON network. The offered traffic load ρ is normalized to the link capacity.

In Figures 10.3 and 10.4 we present the impact of the number of available paths (LSPs) on overall BLP performance under PER and BPR routing strategies respectively.

Firstly, we can see that both PER and BPR outperform SPR under low and moderate traffic loads in each scenario. Moreover, the efficiency of PER under high loads ($\rho > 1$) can be still better than of SPR, whilst BPR has a worsen performance. These results are consistent with the conclusions presented in [ZVR$^+$04]. Particulary, BPR algorithm does not have any limits on the number of deflections performed and it can increase the network load, and so the burst blocking, significantly. On the other hand, the number of deflections in PER is limited, at most, to the number of available paths k. The network is hardly overloaded in such case.

The next conclusion is that more LSPs improves the network performance. It is obvious since there are more possibilities to perform the deflection in case of unavailability of resources on primary paths. The improvement of performance can be really high under BPR strategy and in smaller networks (see Figure 10.4a). BPR with high number of LSPs available behaves like a hot-potato routing; the burst can be sent even to the previous node (loops possible). Nevertheless, the selection of the set of LSPs should be reasonable in order to preserve from the use of too-long paths, as e.g., in Figure 10.4c), where the performance with $k = 10$ is worsen than with $k = 8$.

When comparing the routing strategies we can see that BPR offers better performance than PER (except high-load traffic conditions). Again, it is clear since BPR has more chances for a successful deflection in every intermediate node.

In Figure 10.5 we investigate the distribution of the number of hops the burst, which is successfully delivered to the destination, experiences with BPR in SIMPLE and NSFNET network scenarios. As a reference we provide the similar distribution obtained with SPR; the maximum number of hops here is 3 and 4 for SIMPLE and NSFNET respectively. We can see that BPR can increase the length of the burst routing path significantly, especially, under higher loads (compare Figure 10.5b with Figure 10.5c). On the other hand, under PER strategy the maximum burst routing path in the network is limited by the length of the longest LSP (it results from the behavior of routing algorithm).

10.4 Summary

Isolated alternative routing performs the route selection in consecutive nodes based on local node state information (e.g., link occupancy, available wavelengths), i.e. each

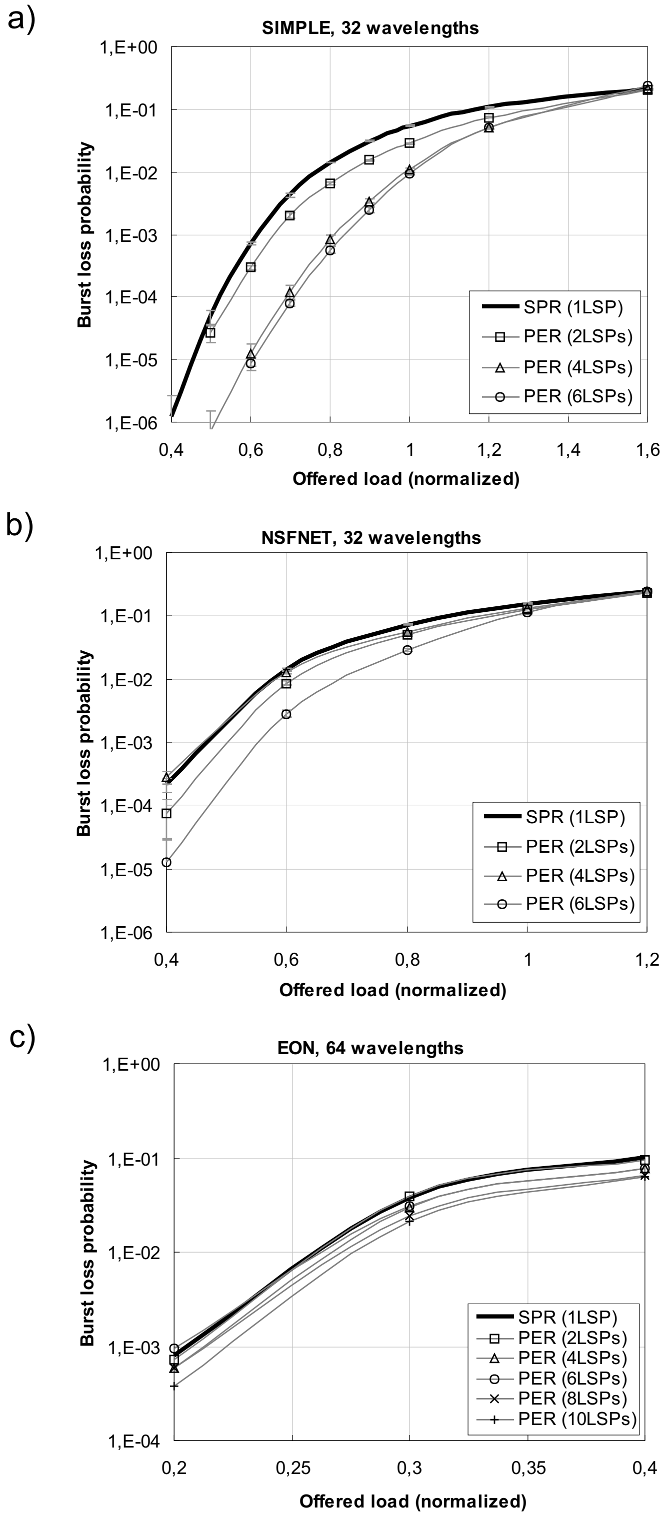


Figure 10.3: Burst loss probability in PER, a) SIMPLE (32λ), b) NSFNET (32λ), and c) EON (64λ).

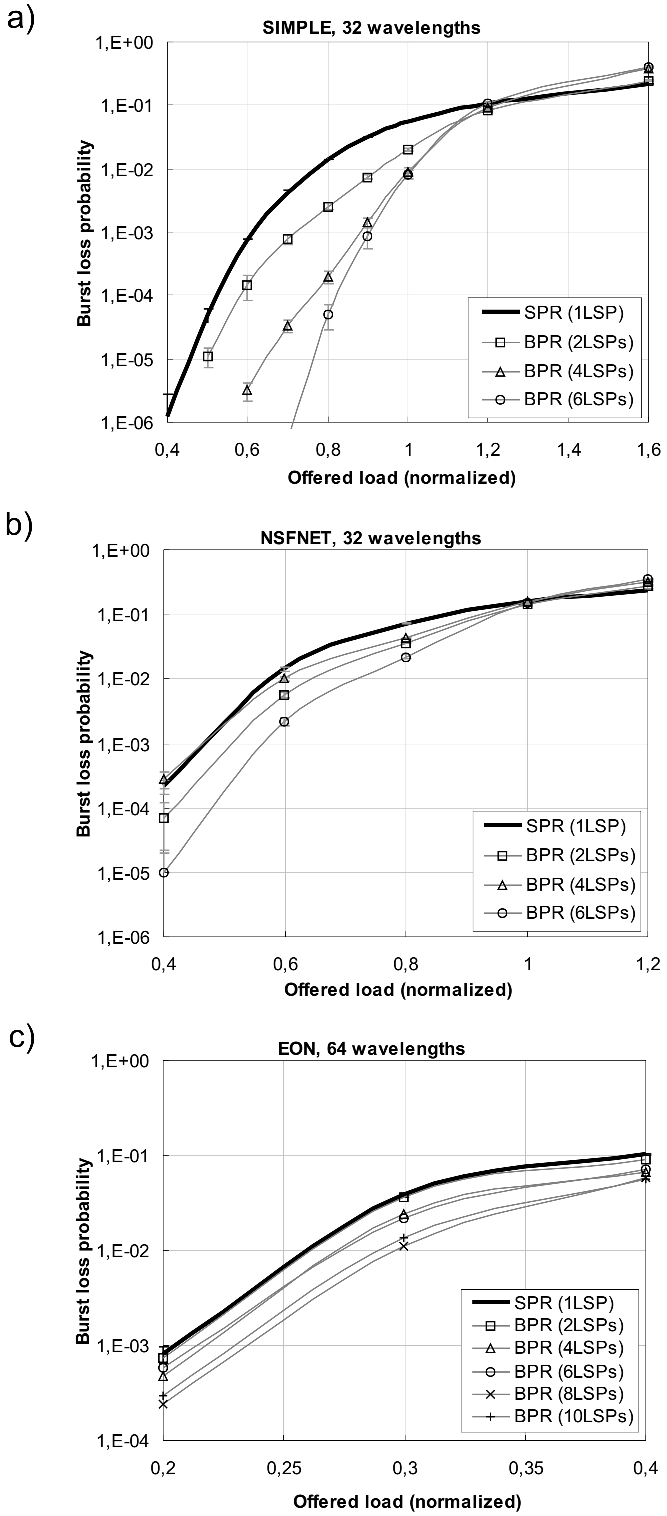


Figure 10.4: Burst loss probability in BPR, a) SIMPLE (32λ), b) NSFNET (32λ), and c) EON (64λ).

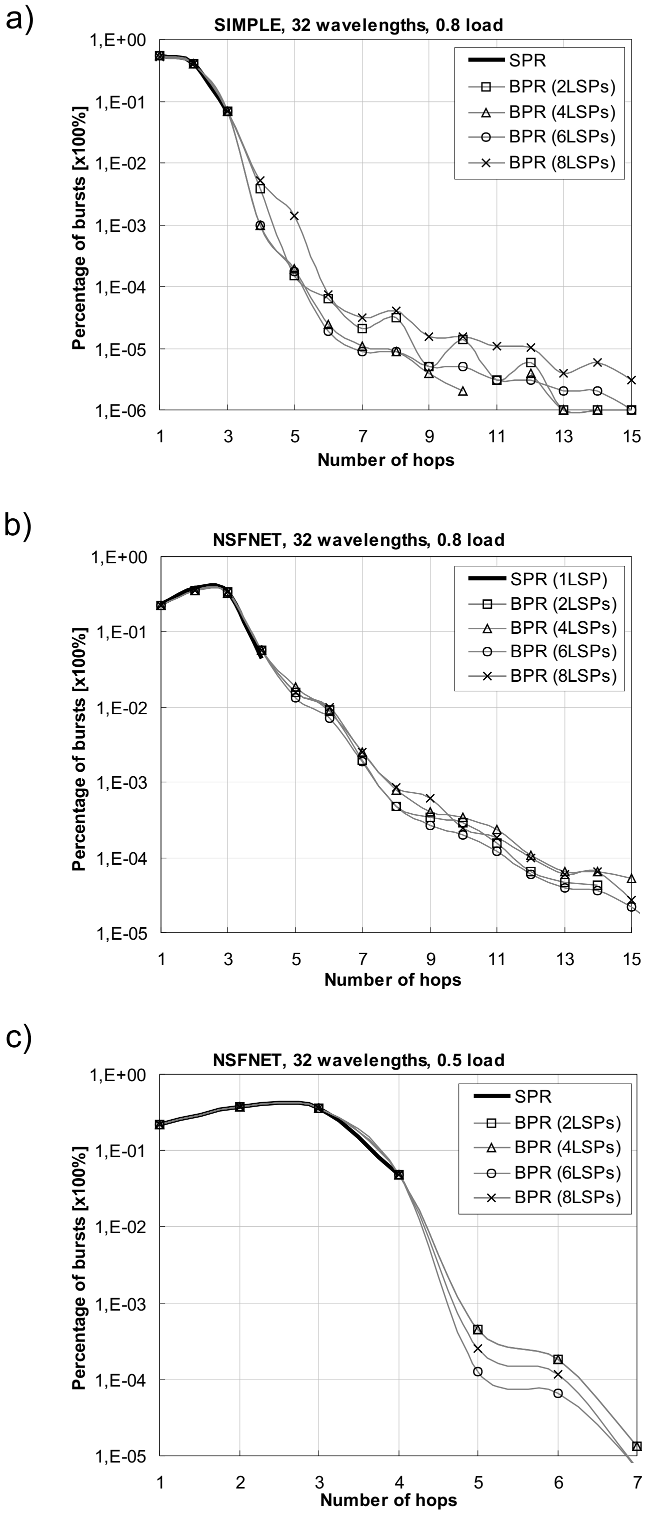


Figure 10.5: Amount of bursts experiencing given number of hops in BPR, a) SIMPLE (32λ, $\rho = 0.8$), b) NSFNET (32λ, $\rho = 0.8$), and c) NSFNET (32λ, $\rho = 0.5$).

node can take a decision according to the state of its own output ports. The route is selected for each burst individually in all nodes. Although the solution is suboptimal, since it only considers local information, still it provides good flexibility as well as no additional signalling is required.

We propose and evaluate two isolated alternative routing algorithms for labelled E-OBS networks, namely the path excluding routing and the bypass routing. The obtained results show that our solutions can help in the burst blocking problem in OBS networks. In particular, BPR can offer a significantly improved performance, with respect to the shortest-path routing, in small and medium-size networks and under low and moderate traffic loads. Although the performance of PER is slightly worsen in such scenarios (comparing to BPR), still, it behaves better under high loads.

An E-OBS architecture gives a special opportunity to BPR as long as there is no restriction imposed by the setup of offset times in the edge node on the length of routing path. Indeed the burst routing path in BPR can be lengthen significantly due to the deflection operation. Thus the application of this routing strategy in C-OBS might be difficult.

Alternative routing strategies introduce the problem of out-of-order burst arrival. Indeed the bursts which are deflected over the paths of different length may arrive to the destination in an unsettled sequence. The BPR algorithm, which introduce an unlimited deflection, is particularly sensitive to this problem. Another important issue is the increase of burst delay; we have already commented that the propagation delay is a dominant delay factor in OBS. For all these reasons BPR might require some additional constraints on the maximum number of deflections allowed. As well, the application of BPR might be reasonable only in low-loaded networks, where the percentage of deflected bursts is small.

In order to support the PER algorithm in the out-of-order burst arrival problem we could try to establish the LSPs of similar lengths. In this way the deflected bursts would experience comparative transmission delays as on the primary paths.

Chapter 11

Optimization of multi-path routing

This Chapter addresses the problem of routing optimization in OBS networks. We use a simplified analytical model of OBS network with overall burst loss probability as the primary metric of interest. An approximated form of the overall burst loss probability, which can be found e.g., in [RVZW03], has a nonlinear character and it may produce some difficulties in formulating an optimization problem. Indeed the routing solutions presented in [ZLW$^+$04], [HN04], or [TR05] use a linear programming (LP) formulation, which either does not consider the overall burst loss probability as a metric of interest or it takes an approximated form of this metric.

In this Chapter we formulate a non-linear optimization problem for multi-path source routing in OBS and we propose two different methods to solve it. First approach is based on a non-reduced link load calculation with strict partial derivatives given. The second method is designed for an OBS network model with a reduced link load calculation. This approach applies a routing optimization framework considered initially for circuit-switched (CS) networks.

In our routing scenario we assume that there is a pre-established virtual path topology consisting of a limited number of paths between each pair of source-destination nodes. We calculate a *traffic splitting vector* that determines the distribution of traffic over these paths. The proposed solution can be used, in particular, for a static (pre-planed) routing, where the traffic distribution is calculated based on a given (long-term) matrix of demands. Then either a periodic or a threshold-triggered update of the splitting vector can be performed if the demand matrix is changed.

11.1 Routing scenario

Consider an OBS network such as that illustrated in Figure 11.1. There are K links, labelled $e = 1, 2, \dots K$, and link e comprises C_e wavelengths. A subset $p \subseteq \{1, 2, \dots K\}$ identifies a path; we define an incidence coefficient α_{ep} such that $\alpha_{ep} = 1$ if link e belongs to path p, and $\alpha_{ep} = 0$ otherwise. In the network there is a set P of paths pre-established between sources (s) and destinations (d). A subset $P_{sd} \subseteq P$ identifies all paths from node s to node d (later $|P|$ indicates the number of paths in set P).

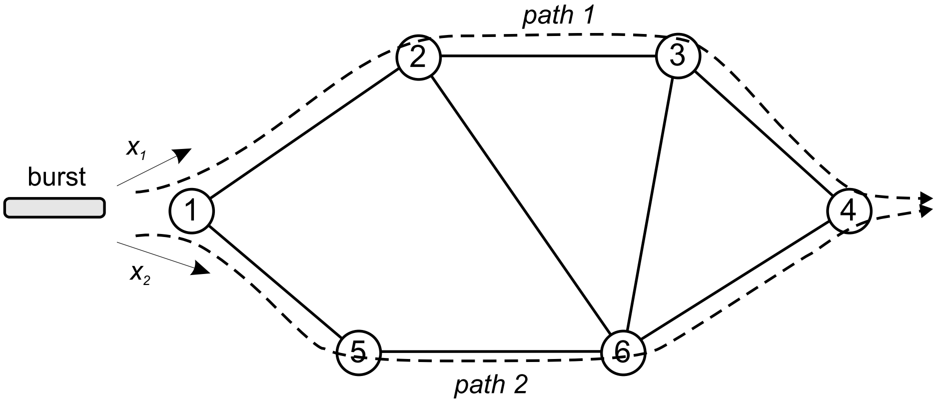


Figure 11.1: Example of OBS network with multi-path source-based routing; x_1 and x_2 are the splitting factors and $x_1 + x_2 = 1$.

We assume that the routing decision is source-based so that the source node determines the path of a burst that enters the network (see Figure 11.1). Moreover, the network applies multi-path routing strategy, i.e., each subset P_{sd} comprises a small number of paths and a burst can take one of those paths. The path selection is performed according to a given splitting factor x_p, such that the sum of x_p of all the paths p belonging to a given subset P_{sd} is equal to 1.

We assume that the nodes are capable to perform a full wavelength conversion according to the random wavelength-selection algorithm. A burst going over a path p is blocked and lost if on a given link k that belongs to p there are no free wavelengths. Otherwise a wavelength in the link is reserved for the burst duration and then released immediately after the burst transmission.

The reservation (holding) periods on each link are i.i.d random variables with the mean equal to the mean burst duration l; for simplicity we assume $l = 1$. The demand traffic pattern is described by matrix $[t_{sd}]$ and bursts destined to a given node d arrive to a node s as a Poisson process of (long-term) rate $t_{sd}/l = t_{sd}$. Let $t_p = t_{sd}$ for each $p \subset P_{sd}$. Thus traffic v_p offered to path p can be calculated as

$$v_p = x_p t_p. \tag{11.1}$$

Here vector $\bar{x} = (x_1, \ldots, x_{|P|})$ determines the distribution of traffic over the network and it should be selected so that to reduce congestion and to improve overall performance.

11.2 Formulation

11.2.1 Loss models of OBS network

A loss model of OBS network based on the *Erlang fixed-point approximation* was proposed by Zukerman in [RVZW03]. In particular, the traffic offered to link e is obtained as a sum of the traffic offered to all the paths that cross this link diminished by the traffic lost in the preceding links along these paths,

$$\rho_e = \sum_{p \in P} \alpha_{ep} v_p \prod_{g=1}^{K} \left(1 - \beta_{pge} E_g\right), \tag{11.2}$$

where β_{pge} equals 1 or 0 depending whether or not link g precedes link e along path p, respectively. We call this model a **reduced link load** (R-LL) model.

The Zukerman formulation may bring some difficulty in the context of computation of partial derivatives (for optimization purposes). Therefore we propose a simplified link load model, later called a **non-reduced link load** (NR-LL) model, where the traffic offered to link e is calculated as a sum of the traffic offered to all the paths that cross this link,

$$\rho_e = \sum_{p \in P} \alpha_{ep} v_p. \tag{11.3}$$

The rationale for this proposal lays behind the fact that under low link losses E_g, as one can expect in a well dimensioned network, model (11.2) can be approximated to (11.3).

The main modelling steps include the calculation of burst loss probabilities in links, paths and entire network, successively.

1. We assume that the offered burst traffic at each link is the aggregation of a large number of independent traffic flows. Hence, the link range dependence within the aggregate traffic will be reduced to zero or to very short range dependent and the traffic arrival at each link in the network can be approximated by the Poisson process [LLGC06]. Then burst loss probabilities E_e in links are given by the Erlang loss formula

$$E_e = E(\rho_e, C_e) = \frac{\rho_e^{C_e}}{C_e!} \left[\sum_{i=0}^{C_e} \frac{\rho_e^i}{i!} \right]^{-1}. \tag{11.4}$$

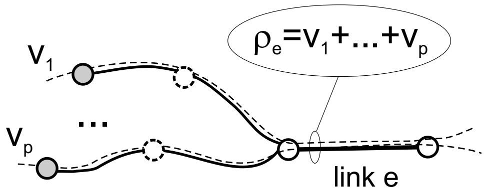
2. Given the difficulty in obtaining an exact path-level blocking formulas we have assumed that each blocking event occurs independently from link to link along any path inside the network. Then loss probabilities L_p of bursts offered to paths are calculated taking into account the losses in each link that is crossed by given path, according to the formula

$$L_p = 1 - \prod_{e=1}^{K} \left(1 - \alpha_{ep} E_e\right). \tag{11.5}$$

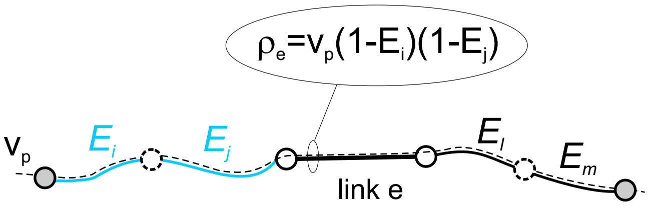
3. The overall burst loss probability B, which is the sum of traffic lost in the network normalized by the traffic offered to the network, is obtained as

$$B = \sum_{p \in P} v_p L_p \left[\sum_{p \in P} v_p \right]^{-1}. \tag{11.6}$$

a) non-reduced OBS link load model



b) reduced OBS link load model



c) reduced CS link load model

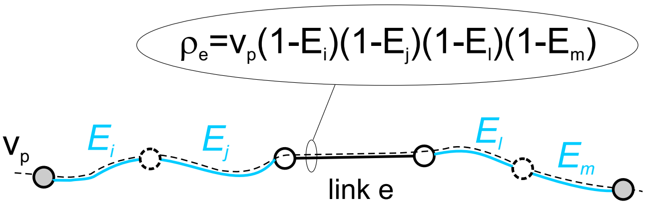


Figure 11.2: Link load models: a) non-reduced OBS, b) reduced OBS, and c) reduced CS.

Later we will also refer to a reduced link load model of the circuit-switching network, so here we introduce it. The only difference between this model and the OBS network ones is the way the link load is calculated. Particularly, the traffic offered to link e is obtained as a sum of the traffic offered to all the paths that cross this link diminished by the traffic lost in both the preceding and the succeeding links along these paths,

$$\begin{aligned}\rho_e &= \sum_{p\in P} \alpha_{ep} v_p \prod_{g=1, g\neq e}^{K} (1 - \alpha_{ep} E_g) \\ &= (1 - E_e)^{-1} \sum_{p\in P} \alpha_{ep} v_p (1 - L_p).\end{aligned} \tag{11.7}$$

The calculation of link, path and overall blocking probabilities is the same as for

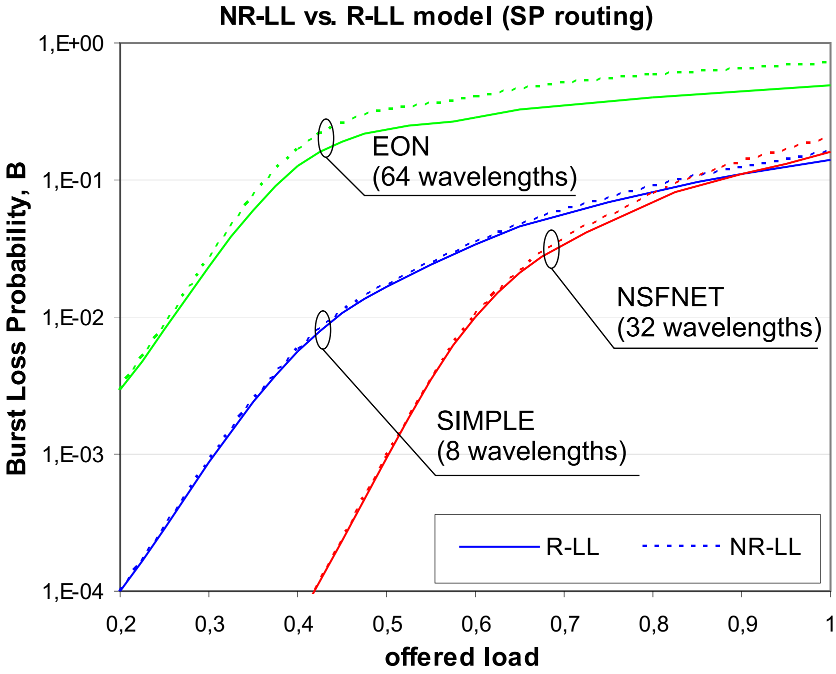


Figure 11.3: Accuracy of NR-LL model, in SIMPLE, NSFNET, and EON topologies, with 8, 32, and 64 wavelengths per link, respectively.

the OBS loss model, using formulae (11.4)-(11.6).

Figures 11.2a-c present illustrative examples of link load calculations in all of the introduced OBS and CS network loss models.

Figure 11.3 compares the overall burst loss probability B of both OBS network loss models, calculated in the function of traffic load, in different network scenarios (see Section 10.1 for more details on the evaluation scenario). The network routing we use here is based on a single-path shortest-hop routing algorithm. We can see that the accuracy of NR-LL model is very strict for B below 10^{-2}.

11.2.2 Optimization problem

From equations (11.1) and (11.6) we define a cost function to be the subject of optimization:

$$B(\bar{x}) = \sum_{p \in P} x_p t_p L_p \tag{11.8}$$

The optimization problem is formulated as follows:

$$\min B(\bar{x}) \tag{11.9}$$

subject to:

$$\sum_{p \in P_{sd}} x_p = 1, \forall P_{sd}, \tag{11.10}$$

$$0 \leq x_p \leq 1, \forall p \in P. \tag{11.11}$$

Since the overall BLP is a non-linear function of vector $\bar{x}$ the cost function is non-linear as well. According to [PW85] for solving such optimization problem we can use for instance the modified reduced gradient method described in [Har76].

11.3 Partial derivatives

Gradient methods need to employ the calculation of partial derivatives of the cost function. The partial derivatives indicate a direction which leads to the reduction in total cost.

11.3.1 NR-LL model

Direct calculation

In NR-LL model the partial derivatives can be derived from formulae (11.3)-(11.6), successively, starting from

- partial derivatives of the offered traffic load function in link e

$$\frac{\partial \rho_e(x)}{\partial x_i} = A_{ei} t_i = \left\{ \begin{array}{ccc} t_i & ,if & \alpha_{ei} = 1 \\ 0 & ,if & \alpha_{ei} = 0 \end{array} \right\}, \tag{11.12}$$

- partial derivatives of the burst loss probability function in link e

$$\frac{\partial E_e(x)}{\partial x_i} = E_e \left(E_e + \frac{C_e - \rho_e}{\rho_e} \right) \frac{\partial \rho_e(x)}{\partial x_i}, \tag{11.13}$$

- partial derivatives of the burst loss probability function in path p

$$\frac{\partial L_p(x)}{\partial x_i} = \sum_{e=1}^{K} \alpha_{ep} \prod_{g=1, g \neq e}^{K} (1 - \alpha_{gp} E_g) \frac{\partial E_e(x)}{\partial x_i}, \tag{11.14}$$

- and finally, partial derivatives of the cost function

$$\frac{\partial B(\bar{x})}{\partial x_i} = t_i L_i + \sum_{p \in P} t_p x_p \frac{\partial L_p(x)}{\partial x_i}. \tag{11.15}$$

Fast calculation

This direct calculation of partial derivatives may be time-consuming. Indeed we should find $|P|$ partial derivatives (11.14) to calculate (11.15); notice that $|P|$ may be really big in larger networks. Instead we provide the following exact derivation, similar to the one proposed by Kelly in [Kel88] for circuit-switching networks.

We shall write $f = f(v; C)$ when we wish to emphasize the functional dependence of a function f on the system parameters $v = (v_p, p \in P)$ and $C = (C_1, \ldots, C_K)$. In summations, products and the definitions of matrices i and p ranges over P, and e or g range over $\{1, 2, \ldots, K\}$.

For each link e define

$$\eta_e = E(\rho_e, C_e - 1) - E(\rho_e, C_e). \tag{11.16}$$

From (11.4) it follows that

$$\frac{d}{d\rho_e} E(\rho_e, C_e) = [1 - E(\rho_e, C_e)]\, \eta_e. \tag{11.17}$$

From this and (11.3)

$$\frac{d}{dv_i} E_e(v; C) = \alpha_{ei}(1 - E_e)\eta_e. \tag{11.18}$$

Define the form

$$B(v; E; C) = \sum_p v_p \left(1 - \prod_g (1 - \alpha_{gp} E_g)\right). \tag{11.19}$$

We find that

$$\frac{\partial}{\partial v_i} B(v; E; C) = 1 - \prod_g (1 - \alpha_{gi} E_g) = L_i \tag{11.20}$$

and

$$\frac{\partial}{\partial E_e} B(v; E; C) = (1 - E_e)^{-1} \sum_p \alpha_{ep} v_p (1 - L_p). \tag{11.21}$$

For each link e define c_e such that

$$c_e = \eta_e \sum_p \alpha_{ep} v_p (1 - L_p). \tag{11.22}$$

From the above

$$\begin{aligned}\frac{\partial}{\partial v_i}B(v;C) &= \frac{d}{dv_i}B(v;E(v;C);C) \qquad (11.23)\\ &= \left[\frac{\partial}{\partial v_i} + \sum_e \frac{d}{dv_i}E_e(v;C)\frac{\partial}{\partial E_e}\right]B(v;E;C)\\ &= L_i + \sum_e \alpha_{ei}(1-E_e)\eta_e\frac{\partial}{\partial E_e}B(v;E;C)\\ &= L_i + \sum_e \alpha_{ei}\eta_e\sum_p \alpha_{ep}v_p(1-L_p)\\ &= L_i + \sum_e \alpha_{ei}c_e.\end{aligned}$$

Finally, since (11.1) we have

$$\frac{d}{dx_i}B(\bar{x}) = t_i\left[L_i + \sum_e \alpha_{ei}c_e\right]. \qquad (11.24)$$

This calculation of partial derivatives is straightforward. Indeed once K unknowns (c_e) are pre-calculated then they can be used in (11.24) to obtain the partial derivatives of B.

11.3.2 R-LL model

Computing of partial derivatives in the R-LL model is even more complex than in the NR-LL model. Therefore in order to find them we take the approach considered by Kelly in [Kel88] for the circuit-switched network model (see (11.7)) and use it as a rough approximation.

Let $c = (c_1, c_2, \ldots, c_K)$ be the (unique) solution to the equation:

$$c_e = \eta_e\,(1-E_e)^{-1}\sum_{p:e\in p} x_p t_p\,(1-L_p)\left(1 - \sum_{g\in p-\{e\}} c_g\right) \qquad (11.25)$$

Then

$$\frac{d}{dx_i}B(\bar{x}) \approx t_i\left[1 - (1-L_i)\left(1 - \sum_{e\in i} c_e\right)\right] \qquad (11.26)$$

Notice that the formula 11.26 corresponds strictly to the CS network case.

11.3.3 Remarks

Although the correctness of our approximation of partial derivatives for R-LL model is not confirmed theoretically, still our numerical results show that these derivatives lead us to an optimal solution of the optimization problem. Some explanation of this fact could be the similarity of both OBS and CS reduced link loss models (see Figure 11.2). Indeed the only difference is that the link load reduction in CS networks is higher by the traffic lost in succeeding links when comparing to OBS networks.

Regarding the NR-LL model, although we are not able to prove that the (not unique) solution is optimal in a global sense, numerical results show that several repetitions of the optimization of (11.8) using formula (11.24) always give us the same (with a finite numerical precision) near-optimal value of B.

In order to get insight into the character of function B in NR-LL model we calculate it for vector $\bar{x}_0(\gamma)$, such that:

$$\bar{x}_0(\gamma) = \gamma \bar{x}_1 + (1 - \gamma)\bar{x}_2 \tag{11.27}$$

where $\bar{x}_1$ and $\bar{x}_2$ are two (different) near-optimal vectors, and $\gamma \in [0, 1]$.

Numerical results show that $B(\bar{x}_0(\gamma))$ is a monotonic function of near-horizontal character.

11.4 Implementation issues

The proposed optimization framework can be used to calculate a traffic splitting vector that determines the distribution of traffic over the network in a multi-path source based routing scenario. We assume that there is a virtual path topology pre-established that comprise, for instance, a limited number of shortest paths between each pair of source-destination nodes. Such virtual topology can be established, e.g., in a labelled OBS network (see Chapter 9).

Centralized routing

Centralized routing optimization can be applied, for instance, in a pre-planed routing, where the traffic distribution is calculated based on a given (long-term) matrix of demands. Then either a periodic or a threshold-triggered update of the splitting vector can be performed if the matrix of demands changes. In principle, both NR-LL model and R-LL model can be used for such a centralized routing optimization. Nevertheless, as long as the accuracy of NR-LL model is very strict at a low burst-loss working point and the calculation of its partial derivatives is straightforward this model is a preferable candidate for centralized routing optimization.

Distributed routing

A distributed routing should react rapidly to a local disturbance at the point of the disturbance, with slower adjustments in the rest of the network. Similarly like it was proposed for circuit-switched networks [Kel88] the R-LL model could potentially be used in a distributed adaptive routing algorithm in OBS networks (some similar study was presented in [LLGC06]).

In such distributed adaptive routing the network should offer the possibility of limited communication between the nodes. The nodes should be capable to measure the loads carried through the links and the source nodes should be able to measure the loads carried on the paths. Moreover, such routing requires a (limited) arithmetical processing ability for each link and route, which may be distributed over the nodes of the network; for example the processing for routes might be carried out at the sources nodes. Then the measurements of actual loads together with computing of partial derivatives could be used to implement a decentralized hill-climbing search procedure able gradually to vary routing patterns in response to changes in the demands placed on the network (as in [Kel88]).

The design of an optimized distributed routing algorithm is left for future study.

11.5 Performance

We evaluate the performance of our optimized multi-path routing in the simulation scenario described in Section 10.1. In order to find a splitting vector $\bar{x}$ that yields to a near-optimal routing we use a solver *fmincon*, for constrained nonlinear multi-variable functions, which is available in the *Matlab* environment. Then we apply this vector in the simulator. The optimized routing, OR-NR and OR-R respectively for NR-LL model and R-LL model, is compared with simple shortest-hop routing (SPR). We consider 2 shortest paths per each source-destination pair of nodes; they are not necessarily disjoint. In SPR only 1 path is available. Uniform traffic matrix as well as exponential burst inter-arrivals and durations are considered.

In Figure 11.4 we show the overall BLP in the function of offered traffic load ρ normalized to the link capacity. We evaluate 3 scenarios of small (SIMPLE), medium (NSFNET) and large (EON) network dimension. We can see that optimized routing achieves significantly lower losses than SPR in each scenario. Moreover, we can also observe that both the OR-NR and the OR-R offer the same routing performance. Finally, we validate that the analytical results (OR-NR (an) in the Figure) calculated from the model match very well the simulation ones (OR-NR (sim)).

11.5.1 Comparison of routing schemes

Having validated the optimized, with NR-LL model, multi-path routing (OR) we compare it with PER and BPR isolated alternative routing strategies proposed in Chapter 10. We consider $k = 2$ LSPs per each pair of source-destination nodes in OR, whilst $k = 2$, or $k = 6$ in the case of alternative routing.

In Figure 11.5 we evaluate the overall BLP performance in the function of offered, normalized traffic load. Our first observation is that with the same number of paths k available and under either low or high load conditions the OR performs better than the corresponding alternative routing algorithms. The fact can be explained by a better global knowledge of the network congestion state in the optimized multi-path routing then in the isolated alternative routing. This knowledge allows to distribute the traffic over the paths that traverse underutilized links of the network, and so it preserves from the use of overloaded links.

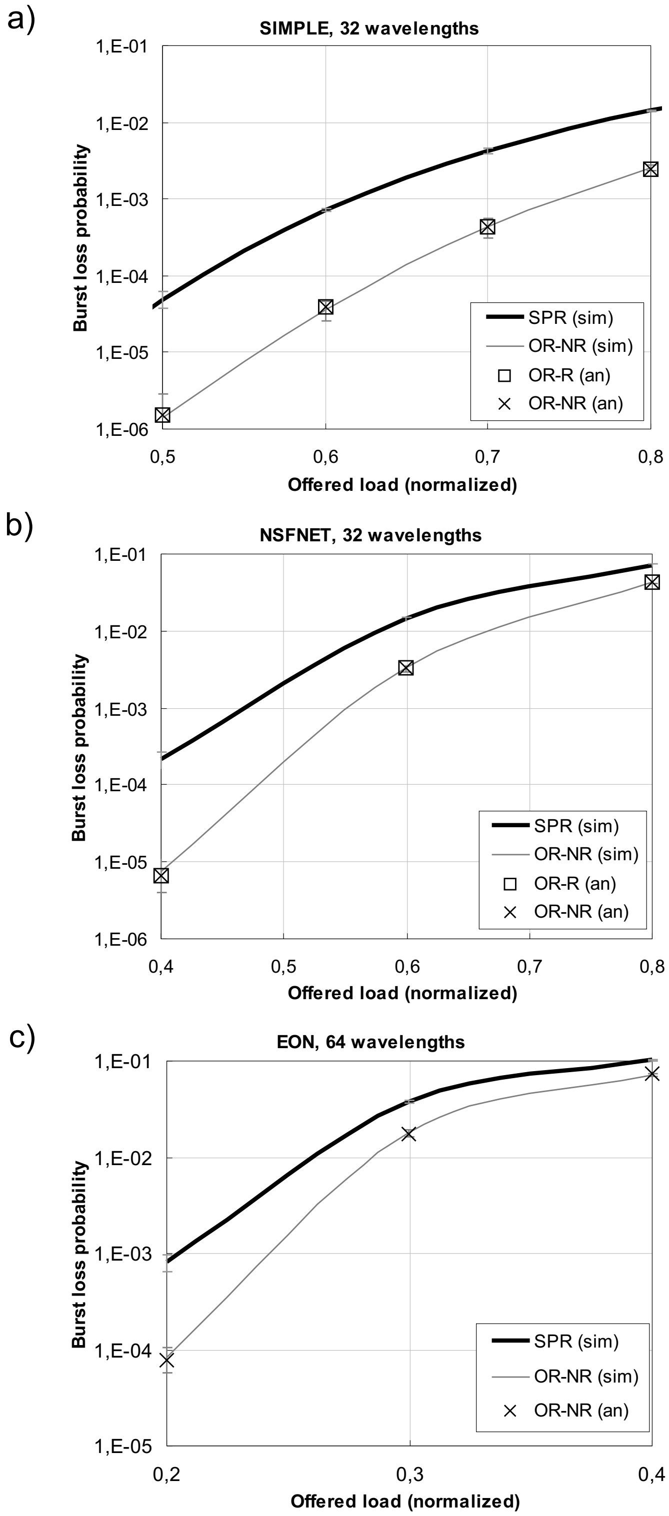


Figure 11.4: Burst loss probability in OR, a) SIMPLE (32λ), b) NSFNET (32λ), and c) EON (64λ).

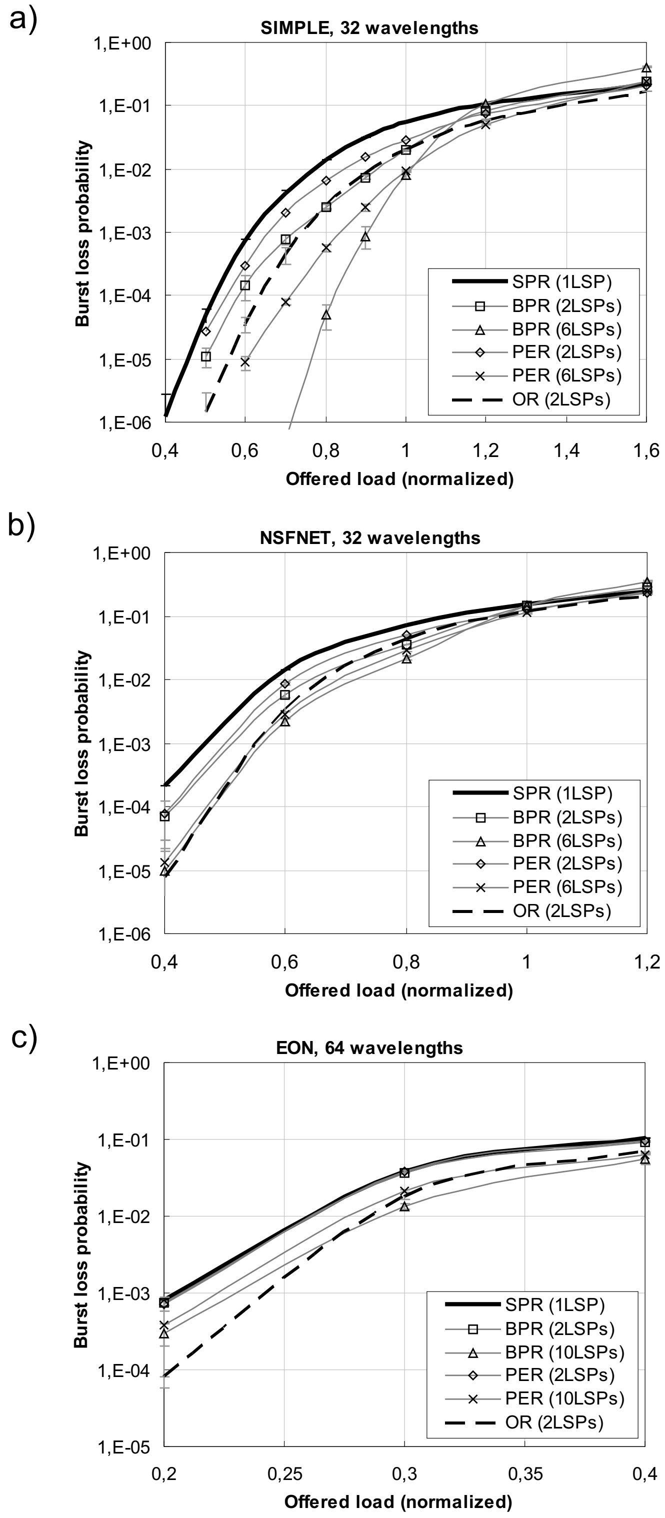


Figure 11.5: Comparison of optimized multipath source routing with isolated alternative routing strategies, a) SIMPLE (32λ), b) NSFNET (32λ), and c) EON (64λ).

In a small network (Figure 11.5a), both alternative routing algorithms can take advantage of their reactive contention resolution feature if the number of LSP they can access is high ($k = 6$). On the contrary, isolated alternative routing might have some difficulty with the reduction of burst blocking in larger networks (Figures 11.5b-c).

11.6 Summary

In this Chapter we propose a non-linear optimization method for multi-path source routing problem in OBS networks. In our proposal we calculate the traffic splitting vector that determines a near-optimal distribution of traffic over routing paths. The formulas for partial derivatives we present are straightforward and very fast in computing; it makes the proposed non-linear optimization method a viable alternative for linear programming formulations.

The simulation results demonstrate that in a static traffic scenario our optimization method effectively distributes the traffic over the network. As a result the network-wide burst loss probability is reduced compared to the shortest path routing. Moreover, the optimized multi-path routing outperforms alternative routing strategies if the same number of routing paths is considered.

Our optimization method can be possibly extended to a distributed routing scenario. The design of a distributed adaptive routing algorithm based on the reduced link load model, and adequate to the one proposed by Kelly for the circuit-switched networks, is left for future study.

PART V

Conclusions and future works

Chapter 12

Conclusions and future works

The tremendous growth of the Internet, together with the fact that it is a packet-based network, is the main drivers to develop a data-centric transport network. In this context the optical burst switching architecture is considered as a promising network solution. The advantage of having small switching granularities is in the conventional OBS architectures counterbalanced by high burst blocking probabilities. Therefore, there is a strong requirement for dedicated hardware and control solutions in order to enable both feasible and effective operation in such networks. In this dissertation we propose an architectural solution for OBS networks and we addresses several issues related to QoS provisioning and routing as well as network modelling and performance evaluation.

E-OBS architecture: characteristics and modelling

We confront the conventional OBS (C-OBS) architecture with the offset time-Emulated OBS (E-OBS) scheme. E-OBS introduces the offset times artificially, by means of additional fiber delay elements used in core nodes. We show that C-OBS architectures possess several drawbacks, such as the problem of unfairness in access to transmission resources, constraints in alternative routing, a need for complex void filling-based resources reservation algorithms, some difficulties in QoS provisioning, etc. On the contrary, the E-OBS can bring significant facilities to the mentioned problems. Taking into account all the arguments provided in this thesis, there is a motivation for recognizing the E-OBS architecture as an efficient and functional alternative to conventional OBS networks.

Since E-OBS architectures need for additional fiber delay elements, we provide a study on their feasibility in relation to other key system parameters. In particular, we address the problem of congestion in the control plane and the resulting insufficient offset effect. In order to approach this issue a queuing model of OBS control plane operation is studied. We give some preliminary results for an exemplary E-OBS system with a single processor performing at the node controller. Depending on the distribution of processing times we model such system either as M/M/1 queue with reneging or as M/D/1/K queue without reneging. The obtained results show that an appropriate setup of burst lengths may effectively limit the congestion in control plane. Moreover for the analyzed node controller of moderate processing times we show that the offset times can be provisioned effectively in core nodes, and still, the performance is preserved.

QoS provisioning

We study the performance of the most addressed mechanisms providing relative QoS differentiation in OBS networks. In particular, we show that the burst preemptive mechanism concurrently achieves efficient resources utilization and offers highly effective QoS differentiation. The offset time differentiation mechanism, which is frequently invoked in literature, provides high HP class performance as well, however its scheduling efficiency, and so the throughput, is aggravated by the variation of offset-times. Finally, the wavelength threshold-based mechanism, which is however able to perform class differentiation, is characterized by the poorest overall performance that significantly depends on its threshold value. The application of this mechanism may be reasonable only in highly dimensioned networks where the wavelength threshold is relatively high (in order to service efficiently the LP traffic) and it adapts to traffic changes.

The high performance of burst preemption mechanism designates it to be a suitable mechanism for QoS differentiation in OBS. Although in this thesis we concern on relative quality guarantees, still, the preemption scheme can be extended to absolute QoS provisioning. Such a study can be found e.g. in [OS06]. There the superiority of a preemptive-based burst dropping mechanism over other mechanisms, also over the scheme with intentional packet dropping, was demonstrated again. The main drawback of the burst preemption mechanism in OBS is the overbooking of resources in case of a successful preemption. Nevertheless, as we have discussed in this thesis, such a problem can be avoided in E-OBS with the preemption window mechanism introduced.

Routing

An E-OBS architecture gives a special opportunity to the alternative routing as long as there is no restriction imposed by the setup of offset times in the edge node on the length of routing path. As a result, the burst can be freely deflected in intermediate nodes with any routing algorithm. In this thesis we propose and evaluate two isolated alternative routing algorithms for labelled E-OBS networks, namely the path excluding routing and the bypass routing. The obtained results show that our solutions can help in the burst blocking problem in OBS networks.

Another routing approach that we address in this thesis is multi-path routing. In this context we propose a novel approach to its optimization. Our proposal is based on the theory of non-linear optimization with a straightforward calculation of partial derivatives. Simulation results demonstrate that the optimized routing effectively reduces the overall burst loss probability with respect to the shortest path routing. Moreover, if we consider the same number of routing paths available, it outperforms the alternative routing as well.

As a final remark, we would like to mention that there is one more great benefit of E-OBS. In particular E-OBS can be seen as an immediate migration step towards OPS. Indeed due to the application of fiber delay elements the operation of E-OBS and OPS are very similar. Still, the differences between both technologies lay in the length of transmission units (burst vs. packet) and signalling mode (out-of-band vs. in-band), and thus, higher hardware and processing requirements of the OPS. Nevertheless, as the progress in optical technologies will continue these dissimilarities should disappear in the future. As a result, the application of E-OBS may facilitate the migration from OBS networks to OPS networks.

Some particular conclusions of this thesis are the following:

- The problem of unfairness in C-OBS networks starts to play role if the length of bursts is short, when comparing to the length of offset times. E-OBS is free of the unfairness problem.
- The burst loss and the delay performance of both C-OBS and E-OBS is (almost) the same.
- In an exemplary E-OBS node with a single-processor controller, a feasible fiber delay coil ($25\mu s$ of delay), and with fast processing times ($T_p = 200ns$), the minimum value of average burst length, which is equal to several $kbytes$, is very close to the one determined by the control-plane stability constraint. Under moderate processing times ($T_p = 1\mu s$), the length of burst is more restricted, and it should be $100kbytes$ at least (instead of $77kbytes$ obtained from the stability constraint).
- Effective throughput of the burst preemption mechanism is higher than of the offset time differentiation mechanism. Under high loads ($\rho = 0.8$) the difference might be even of $2 \div 3\%$.
- The amount of (useless) phantom bursts generated in a single node, which is enhanced with a burst preemption mechanism, is of about 4% in an exemplary system of 16 wavelengths, $\rho = 0.8$, HP load of 30%.
- Application of the preemption window mechanism in an E-OBS node allows to avoid the problem of phantom bursts. In an exemplary OBS system of 32 wavelengths, $\rho = 0.8$, $\alpha = 30\%$, and the preemption window equal to $15\mu s$ ($3km$), the HP burst loss probabilities of about 10^{-5} can be achieved.
- When using isolated alternative routing strategies in highly loaded, small networks, the improvement of burst loss performance can be even of 2 orders of magnitude, comparing to shortest path routing. In larger networks, the improvement is not so high (below 1 order of magnitude).
- Under the same number of paths available, optimized multi-path routing can perform better than isolated alternative routing algorithms.

Concluding, E-OBS was shown to be a functional and feasible alternative for OBS networks, with a support for highly effective QoS provisioning and facilitated routing management.

We can distinguish several issues that would be addressed in future work.

- Firstly, the modelling of multi-processor switch controller architectures, which is desired in the context of system design and dimensioning. Such study would allow to find the trade-offs between performance and complexity/cost of different controller architectures.
- Another topic started in this thesis is optimization of routing in OBS networks. In this context, both distributed multi-path routing and single-path routing strategies, possibly with QoS constraints, will be studied. Since multi-path routing introduces the problem of out-of-order burst arrival this issue has to be addressed as well.
- An important issue which is a hot topic of current research activity is the deployment of control plane in OBS networks. As a solution, we have initiated to consider the generalized MPLS protocol (GMPLS). Adaptation of GMPLS to OBS might be desired, in particular, in the context of the network migration as long as GMPLS is an accepted solution in OCS networks. While GMPLS should facilitate the coexistence of OCS and OBS, the concept of E-OBS enables the migration towards OPS networks. Therefore, the loop can be closed allowing the continuous deployment of ASON, OBS and OPS.

Acronyms

ABT	ATM Block Transfer
ADSL	Asymmetric Digital Subscriber Line
ASON	Automatic Switched Optical Network
ATM	Asynchronous Transfer Mode
BCP	Burst CP
BD-W	Burst Dropping with Wavelength threshold
BLP	Burst Loss Probability
BP	Burst Preemption
BPR	Baypass Path Routing
CC	Control Channel
C-OBS	Conventional OBS
CP	Control Packet
CPU	Control Processor Unit
CS	Circuit Switching
DWDM	Dense WDM
E-OBS	Offset Time Emulated OBS
FDC	Fiber Delay Coil
FDL	Fiber Delay Line
FTTH	Fiber to the Home
GMPLS	Generalized MPLS
HP	High Priority
IP	Internet Protocol
LP	Low Priority Class
LSP	Label Switched Path
MEMS	Micro-Electro-Mechanical Systems
MPLS	Multi-Protocol Label Switching
NG-SDH	Next-Generation SDH
NLP	Non-Linear Programming
NR-LL	Non-Reduced Link Load
OBS	Optical Burst Switching
OCS	Optical Circuit Switching
ODM	Optical Drop Multiplexer
OPS	Optical Packet Switching
OR	Optimized Routing
ORION	Ontario Research and Innovation Optical Network

OR-NR	OR with NR-LL model
OR-R	OR with R-LL model
OT	Offset Time
OTD	Offset Time Differentiation
OXC	Optical Cross-connect
P2P	Pear to Pear
PER	Path Excluding Routing
PW	Preemption Window
QoS	Quality of Service
RAM	Random Access Memory
RED	Random Early Detection
R-LL	Reduced Link Load
RWA	Routing and Wavelength Assignment
SDH	Synchronous Digital Hierarchy
SOA	Semiconductor Optical Amplifier
SONET	Synchronous Optical Networking
SP	Shortest Path
SPR	SP Routing
TAG	Tell-and-Go
TAW	Tell-and-Wait
TCP	Transmission Control Protocol
TE	Traffic Engineering
UDP	User Datagram Protocol
WDM	Wavelength Division Multiplexing
WLAN	Wireless Local Area Network
WR-OBS	Wavelength-Routed OBS
WS	Wavelength Conversion/Wavelength Converter

Bibliography

[ACP04] A. Agusti and C. Cervello-Pastor. A new contentionless dynamic routing protocol for obs using wavelength occupation. In *Proceedings of the 12th IEEE Mediterranean Electrotechnical Conference (MELECON 2004)*, Dubrovnik, Croatia, May 2004.

[AdDA07] C. G. Argos, O. Gonzlez de Dios, and J. Aracil. Adaptive multi-path routing for obs networks. In *Proceedings of the 9th IEEE International Conference on Transparent Optical Networks (ICTON2007)*, Rome, Italy, July 2007.

[AEBS05] A. Agrawal, T. S. El-Bawab, and L. B. Sofman. Comparative account of bandwidth efficiency in optical burst switching and optical circuit switching networks. *Photonic Network Communications*, 9(3):297309, 2005.

[AST$^+$06] A. Al Amin, K. Shimizu, M. Takenaka, T. Tanemura, R. Inohara, and K. Nishimura et al. 40/10 gbps bit-rate transparent burst switching and contention resolving wavelength conversion in an optical router prototype. In *Proceedings of 32nd European Conference on Optical Communications (ECOC 2006)*, Cannes, France, October 2006.

[Bar57] D. Y. Barrer. Queuing with impatient customers and ordered service. *Journal Operations Res. Soc. Amer.*, 5:650–656, 1957.

[BBE$^+$05] I. Baldine, A. Bragg, G. Evans, M. Pratt, M. Singhai, D. Stevenson, and R. Uppalli. Jumpstart deployments in ultra-high-performance optical networking testbeds. *IEEE Optical Communications*, 43(11):S18–S25, November 2005.

[BD06] N. Barakat and T. E. Darcie. Control-plane congestion in obs networks. In *Proceedings of the 4th Workshop on OBS (WOBS)*, San Jose, CA (USA), October 2006.

[BD07] N. Barakat and T. E. Darcie. The control-plane stability constraint in optical burst switching networks. *IEEE Communications Letters*, 11(3):267–269, March 2007.

[BIPe99] D. Blumenthal, T. Ikegami, P. R. Prucnal, and (editors). Special issue on photonic packet switching technologies, techniques, and systems. *IEEE Journal of Lightwave Technology*, 17(12), December 1999.

[BNH$^+$03] S. Bjornstad, M. Nord, D.R. Hjelme, N. Stol, F. Callegati, W. Cerroni, C. Raffaelli, P. Zaffoni, C. M. Gauger, C. Develder, J. Cheyns, E. Van Breusegem, E. Baert, D. Colle, M. Pickavet, P. Demeester, M. Lackovic, D. Careglio, G. Junyent, M. Klinkowski, M. Marciniak, and M. Kowalewski. Optical burst and packet switching: Node and network design, contention resolution and quality of service. In *Proceedings of the 7th IEEE International Conference on Telecommunications (ConTEL2003)*, Zagreb, Croatia, June 2003.

[Boc05] S. Bocquet. Queueing theory with reneging. *Technical Report of DSTO*, September 2005.

[Bou03] N. Boudriga. Optical burst switching protocols for supporting qos and adaptive routing. *Computer Communication Journal*, 26(15):1804–1812, 2003.

[BP03] T. Battestilli and H. Perros. An introduction to optical burst switching. *IEEE Optical Communications*, 41(8):S10–S15, August 2003.

[BS04] N. Barakat and E.H. Sargent. On optimal ingress treatment of delay-sensitive traffic in multi-class obs systems. In *Proceedings of the 3rd International Workshop on Optical Burst Switching (WOBS 2004), co-located with BroadNets 2004*, San Jose, CA (USA), October 2004.

[Cav00] D. Cavendish. Evolution of optical transport technologies: from sonet/sdh to wdm. *IEEE Communications Magazine*, 38:164–172, June 2000.

[CC01] F. Callegati and W. Cerroni. Wavelength allocation algorithms in optical buffers. In *Proceedings of IEEE International Conference on Communications (ICC 2001)*, Helsinki, Finland, June 2001.

[CCRZ04] F. Callegati, W. Cerroni, C. Raffaelli, and P. Zaffoni. Wavelength and time domain exploitation for qos management in optical packet switches. *Computer Networks*, 44(1):569–582, January 2004.

[CCXV99] F. Callegati, H. Cankaya, Y. Xiong, and M. Vandenhoute. Design issues of optical ip routers for internet backbone applications. *IEEE Communications Magazine*, 37(12):124–128, December 1999.

[CDB$^+$03] J. Cheyns, C. Develder, E. Van Breusegem, A. Ackaert, M. Pickavet, and P. Demeester. Routing in an awg-based optical packet switch. *Photonic Network Communications*, 5(1):69–80, 2003.

[CEJ05a] T. Coutelen, H. Elbiaze, and B. Jaumard. An efficient adaptive offset mechanism to reduce burst losses in obs networks. In *Proceedings of IEEE Global Communications Conference (GLOBECOM 2005)*, St. Louis, MO (USA), December 2005.

[CEJ05b] T. Coutelen, H. Elbiaze, and B. Jaumard. An efficient adaptive offset mechanism to reduce burst losses in obs networks. In *Proceedings of Proceedings of IEEE Global Communications Conference (GLOBECOM 2005)*, St. Louis, MO (USA), December 2005.

[CEJ05c] T. Coutelen, H. Elbiaze, and B. Jaumard. Performance comparison of ocs and obs switching paradigms. In *Proceedings of the 7th IEEE International Conference on Transparent Optical Networks (ICTON2005)*, , Barcelona, Spain, July 2005.

[CGK92] I. Chlamtac, A. Ganz, and G. Karmi. Lightpath communications: An approach to high-bandwidth optical wans. *IEEE Transactions on Communications*, 40(7):1171–1182, July 1992.

[Chi95] D. Chiaroni. A novel photonic architecture for high capacity atm switch applications. In *Proceedings of the Photonics in Switching conference*, Salt Lake City, UT, April 1995.

[Chi01] D. Chiaroni. Status and applications of optical packet switching. In *Proceedings of 27th European Conference on Optical Communications (ECOC 2001)*, Amsterdam, Netherlands, October 2001.

[CLP02] P.B Chu, S.-S. Lee, and S. Park. Mems: The path to large optical crossconnects. *IEEE Communications Magazine*, 40(3):80–87, March 2002.

[CMC06] Q. Chen, G. Mohan, and K.C. Chua. Route optimization for efficient failure recovery in optical burst switched networks. In *Proceedings of IEEE High Performance Switching and Routing workshop (HPSR 2006)*, Poznan, Poland, June 2006.

[CR06] M. Casoni and C. Raffaelli. Analytical framework for end-to-end design of optical burst-switched networks. *Optical switching and Networking*, (4):33–43, August 2006.

[CVG$^+$98] A. Carena, M. D. Vaughn, R. Gaudino, M. Shell, and D. J. Blumenthal. Opera: An optical packet experimental router architecture with label swapping capability. *IEEE Journal of Lightwave Technology*, 16(12), December 1998.

[CWXQ03] Y. Chen, H. Wu, D. Xu, and C. Qiao. Performance analysis of optical burst switched node with deflection routing. In *Proceedings of IEEE International Conference on Communication (ICC 2003)*, Seattle, USA, May 2003.

[CZZ04] C. Cameron, A. Zalesky, and M. Zukerman. Shortest path prioritized random deflection routing (sp-prdr) in optical burst switched networks. In *Proceedings of ICST International Workshop on Optical Burst Switching (WOBS)*, San Jose, USA, 2004.

[DG01] K. Dolzer and C.M. Gauger. On burst assembly in optical burst switching networks - a performance evaluation of just-enough-time. In *Proceedings of the 17th International Teletraffic Congress (ITC 17)*, Salvador, Brazil, December 2001.

[DHL+03] H. J. S. Dorren, M. T. Hill, Y. Liu, N. Calabretta, A. Srivatsa, F. M. Huijskens, H. de Waardt, and G. D. Khoe. Optical packet switching and buffering by using all-optical signal processing methods. *Journal of Lightwave Technology*, 21(1):2–12, January 2003.

[DKKB00] M. Duser, E. Kozlovski, R.I. Killey, and P. Bayvel. Design tradeoffs in optical burst switched networks with dynamic wavelength allocation. In *Proceedings of 26th European Conference on Optical Communications (ECOC 2000)*, Munich, Germany, September 2000.

[DPZQ06] Y. Du, T. Pu, H. Zhang, and Y. Quo. Adaptive load balancing routing algorithm for optical burst-switching networks. In *Proceedings of Optical Fiber Communication Conference (OFC 2006)*, Anaheim, CL (USA), March 2006.

[ELP03] V. Eramo, M. Listanti, and P. Pacifici. A comparison study on the number of wavelength converters needed in synchronous and asynchronous all-optical switching architectures. *Journal of Lightwave Technology*, 21(2):340–355, February 2003.

[Fib07a] Compact time delay coils, 2007. http://www.newport.com/.

[Fib07b] Fiber delay coils, 2007. http://www.generalphotonics.com/.

[FJ03a] F. Farahmand and J. Jue. Look-ahead window contention resolution in optical burst switched networks. In *Proceedings of IEEE High Performance Switching and Routing workshop (HPSR 2003)*, Torino, Italy, June 2003.

[FJ03b] F. Farahmand and J. P. Jue. Supporting qos with look-ahead window contention resolution in optical burst switched network. In *Proceedings of IEEE Global Communications Conference (GLOBECOM 2003)*, San Francisco, CA (USA), December 2003.

[G.700] ITU-T Recommendation G.707. Network node interface for the synchronous digital hierarchy (sdh), October 2000.

[Gau02] C. Gauger. Dimensioning of fdl buffers for optical burst switching nodes. In *Proceedings of the 6th IFIP Working Conference on Optical Network Design and Modelling (ONDM 2002)*, Torino, Italy, February 2002.

[Gau03] C. Gauger. Trends in optical burst switching. In *Proceedings of the SPIE/ITCOM 2003*, volume 5247, pages 115–125, Orlando, FL (USA), September 2003.

[GBIQ04] S. Ganguly, S. Bhatnagar, R. Izmailov, and C. Qiao. Multi-path adaptive optical burst forwarding. In *Proceedings of IEEE High Performance Switching and Routing workshop (HPSR 2004)*, Phoenix, AR (USA), April 2004.

[GDW03] N. Ghani, S. Dixit, and T.-S. Wang. On ip-wdm integration: a retrospective. *IEEE Communications Magazine*, 41(9):42–45, September 2003.

[GKS04] C. Gauger, M. Kohn, and J. Scharf. Performance of contention resolution strategies in obs network scenarios. In *Proceedings of the 3rd International Conference on the Optical Internet (COIN2004)*, Yokohama, Japan, July 2004.

[GKZM05] C. Gauger, M. Kohn, J. Zhang, and B. Mukherjee. Network performance of optical burst/packet switching: The impact of dimensioning, routing and contention resolution. In *ITG-Fachtagung Photonic Networks*, Leipzig, Germany, May 2005.

[Gun07] S. Gunreben. Multi-layer analysis to quantify the impact of optical burst reordering on tcp performance. In *Proceedings of the 9th IEEE International Conference on Transparent Optical Networks (ICTON2007)*, Rome, Italy, July 2007.

[GWL$^+$05] H. Guo, J. Wu, X. Liu, J. Lin, and Y. Ji. Multi-qos traffic transmission experiments on obs network testbed. In *Proceedings of 31nd European Conference on Optical Communications (ECOC 2005)*, Glasgow, Scotland, September 2005.

[GZ06] D. Gao and H. Zhang. Information sharing based optimal routing for optical burst switching (obs) network. In *Proceedings of Optical Fiber Communication Conference (OFC 2006)*, Anaheim, CL (USA), March 2006.

[HAM$^+$05] Pan H., T. Abe, Y. Mori, Ch. Young-Bok, and H. Okada. Feedback-based load balancing routing for optical burst switching networks. In *Proceedings of the 11th Asia-Pacific Conference on Communications (APCC 2005)*, Perth, Western Australia, October 2005.

[Har76] R.J Harris. The modified reduced gradient method for optimally dimensioning telephone networks. *Australian Telecom. Research*, 10(1):30–35, 1976.

[HCA98] D. K. Hunter, M. C. Chia, and I. Andonovic. Buffering in optical packet switches. *Journal of Lightwave Technology*, 16(12):2081–2094, December 1998.

[HHM05] Y. Huang, J.P. Heritage, and B. Mukherjee. Dynamic routing with preplanned congestion avoidance for survivable optical burst-switched (obs) networks. In *Proceedings of Optical Fiber Communication Conference (OFC 2005)*, Anaheim, CL (USA), March 2005.

[HLH02a] C. Hsu, T. Liu, and N. Huang. Performance analysis of deflection routing in optical burst-switched networks. In *Proceedings of the 21st Joint Conference of IEEE Computer and Communications Societies (INFOCOM 2002)*, New York, NY (USA), June 2002.

[HLH02b] C. F. Hsu, T. L. Liu, and N. F. Huang. Performance analysis of deflection routing in optical burst-switched networks. In *Proceedings of the 21st Conference of the IEEE Communications Society (INFOCOM 2002)*, New York, NY (USA), June 2002.

[HN04] E. Hyytia and L. Nieminen. Linear program formulation for routing problem in obs networks. In *Proceedings of the 9th IEEE Symposium on Computers and Communications (ISCC 2004)*, Alexandria, Egypt, June - July 2004.

[HSE04] H.Mouftah, S.Said, and H. Elbiaze. A qos-based restoration mechanism for obs networks. In *Proceedings of the 8th IEEE International Conference on Transparent Optical Networks (ICTON 2006)*, volume 3, pages 1853–1863, Hong Kong, China, March 2004.

[HTM06] Y. Hirota, H. Tode, and K. Murakami. Cooperation method considering wavelength assignment and routing problem in optical burst switched networks. In *Proceedings of Optical Fiber Communication Conference (OFC 2006)*, Anaheim, CL (USA), March 2006.

[I.300] ITU-T Recommendation I.371. Traffic control and congestion control in b-isdn, March 2000.

[IA01] M. Izal and J. Aracil. On the influence of self-similarity on optical burst switching traffic. In *Proceedings of IEEE Global Communications Conference (GLOBECOM 2001)*, San Antonio, TX (USA), November 2001.

[IYS05] D. Ishii, N. Yamanaka, and I. Sasase. Self-learning route selection scheme using multipath searching packets in an obs network. *Journal of Optical Networking*, 4(7):432–445, 2005.

[JECA03] J.Cheyns, E.V.Breusegem, C.Develder, and A.Ackaert. Performance improvement of an internally-blocking optical/packet switch. In *Proceedings of IEEE International Conference on Communications (ICC 2003)*, Anchorage, AK (USA), May 2003.

[JQX00] M. Jeong, C. Qiao, and Y. Xiong. Reliable wdm multicast in optical burst-switched networks. In *Proceedings of Optical Networking and Communications Conference (Opticomm 2000)*, Dallas, TX (USA), October 2000.

[JXC$^+$00] M. Jeong, Y. Xiong, H. C. Cankaya, M. Vandenhoute, and C. Qiao. Efficient multicast schemes for optical burst-switched wdm networks. In *Proceedings of IEEE International Conference on Communication (ICC 2000)*, New Orleans, LA (USA), June 2000.

[KA98] E. Karasan and E. Ayanoglu. Performance of wdm transport networks. *IEEE Journal on Selected Areas in Communications*, 16(9):1081–1096, September 1998.

[KA03] A. Kaheel and H. Alnuweiri. A strict priority scheme for quality-of service provisioning in optical burst switching networks. In *Proceedings of IEEE Symposium on Computers and Communications (ISCC 2003)*, Turkey, June 2003.

[KA04] A. Kaheel and H. Alnuweiri. Quantitative qos guarantees in labeled optical burst switching networks. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM 2004)*, Dallas, TX, November 2004.

[Kar02] S. V. Kartalopoulos. *DWDM: Networks, Devices, and Technology.* John Wiley & Sons, October 2002.

[KB02] E. Kozlovski and P. Bayvel. Qos performance of wr-obs network architecture with request scheduling. In *Proceedings of IFIP 6th Working Conference on Optical Networks Design and Modelling (ONDM 2002)*, Turin, Italy, February 2002.

[KCK04] J. Kim, J. Choi, and M. Kang. Offset-time based scheduling algorithm for burst control packet in optical burst switching networks. In *Proceedings of (ICOIN)*, Busan, Korea, February 2004.

[KCM04] B. C. Kim, Y. Z. Cho, and D. Montgomery. An efficient optical burst switching technique for multi-hop networks. *IEICE Transactions on Communications*, E87-B(6):1737–1740, June 2004.

[KCMSP06] M. Klinkowski, D. Careglio, D. Morato, and J. Sole-Pareta. Effective burst preemption in obs network. In *Proceedings of IEEE High Performance Switching and Routing workshop (HPSR 2006)*, Poznan, Poland, June 2006.

[KCSSP05] M. Klinkowski, D. Careglio, S. Spadaro, and J. Solé-Pareta. Impact of burst length differentiation on qos performance in obs networks. In *Proceedings of the 7th IEEE International Conference on Transparent Optical Networks (ICTON 2005)*, Barcelona, Spain, July 2005.

[Kel88] F. P. Kelly. Routing in circuit-switched networks: Optimization, shadow prices and decentralization. *Advanced Applied Probability*, 20:112–144, 1988.

[KG03a] M. Khn and C. Gauger. Contention-based limited deflection routing in obs networks. In *Proceedings of Global Telecommunications Conference (GLOBECOM 2003)*, San Francisco, CL (USA), December 2003.

[KG03b] M. Khn and C. Gauger. Dimensioning of sdh/wdm multilayer networks. In *Proceedings of the 4th ITG Workshop on Photonic Networks*, Leipzig, Germany, 2003.

[KHCSP05] M. Klinkowski, F. Herrero, D. Careglio, and J. Sol-Pareta. Adaptive routing algorithms for optical packet switching networks. In *Proceedings of the 9th IFIP Working Conference on Optical Networks Design and Modelling (ONDM 2005)*, Milan, Italy, February 2005.

[KKK02] S. Kim, N. Kim, and M. Kang. Contention resolution for optical burst switching networks using alternative routing. In *Proceedings of IEEE International Conference on Communications (ICC 2002)*, New York, NY (USA), April-May 2002.

[KM01] M. Klinkowski and M. Marciniak. Development of ip/wdm optical networks. In *Proceedings of 3th International Workshop on Laser and Fiber optic Network Modeling (LFNM 2001)*, Kharkov, Ukraine, May 2001.

[KMM+05] K. Kitayama, T. Miki, T. Morioka, H. Tsushima, M. Koga, K. Mori, S. Araki, K. Sato, H. Onaka, S. Namiki, and T. Aoyama. Photonic network r&d activities in japan-current activities and future perspectives. *IEEE Journal of Lightwave Technology*, 23(10):2404–3418, October 2005.

[LES00] M. Listanti, V. Eramo, and R. Sabella. Architectural and technological issues for future optical internet networks. *IEEE Communications Magazine*, 38(9):82–92, September 2000.

[LKSG03] S. Lee, H. Kim, J. Song, and D. Griffith. A study on deflection routing in optical burst-switched networks. *Photonic Network Communications*, 6(1):51–59, 2003.

[LLGC06] J. Lu, Y. Liu, M. Gurusamy, and K.C. Chua. Gradient projection based multi-path traffic routing in optical burst switching networks. In *Proceedings of IEEE High Performance Switching and Routing workshop (HPSR 2006)*, Poznan, Poland, June 2006.

[LMC05] J. Li, G. Mohan, and Kee Chaing Chua. Dynamic load balancing in ip-over-wdm optical burst switching networks. *Computer Networks*, 47(3):393–408, 2005.

[LQXX04] J. Li, C. Qiao, J. Xu, and D. Xu. Maximizing throughput for optical burst switching networks. In *Proceedings of the 23rd Conference of the IEEE Communications Society (INFOCOM 2004)*, volume 3, pages 1853–1863, Hong Kong, China, March 2004.

[LQYG06] Xin Liu, Chunming Qiao, Xiang Yu, and Weibo Gong. A fair packet-level performance comparison of obs and ocs. In *Proceedings of Optical Fiber Communication (OFC 2006)*, Anaheim, CL (USA), March 2006.

[LTWW94] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic. *IEEE/ACM Transactions on Networking*, 2(1):1–15, 1994.

[LY06] J. Li and K. L. Yeung. Burst cloning with load balancing. In *Proceedings of Optical Fiber Communication Conference (OFC 2006)*, Anaheim, CL (USA), March 2006.

[LYH$^+$06] K. Long, X. Yang, S. Huang, Q. Chen, and R. Wang. Adaptive parameter deflection routing to resolve contentions in obs networks. In *Proceedings of the 5th International Conference on Networking (Networking 2006)*, Coimbra, Portugal, May 2006.

[MGK$^+$04] I. De Miguel, J. C. Gonzalez, T. Koonen, R. Duran, P. Fernandez, and I. T. Monroy. Polymorphic architectures for optical networks and their seamless evolution towards next generation networks. *Photonic Network Communications*, 8(2):177–189, 2004.

[MRZ04] G. Muretto, C. Raffaelli, and P. Zaffoni. Effective implementation of void filling in obs networks with service differentiation. In *Proceedings of the 3rd International Workshop on Optical Burst Switching (co-located with BroadNets 2004)*, San Jose, CA (USA), October 2004.

[Nor03] M. Nord. Node design in optical packet and optical burst switching. In *Proceedings of 5th International Conference on Transparent Optical Networks (ICTON 2003)*, Warsaw, Poland, June-July 2003.

[Nsf] Nsfnet-the national science foundation network. http://moat.nlanr.net/.

[NTL+05] K. Nashimoto, N. Tanaka, M LaBuda, D. Ritums, J. Dawley, M. Raj, D. Kudzuma, and T. Vo. High-speed plzt optical switches for burst and packet switching. In *Proceedings of the 2nd International Conference on Broadband Networks (Broadnets 2005)*, Boston, MA (USA), October 2005.

[OA05] N. Ogino and N. Arahata. A decentralized optical bursts routing based on adaptive load splitting into pre-calculated multiple paths. *IEICE Transactions on Communications*, E88-B(12):4507–4516, 2005.

[Odl04] A. Odlyzko. Pricing and architecture of the internet: Historical perspectives from telecommunications and transportation. In *Proceedings of the 32nd Research Conference on Communication, Information and Internet Policy (TPRC 2004)*, Arlington, Virginia, October 2004.

[OS06] Harald Overby and Norvald Stol. Qos differentiation in asynchronous bufferless optical packet switched networks. *Wireless Networks*, 12(3):383–394, June 2006.

[OTYC05] L. Ou, X. Tan, H. Yao, and W. Cheng. A study on dynamic load balanced routing techniques in time-slotted optical burst switched networks. In *Proceedings of the 3rd International Conference on Networking and Mobile Computing, (ICCNMC 2005)*, Zhangjiajie, China, August 2005.

[PMP07] J. M. Pedro, P. Monteiro, and J. J. O. Pires. Efficient multi-path routing for optical burst-switched networks. In *Proceedings of Conference on Telecommunications (ConfTele)*, Peniche , Portugal, May 2007.

[PPP03] G. I. Papadimitriou, Ch. Papazoglou, and A. S. Pomportsis. Optical switching: Switch fabrics, techniques, and architectures. *IEEE Journal of Lightwave Technology*, 21(2):384–405, February 2003.

[PPP06] G. I. Papadimitriou, Ch. Papazoglou, and A. S. Pomportsis. *Optical Switching*. Wiley-Interscience, December 2006.

[Pro05] FP6-506760 IP Nobel Project. Deliverable d16 - preliminary definition of burst/packet network and node architectures and solutions, March 2005.

[PSPCK07] P. Pedroso, J. Sol-Pareta, D. Careglio, and M. Klinkowski. Integrating gmpls in the obs networks control plane. In *Proceedings of the*

9th IEEE International Conference on Transparent Optical Networks (ICTON2007), Rome, Italy, July 2007.

[PW85] M. Pioro and B. Wallstrom. Multihour optimization of non-hierarchical circuit switched communication networks with sequential routing. In *Proceedings of the 11th International Teletraffic Congress (ITC-11)*, 1985.

[Qia00] C. Qiao. Labeled optical burst switching for ip-over-wdm integration. *IEEE Communications Magazine*, 38(9):104–114, September 2000.

[QY99] C. Qiao and M. Yoo. Optical burst switching (obs) - a new paradigm for an optical internet. *Journal of High Speed Networks*, 8(1):69–84, March 1999.

[RFL05] J. Rodrigues, M.M. Freire Freire, and P. Lorenz. Performance implications of meshing degree for optical burst switched networks using one-way resource reservation protocols. *Telecommunication Systems*, 30(1-3):35–47, November 2005.

[RG04] M. De Vega Rodrigo and J. Gotz. An analytical study of optical burst switching aggregation strategies. In *Proceedings of IEEE/SPIE Third International Workshop on Optical Burst Switching (WOBS 2004)*, San Jose, CA (USA), October 2004.

[RI03] B. Mikac R. Inkret, A. Kuchar. Advanced infrastructure for photonic networks. extended final report of cost action 266. Technical Report ISBN 953-184-064-4, Published by the Faculty of Electical Engineering and Computing, University of Zagreb, Croatia, September 2003. Available also at http://www.ufe.cz/dpt240/cost266/index.html under item 'Reports'.

[Ros01] E. Rosen. Multiprotocol label switching architecture. IETF RFC 3031, January 2001.

[RVZW03] Z. Rosberg, H. L. Vu, M. Zukerman, and J. White. Blocking probabilities of optical burst switching networks based on reduced load fixed point approximations. In *Proceedings of IEEE INFOCOM 2003*, New York, NY (USA), March-April 2003.

[SC05] J. M. Smith and F. Cruz. The buffer allocation problem for general finite buffer queueing networks. *IIE Transactions*, 37(4):343–366, September 2005.

[SPG05] M. Schlosser, E. Patzak, and P. Gelpke. Impact of deflection routing on tcp performance in optical burst switching networks. In *Proceedings of the 7th IEEE International Conference on Transparent Optical Networks (ICTON2005)*, Barcelona, Spain, July 2005.

[Tan88] A. S. Tanenbaum. *Computer Networks (2nd ed.)*. Prentice Hall, 1988.

[TR05] J. Teng and G. N. Rouskas. Traffic engineering approach to path selection in optical burst switching networks. *Journal of Optical Networking*, 4(11):759–777, 2005.

[Tur99] J.S. Turner. Terabit burst switching. *Journal of High Speed Networks*, 8(1):3–16, March 1999.

[TVJ03] G.P. Thodime, V.M. Vokkarane, and J.P. Jue. Dynamic congestion-based load balanced routing in optical burst-switched networks. In *Proceedings of IEEE Global Communications Conference (GLOBECOM 2003)*, San Francisco, CA (USA), November 2003.

[TZ99] R. S. Tucker and W. De Zhong. Photonic packet switching: An overview. *IEICE Trans. commun.*, E82-B(2):254–264, February 1999.

[VGPMGH$^+$07] J. Veiga-Gontan, P. Pavon-Marino, J. Garcia-Haro, M. Rodelgo, C. Lopez-Bravo, and F.J. Gonzalez-Castano. Network processors: a practical approach for achieving wire-speed packet processing in the emerging optical backbone networks. In *Proceedings of the 9th IEEE International Conference on Transparent Optical Networks (ICTON2007)*, Rome, Italy, July 2007.

[VJ02a] V. Vokkarane and J. Jue. Burst segmentation: an approach for reducing packet loss in optical burst switched networks. In *Proceedings of IEEE International Conference on Communications (ICC 2002)*, New York, NY (USA), April-May 2002.

[VJ02b] V. Vokkarane and J. Jue. Prioritized routing and burst segmentation for qos in optical burst-switched networks. In *Proceedings of Optical Fiber Communication Conference (OFC 2002)*, Anaheim, CA (USA), March 2002.

[VJ03] V. M. Vokkarane and J. P. Jue. Prioritized burst segmentation and composite burst-assembly techniques for qos support in optical burst-switched networks. *IEEE Journal on Selected Areas in Communications (JSAC)*, 21(7):1198–1209, September 2003.

[VKM$^+$01] M. Veeraraghavan, R. Karry, T. Moors, M. Karol, and R. Grobler. Architectures and protocols that enable new applications on optical networks. *IEEE Communications Magazine*, 39(3):118–127, March 2001.

[VS97] E. Varvarigos and V. Sharma. The ready-to-go virtual-circuit protocol: A loss-free protocol for multigigabit networks using fifo buffers. *IEEE/ACM Transactions on Networking*, 5(5):705–718, 1997.

[Whi83] W. Whitt. The queueing network analyzer. *Bell Systems Technical Journal*, 62(9):2779–2815, November 1983.

[Wid95] I. Widjaja. Performance analysis of burst admission-control protocols. *IEEE Proceeding of Communications*, 142:7–14, 1995.

[WM00] J. Y. Wei and R. I. McFarland. Just-in-time signaling for wdm optical burst switching networks. *IEEE Journal of Lightwave Technology*, 18(12):2019–2037, December 2000.

[WMA02] X. Wang, H. Morikawa, and T. Aoyama. Burst optical deflection routing protocol for wavelength routing wdm networks. *Optical Networks Magazine*, 3(6), November/December 2002.

[WR04] Y. Wang and B. Ramamurthy. Cpq: A control packet queuing optical burst switching protocol for supporting qos. In *Proceedings of 3rd International Workshop on Optical Burst Switching (WOBS 2004), co-located with BroadNets 2004*, San Jose, CA (USA), October 2004.

[WZSZ03] J. Wan, Y. Zhou, X. Sun, and M. Zhang. Guaranteeing quality of service in optical burst switching networks based on dynamic wavelength routing. *Optics Communications*, 220(1-3):85–95, May 2003.

[WZV02] J. White, M. Zukerman, and H. L. Vu. A framework for optical burst switching network design. *IEEE Communications Letters*, 6(6):268–270, June 2002.

[XPR01] L. Xu, H.G. Perros, and G. Rouskas. Techniques for optical packet switching and optical burst switching. *IEEE Communications Magazine*, 39(1):136–142, January 2001.

[XQLX03] J. Xu, C. Qiao, Jikai Li, and G. Xu. Efficient channel scheduling algorithms in optical burst switched networks. In *Proceedings of IEEE INFOCOM 2003*, New York, NY (USA), March-April 2003.

[XTGE$^+$04] Y. Xin, J. Teng, G.Karmous-Edwards, G. Rouskas, and D. Stevenson. Fault management with fast restoration for optical burst switched networks. In *Proceedings of the 1st Annual International Conference on Broadband Networks (BROADNETS 2004)*, San Jose, CA (USA), October 2004.

[XVC99] Y. Xiong, M. Vandenhoute, and H. Cankaya. Design and analysis of optical burst-switched networks. In *Proceedings of SPIE99 Conference on All Optical Networking*, Boston, MA (USA), September 1999.

[XVC00] Y. Xiong, M. Vanderhoute, and C. Cankaya. Control architecture in optical burst-switched wdm networks. *IEEE Journal of Selected Areas in Communications*, 18(10):1838–1851, October 2000.

[YJJ03] L. Yang, Y. Jiang, and S. Jiang. A probabilistic preemptive scheme for providing service differentiation in obs networks. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM 2003)*, Singapore, December 2003.

[YLC$^+$04] X. Yu, J. Li, X. Cao, Y. Chen, and C. Qiao. Traffic statistics and performance evaluation in optical burst switched networks. *IEEE Journal of Lightwave Technology*, 22(12):2722–2738, December 2004.

[YMY01] S. Yao, B. Mukherjee, and S.J.B. Yoo. A comparison study between slotted and unslotted all-optical packet-switched network with priority-based routing. In *Proceedings of OFC 2001*, Anaheim, CA (USA), March 2001.

[YQ97] M. Yoo and C. Qiao. Just-enough-time (jet): a high speed protocol for bursty traffic in optical networks. In *Proceedings of IEEE/LEOS Technologies for a Global Information Infrastructure (IEEE 1997)*, Montreal, Canada, August 1997.

[YQ98] M. Yoo and C. Qiao. New optical burst switching (obs) protocol for supporting quality of service. *SPIE Proceedings, All Optical Networking: Architecture, Control and Management Issues*, 3531:396–405, November 1998.

[YQD01] M. Yoo, C. Qiao, and S. Dixit. Optical burst switching for service differentiation in the next-generation optical internet. *IEEE Communications Magazine*, 39(2):98–104, February 2001.

[YR06a] L. Yang and G. N. Rouskas. Adaptive path selection in optical burst switched networks. *IEEE/OSA Journal of Lightwave Technology*, 24(8):3002–3011, August 2006.

[YR06b] Li Yang and George N. Rouskas. A framework for absolute qos guarantees in optical burst switched networks. In *Proceedings of IEEE Broadnets 2006*, San Jose, CA (USA), October 2006.

[ZCC$^+$04] P. Zaffoni, F. Callegati, W. Cerroni, G. Muretto, and C. Raffaelli. Qos routing in dwdm optical packet networks. In *Proceedings of WQoSR2004 co-located with QoFIS 2004*, Barcelona, Spain, September 2004.

[ZJM00] H. Zang, J. Jue, and B. Mukherjee. A review of routing and wavelength assignment approaches for wavelength-routed optical wdm networks. *Optical Networks Magazine*, 1(1):47–60, January 2000.

[ZLW$^+$04] J. Zhang, H. J. Lee, S. Wang, X. Qiu, K. Zhu, Y. Huang, D. Datta, Y. C. Kim, and B.Mukherjee. Explicit routing for traffic engineering in labeled optical burst-switched wdm networks. In *Proceedings of*

International Conference on Computational Science (ICCS 2004), Krakow, Poland, June 2004.

[ZVJC04] Q. Zhang, V. M. Vokkarane, J. P. Jue, and B. Chen. Absolute qos differentiation in optical burst-switched networks. *IEEE Journal on Selected Areas in Communications*, 22(9):1781–1795, November 2004.

[ZVR+04] A. Zalesky, H.L. Vu, Z. Rosberg, E.W.M. Wong, and M. Zukerman. Modelling and performance evaluation of optical burst switched networks with deflection routing and wavelength reservation. In *Proceedings of the 23rd Joint Conference of IEEE Computer and Communications Societies (INFOCOM 2004)*, Hong Kong, China, March 2004.

[ZWZ+04] J. Zhang, S. Wang, K. Zhu, D. Datta, Y.-C. Kim, and B. Mukherjee. Pre-planned global rerouting for fault management in labeled optical burst-switched wdm networks. In *Proceedings of Global Telecommunications Conference (GLOBECOM 2004)*, Dallas, TX (USA), December 2004.

Appendix A

Related publications

A.1 Papers

1. **M. Klinkowski**, M. Pioro, D. Careglio, M. Marciniak and J. Sole-Pareta, “Nonlinear Optimization for Multipath Source-Routing in OBS Networks”, IEEE Communications Letters, vol. 11, no. 12, December 2007.
2. **M. Klinkowski**, M. Pioro and M. Marciniak, ”Optimization of routing in optical burst switching networks: a multi-path routing approach”, Chapter of the COST 293 book (being edited).
3. **M. Klinkowski**, D. Careglio and J. Solé-Pareta, ”Performance Overview of QoS Mechanisms for OBS”, Chapter of the book ”Current Research Progress of Optical Networks” (being edited).
4. **M. Klinkowski**, D.Careglio and J. Solé-Pareta, “Modelling of Control Plane in OBS Networks”, in Proceedings of the 9th IEEE International Conference on Transparent Optical Networks (ICTON2007), Rome, Italy, July 2007.
5. **M. Klinkowski**, M. Pioro, D.Careglio, M. Marciniak and J. Solé-Pareta, “Routing Optimization in Optical Burst Switching Networks”, in Proceedings of the 11th Conference on Optical Network Design and Modelling (ONDM2007), Athens, Greece, May 2007.
6. O. González de Dios, **M. Klinkowski**, C. García Argos, D. Careglio, J. Solé-Pareta, “Performance Analysis of Routing Algorithms for Optical Burst Switching”, in Proceedings of the 11th Conference on Optical Network Design and Modelling (ONDM2007), Athens, Greece, May 2007.
7. **M. Klinkowski**, M. Marciniak, D.Careglio and J. Solé-Pareta, ”Evaluation of Quality of Service Mechanisms in Optical Burst Switched Networks”, the 4rd Workshop on Optimization of Optical Networks (OON2007), Montreal, Canada, May 2007.

8. **M. Klinkowski**, M. Pioro, D.Careglio, M. Marciniak and J. Solé-Pareta, “Routing Optimization in OBS networks”, COST 293 GRAAL and COST 295 DYNAMO Discussion Workshop, Maribor, Slovenia, January/February 2007.

9. J. Aracil, N. Akar, S. Bjørnstad, M. Casoni, K. Christodoulopoulos, D. Careglio, J. Fdez-Palacios, C. Gauger, O. Gonzalez de Dios, G. Hu, E. Karasan, **M. Klinkowski**, D. Morato, R. Nejabati, H. Øverby, C. Raffaelli, D. Simeonidou, N. Stol, G. Tosi-Beleffi and K. Vlachos, ”Research in Optical Burst Switching within the e-Photon/ONe Network of Excellence”, Elsevier Optical Switching and Networking (OSN) journal, vol. 4, no. 1, pp. 1-19, February 2007.

10. **M. Klinkowski**, D. Careglio and J. Solé-Pareta, “Offset Time Emulated OBS Control Architecture”, in Proceedings of the 32nd European Conference on Optical Communication (ECOC2006), Cannes, France, September 2006.

11. **M. Klinkowski**, D. Careglio, M. Marciniak and J. Solé-Pareta, “Comparative Study of QoS Mechanisms in OBS Networks” , in Proceedings of the 11th European Conference on Networks and Optical Communications (NOC2006), Berlin, Germany, July 2006.

12. **M. Klinkowski**, D. Careglio and J. Solé-Pareta, “Comparison of Conventional and Offset Time-Emulated Optical Burst Switching Architectures”, in Proceedings of the 8th IEEE International Conference on Transparent Optical Networks (ICTON2006), Nottingham, UK, June 2006.

13. Javier Aracil and S. Bjornstad, M. Casoni, K. Christodoulopoulos, J. Fdez-Palacios, C. Gauger, O. Gonzalez, G. Hu, E. Karasan, **M. Klinkowski**, D. Morato, R. Nejabati, H. Overby, C. Raffaelli, D. Simeonidou, J. Sole-Pareta, N. Stol, G.M. Tosi Beleffi and K. Vlachos, “The Research Agenda in Optical Burst Switching in e-Photon/ONe (Tutorial)”, in Proceedings of the 8th IEEE International Conference on Transparent Optical Networks (ICTON2006), Nottingham, UK, June 2006.

14. **M. Klinkowski**, D. Careglio, D. Morató and J. Solé-Pareta, “Effective Burst Preemption in OBS Network”, in Proceedings of IEEE International Workshop on High Performance Switching and Routing (HPSR 2006), Poznan, Poland, June 2006.

15. F. Callegati, J. Aracil, L. Wosinska, N. Andriolli, D. Careglio, **M. Klinkowski**, A. Giorgetti, J. Fdez-Palacios, C. Gauger, O. Gonzáles de Dios, G. Hu, E. Karasan, F. Matera, H. Overby, C. Raffaelli, L. Rea, N. Sengezer, M. Tornatore and K. Vlachos, “Research on Optical Core Networks in the e-Photon/ONe Network of Excellence”, in Proceedings of IEEE INFOCOM High Speed Networking Workshop: The Terabits Challenge (co-located with INFOCOM2006), Barcelona, Spain, April 2006.

16. E. Hortas, D. Careglio, **M. Klinkowski** and J. Solé-Pareta, “Análisis de Prestaciones de Algoritmos de Encaminamiento Adaptativos para la Conmutación de

Ráfagas Ópticas", XV Jornadas Telecom I+D, Madrid / Barcelona / Valencia, Spain, November 2005.

17. **M. Klinkowski**, D. Careglio, E. Hortas and J. Solé-Pareta, "Performance Analysis of Isolated Adaptive Routing Algorithms in OBS Networks", in Proceedings of e-Photon/ONe Summer School workshop, Rimini, Italy, August 2005.

18. **M. Klinkowski**, D. Careglio, S. Spadaro and J. Solé-Pareta, "Impact of Burst Length Differentiation on QoS Performance in OBS Networks", in Proceedings of the 7th IEEE International Conference on Transparent Optical Networks (ICTON2005), Barcelona, Spain, July 2005.

19. **M. Klinkowski**, D. Careglio and J. Solé-Pareta, "Wavelength vs Burst vs Packet Switching: Comparison of Optical Network Models", in Proceedings of e-Photon/ONe Winter School workshop, Aveiro, Portugal, February 2005.

20. **M. Klinkowski**, F. Herrero, D. Careglio and J. Solé-Pareta, "Adaptive Routing Algorithms for Optical Packet Switching Networks", in Proceedings of 9th IFIP Working Conference on Optical Network Design and Modelling (ONDM2005), Milan, Italy, February 2005.

21. F. Herrero, D. Careglio, J. Solé-Pareta and **M. Klinkowski**, "Algoritmos de Enrutamiento para la Conmutación de Paquetes Ópticos", XIV Jornadas Telecom I+D, Madrid/Barcelona/Valencia, November 2004.

22. **M. Klinkowski**, D. Careglio, X. Masip-Bruin, S. Spadaro, S. Sanchez-López and J. Solé-Pareta, "A Simulation Study of Combined Routing and Contention Resolution Algorithms in Connection-Oriented OPS Network Scenario", in Proceedings of the 6th IEEE International Conference on Transparent Optical Networks (ICTON2004), Wroclaw, Poland, July 2004.

23. S. Bjornstad, C. M. Gauger, M. Nord, E. Baert, F. Callegati, D. Careglio, **M. Klinkowski**, M. Marciniak, J. Solé-Pareta et al, "Optical Burst Switching and Optical Packet Switching". Chapter 4 of Book Advanced Infrastructure for Photonic Networks – Extended Final Report of COST Action 266, pp. 115-154. R. Inkret et al., Editorial Faculty of Electrical, Engineering and Computing, University of Zagreb, 2003, ISBN 953-184-064-4.

24. S. Bjornstad, M. Nord, D.R. Hjelme, N. Stol, F. Callegati, W. Cerroni, C. Raffaelli, P. Zaffoni, C. M. Gauger, C. Develder, J. Cheyns, E. Van Breusegem, E. Baert, D. Colle, M. Pickavet, P. Demeester, M. Lackovic, D. Careglio, G. Junyent, **M. Klinkowski**, M. Marciniak and M. Kowalewski, "Optical Burst and Packet Switching: Node and Network Design, Contention Resolution and Quality of Service", in Proceedings of the 7th IEEE International Conference on Telecommunications (ConTEL2003), Zagreb, Croatia, June 2003.

A.2 Papers under submission

1. **M. Klinkowski**, D. Careglio and J. Solé-Pareta, “Reactive and Proactive Routing in Labelled OBS Networks”, submitted to IET Journal.
2. **M. Klinkowski**, D. Careglio, Daniel Morató and J. Solé-Pareta, “Preemption Window for Burst Differentiation in OBS”, submitted to OFC 2008 Conference.

A.3 Contribution to European projects

1. Deliverable D3.1, “Architectures and preliminary definition of specific CP and MP functions for hybrid opto-electronic burst/packet networks”, FP6-027305 IP Nobel 2 Project, October 2006.
2. Deliverable D32, “Preliminary report on feasibility studies on opto-electronic burst/packet switching nodes”, FP6-506760 IP Nobel Project, January 2006.
3. Deliverable D23, “Definition of hybrid opto-electronic burst/packet switching node structures and related management functions”, FP6-506760 IP Nobel Project, September 2005.
4. Deliverable D16, “Preliminary definition of burst/packet network and node architectures and solutions”, FP6-506760 IP Nobel Project, March 2005.
5. M. Klinkowski, M. Marciniak et al, “Optical Packet Router with QoS Capabilities: Introductory Study of Computer Simulator Design”. In Progress Report of COST 266 Action, June 2002.

A.4 Other publications

1. P. Pedroso, J. Solé-Pareta, D. Careglio and **M. Klinkowski**, “Integrating GMPLS in the OBS Networks Control Plane”, in Proceedings of the 9th IEEE International Conference on Transparent Optical Networks (ICTON2007), Rome, Italy, July 2007.
2. E. Bonada, F. Callegati, D. Careglio, W. Cerroni, **M. Klinkowski**, G. Muretto, C. Raffaelli and J. Solé-Pareta, “SCWS Technique for QoS Support in Connection-Oriented Optical Packet Switching Network”, in Proceedings of the 8th IEEE International Conference on Transparent Optical Networks (ICTON2006), Nottingham, UK, June 2006.
3. D. Careglio, J. Solé-Pareta, **M. Klinkowski** and S. Spadaro, “Modelling and Optimisation of the IST DAVID Metro Networks”, in Proceedings of the 10th European Conference on Networks and Optical Communications (NOC2005), invited paper, London, UK, July 2005.

4. **M. Klinkowski**, D. Careglio, M. Marciniak and J. Solé-Pareta, “Performance Analysis of the Simple Prioritized Buffering Algorithm in Optical Packet Switch for DiffServ Assured Forwarding”, in Proceedings of the 5th IEEE International Conference on Transparent Optical Networks (ICTON2003), Warsaw, Poland, June 2003.

5. **M. Klinkowski** and M. Marciniak, “Services Differentiation in MPLS Photonic Packet Networks”, in Proceedings of the 7th IFIP Working Conference on Optical Network Design and Modeling (ONDM2003) - COST266 IST OPTIMIST Workshop, Budapest, Hungary, February 2003.

6. **M. Klinkowski** and M. Marciniak, “Optical Packet Switching: Approach to Performance Modeling and Simulation”, in Proceedings of the 27th triennial General Assembly of the International Union of Radio Science (URSI2002), Maastricht (NL), August 2002.

7. M.Marciniak and **M.Klinkowski**, “Advanced Optical Infrastructure for the Emerging Optical Internet Services”, in Proceedings of the 3th International Conference on Advances in Infrastructure for e-Business, e-Education, e-Science, and e-Medicine on the Internet (SSGRR2002), L’Aquila, Italy, July/August 2002.

8. **M.Klinkowski** and M.Marciniak, “Optical Packet Router with QoS Capabilities Introductory Study of Computer Simulator Design”, in Proceedings of the 4th IEEE International Conference on Transparent Optical Networks (ICTON2002), Warsaw, Poland, April 2002.

9. **M. Klinkowski** and M. Marciniak, “QoS Guarantees in IP Optical Networks using MPLS/ MPLambdaS”, in Proceedings of the 3rd IEEE International Conference on Transparent Optical Networks (ICTON2001), Cracow, Poland, June 2001.

10. **M. Klinkowski** and M. Marciniak, “Development of IP/WDM Optical Networks”, in Proceedings of the 3th International Workshop on Laser and Fiber optic Network Modeling (LFNM2001), Kharkov, Ukraine, May 2001.

11. **M. Klinkowski** and M. Marciniak, “IP over Optical Network: Strategy of Deployment”, Journal of Telecommunications and Information Technology, NIT, no. 2, pp. 51-56, April 2001.